



UNIVERSIDADE ESTADUAL DE CAMPINAS

Faculdade de Engenharia Química

CARLOS EDUARDO LIMA DE OLIVEIRA

DETERMINAÇÃO DE GRANDEZAS TERMODINÂMICAS DE COMPOSTOS
TERPÊNICOS PRESENTES EM ÓLEOS ESSENCIAIS POR CROMATOGRAFIA GASOSA
E PROPOSIÇÃO DE MODELO DE CONTRIBUIÇÃO DE GRUPOS PARA ESTIMAÇÃO
DE ENTALPIAS DE VAPORIZAÇÃO

CAMPINAS
2017

CARLOS EDUARDO LIMA DE OLIVEIRA

DETERMINAÇÃO DE GRANDEZAS TERMODINÂMICAS DE COMPOSTOS
TERPÊNICOS PRESENTES EM ÓLEOS ESSENCIAIS POR CROMATOGRAFIA GASOSA
E PROPOSIÇÃO DE MODELO DE CONTRIBUIÇÃO DE GRUPOS PARA ESTIMAÇÃO
DE ENTALPIAS DE VAPORIZAÇÃO

*Tese apresentada à Faculdade de Engenharia
Química da Universidade Estadual de
Campinas como parte dos requisitos exigidos
para a obtenção do título de Doutor em
Engenharia Química.*

Orientador: PROF. DR. MARCO AURÉLIO CREMASCO

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL TESE
DEFENDIDA PELO ALUNO CARLOS EDUARDO LIMA DE
OLIVEIRA, E ORIENTADA PELO PROF. DR. MARCO
AURÉLIO CREMASCO.

CAMPINAS
2017

Agência(s) de fomento e nº(s) de processo(s): CNPq, 145500/2012-8; FUNCAMP
ORCID: <http://orcid.org/0000-0003-0035-7192>

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Luciana Pietrosanto Milla - CRB 8/8129

Oliveira, Carlos Eduardo Lima de, 1985-
OL4d Determinação de grandezas termodinâmicas de compostos terpênicos presentes em óleos essenciais por cromatografia gasosa e proposição de modelo de contribuição de grupos para estimação de entalpias de vaporização / Carlos Eduardo Lima de Oliveira. – Campinas, SP : [s.n.], 2017.

Orientador: Marco Aurélio Cremasco.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Química.

1. Terpenos. 2. Óleos essenciais. 3. Termodinâmica. 4. Cromatografia gasosa. 5. Python (Linguagem de programação de computador). I. Cremasco, Marco Aurélio, 1962-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Química. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Determination of thermodynamic properties of terpene compounds present in essential oils by gas chromatography and proposition of group contribution method for estimation of vaporization enthalpies

Palavras-chave em inglês:

Terpenes
Essential oils
Thermodynamic
Gas chromatography
Python (Computer programming language)

Área de concentração: Engenharia de Processos

Titulação: Doutor em Engenharia Química

Banca examinadora:

Marco Aurélio Cremasco [Orientador]
Silvana Mattedi e Silva
Marlus Pinheiro Rolemberg
Roberta Ceriane
Edson Tomaz

Data de defesa: 10-04-2017

Programa de Pós-Graduação: Engenharia Química

A Banca Examinadora composta pelos membros abaixo aprovou esta tese:

Prof. Dr. Marco Aurélio Cremasco

Instituição: Universidade Estadual de Campinas (UNICAMP)

Profa. Dra. Silvana Mattedi e Silva

Instituição: Universidade Federal da Bahia (UFBA)

Prof. Dr. Marlus Pinheiro Rolemberg

Instituição: Universidade Federal de Alfenas (UNIFAL)

Profa. Dra. Roberta Ceriani

Instituição: Universidade Estadual de Campinas (UNICAMP)

Prof. Dr. Edson Tomaz

Instituição: Universidade Estadual de Campinas (UNICAMP)

A Ata de Defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do aluno.

*Aos meus pais, Eurides e Lauro (in memoriam),
sinônimos de luta e de sabedoria, e aos meus
familiares e amigos.*

Agradecimentos

Nenhum grande trabalho pode ser feito sem a colaboração de outras pessoas. Com este trabalho, que requer muita dedicação, além de abrir mão de desejos de curto prazo, não foi diferente. Por isso, expresso meus sinceros agradecimentos,

Ao nosso senhor Deus, em primeiro lugar, a quem sempre recorri quando as dificuldades apareciam. Obrigado, Senhor, pelo apoio incondicional!

Aos meus familiares os quais sempre me incentivaram e acreditaram na minha capacidade e no meu potencial.

Ao meu orientador, professor Dr. Marco A. Cremasco, que abriu as portas da Universidade Estadual de Campinas (UNICAMP) e da Faculdade de Engenharia Química (FEQ) para mim e que, ao longo deste trabalho, acreditou em mim e me deu a oportunidade de desenvolver um trabalho através do qual possibilitou meu crescimento tanto profissional como pessoal.

Ao professor Dr. Edson Tomaz, do LPDTA/DEPro/FEQ/UNICAMP, desta faculdade, que gentilmente reservou um pouco do seu tempo para me ouvir e cedeu suas instalações e todo suporte para que eu pudesse realizar as análises cromatográficas.

A professora Dra. Roberta Ceriani, desta faculdade, que muito solícitamente me recebeu em sua sala para tirar minhas dúvidas sobre modelos de contribuição de grupos e cujas conversas foram fundamentais para desenvolver minhas ideias.

Aos professores da FEQ/UNICAMP que participaram da minha trajetória e foram os grandes responsáveis pelo meu progresso pessoal e profissional.

A Unicamp pela estrutura e a todos os funcionários da FEQ porque sem eles eu não teria um ambiente adequado para desenvolver a pesquisa. Em especial a Celso Camargo por estar disposto a ajudar sempre que precisei.

A Gláucia T. Corrêa, em especial, por estar comigo neste momento singular e por compartilhar comigo seus sonhos e suas alegrias.

Aos meus amigos que estiveram comigo desde o início da minha trajetória na Unicamp: Hugo Valença, Raphael Ribeiro, Ivanei Pinheiro, Vicente Franco, Bruno (Tefê), João Marques, Audirene A. Santana, Loyrival Paixão, Diogo Moraes, Adilto Cunha, Clairon Pinheiro, Jean Moreira, Ana Carolina Nascimento, Murilo Ferrari, Rhelvis Oliveira, Alex Ramos, Eriksen Miyasak Rafael Scatena, Victor Backer e, em especial, Helder Moura e Rodrigo Amaral, pela grande parceria que tivemos. Obrigado, amigos, pelas conversas, pelas brincadeiras e pela troca de conhecimentos.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de doutorado e à Fundação de Desenvolvimento da Unicamp (FUNCAMP) pelo apoio financeiro na reta final deste trabalho.

Meu muito obrigado a todos!

Carlos Eduardo Lima de Oliveira

Campinas, 10 de abril de 2017.

*Acorde! Passarinho não deve nada
a ninguém e já amanhece voando.*

L. A. de Oliveira

Resumo

Terpenos são os principais componentes de óleos essenciais e apresentam-se em várias classes orgânicas, e são substâncias de grande importância para as indústrias químicas, farmacêuticas e alimentícias. Para que possam ser utilizadas nos vários campos de aplicação, é extremamente necessário o conhecimento de suas propriedades termodinâmicas. A obtenção de dados confiáveis dessas grandezas, tanto por meio experimental como por predição, tem sido desafio constante para a comunidade científica.

A presente Tese de Doutorado tem como objetivo a obtenção de dados de pressão de vapor (P_{vap}), entalpia de vaporização (ΔH_{vap}) e temperatura normal de ebulição (T_{nb}) de compostos terpênicos existentes em óleos essenciais usando a técnica de cromatografia gasosa (CG), além de propor um modelo para predição de ΔH_{vap} , visando melhor desempenho de predição para terpenos de natureza bicíclica.

A determinação de dados termodinâmicos usando CG envolve o uso do índice de retenção cromatográfica em conjunto com equações do tipo Kirchhoff–Rankin. Já o modelo de predição de ΔH_{vap} proposto utiliza conceitos de contribuição de grupos e foi obtido pela técnica de regressão linear múltipla e a otimização foi feita por gradiente descendente, implementado em Python 2 (Versão 2.7.11), que é uma ferramenta de programação orientada a objetos e totalmente gratuita.

Os experimentos de cromatografia gasosa foram realizados com intervalo de temperatura definido conforme a natureza dos compostos estudados. Ao todo, foram avaliados 25 compostos terpênicos, divididos conforme sua volatilidade e analisados em três misturas distintas. Além disso, nove hidrocarbonetos alifáticos (ramificados e de cadeia aberta e cíclica) foram incluídos na análise CG para dar suporte à discussão e avaliar a sensibilidade do método em volatilidades diferentes.

Os dados de P_{vap} e ΔH_{vap} foram determinados a 298,15 K. Os resultados obtidos ficaram em concordância com a literatura, apresentando pequenos desvios relativos médios (d_M). Grande

parte dos compostos com dados de P_{vap} disponíveis na literatura teve d_M entre 0,28 e 7,56 %, e apenas 5 compostos com d_M na faixa de 11,73 a 16,27 %. Quanto aos dados de ΔH_{vap} , apenas o isosafrol apresentou desvio relativo médio elevado de 15,06 %. Para os outros compostos o d_M ficou na faixa de 0,5 – 8,33 %. Os dados de T_{nb} apresentaram desvios ainda menores, ficando entre 0,02 – 3,15 %, demonstrando que a técnica CG pode ser uma boa ferramenta para obtenção das grandezas termodinâmicas estudadas.

Em relação ao modelo preditivo proposto, o desempenho de predição foi comparado com três modelos da literatura: Joback e Reid (1987), Chickos *et al.* (1998) e Kolská *et al.* (2005). Tanto o modelo proposto neste trabalho como os três modelos da literatura avaliados foram usados para predição das entalpias de vaporização dos 25 terpenos analisados no presente trabalho por CG e de outros 30 dados de terpenos da literatura. Para o primeiro conjunto de dados, o modelo proposto apresentou desvio relativo médio (d_M) de 9,87 %, ao passo que o de Joback e Reid (1987) apresentou $d_M = 15,72$ %, o de Kolská *et al.* (2005), $d_M = 20,50$ % e o Chickos *et al.* (1998), $d_M = 14,31$ %. Com os 30 dados da literatura, o modelo deste trabalho apresentou $d_M = 9,51$ %, enquanto os modelos de Joback e Reid (1987), Chickos *et al.* (1998) e Kolská *et al.* (2005) mostraram, respectivamente, d_M igual a 19,21 %, 13,85 % e 15,79 %.

Com respeito aos compostos bicíclicos, as predições de ΔH_{vap} feitas pelo modelo proposto, tanto para os compostos analisados por CG neste trabalho quanto com os dados da literatura, foram melhores que o modelo de Joback e Reid (1987), e apresentou resultados em concordância com os modelos de Chickos *et al.* (1998) e Kolská *et al.* (2005). No entanto, esses dois últimos modelos apresentaram desvio relativo (d_R) elevados para alguns bicíclicos, como por exemplo $d_R = 53,8$ % para o álcool fenchil, apresentando pelo modelo de Chickos *et al.* (1998) e $d_R = 32,07$ % para (-)-mirtenal, mostrado pelo modelo de Kolská *et al.* (2005).

Deve-se destacar que além do modelo de predição proposto neste trabalho, foram apresentados dados de P_{vap} a 298,15 K inéditos para os compostos 5-vinil-2-norborneno, álcool fenchil, (-)-borneol, isosafrol, 5-etilideno-2-norborneno, biciclo[2.2.1]hepta-2,5-dieno, 3-BHA, timol metil éter, (-)-*trans*-cariofileno, biciclo[2.2.2]octano-2,5-diona, e biciclo[4.1.0]hepta-7-ácido carboxílico, além de dados inéditos de ΔH_{vap} a 298,15 K e de T_{nb} para os sete últimos compostos.

Palavras-chave: compostos terpênicos; dados termodinâmicos; contribuição de grupos; python; gradiente descendente.

Abstract

Terpenes are key components of essential oils and can be found in a variety of organic classes. They are substances of great importance for chemical, pharmaceutical, and food industries. Thus, because of their applications in various fields, it is extremely necessary to know the thermodynamics properties and to obtain reliable data of these quantities, both experimentally and by prediction. However, it has been a constant challenge for the scientific community to do so. This Thesis aims to determinate experimentally the vapor pressure (P_{vap}), enthalpy of vaporization (ΔH_{vap}) and normal boiling temperature (T_{nb}) of terpene compounds present in essential oils by the gas chromatography technique (GC), besides proposing a model for predicting ΔH_{vap} , seeking a better performance for bicyclic terpenes. The experimental determination of thermodynamic data using GC involves the use of the chromatographic retention index in combination with Kirchhoff-Rankin equations type. The proposed ΔH_{vap} prediction model uses group contribution concepts and was obtained by the multiple linear regression technique. The optimization was done by gradient descent, implemented in Python 2 (Version 2.7.11), which is a totally free of cost programming tool. Gas chromatography experiments were carried out with defined temperature range per the nature of the compounds studied. In all, 25 terpene compounds were evaluated, divided per their volatility and analyzed in three different mixtures. Moreover, nine aliphatic hydrocarbons (branched and open chain and cyclic) were included in the GC analysis to support the discussion and to evaluate the sensitivity of the method at different volatilities.

The P_{vap} and ΔH_{vap} data were determined at 298.15 K. The results obtained agreed with the literature, showing small average relative deviations (d_M). Most of the compounds with P_{vap} data available in the literature showed d_M between 0.28 and 7.56 %, and only 5 compounds with d_M in the range of 11.73 to 16.27 %. As for the ΔH_{vap} data, only the isosafrole had elevate average relative deviation of 15.06 %. For the other compounds, the d_M was in the range of 0.5 - 8.33 %. The T_{nb} data showed even smaller deviations, in the range 0.02 - 3.15 %, demonstrating that the CG technique can be a good tool to obtain the thermodynamic quantities studied.

Prediction performance was compared to three models from literature: Joback e Reid (1987), Chickos *et al.* (1998) e Kolská *et al.* (2005). Both the model proposed in this paper and the three models of the literature evaluated were used to predict the enthalpy of vaporization of the 25 terpenes analyzed in the present work by GC and 30 terpene data from the literature. For the first set of data, the proposed model presented average relative deviation (d_M) of 9.87 %, while that of Joback and Reid (1987) presented $d_M = 15.72$ %, that of Kolská *et al.* (2005), $d_M = 20.50$ %, and Chickos *et al.* (1998), $d_M = 14.31$ %. With the 30 data of the literature, the model of this work presented $d_M = 9.51$ %, while the models of Joback and Reid (1987), Chickos *et al.* (1998), and Kolská *et al.* (2005) show, respectively, d_M equal to 19.21%, 13.85 %, and 15.79 %.

With respect to the bicyclic compounds, the predictions of ΔH_{vap} made by the proposed model, both for the compounds analyzed by GC in this work and with the literature data, were better than the Joback and Reid (1987) model, and presented results in agreement with the models of Chickos *et al.* (1998) and Kolská *et al.* (2005). However, these last two models presented high relative deviation (d_R) for some bicycles, such as $d_R = 53.8$ % for the alcohol phenol, presented by Chickos *et al.* (1998) model and $d_R = 32.07$ % for (-)-myrtenal, shown by Kolská *et al.* (2005) model.

It should be highlighted that simultaneously to the prediction model, unpublished data of P_{vap} at 298.15 K were presented for terpenes 5-vinyl-2-norbornene, fenchyl alcohol, (-)-borneol, isosafrol, 5-ethylidene-2-norbornene, bicyclo[2.2.1]hepta-2,5-diene, 3-BHA, thymol methyl ether, (-)-*trans*-caryophyllene, bicyclo[2.2.2]octane-2,5-dione, and bicyclo[4.1.0]hepta-7-carboxylic acid, in addition to unpublished data of ΔH_{vap} at 298.15 K and T_{nb} for the last seven compounds.

Keywords: terpene compounds; thermodynamic data; group contribution; python; gradient descent.

Lista de Figuras

- Figura 1.1** - Volume em mil toneladas e valor FOB em milhões de US\$ obtido com exportações de óleos essenciais e produtos de uso nas indústrias de fragrâncias e higiene pessoal. Dados portal Aliceweb do Sistema de Análise das Informações de Comércio Exterior e do Sistema Integrado de Comércio Exterior (SISCOMEX).
..... 31
- Figura 2.1** - Monoterpenos frequentemente encontrados em óleos essenciais: (1) mirceno, (2) limoneno, (3) triceno, (4) 3-careno, (5) α -thujeno, (6) α -pineno, (7) canfeno. Monoterpenóides (presença de grupo funcional): (8) linalol, (9) ácido tújico, (10) tujona, (11) safanal. 37
- Figura 2.2** - Sesquiterpenos encontrados em óleos essenciais; (12) nerolidol, (13) cadalene, (14) longefoleno. 38
- Figura 2.3** - Orbitais moleculares e a hibridização do carbono (McMURRY, 2011). 39
- Figura 2.4** - Orbitais moleculares e os tipos de hibridização sp^3 (a) e sp^2 (b) (baseado em McMURRY, 2011). 40
- Figura 2.5** - Ilustração de uma reação de moléculas de eteno e menteno com uma molécula hipotética A_2 41
- Figura 2.6** - Ilustração do processo fenomenológico do surgimento das forças de dispersão (McMURRY, 2011). 43
- Figura 2.7** - Presença de ligação de hidrogênio em moléculas polares contendo átomo de hidrogênio ligado a átomo altamente eletronegativo. 44
- Figura 2.8** - Formato molecular do n-pentano (linear) e do neopentano (esférico) (baseado em McMURRY, 2011). 45

Figura 2.9 - Efeito da ligação de hidrogênio em moléculas polares diferentes.....	46
Figura 2.10 - Descrição correta do modelo de Kolská et al. (2005) para: (a) grupos de primeira ordem; (b) grupos de segunda ordem; (c) grupos de terceira ordem.	58
Figura 2.11 - Esquema simples de uma tarefa de regressão.....	60
Figura 2.12 - Fluxograma de um processo de aprendizagem de máquina por regressão.....	61
Figura 2.13 - Ilustração geral de uma função convexa.....	64
Figura 2.14 - Variação do desvio de treinamento em função; (a) do número de pontos de dados e; (b) aumento da complexidade.	65
Figura 2.15 - Ilustração de um processo de validação cruzada.	67
Figura 3.1 - Esquema de um sistema cromatográfico a gás usado na análise dos compostos examinados.....	72
Figura 3.2 - Molécula hipotética ilustrando os grupos utilizados para o desenvolvimento do modelo proposto.	77
Figura 3.3 - Macro para aquisição de dados.....	79
Figura 3.4 - Macro para visualizar as informações adicionadas.	79
Figura 4.1 - Ilustração da dependência do índice de retenção de Kováts em função da temperatura para o 2,2-dimetilbutano.	84
Figura 4.2 - Ilustração da dependência do índice de retenção de Kováts em função da temperatura para o carvacrol.	87
Figura 4.3 - Ilustração da dependência do índice de retenção de Kováts em função da temperatura para o timol.	87
Figura 4.4 - Ilustração da dependência do índice de retenção de Kováts em função da temperatura para o metilciclohexano.	88
Figura 4.5 - Ajuste com a série de Taylor para o 2,2-dimetilbutano usando a Equação 4.1...	89

Figura 4.6 - Ajuste com a série de Taylor para o 2,2-dimetilbutano usando a Equação 2.16.	89
Figura 4.7 - Ajuste com a série de Taylor para o 2,2-dimetilbutano usando a Equação 4.2...	90
Figura 4.8 - Curvas de pressão de vapor em função da temperatura obtidas pela Equação 2.14 e ajuste da equação de Antoine (Equação 4.4) para os hidrocarbonetos alifáticos analisados.	94
Figura 4.9 - Curvas de pressão de vapor em função da temperatura obtidas pela Equação 2.14 e ajuste da equação de Antoine (Equação 4.4) para os terpenos da mistura 1.	95
Figura 4.10 - Curvas de pressão de vapor em função da temperatura obtidas pela Equação 2.14 e ajuste da equação de Antoine (Equação 4.4) para os terpenos da mistura 2.	95
Figura 4.11 - Curvas de pressão de vapor em função da temperatura obtidas pela Equação 2.14 e ajuste da equação de Antoine (Equação 4.4) para os terpenos da mistura 3.	96
Figura 4.12 - Valores de pressão de vapor calculados a 298,15 K <i>versus</i> valores médios da literatura (Equação 4.5).	104
Figura 4.13 - Valores de pressão de vapor calculados a 298,15 K <i>vs.</i> valores médios da literatura (Equação 4.8).	104
Figura 4.14 - Ajuste $\ln P$ <i>vs.</i> T da Equação 4.13 para o: (a) 2,2-dimetilbutano; (b) 3-metilpentano; (c) 3-BHA e; (d) (-)-trans-cariofileno.	106
Figura 4.15 - Gráfico de dispersão dos dados de entalpias incluídos no modelo.....	120
Figura 4.16 - Histograma das frequências absolutas dos 3.643 dados de entalpia do conjunto de dados analisado.....	122
Figura 4.17 - Gráfico de caixa para o conjunto de 3.643 dados de entalpia de vaporização.	123
Figura 4.18 - Desvios de treinamento e validação <i>vs.</i> complexidade do modelo.....	125
Figura 4.19 - Desvio de validação cruzada <i>vs.</i> complexidade do modelo.....	126
Figura 4.20 - Variação do SQR_m em função do valor inicial dos pesos.	127
Figura 4.21 - Variação do SQR_m em função do valor da tolerância.	127

Figura 4.22 - Variação do SQR_m em função do tamanho do passo η	128
Figura 4.23 - Ilustração do efeito do parâmetro de regularização λ_2 para o modelo treinado apenas com um atributo (-CH3).	130
Figura 4.24 - Efeito do aumento do parâmetro de regularização λ_2 sobre o valor dos coeficientes.	131
Figura 4.25 - Seleção do parâmetro de regularização λ_2 por validação cruzada.	132
Figura 4.26 - Valores preditos vs. literatura sobre os dados de teste.	135
Figura 4.27 - Valores preditos vs. literatura sobre os dados de teste organizados em ordem crescente do valor de entalpia de vaporização.	136
Figura 4.28 - Comparação entre os resultados preditos vs. experimentais.	136

Lista de Tabelas

Tabela 2.1 - Principais óleos essenciais exportados pelo Brasil (baseado em dados do portal Aliceweb, códigos 33 do NCM)*.....	36
Tabela 2.2 - Grupos funcionais frequentemente encontrados em óleos essenciais.	42
Tabela 2.3 - Vantagens e desvantagens dos principais métodos convencionais de obtenção de dados termodinâmicos (OLIVEIRA, 2012).	49
Tabela 3.1 - Compostos terpênicos analisados neste trabalho.....	70
Tabela 3.2 - Hidrocarbonetos alifáticos analisados neste trabalho.....	71
Tabela 3.3 - <i>n</i> -Alcanos usados como padrão de referência neste trabalho.....	71
Tabela 3.4 - Dados termodinâmicos retirados da literatura para os compostos analisados.....	75
Tabela 3.5 - Dados termodinâmicos preditos pelo NIST/TDE para os compostos analisados com dados não disponíveis na literatura.....	76
Tabela 3.6 - Grupos usados para o desenvolvimento do modelo proposto.	78
Tabela 3.7 - Classe dos compostos inseridos no banco de dados.....	80
Tabela 4.1 - Índice de retenção obtidos por meio da Equação 2.11a para os hidrocarbonetos estudados.	84
Tabela 4.2 - Índice de retenção para os terpenos estudados.	85
Tabela 4.3 - Resultados de ajustes da série de Taylor com as Equações 2.15, 4.1 e 4.3 para o 2,2-dimetilbutano.	91

Tabela 4.4 - Coeficientes da Equações 2.15 (a_0 , a_1 e a_2) e 4.1 (a_0 e a_1) para os hidrocarbonetos alifáticos estudados.....	92
Tabela 4.5 - Coeficientes da Equações 2.15 (a_0 , a_1 e a_2) e 4.1 (a_0 e a_1) para compostos terpênicos estudados.....	93
Tabela 4.6 - Constantes de Antoine para todos os compostos analisados obtidas por meio de ajustes das curvas de pressão de vapor das Figuras 4.8, 4.9, 4.10 e 4.11.....	96
Tabela 4.7 - Dados de pressão de vapor para os compostos analisados calculados através da Equação 4.5 e uma comparação com os dados da literatura.	98
Tabela 4.8 - Número de ligação torsional efetivo ($\tau_{efet.}$), número de ligação de hidrogênio (N_{LH}) e estimacão de $\Delta S_{b,X}^{vap}$ ara os compostos analisados.....	101
Tabela 4.9 - Dados de pressão de vapor a 298,15 K para os compostos analisados obtidos pela Equação 4.9 e comparação com dados da literatura e com dados advindos da Equação 4.5.	102
Tabela 4.10 - Coeficientes b_0 , b_1 e b_2 da Equação 4.13 para todos os compostos analisados.	107
Tabela 4.11 - Valores de entalpia de vaporização a 298,15 K (em kJ mol^{-1}) para os hidrocarbonetos alifáticos e os terpenos estudados e comparação com os dados da literatura.....	108
Tabela 4.12 - Valores de temperatura normal de ebulição (em K) para os hidrocarbonetos alifáticos e os terpenos estudados e comparação com os dados da literatura...	109
Tabela 4.13 - Principais grupos extraídos de Joback e Reid (1987) utilizados neste trabalho.	111
Tabela 4.14 - Estimacão da entalpia de vaporização do (+)-limoneno pelo método de Joback e Reid (1987).....	112
Tabela 4.15 - Estimacão da entalpia de vaporização do timol pelo método de Joback e Reid (1987).	112

Tabela 4.16 - Termos envolvidos na predição pelo modelo de Chickos <i>et al.</i> (1998).	113
Tabela 4.17 - Predição da entalpia de vaporização para o α -terpineol pelo modelo de Chickos <i>et al.</i> (1998).	114
Tabela 4.18 - Predição da entalpia de vaporização para o 3-BHA pelo modelo de Chickos <i>et al.</i> (1998).	114
Tabela 4.19 - Grupos de primeira ordem e suas contribuições empregados na predição de entalpia de vaporização pelo modelo de Kolská <i>et al.</i> (2005).	115
Tabela 4.20 - Grupos de segunda ordem e suas contribuições empregados na predição de entalpia de vaporização pelo modelo de Kolská <i>et al.</i> (2005).	116
Tabela 4.21 - Predição da entalpia de vaporização para o (+)- α -pineno pelo modelo de Kolská <i>et al.</i> (2005).	117
Tabela 4.22 - Predição da entalpia de vaporização para o (+)- β -pineno pelo modelo de Kolská <i>et al.</i> (2005).	117
Tabela 4.23 - Desempenho dos modelos de predição testados neste trabalho. Em parênteses encontra-se o desvio relativo (em %) em relação aos dados obtidos por cromatografia gasosa.	119
Tabela 4.24 - Distribuição de frequência para os 3.643 dados de entalpia de vaporização. .	122
Tabela 4.25 - Desempenho de regressão: validação tradicional vs. validação cruzada.	125
Tabela 4.26 - Comparação da soma dos quadrados residuais médio (SQR_m) para escolha dos melhores parâmetros de entrada.	129
Tabela 4.27 - Desempenho de regressão por gradiente descendente (GD) vs. GraphLab.	129
Tabela 4.28 - Desempenho da regressão regularizada e sua comparação com a regressão regularizada do pacote GraphLab Create e com a regressão sem regularização.	133
Tabela 4.29 - Organização do conjunto de dados após exclusão dos outliers acima de	115
kJ mol ⁻¹	134

Tabela 4.30 - Desempenho do modelo após a exclusão dos outliers acima de 115 kJ mol ⁻¹ .	134
Tabela 4.31 - Grupos selecionados e as suas respectivas contribuições.....	137
Tabela 4.32 - Desempenho do modelo desenvolvido neste trabalho em comparação com os dados experimentais obtidos por cromatografia gasosa e com os modelos preditivos analisados da literatura. Em parênteses encontra-se o desvio relativo (%) em relação aos dados experimentais.	139
Tabela 4.33 - Desempenho geral do modelo desenvolvido e uma comparação com os modelos preditivos da literatura analisados.....	141
Tabela 4.34 - Predição de ΔH_{vap} para compostos terpênicos com dados experimentais da literatura. Em parênteses encontram-se os desvios relativos.....	143
Tabela 4.35 - Distribuição dos desvios relativos (d_R) obtidos para todos compostos analisados da literatura.	144

Lista de Siglas

ALICEWEB	Sistema de Análise das Informações de Comércio Exterior
AM	Aprendizagem de máquina
CG	Cromatografia gasosa
DA	Desvio médio absoluto
DM	Desvio médio relativo
DSC	Calorimetria de Varredura Diferencial
FID	Detector de Ionização de Chamas
FOB	<i>Free on Board</i>
GC/FID	Cromatografia Gasosa com Detector de Ionização de Chamas
GD	Gradiente descendente
GL	Graphlab Create
MM	Massa molar
MMA	Ministério do Meio Ambiente
NCM	Nomenclatura Comum do Mercosul
NIST	National Institute of Standards and Technology
OE	Óleo essencial
PEC	Princípio dos Estados Correspondentes
RDQM	Raiz do desvio quadrado médio
SISCOMEX	Sistema Integrado de Comércio Exterior
SQR	Soma dos quadrados residuais
NIST/TDE	<i>National Institute of Standards and Technology/ThermoData Engine</i>

Lista de Símbolos

LETRAS LATINAS

A	Parâmetro da Equação 2.21 do método de Basarová e Svoboda (1995)
a_0, a_1, a_2	Coefficientes da Equação 2.20
A_T	Amplitude total
b_0, b_1, b_2	Coefficientes da Equação 2.20
d_R	Desvio relativo, [%]
d_M	Desvio relativo médio, [%]
F_i	Frequência
f_i	Frequência relativa
f_{i_ac}	Frequência relativa acumulada
f_L, f_U	Limites inferior e superior
F_{Q1} e F_{Q3}	Frequências absolutas de suas respectivas classes
G^g	Energia de Gibbs da fase vapor, [J mol ⁻¹]
G^l	Energia de Gibbs da fase líquida, [J mol ⁻¹]
h	intervalo dentre classes
I_R	Índice de retenção de Kováts
$I_X^{298,15K}$	Índice de retenção de Kováts a 298,15 K
k	Número de classes
l_{Q1} e l_{Q3}	Limite inferior da classe que contém o elemento que representa Q_1 e Q_3
n	Número de elementos do conjunto de dados na Equação de Sturges
N_{LH}	Número de ligação de hidrogênio
p_s	Pressão de saturação [Pa, kPa]
P_{vap}	Pressão de vapor, [Pa, kPa]
$P_{vap}^{298,15K}$	Pressão de vapor a 298,15 K, [Pa, kPa]

$P_{vap}^{calc.}$	Pressão de vapor calculada, [Pa, kPa]
P_i	Pressão de vapor da espécie i , [Pa, kPa]
$P_{vap}^{lit.}$	Pressão de vapor da literatura, [Pa, kPa]
Q_1, Q_3	Primeiro e segundo quartis
r^2	Coefficiente de determinação
S^g	Entropia da fase vapor, [J mol ⁻¹ K ⁻¹]
S^l	Entropia da fase líquida, [J mol ⁻¹ K ⁻¹]
$t'_{R,X}$	Tempo de retenção ajustado de uma substância X , [min]
$t'_{R,z}$	Tempo de retenção ajustado de um n -alcano de referência z , [min]
$t'_{R,z+1}$	Tempo de retenção ajustado de um n -alcano de referência $z+1$, [min]
T_{nb}	Temperatura normal de ebulição, [K]
T_r	Temperatura reduzida, [K]
t_R	Tempo de retenção, [min]
$T_{nb}^{calc.}$	Temperatura normal de ebulição calculada [K]
$T_{nb}^{lit.}$	Temperatura normal de ebulição da literatura, [K]
$V'_{R,X}$	Volume de retenção de uma substância X , [μL]
$V'_{R,z}$	Volume de retenção de um n -alcano de referência z , [μL]
$V'_{R,z+1}$	Volume de retenção de um n -alcano de referência $z+1$, [μL]
w_0, w_1	Coefficientes linear e angular (Equação 2.28) de modelo de regressão linear
\hat{w}_0, \hat{w}_1	Valores estimados dos parâmetros de um modelo linear
x_i	Fração molar da fase líquida
y_i	Fração molar da fase vapor

LETRAS GREGAS

α	Parâmetro da Equação 2.21 do método de Basarová e Svoboda (1995)
γ_i^∞	Coefficiente de atividade na diluição infinita da espécie i
γ_z^∞	Coefficiente de atividade na diluição infinita de um n -alcano de referência z
γ_{z+1}^∞	Coefficiente de atividade na diluição infinita de um n -alcano de referência $z+1$

γ_X^∞	Coeficiente de atividade na diluição infinita de uma substância X
ΔG	Varição da energia de Gibbs, [J mol ⁻¹]
ΔH	Varição de entalpia, [kJ mol ⁻¹]
ΔH_{vap}	Entalpia de vaporização, [kJ mol ⁻¹]
$\Delta H_{vap}^{298,15K}$	Entalpia de vaporização a 298,15 K, [J mol ⁻¹ K ⁻¹]
$\Delta H_{vap}^{calc.}$	Entalpia de vaporização calculada, [J mol ⁻¹ K ⁻¹]
$\Delta H_{vap}^{lit.}$	Entalpia de vaporização da literatura, [J mol ⁻¹ K ⁻¹]
ΔS	Varição de entropia, [J mol ⁻¹ K ⁻¹]
ΔS_{vap}^b	Entropia na temperatura normal de ebulição, [J mol ⁻¹ K ⁻¹]
ΔV	Varição de volume
η	Tamanho do passo usado em gradiente descendente
λ_1	Coeficiente de regularização <i>Lasso regression</i>
λ_2	Coeficiente de regularização <i>Ridge regression</i>
ΣF_i	Frequência acumulada
Σf_i	Frequência relativa acumulada
τ	Número de ligação torsional
$\tau_{efet.}$	Número de ligação torsional efetivo

Sumário

CAPÍTULO 1 - INTRODUÇÃO	30
1.1 Apresentação do problema	30
1.2 Objetivo geral	33
1.2.1 Objetivos específicos.....	33
1.3 Organização da tese	34
CAPÍTULO 2 – REVISÃO DA LITERATURA	35
2.1 Óleos essenciais	35
2.2 Compostos terpênicos	37
2.3 Natureza das substâncias orgânicas	38
2.3.1 Hibridização de carbonos	39
2.3.2 Grupos funcionais	41
2.4 Influência das forças atrativas sobre as propriedades termodinâmicas	41
2.4.1 Forças atrativas.....	42
2.4.2 Atração dipolo-dipolo	44
2.4.3 Ligação de hidrogênio.....	44
2.5 Importância das propriedades termodinâmicas de substâncias orgânicas	46
2.5.1 Equação de Clausius-Clapeyron e a relação da entalpia de vaporização com a pressão de vapor.....	47
2.6 Obtenção de grandezas termodinâmicas	48
2.6.1 Obtenção de dados termodinâmicos por cromatografia gasosa (CG).....	50

2.6.1.1 Teoria.....	51
2.6.2 Obtenção de entalpia de vaporização por métodos de predição.....	54
2.6.2.1 Modelos baseados no princípio dos estados correspondentes	55
Método de Basarová e Svoboda (1995).....	55
Método de Tu e Liu (1996).....	55
2.6.2.2 Modelos baseados somente na estrutura química.....	56
Modelo de Joback e Reid (1987)	56
Modelo de Chickos et al. (1998).....	56
Modelo de Kolská et al. (2005).....	57
2.7 Desenvolvimento de modelos preditivos	58
2.7.1 Aprendizagem de máquina.....	59
2.7.1.1 Regressão.....	59
2.7.1.2 Regressão linear simples	60
2.7.1.3 Regressão ordinária por mínimos quadrados	61
2.7.1.4 Regressão linear múltipla	62
2.7.1.5 Regressão regularizada por mínimos quadrados	62
2.7.1.6 Algoritmos de otimização.....	63
Otimização por gradiente descendente	63
2.8 Avaliação de desempenho	64
2.8.1 Treinamento, teste e validação	65
2.8.2 Validação cruzada	66
CAPÍTULO 3 – MATERIAL E MÉTODOS.....	69
3.1 PARTE EXPERIMENTAL – GC/FID	69
3.1.1 Material	69
3.1.1.1 Compostos terpênicos analisados	69
3.1.1.2 Hidrocarbonetos alifáticos.....	70

3.1.1.3 n-Alcanos.....	71
3.1.1.4 Sistema cromatográfico a gás.....	72
3.1.2 Métodos.....	72
3.1.2.1 Preparo das soluções.....	73
3.1.2.2 Análises cromatográficas.....	73
3.1.2.3 Ajuste dos tempos de retenção	74
3.1.2.3 Dados termodinâmicos disponíveis na literatura.....	74
3.2 PROPOSIÇÃO DE MODELO DE PREDIÇÃO	76
3.2.1 Seleção de modelos de predição de ΔH_{vap} da literatura	76
3.2.2 Proposição de modelo de predição de ΔH_{vap}	76
3.2.2.1 Seleção dos compostos	76
3.2.2.2 Fragmentação dos grupos	77
3.2.2.3 Aquisição dos dados de entalpia de vaporização	78
3.2.2.4 Formação do conjunto de dados	80
3.2.2.5 Procedimento para estimação dos parâmetros do modelo.....	80
3.2.2.6 Divisão do conjunto de dados.....	81
3.2.2.7 Regressão ordinária por mínimos quadrados	81
3.2.2.8 Regressão regularizada por mínimos quadrados	81
3.2.2.9 Medida da qualidade da regressão.....	82
CAPÍTULO 4 – RESULTADOS E DISCUSSÃO	83
4.1 RESULTADOS EXPERIMENTAIS DE GC/FID	83
4.1.1 Obtenção dos índices de retenção	83
4.1.2 Obtenção da pressão de vapor.....	91
4.1.2.1 Extrapolação da Equação 2.14 para a obtenção de dados de pressão de vapor a 298,15 K	97

4.1.3 Obtenção de dados de entalpias de vaporização e temperatura normal de ebulição	105
4.2 DESEMPENHO PREDITIVO DOS MODELOS DA LITERATURA	110
4.2.1 Modelo de Joback e Reid (1987).....	110
4.2.2 Modelo de Chickos et al. (1998).....	113
4.2.3 Modelo de Kolská et al. (2005).....	115
4.2.4 – Desempenho geral dos modelos da literatura na determinação de ΔH_{vap} dos terpenos analisados por CG	118
4.3 PROPOSTA DE MODELO DE CONTRIBUIÇÃO DE GRUPOS PARA PREDIÇÃO DE ENTALPIAS DE VAPORIZAÇÃO	120
4.3.1 Comportamento dos dados	120
4.3.2 Procedimentos de regressão para obtenção do modelo proposto	123
4.3.2.1 Regressão ordinária por mínimos quadrados usando o GL.....	124
Seleção dos parâmetros do modelo.....	124
4.3.2.3 Regressão ordinária por mínimos quadrados usando o GD	126
Seleção dos parâmetros de entrada	126
Treinamento do modelo com regressão ordinária por mínimos quadrados.....	129
Treinamento do modelo com regressão regularizada por mínimos quadrados	130
Seleção do parâmetro de regularização λ_2	131
4.3.2.4 Desempenho de regressão regularizada por mínimos quadrados usando GD. 132	
4.3.2.5 Desempenho do modelo sem os dados considerados outliers	133
4.3.2.9 Construção do modelo final.....	134
4.3.3 – Avaliação de desempenho do modelo de predição proposto	138
4.3.4 Avaliação com dados da literatura	142
CAPÍTULO 5 – CONCLUSÃO E SUGESTÕES	145

5.1 Conclusões finais	145
5.2 Sugestões para trabalhos futuros	147
Referências	148
Apêndice A - Estruturas químicas dos compostos estudados	161
Apêndice B - Propriedades dos <i>n</i>-alcanos	164
Apêndice C - Classe dos compostos incluídos no conjunto de dados	165
Apêndice D - Ajustes <i>I vs. T</i>	167
Apêndice E - Ajustes da equação de Kirchhoff–Rankin	172
Apêndice F - Compostos <i>outliers</i>	176
Apêndice G - Dados termodinâmicos para os hidrocarbonetos alifáticos	182
Apêndice H - Algoritmos usado para estimação dos parâmetros do modelo	185
Apêndice I - Cromatogramas	186

Capítulo 1

INTRODUÇÃO

A presente Tese de Doutorado envolve a obtenção de grandezas termodinâmicas de compostos orgânicos terpênicos de ocorrência em óleos essenciais, bem como a proposição de modelo preditivo para a estimativa da entalpia de vaporização, visando melhor desempenho de predição em compostos terpenos de natureza bicíclica. Neste capítulo é abordada a posição do Brasil no cenário mundial na produção de óleos essenciais e a importância dessas substâncias para a sociedade e como grandezas termodinâmicas, especificamente a pressão de vapor, entalpia de vaporização e temperatura normal de ebulição, podem ser úteis nas indústrias, além dos motivos que impulsionaram a realização deste trabalho. Ao final do capítulo são listados os objetivos gerais e específicos.

1.1 Apresentação do problema

O Brasil é considerado o país com a maior biodiversidade do mundo. Segundo dados do Ministério do Meio Ambiente (MMA), são pelo menos 103.870 espécies animais e 43.020 espécies vegetais conhecidas no país. Todas essas espécies estão distribuídas em seis biomas de características diferentes (Amazônia, Cerrado, Mata Atlântica, Caatinga, Pantanal e Pampa), onde cada um possui fauna e flora de tipos distintos (MMA, 2016).

Conhecido como a savana mais rica existente no mundo, o Cerrado representa o segundo maior bioma da América do Sul. Contém mais de 11.627 espécies de plantas nativas catalogadas. Nesse sentido, destaca-se a importância do Estado do Maranhão, que corresponde à cerca de 9,8 milhões dos quase de 207 milhões de hectares de Cerrado do território brasileiro (CONCEIÇÃO e CASTRO, 2009).

Devido sua grande riqueza natural, o Brasil destaca-se como um dos principais exportadores de óleos essenciais e seus derivados para todo o mundo, ao lado de países como França, Índia, Argentina e China (BIZZO *et al.*, 2009). Informações do Sistema de Análise das

Informações de Comércio Exterior (portal Aliceweb) revelam que o setor de óleos essenciais, incluindo resinoides, produtos de perfumaria e preparações cosméticas, têm sido importantes e movimentado milhões de dólares com exportações. Por exemplo, a Figura 1.1 mostra o valor FOB (*Free on Board*), em dólares, movimentado no período de 2005 a 2015. Pode-se notar em tal figura que o valor obtido com exportações subiu mais de 100 % de 2005-2011, saindo de cerca de US\$ 400 milhões em 2005 para mais de US\$ 800 milhões em 2011. No ano de 2015, o valor movimentado foi cerca de US\$ 706 milhões (Figura 1.1). Esses dados são relativos à pesquisa feita com os códigos 33 (que classificam os produtos advindos de óleos essenciais e resinoides; produtos de perfumaria ou de toucador preparados e preparações cosméticas), da Nomenclatura Comum do Mercosul, NCM. Vale destacar que, embora o volume exportado no período em questão apresente-se praticamente constante, Speziali (2012) reportou que esta quantidade era bem menor na década de 1990 (próximo à 75 mil toneladas), tendo um aumento considerável de 2000 até 2008.

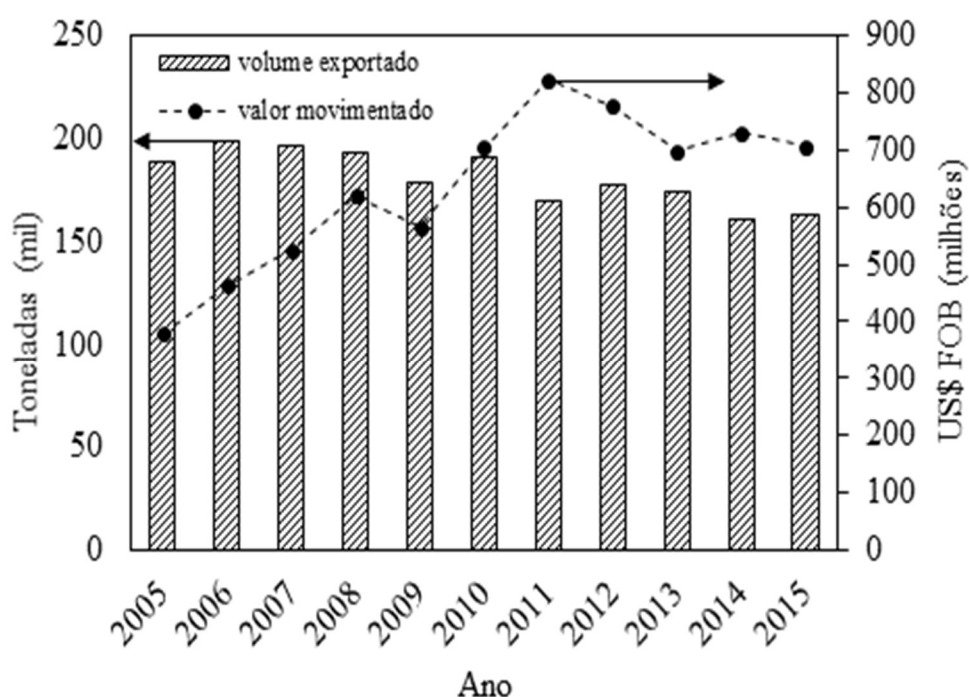


Figura 1.1 - Volume em mil toneladas e valor FOB em milhões de US\$ obtido com exportações de óleos essenciais e produtos de uso nas indústrias de fragrâncias e higiene pessoal. Dados portal Aliceweb do Sistema de Análise das Informações de Comércio Exterior e do Sistema Integrado de Comércio Exterior (SISCOMEX).

Esses dados são importantes para frisar o grande potencial do Brasil na produção de óleos essenciais cujas substâncias são ricas em compostos terpênicos, que têm grande aplicabilidade nas indústrias químicas, farmacêuticas e alimentícias. Muitos compostos extraídos de óleos essenciais são usados na forma *in natura* ou na síntese de vários outros produtos empregados na indústria química em geral. Por exemplo, a condensação do citral (presente no óleo essencial de capim-limão, *Lemongrass*) pode gerar o α -metilionona e o isometilionana, que são muito utilizados na indústria de alimentos, principalmente em bebidas não alcoólicas, sorvetes, gelatina, dentre outros. Outro exemplo é o eugenol (*Eugenia caryophyllata*), que é precursor de várias substâncias, dentre elas a vanilina, amplamente utilizada como aromatizante e em formulações farmacêuticas (CRAVEIRO e QUEIROZ, 1993).

É evidente, portanto, o grande desafio que está à frente em termos de conhecimento dessas substâncias. Tão importante como descobrir uma nova substância potencial, é gerar dados físicos e químicos para que possam ser empregados nas indústrias. Dentre muitos dados que envolvem os aspectos termodinâmicos de um composto, os dados de pressão de vapor (P_{vap}), entalpia de vaporização (ΔH_{vap}) e temperatura normal de ebulição (T_{nb}) são grandezas extremamente úteis. Os dados de P_{vap} e ΔH_{vap} , especificamente, têm sido um dos parâmetros mais importantes em processos multicomponentes e multiestágios baseados em equilíbrio líquido-vapor, tais como destilação e evaporação (KOLSKÁ *et al.*, 2012).

Vários esforços têm sido feitos para determinar esses dados. As técnicas de determinação direta consideradas primitivas como o isoteniscópio e a efusão gasosa vêm perdendo espaço para técnicas indiretas, devido a facilidade de utilização e por requerem pequena quantidade de amostra, como é o caso da cromatografia gasosa e da análise termogravimétrica. As técnicas experimentais possuem grande vantagem porque podem apresentar resultados mais exatos e precisos, porém podem demandar materiais dispendiosos e necessitar de boa prática experimental, além de conhecimento teórico.

Desse modo, surge outro meio de obtenção dessas propriedades que são os modelos preditivos. Esses métodos possuem inúmeras vantagens em relação às técnicas experimentais, uma vez que descartam quaisquer custos com parte experimental. Como será visto adiante, existem muitos métodos preditivos para determinar dados termodinâmicos tais como pressão de vapor, entalpia de vaporização e temperatura normal de ebulição. Cada método tem sua particularidade, possuindo vantagens e desvantagens, porém, existem aqueles capazes de representar melhor determinada classe de compostos.

Diante deste contexto, alguns fatores que motivaram a realização dessa Tese de Doutorado merecem ser destacados:

- ✚ Forte potencial do Brasil em biodiversidade e sua grande capacidade de produção de compostos de origem natural, que são de suma importância para a indústria e a sociedade;
- ✚ Grande interesse das indústrias químicas por dados termodinâmicos confiáveis;
- ✚ Contribuição com dados termodinâmicos de compostos terpênicos, que são bastante úteis para a indústria química;
- ✚ Possibilidade de contribuir com a comunidade acadêmica com a inclusão de um novo modelo de predição, atendendo aqueles que trabalham com óleos essenciais.

1.2 Objetivo geral

O objetivo geral da presente Tese de Doutorado é determinar dados termodinâmicos de pressão de vapor, entalpias de vaporização e temperatura normal de ebulição de compostos orgânicos terpênicos presentes em óleos essenciais extraídos de plantas naturais, pela técnica de cromatografia gasosa e desenvolver um modelo preditivo, baseado no conceito de contribuição de grupos, para estimar dados de entalpia de vaporização, enfatizando os compostos de natureza bicíclica. Para isso, foi necessário obter os seguintes objetivos específicos:

1.2.1 Objetivos específicos

- ✚ Fazer uma análise para encontrar a temperatura em que as soluções podem ser submetidas sem que o cromatograma apresente sobreposição de picos;
- ✚ Obter os dados de retenção cromatográfica de cada composto analisado para o intervalo de temperatura encontrado;
- ✚ Obter os índices de retenção de Kováts para cada composto analisado e em cada intervalo de temperatura, utilizando os dados de retenção cromatográfica;
- ✚ Aplicar equações do tipo Kirchhoff-Rankin para determinar as grandezas termodinâmicas avaliadas;

- ✚ Coletar uma quantidade significativa de dados de entalpia de vaporização da literatura para compostos orgânicos de ocorrência em óleos essenciais e organiza-los em uma tabela (banco ou conjunto de dados);
- ✚ Obter a estrutura química de cada composto incluso no banco de dados;
- ✚ Analisar a estrutura química de cada composto do conjunto de dados, para que sejam extraídos seus grupos representativos;
- ✚ Implementar um algoritmo para encontrar a melhor relação entre os grupos e entalpia de vaporização dos compostos utilizados;
- ✚ Selecionar alguns modelos de predição de entalpia de vaporização da literatura e confrontar os resultados com os obtidos neste trabalho.

1.3 Organização da tese

O presente trabalho de tese está organizado da seguinte forma:

- ✚ O Capítulo 2 apresenta os conceitos fundamentais que envolvem a obtenção de dados termodinâmicos por meio experimental e por predição, bem como os conceitos necessários para compreensão do desenvolvimento do modelo proposto;
- ✚ O Capítulo 3 descreve os detalhes experimentais realizados para a obtenção dos dados termodinâmicos estudados pela técnica de cromatografia gasosa e apresenta a proposta para a obtenção do modelo de predição de entalpia de vaporização proposto;
- ✚ O Capítulo 4 mostra os resultados obtidos neste trabalho e está dividido em duas partes. Na primeira parte são apresentados via cromatografia gasosa e a segunda traz os resultados por modelos de contribuição de grupos da literatura e os resultados obtidos pelo modelo proposto;
- ✚ O Capítulo 5 contém as conclusões extraídas dos resultados encontrados, bem como as propostas para a continuidade deste trabalho.

Capítulo 2

REVISÃO DA LITERATURA

Neste capítulo são apresentados os aspectos relevantes sobre a obtenção das grandezas termodinâmicas e do desenvolvimento do modelo preditivo que é proposto. É feita uma discussão sobre a natureza dos óleos essenciais e como a estrutura química pode afetar as propriedades termodinâmicas de compostos orgânicos. Também são apresentados os aspectos teóricos que envolvem a definição de entalpia de vaporização e a pressão de vapor e como tais grandezas podem ser determinadas experimentalmente. Além disso, é apresentado um tópico sobre alguns modelos de estimativa de entalpia de vaporização existentes na literatura. Por fim, são abordadas as etapas envolvidas no desenvolvimento de um modelo de predição.

2.1 Óleos essenciais

O Brasil é detentor da maior biodiversidade do mundo e destaca-se por possuir o maior acervo em plantas aromáticas. As plantas possuem substâncias de grande interesse industrial, popularmente conhecidas como óleos essenciais (OE). Os óleos essenciais podem ser extraídos de várias partes de uma planta, como folhas, frutos e caules, e são formados basicamente por um grupo de compostos denominados terpenos, que por sua vez são divididos em duas classes dominantes: os hidrocarbonetos terpênicos e os terpenoides, que são derivados de hidrocarbonetos que apresentam funções orgânicas como cetonas, álcoois, éteres, ésteres e fenóis (MOUCHREK FILHO, 2001; MONTEIRO, 2008).

As propriedades dos OE dependem das características do ambiente em que as espécies são cultivadas. Assim, em cada parte do Brasil os OE podem apresentar variados tipos de substâncias. A região Sudeste, por exemplo, notabiliza-se com o óleo essencial de eucalipto. Espécies como *Eucalyptus globulus*, *E. citriodora* e *E. staigeriana* produzem óleos essenciais com emprego nas indústrias de perfumaria e farmacêutica, devido a presença de grandes quantidades de eucaliptol, citral e citronelol (SOUZA *et al.*, 2010). O perfil ambiental do

Nordeste brasileiro favorece a existência de espécies de menor porte. As espécies de *Lippia*, por exemplo, arbusto da família Verbenaceae, típica do Cerrado, produzem óleos essenciais utilizados nas indústrias farmacêuticas por causa da presença de timol e carvacrol como componentes majoritários, que têm atividade antimicrobiana (OLIVEIRA, 2012).

Em relação aos aspectos econômicos que tangem os óleos essenciais, os principais produtos exportados estão apresentados na Tabela 2.1, conforme dados do portal Aliceweb para os códigos 33, do NCM. Os OE mais exportados são originados de espécies cítricas, com maioria proveniente da laranja, e os maiores destinos são os mercados de aromas e fragrâncias.

Tabela 2.1 - Principais óleos essenciais exportados pelo Brasil (baseado em dados do portal Aliceweb, códigos 33 do NCM)*.

Óleo essencial	Exemplo de espécie	Composto majoritário	Exemplo de usos	Referência
bergamota	<i>Citrus aurantium</i>	limoneno	aromas e fragrâncias	Zarrad <i>et al.</i> (2015)
laranja	<i>Citrus sinensis</i>	limoneno	aromas e fragrâncias	Araujo <i>et al.</i> (2016)
limão	<i>Citrus limonum</i>	<i>d</i> -limoneno	domissanitários e química fina	Wang <i>et al.</i> (2015)
lima	<i>Citrus aurantifolia</i>	limoneno	aromas e fragrâncias	Amorim <i>et al.</i> (2016)
gerânio	<i>Pelargonium graveolens</i>	geraniol	perfumes, aromatizantes	Juarez <i>et al.</i> (2016)
lavanda	<i>Lavandula angustifolia</i>	linalol	cosméticos e perfumes	Carrasco <i>et al.</i> (2016)
hortelã-pimenta	<i>Mentha piperita</i>	mentol	bebidas, alimentos	Hashimoto <i>et al.</i> (2016)
menta japonesa	<i>Mentha arvensis</i>	mentol	bebidas, alimentos	Pandey <i>et al.</i> (2003)
menta spearmint	<i>Mentha viridis L.</i>	carvona	aromatizante, repelentes	Mkaddem <i>et al.</i> (2009)
vetiver	<i>Vetiveria zizanoides</i>	khusimol	aromas e fragrâncias	Barros <i>et al.</i> (2009)
citronela	<i>Cymbopogon winterianus Jowitt</i>	citronelal	perfumes, repelentes	Timung <i>et al.</i> (2016)
cedro	<i>Juniperus virginiana</i>	cedrol	aromas e fragrâncias	Mun e Prewitt (2011)
pau-santo	<i>Bulnesia sarmientoi</i>	guaiol	perfumes	Rodilla <i>et al.</i> (2011)
lemongrass	<i>Cymbopogon schoenanthus L.</i>	geraniol	perfumes, cigarros e repelentes	Chagas <i>et al.</i> (2012)

* Códigos 33 da Nomenclatura Comum do Mercosul (NCM) que classificam os produtos advindos de óleos essenciais e resinoides; produtos de perfumaria ou de toucador preparados e preparações cosméticas.

2.2 Compostos terpênicos

Os terpenos podem ser divididos em dois subgrupos: monoterpenos e sesquiterpenos. Os monoterpenos (hidrocarbonetos ou derivados) são formados a partir de duas unidades de isopreno (2-metilbutadieno, C_5H_8) e são os mais abundantes em óleos essenciais. São os compostos mais voláteis e podem compreender cerca de 90 % da composição dos óleos essenciais (SIMÕES e SPITZER, 2007). Em relação às suas estruturas, podem ser encontrados na forma acíclica, monocíclica, bicíclica, aromática, dentre outras. A Figura 2.1 apresenta alguns dos principais monoterpenos e monoterpenóides encontrados em óleos essenciais extraídos de plantas.

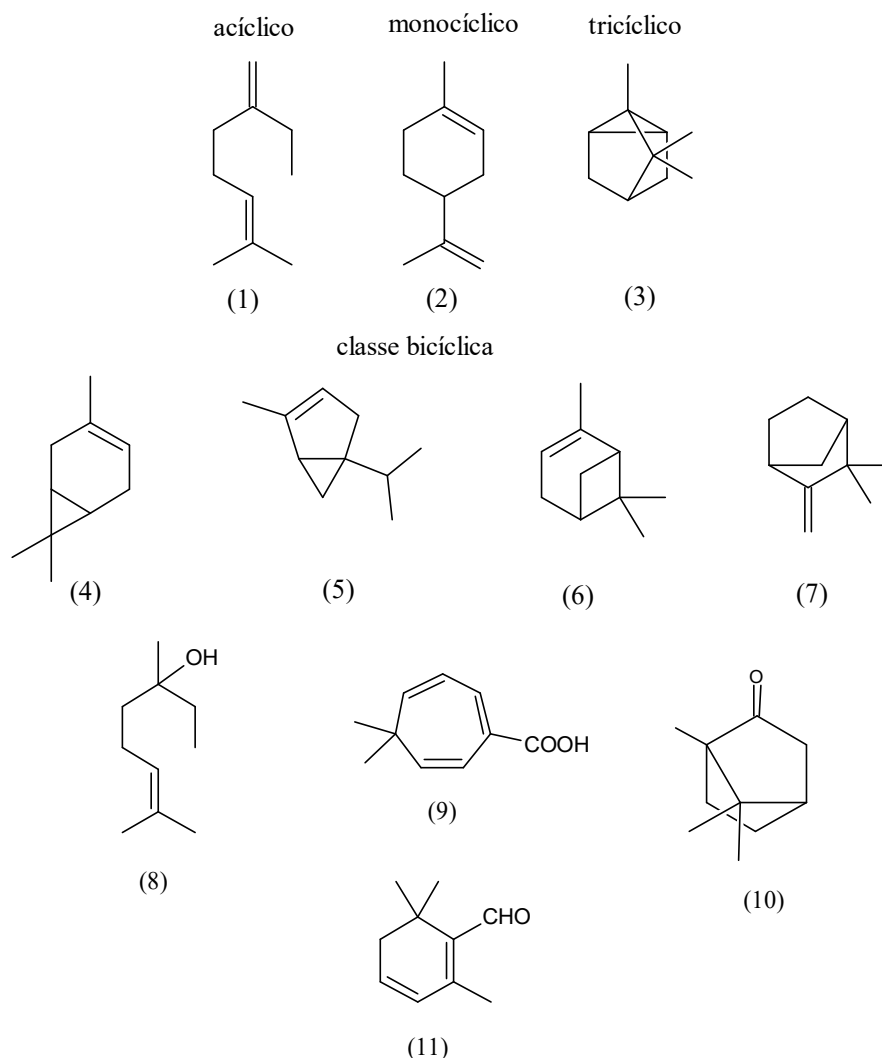


Figura 2.1 - Monoterpenos frequentemente encontrados em óleos essenciais: (1) mirceno, (2) limoneno, (3) triceno, (4) 3-careno, (5) α -tujeno, (6) α -pineno, (7) canfeno. Monoterpenóides (presença de grupo funcional): (8) linalol, (9) ácido tújico, (10) tujona, (11) safanal.

Os sesquiterpenos (hidrocarbonetos ou derivados) contêm geralmente 15 carbonos em sua estrutura química e são formados pelo conjunto de 3 unidades de isopreno. São compostos com volatilidade média (ou semivoláteis). Depois dos monoterpenos, são os compostos mais presentes em óleos essenciais. Suas estruturas podem ser encontradas frequentemente na forma acíclica, mono-, bi-, tri- e tetracíclica. Algumas estruturas químicas dessa classe estão apresentadas na Figura 2.2.

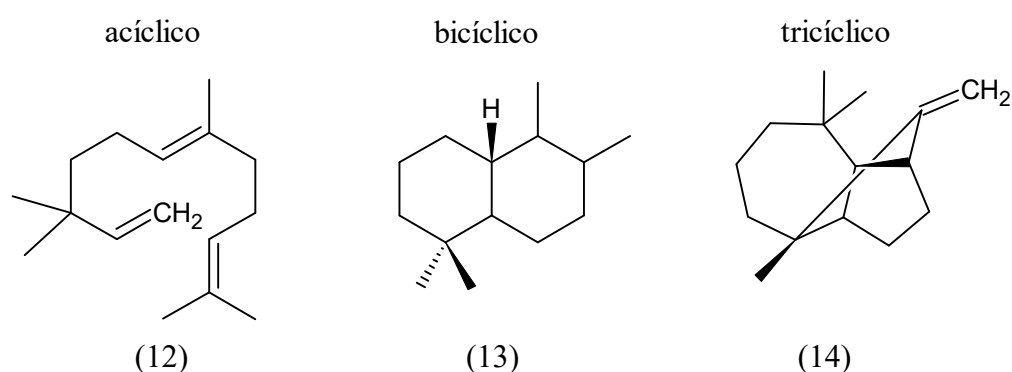


Figura 2.2 - Sesquiterpenos encontrados em óleos essenciais; **(12)** nerolidol, **(13)** cadalene, **(14)** longifoleno.

2.3 Natureza das substâncias orgânicas

Desde meados dos anos 1800, todos os compostos que continham o elemento carbono passaram a ser conhecidos como substâncias orgânicas (com algumas exceções), diferenciando-se das inorgânicas (McMURRY, 2011). Dos milhões de compostos orgânicos existentes, mais de 99 % contém o carbono em sua estrutura. Isso deve-se a sua configuração eletrônica, consequência da sua própria posição na tabela periódica.

O carbono possui número atômico $Z = 6$ e quatro elétrons na sua camada de valência ($2s^2, 2p^2$). Desse modo, pode se ligar facilmente a outros átomos, inclusive consigo mesmo, por meio de ligações covalentes (compartilhamento de elétrons). Por isso a grande existência deste elemento em moléculas orgânicas (McMURRY, 2011).

2.3.1 Hibridização de carbonos

O carbono realiza ligações químicas por meio de seus orbitais $2s$ e $2p$. Esta combinação de orbitais é chamada de hibridização. O conceito de hibridização explica como o carbono forma suas ligações, ângulos de ligações e a distância carbono-carbono. Os orbitais s têm formato esférico, enquanto os orbitais p possuem forma de halteres e simétrica (Figura 2.3a).

Quando ocorre a hibridização, os orbitais s e p unem-se e formam um novo orbital (híbrido). Por exemplo, a hibridização de um orbital $2s$ com três orbitais $2p$ forma quatro orbitais sp^3 de estrutura tetraédrica (ângulo de $109,5^\circ$), que acontece em moléculas de metano (Figura 2.3b) (McMURRY, 2011).

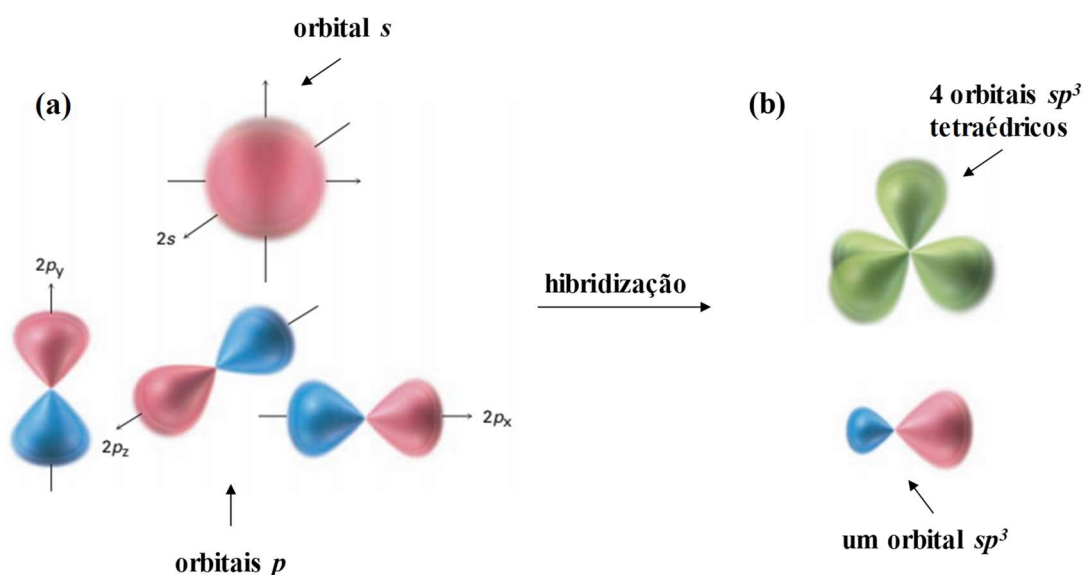


Figura 2.3 - Orbitais moleculares e a hibridização do carbono (McMURRY, 2011).

A hibridização de orbitais ocorre também em ligações carbono-carbono, responsáveis pela formação de cadeias e anéis e o surgimento de milhares de substâncias. A formação da molécula de etano, por exemplo, pode ser entendida imaginando-se a aproximação de dois carbonos sp^3 . A sobreposição de um orbital sp^3 de cada um dos carbonos forma uma ligação denominada de ligação simples (ligação σ). Os três orbitais sp^3 de cada carbono formam

ligações com o orbital s do átomo de hidrogênio (Figura 2.4a). Todos os ângulos de ligação da molécula do etano têm valor próximo ao de um tetraedro.

A hibridização também acontece para formar mais de uma ligação carbono-carbono. Isso ocorre quando o orbital $2s$ combina-se com apenas um ou dois orbitais $2p$ dos três orbitais $2p$ disponíveis. Quando o orbital $2s$ se combina com dois orbitais $2p$, forma o tipo de hibridização chamada sp^2 , existente em moléculas que contêm dupla ligação carbono-carbono. A combinação de um orbital $2s$ com apenas um orbital $2p$ gera a hibridização sp , que são ligações tripla, típica de moléculas de alcinos. A aproximação de dois orbitais hibridizados sp^2 forma uma ligação através da sobreposição dos orbitais (ligação σ sp^2-sp^2). Os dois orbitais $2p$ não hibridizados interagem e a sobreposição desses orbitais geram outra ligação (ligação π $2p-2p$). Esses dois processos dão origem às ligações dupla em alcenos (Figura 2.4b). O mesmo acontece com alcinos, porém, como os alcinos têm dois orbitais $2p$ não hibridizados, fazem duas ligações π $2p-2p$ (McMURRY, 2011).

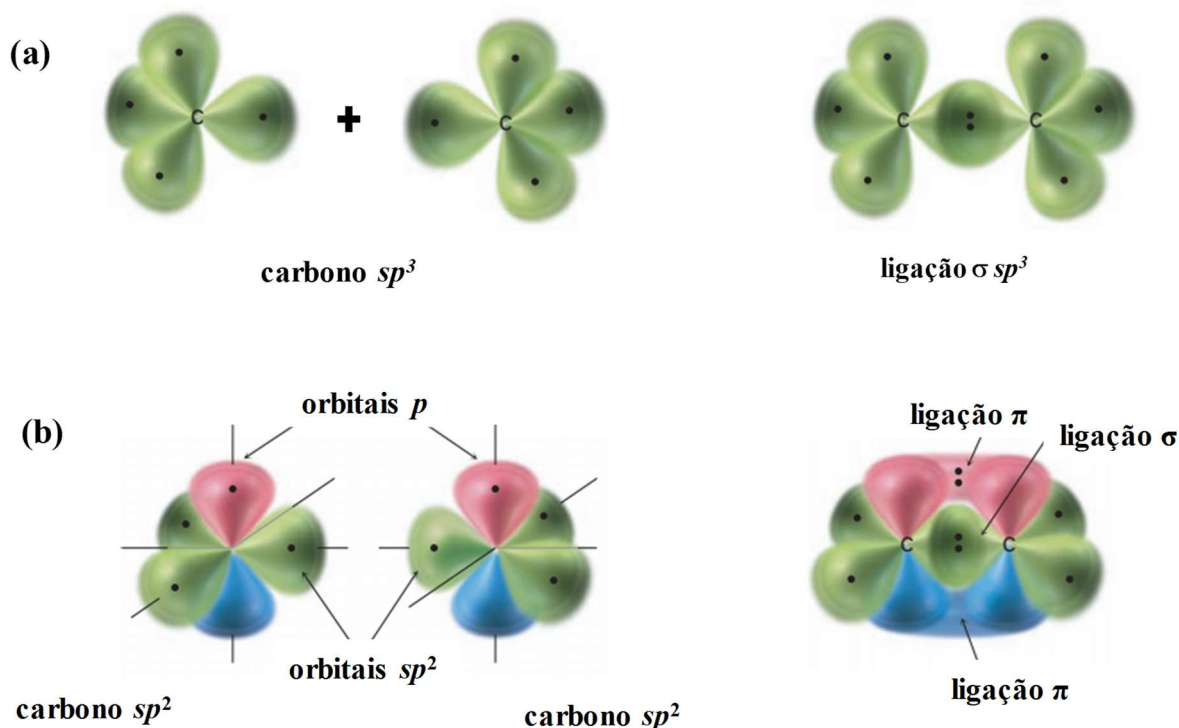


Figura 2.4 - Orbitais moleculares e os tipos de hibridização sp^3 (a) e sp^2 (b) (baseado em McMURRY, 2011).

2.3.2 Grupos funcionais

Um grupo funcional é um grupo de átomos que tem comportamento químico característico, permitindo classificar os compostos por família. De forma geral, desempenham o mesmo papel nas moléculas em que fazem parte. Assim, uma ligação dupla de uma molécula simples de etileno reage da mesma maneira que a ligação dupla do composto orgânico menteno, que é encontrado em óleo essencial de hortelã-pimenta e tem molécula química mais complexa. Numa reação do etileno e menteno com uma molécula hipotética A_2 , por exemplo, ambos reagiriam de maneira semelhante na ligação dupla, dando produtos com um átomo hipotético de A adicionado a cada um dos carbonos com ligação dupla (Figura 2.5). Alguns grupos funcionais mais comuns encontrados em óleos essenciais estão listados na Tabela 2.2 (McMURRY, 2011).

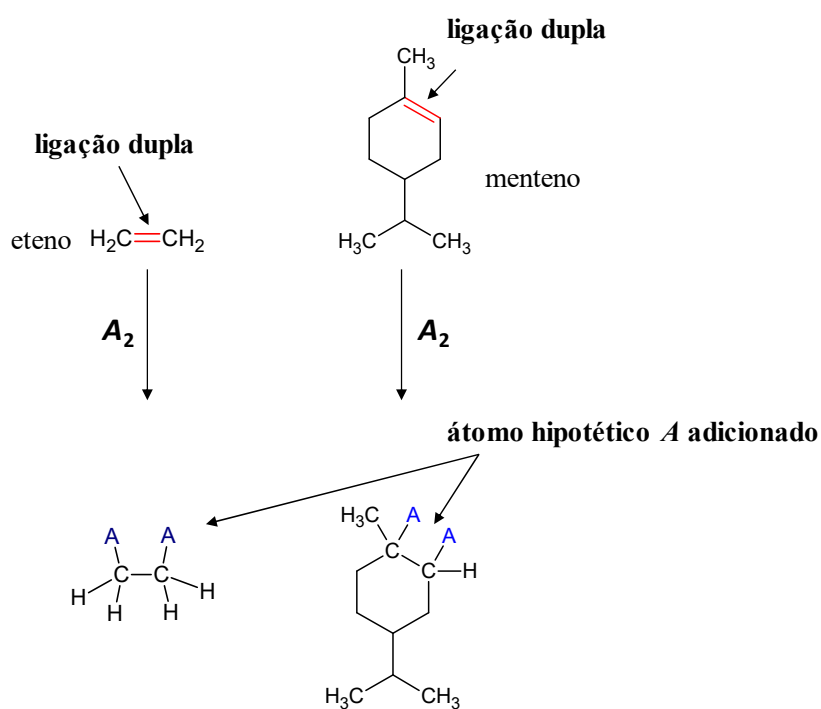


Figura 2.5 - Ilustração de uma reação de moléculas de eteno e menteno com uma molécula hipotética A_2 .

2.4 Influência das forças atrativas sobre as propriedades termodinâmicas

As propriedades físico-químicas de uma substância dependem diretamente de sua estrutura química. A partir da molécula e da natureza das ligações, pode-se deduzir se determinada substância encontra-se no estado sólido, líquido ou gasoso. As substâncias podem

ser encontradas em um desses três estados da matéria devido às forças atrativas intermoleculares, ou somente forças intermoleculares. A magnitude das forças atrativas em sólidos é muito maior que em líquidos, que por sua vez é muito maior que em gases. Muitas propriedades de líquidos, como a temperatura normal de ebulição, refletem as forças intermoleculares envolvidas.

Tabela 2.2 - Grupos funcionais frequentemente encontrados em óleos essenciais.

Estruturas de alguns grupos funcionais comuns		
Nome	Estrutura	Exemplo
Alceno (ligação dupla)		eteno $\text{H}_2\text{C}=\text{CH}_2$
Areno (anel aromático)		benzeno
Álcool		etanol $\text{H}_3\text{C}-\text{CH}_2\text{OH}$
Éter		dimetil éter $\text{H}_3\text{C}-\text{O}-\text{CH}_3$
Aldeído		etanal $\text{H}_3\text{C}-\text{C}(=\text{O})-\text{H}$
Cetona		propanona $\text{H}_3\text{C}-\text{C}(=\text{O})-\text{CH}_3$
Ácido carboxílico		ácido etanoico $\text{H}_3\text{C}-\text{C}(=\text{O})-\text{OH}$
Éster		metil etanoato $\text{H}_3\text{C}-\text{C}(=\text{O})-\text{O}-\text{CH}_3$

2.4.1 Forças atrativas

Em moléculas eletricamente neutras existem três principais forças de atração: forças de dispersão, atração dipolo-dipolo e ligação de hidrogênio. Ambas as forças de dispersão e atração dipolo-dipolo são as conhecidas forças de *van der Waals*.

As forças de dispersão (ou forças de London) explicam a atração de moléculas eletricamente neutras e não polares. Essas forças são fracas e acontecem pelo fenômeno de momento dipolo induzido. O aparecimento dessas forças pode ser explicado da seguinte maneira: em determinado instante, a nuvem eletrônica de um átomo ou molécula (Figura 2.6a) vai para um lado (dipolo instantâneo), que fica com carga parcial negativa, enquanto o outro lado fica parcialmente positivo (Figura 2.6b). Um átomo ou molécula com dipolo instantâneo induz a formação de dipolo de um átomo ou molécula vizinho (dipolo induzido), situação “c” da Figura 2.6.

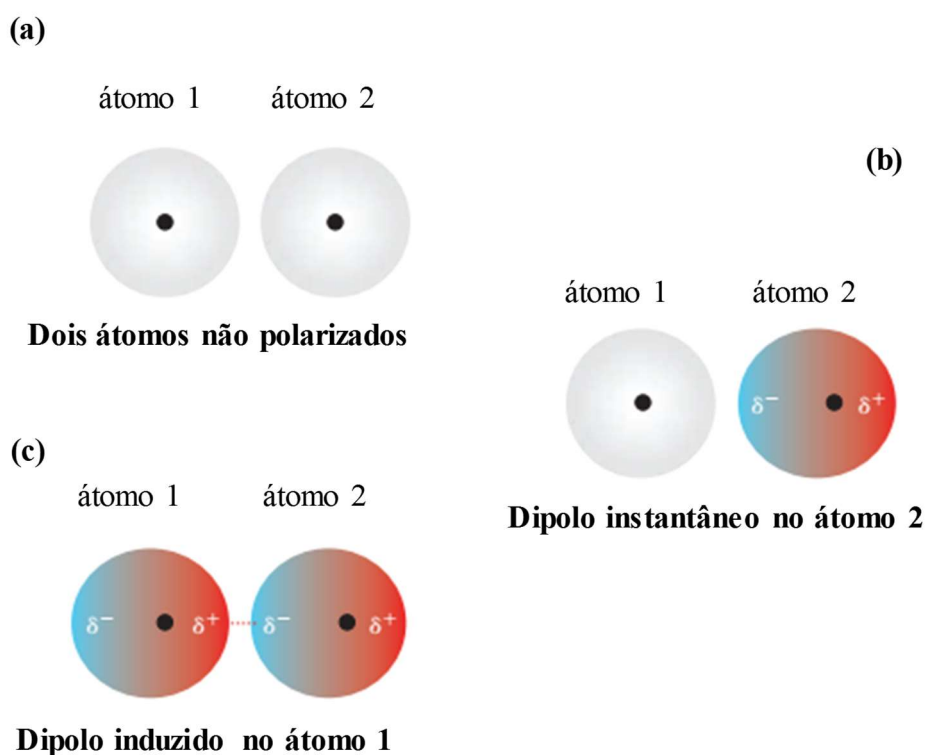


Figura 2.6 - Ilustração do processo fenomenológico do surgimento das forças de dispersão (McMURRY, 2011).

A facilidade com que ocorre esse movimento de elétron ou distorção é chamada de polarizabilidade da molécula. A polarizabilidade aumenta com o aumento do número de elétrons em um átomo ou molécula. Assim, deve-se esperar que a força de dispersão aumente com o aumento da massa molar de uma substância.

2.4.2 Atração dipolo-dipolo

As forças dipolo-dipolo ocorrem em moléculas polares, como é o caso de moléculas contendo o grupo carbonila “C=O”. Essas forças originam-se da atração entre as partes parcialmente positiva e negativa de duas moléculas vizinhas. Em geral, e com algumas exceções, as forças dipolo-dipolo são mais fortes que as forças de dispersão. A comparação, por exemplo, do ponto de ebulição de duas substâncias com massa molar (MM) similar explica o efeito dessas forças: o propano (MM = 44 g mol⁻¹), que é apolar, possui ponto de ebulição (231 K) menor que o da acetonitrila (355 K), que tem massa molar de 41 g mol⁻¹; porém é uma molécula polar e tem a presença de forças dipolo-dipolo.

2.4.3 Ligação de hidrogênio

As forças de ligação de hidrogênio são as forças atrativas mais fortes entre as forças intermoleculares existentes. São típicas em moléculas polares em que há um átomo de hidrogênio (H) ligado a átomos altamente eletronegativos, usualmente flúor (F), oxigênio (O) ou nitrogênio (N). O hidrogênio é um átomo extremamente pequeno e praticamente ausente de elétrons, o que faz com que ele fique com a parte parcial positiva da molécula e sua ligação com um átomo muito eletronegativo de uma molécula vizinha seja forte (Figura 2.7) (McMURRY, 2011).

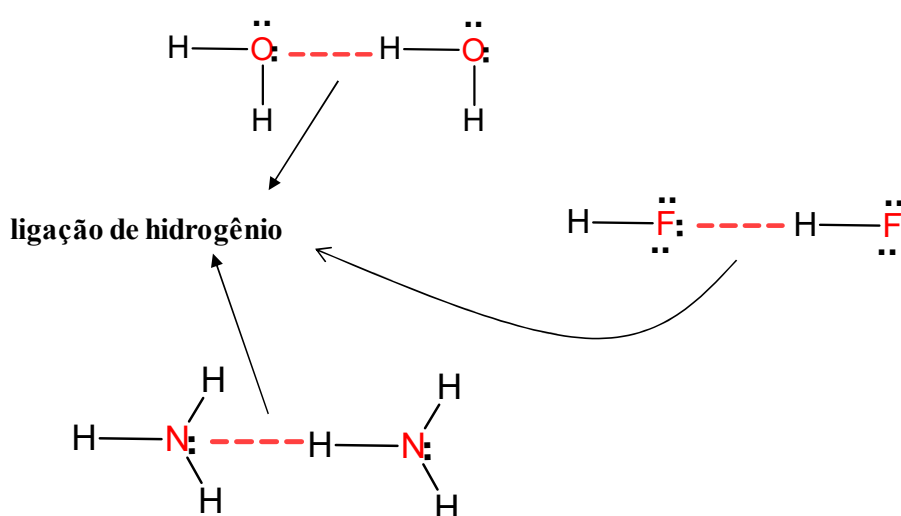


Figura 2.7 - Presença de ligação de hidrogênio em moléculas polares contendo átomo de hidrogênio ligado a átomo altamente eletronegativo.

A geometria molecular de uma substância também influencia a magnitude das forças intermoleculares e, conseqüentemente, as suas propriedades termodinâmicas. Por exemplo, ambos os compostos *n*-pentano e neopentano são apolares e têm fórmula química C_5H_{12} . No entanto, o *n*-pentano possui ponto normal de ebulição (T_{nb}) de 309,4 K, enquanto o neopentano apresenta T_{nb} de 287,2 K. Isso deve-se a forma geométrica das moléculas desses dois compostos. O *n*-pentano tem forma linear, enquanto o neopentano tem formato esférico (Figura 2.8). A geometria do *n*-pentano facilita o contato intermolecular e aumenta a força de dispersão. Um paralelo pode ser feito para moléculas polares, onde moléculas similares podem ter propriedades diferentes por influências de partes apolares e das ramificações.

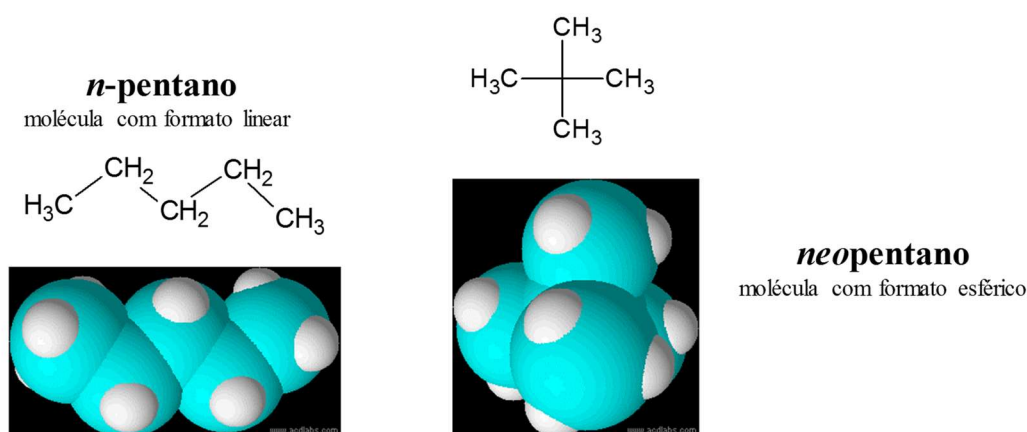


Figura 2.8 - Formato molecular do *n*-pentano (linear) e do neopentano (esférico) (baseado em McMURRY, 2011).

Outro fator importante é o caso de as forças atrativas terem efeito aditivo, ou seja, moléculas que têm mais de uma força atuante devem apresentar propriedades distintas. Por exemplo, tanto o propan-1-ol ($CH_3CH_2CH_2OH$, $MM = 60 \text{ g mol}^{-1}$) como o ácido acético (CH_3COOH , $MM = 60 \text{ g mol}^{-1}$) são polares, fazem ligação de hidrogênio e possuem a mesma massa molar. Porém, o propan-1-ol tem $T_{nb} = 370 \text{ K}$, à medida que o ácido acético apresenta T_{nb} igual a 391 K. Neste caso, o propan-1-ol faz uma ligação de hidrogênio; já o ácido acético faz duas: o hidrogênio e o oxigênio da carbonila ($C=O$) de uma molécula faz ligação de hidrogênio com os mesmos grupos da molécula vizinha (Figura 2.9).

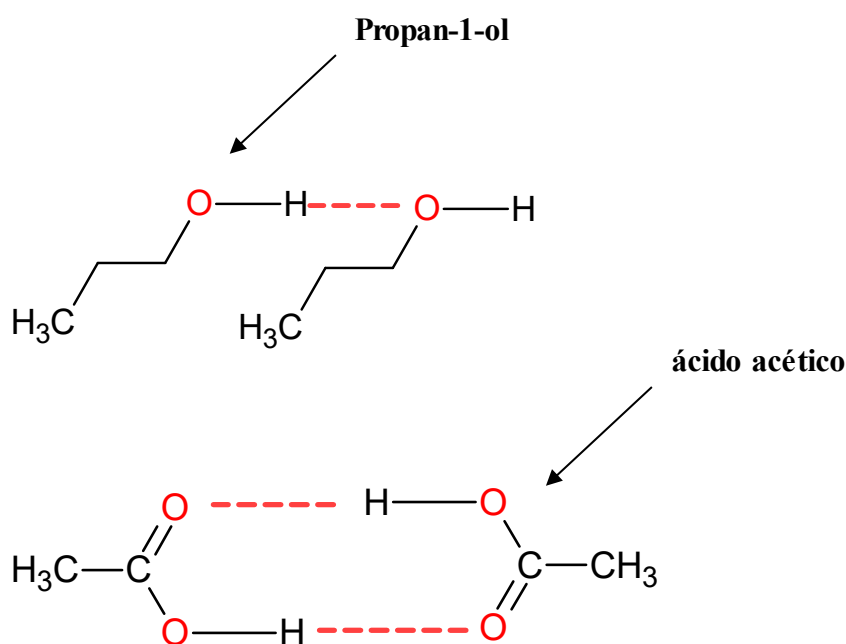


Figura 2.9 - Efeito da ligação de hidrogênio em moléculas polares diferentes.

2.5 Importância das propriedades termodinâmicas de substâncias orgânicas

Assim como um engenheiro civil não pode projetar um edifício sem o conhecimento das características do solo, das propriedades do aço e do concreto, os cientistas e engenheiros químicos frequentemente utilizam propriedades de gases e líquidos para o projeto de muitos tipos de produtos, processos e equipamentos industriais (POLING *et al.*, 2001). Além disso, à medida em que a síntese de novos compostos se torna mais eficiente, a necessidade de conhecimento de suas propriedades físico-químicas é cada vez mais importante em engenharia química e em outras áreas correlatas (LIAN e YALKOWSKY, 2014). Em cálculos de projetos químicos, por exemplo, os engenheiros necessitam de propriedades termodinâmicas como, especificamente, temperatura normal de ebulição, entalpia de vaporização e pressão de vapor.

A temperatura normal de ebulição (T_{nb}) é definida como a temperatura na qual a pressão de vapor de um líquido é igual a pressão atmosférica (POLING *et al.*, 2001). É uma grandeza física constante e inerente a cada substância (VOGEL, 1966). Esta propriedade destaca-se, ao lado do ponto de fusão, como um dos critérios básicos na determinação do grau de pureza de substâncias orgânicas (DIAS *et al.*, 2014). O recente trabalho de Deng *et al.* (2016) e o trabalho de Calm e Hourahan (2011) indicam a importância e a potencial utilidade desta propriedade em aplicações com fluidos de refrigeração.

Assim como a T_{nb} , a pressão de vapor (P_{vap}) também é uma propriedade intrínseca das substâncias e é dependente da temperatura. A pressão de vapor tem sido uma das propriedades termodinâmicas mais difundidas na comunidade acadêmica (VAN HECKE, 1992). Além de aplicações bem conhecidas como em processos de equilíbrio líquido-vapor e destilação, o conhecimento desta grandeza é indispensável para a modelagem do destino de produtos químicos no meio ambiente (RŮŽIČKA *et al.*, 1994). Por exemplo, P_{vap} tem sido combinada com dados de solubilidade aquosa em cálculos das constantes da lei de Henry, na modelagem de equilíbrios de compostos orgânicos dissolvidos em água (GOLDFARB, 2013); e recentemente foi aplicada em cálculos de coeficientes de partição entre água e CO₂ supercrítico (BURANT *et al.*, 2016), além de sua bem estabelecida aplicação em relações PVT (ABDULAGATOV *et al.*, 2016).

A entalpia de vaporização ou entalpia molar de vaporização (ΔH_{vap}) é a quantidade de energia necessária para transformar uma quantidade de massa de um líquido para a sua fase vapor (MAJER *et al.*, 1989). Ao lado de T_{nb} e P_{vap} , a entalpia de vaporização de compostos puros tem sido um parâmetro termodinâmico muito importante para a indústria química, principalmente em processos que envolvem destilação e evaporação (KOLSKÁ *et al.*, 2005). Além disso, as entalpias de vaporização são de suma importância em aplicações como termodinâmica e termoquímica, uma vez que estão relacionadas a propriedades de fases vapor e líquida e podem ser usadas em vários tipos de cálculos de conversão entre essas duas fases (MAJER *et al.*, 1989). Devido a sua grande importância, a determinação de ΔH_{vap} vem sendo tema de pesquisas recentes e tem sido determinada para inúmeras substâncias. Mori *et al.* (2016); e Zeng *et al.* (2014) determinaram as entalpias de vaporização e sublimação de compostos nitrogenados. Gobble e Chickos (2016) e Zaitsau *et al.* (2015) determinaram a pressão de vapor e ΔH_{vap} de compostos oxigenados. Verevkin *et al.* (2014) determinaram ΔH_{vap} de compostos halogenados.

2.5.1 Equação de Clausius-Clapeyron e a relação da entalpia de vaporização com a pressão de vapor

A condição de equilíbrio ($\Delta G = 0$) para um sistema onde a fase vapor de uma substância está em equilíbrio com seu líquido (MAJER *et al.*, 1989; SMITH *et al.*, 2007) é tal que

$$dG_s^l = dG_s^g \quad (2.1)$$

em que o subscrito s indica uma mudança no equilíbrio.

A diferencial total da energia de Gibbs é dada pela relação

$$dG = -S dT + V dp \quad (2.2)$$

Aplicando a Equação 2.2 para as fases vapor e líquida em equilíbrio, tem-se

$$V_s^l dp - S_s^l dT = V_s^g dp - S_s^g dT \quad (2.3)$$

em que, rearranjando-se a Equação 2.3, resulta na equação de Clapeyron

$$\frac{\Delta S}{\Delta V} = \frac{\Delta H}{T \Delta V} = \left(\frac{dp}{dT} \right) \quad (2.4)$$

As variações de entalpia e de volume são sempre positivas. Assim, a pressão de vapor sempre aumenta com o aumento da temperatura.

Em regiões de temperatura moderada (geralmente abaixo do ponto de ebulição), os efeitos do desvio da idealidade da equação de Clapeyron ($\Delta V = \Delta ZRT/p_s$) são minimizados e a Equação 2.4 torna-se a equação de Clausius-Clapeyron

$$\Delta H = RT^2 \left(\frac{d \ln p}{dT} \right)_s \quad (2.5)$$

em que o volume da fase líquida é completamente desprezado comparado ao da fase vapor.

A integração da equação de Clausius-Clapeyron

$$\ln P = -\frac{\Delta H}{RT} + \text{constante} \quad (2.6)$$

na qual ΔH independente de T , permitindo a determinação da entalpia de vaporização por meio do logaritmo da pressão de vapor e o inverso da temperatura.

2.6 Obtenção de grandezas termodinâmicas

Os métodos convencionais de determinação de grandezas termodinâmicas são divididos em dois grupos. O primeiro é considerado como *métodos diretos* e em geral se

baseiam em medidas calorimétricas. O segundo grupo é chamado *métodos indiretos* e as propriedades termodinâmicas são obtidas a partir de outros dados.

Dos vários métodos diretos encontrados na literatura, podem ser destacados os calorimétricos (condensação e vaporização), manométricos e isoteniscópicos (método estático), ebuliométricos (método dinâmico) e as células de Knudsen (método da efusão). Quanto aos principais métodos indiretos, podem ser citadas as determinações termogravimétricas e cromatográficas (MAJER *et al.* 1989). Um resumo das vantagens e desvantagens dos principais métodos convencionais pode ser visto na Tabela 2.3, conforme apresentado por Oliveira (2012).

Tabela 2.3 - Vantagens e desvantagens dos principais métodos convencionais de obtenção de dados termodinâmicos (OLIVEIRA, 2012).

Método	Vantagens	Desvantagens
Estático	São empregados em pressões moderadas e resultados precisos.	As substâncias devem ser completamente puras. A variação do volume que pode ocorrer durante a medida real como resultado do movimento do sensor do manômetro ou do detector nulo pode afetar a pressão de vapor da substância e conduzir a um aumento de temperatura.
Dinâmico	Pode ser acoplado com outro ebuliômetro para eliminar erros relacionados ao vapor superaquecido que não está em equilíbrio na mistura (líquido-vapor).	Medem temperaturas de ebulição e condensação e a diferença entre elas pode chegar a 0,005 K se a substância for completamente pura e se não ocorrer decomposição.
Efusão	A aparelhagem é bem completa, o que evita falhas na medição dos dados.	Requer grande quantidade de amostra. Além disso, o tempo de análise é relativamente longo (3-6h).
Cromatografia	Pode-se analisar substâncias em misturas. Não necessita de amostras completamente puras e a quantidade de amostra necessária é extremamente pequena. É usado em temperaturas elevadas e produz os dados necessários rapidamente. Também é aplicado em substâncias estruturalmente complexas e de médio ou alto ponto de ebulição.	A desvantagem deste método é que ele depende de substâncias de referência com dados de pressão conhecidos. Assim sendo, a precisão dos dados de termodinâmicos da substância estudada vai depender da precisão dos valores das substâncias de referência.
Termogravimetria	Requer pouca quantidade de amostra. O tempo efetivo para a realização dos experimentos é relativamente pequeno e exibe boa reprodutibilidade experimental e os resultados são bastante comparáveis com os métodos convencionais.	Necessita de substância de referência com certa similaridade com a espécie estudada. Substâncias com baixa estabilidade térmica podem dificultar a reprodutibilidade experimental. A quantidade de amostra utilizada e o tipo de atmosfera do sistema podem afetar os padrões dos dados termogravimétricos, o que pode afetar na precisão dos resultados.

As técnicas de medidas direta são geralmente consideradas mais precisas. No entanto, têm aplicações limitadas a compostos de massa molar baixa ou muito voláteis (GRAYSON e FOSBRAEY, 1982). Desse modo, os métodos experimentais indiretos vêm sendo adotados. Dentre os métodos indiretos conhecidos, cita-se o método da *taxa de volatilização relativa*, o qual é aplicado em compostos insolúveis em água (DOBBS, 1984); o método da análise termogravimétrica, onde a pressão de vapor de uma substância é determinada por meio da taxa de perda de massa da amostra em função da temperatura (HAZRA *et al.*, 2004; CREMASCO e BRAGA, 2011); e o método de correlação por cromatografia gasosa, em que grandezas termodinâmicas são obtidas através de dados de retenção cromatográfica (CHICKOS *et al.*, 1995; SVOBODA e KOUTEK, 2002; HOSKOVEC *et al.*, 2005; RŮŽIČKA *et al.*, 2012; GOODRICH *et al.*, 2016; GOBBLE *et al.*, 2016).

Encontram-se na literatura vários trabalhos que usaram métodos diretos e indiretos para a obtenção de propriedades termodinâmicas. Farritor e Tao (1970); e Etzler e Connors (1991) determinaram a entalpia de vaporização de líquidos voláteis diretamente por calorimetria de varredura diferencial (DSC). Kruif *et al.* (1981) determinaram dados de P_{vap} e ΔH_{vap} do naftaleno nos estados líquido e sólido utilizando o método estático empregando manômetros. Silva e Monte (1990) determinaram as entalpias de sublimação e pressões de vapor de compostos organometálicos sólidos e orgânicos por meio do método da efusão em célula de Knudsen.

Em relação às técnicas indiretas, Nichols *et al.* (2006) avaliaram entalpias de vaporização, fusão e sublimação de uma série de alcanos por correlações de cromatografia gasosa (método indireto). Dollimore *et al.* (1992), Shen e Alexander (1999), Chatterjee *et al.* (2001), Hazra *et al.* (2002), Wright *et al.* (2004), Oliveira e Cremasco (2014) utilizaram a análise termogravimétrica (método indireto) como ferramenta de obtenção de grandezas termodinâmicas de várias substâncias.

2.6.1 Obtenção de dados termodinâmicos por cromatografia gasosa (CG)

Conforme mostrado na Tabela 2.3, a técnica CG possui algumas vantagens relevantes em relação aos outros métodos, como a possibilidade de se trabalhar com misturas, o que pode diminuir o consumo de reagentes e o tempo de análise, de não necessitar de substâncias completamente puras e de aplicar ampla faixa de temperatura, que dependerá das

características da coluna, como tipo e espessura da fase estacionária. Além disso, existem diversas colunas capilares comercialmente disponíveis, aplicáveis a uma gama de substâncias orgânicas (GROB, 1977; COLLINS *et al.*, 2006).

Desse modo, a técnica CG tem tido grande destaque e tem sido demonstrado por alguns pesquisadores como uma boa ferramenta para determinação de propriedades termodinâmicas de compostos orgânicos (KOUTEK *et al.*, 1997; van ROON *et al.*, 2002; LETCHER e NAICHER, 2004). Hoskovec *et al.* (2005) determinaram dados de pressão de vapor, entalpia de vaporização e temperatura normal de ebulição de 32 compostos orgânicos terpênicos por CG. Os autores encontraram pequenos desvios comparando seus resultados com os dados da literatura.

Dessa maneira, na presente Tese de Doutorado, a técnica CG foi utilizada para obtenção de dados termodinâmicos (P_{vap} , ΔH_{vap} e T_{nb}) para alguns terpenos selecionados. A teoria que relaciona dados de retenção cromatográfica com quantidades termodinâmicas está descrita a seguir.

2.6.1.1 Teoria

Os processos de separação por cromatografia gasosa (CG) envolvem equilíbrio químico entre duas fases. As análises clássicas em CG envolvem a comparação dos dados de retenção de uma amostra desconhecida com os de uma amostra conhecida (GROB, 1977).

Um pico cromatográfico traz informações importantes como tempo de residência ou a diferença entre a eluição entre dois picos, forma e tamanho. Geralmente, o tempo de retenção ajustado ou volume de retenção ajustado são utilizados em análises qualitativas.

Uma vez que os tempos de retenção de um componente desconhecido e conhecido podem ser o mesmo, as separações cromatográficas são utilizadas o índice de retenção de Kováts (KAISER, 1977),

$$I_R = 100 \cdot z + 100 \left[\frac{\ln t'_{R,X} - \ln t'_{R,z}}{\ln t'_{R,z+1} - \ln t'_{R,z}} \right] \quad (2.7a)$$

ou em termos de volume de retenção

$$I_R = 100 \cdot z + 100 \left[\frac{\ln V'_{R,X} - \ln V'_{R,z}}{\ln V'_{R,z+1} - \ln V'_{R,z}} \right] \quad (2.7b)$$

nas quais o subscrito X representa a substância de interesse e os subscritos z e $z+1$ referem-se aos n -alcanos de referência.

Analisando a Equação 2.7a, pode-se relacionar o volume de retenção com a equação dos gases ideais,

$$PV = RT \quad (2.8)$$

A pressão parcial de cada componente pode ser obtida aplicando a lei de Raoult (SMITH *et al.*, 2007),

$$y_i P = x_i P_i \quad (2.9)$$

na qual x_i é a fração molar da fase líquida, y_i é a fração molar da fase vapor, e P_i é a pressão de vapor da espécie i pura na temperatura de trabalho. O lado esquerdo da Equação 2.9 ($y_i P$) refere-se a pressão parcial da espécie i .

Fazendo-se a somatória das frações molares na fase vapor igual a 1 ($\sum_i y_i = 1$) e a massa molar da mistura ou da solução como o produto da fração molar da fase líquida de cada componente por sua massa molar ($M = \sum_i x_i M_i$) e substituindo a Equação 2.8 na Equação 2.9, obtém-se

$$\frac{RT}{V} = MP_i \quad (2.10)$$

Visto que poucas aplicações têm condições ideais, é necessário levar em conta os desvios da idealidade, γ_i^∞ , na Equação 2.10. Resolvendo tal equação V , obtém-se

$$V = \frac{RT}{\gamma_i^\infty P_i M} \quad (2.11)$$

em que γ_i^∞ é o coeficiente de atividade na diluição infinita da espécie i .

Aplicando a Equação 2.11 para o volume de retenção relativo (ou tempo) entre uma substância de referência (n -alcano com z átomos de carbono) e uma substância de interesse X , ou seja, $V_{R,z} / V_{R,X}$ (HINCKLEY *et al.*, 1990),

$$\frac{V_{R,z}}{V_{R,X}} = \frac{RT}{P_z \gamma_z M} \cdot \frac{P_X \gamma_X M}{RT} = \frac{P_X \gamma_X}{P_z \gamma_z} \quad (2.12)$$

Substituindo a Equação 4.12 na Equação 2.7a ou 2.7b e fazendo o mesmo tratamento matemático para uma substância de referência número de carbonos z e $z+1$, resulta em

$$I_R = 100 \cdot z + 100 \left[\frac{\ln(\gamma_z^\infty P_z / \gamma_X^\infty P_X)}{\ln(\gamma_z^\infty P_z / \gamma_{z+1}^\infty P_{z+1})} \right] \quad (2.13)$$

Assumindo $\gamma_X^\infty = \gamma_z^\infty = \gamma_{z+1}^\infty = \gamma$ e expressando a Equação 4.13 em termos de P_X , obtém-se (HOSKOVEC *et al.*, 2005)

$$\ln P_X = \ln P_z + 100 \left[\frac{(100 \cdot z - I_R) \cdot \ln(P_z / P_X)}{\ln(P_z / P_{z+1})} \right] \quad (2.14)$$

a qual relacionada a pressão de vapor de uma substância de interesse X (P_X) com o índice de retenção de Kováts (I_R) e as pressões de vapor de duas substâncias de referência (P_z e P_{z+1}), e tem sido utilizada calcular a pressão de vapor de diversas substâncias orgânicas (FISCHER *et al.*, 1992; SPIEKSMÁ *et al.*, 1994; FISCHER e BALLCHNITER, 1998).

Vários autores procuraram expressar a dependência do índice de retenção de Kováts com a temperatura. Sob este aspecto, têm-se informações desde relações lineares (ETTRE e BILLEB, 1967), modelos logarítmicos (TAKÁCS *et al.*, 1969), bem como modelos hiperbólicos (TUDOR, 1999) para expressar tal dependência. Hoskovec *et al.* (2005) exploraram uma relação baseada em uma série de Taylor, originalmente descrita por Clarke e Glew (1965), em que a variação de entalpia de um processo pode ser expressa como uma perturbação do valor da entalpia padrão a uma temperatura de referência, levando a uma equação geral do tipo

$$\ln K(T) = \alpha + \frac{\beta}{T} + \gamma \ln(T) + \sigma T + \xi T^2 \dots \quad (2.15)$$

A Equação 2.15 está relacionada com grandezas termodinâmicas tais como capacidade calorífica, entropia e entalpia, podendo ser obtidas via análise cromatográfica por meio de ajuste de medidas de volumes de retenção (ou tempo), pressão de vapor, constantes de equilíbrio, solubilidades, bem como pelo próprio índice de Kováts, em função da temperatura (CLARKE e GLEW, 1965). Além disso, esta equação tem sido utilizada até ao terceiro termo

e demonstrou-se eficiente na determinação de grandezas termodinâmicas de aminas e cetonas (CIAZYNSKA-HALAREWICZ e KOWALSKA, 2003), aldeídos (HÉBERGER *et al.*, 2002) e compostos monoterpênicos e sesquiterpênicos (HOSKOVEC *et al.*, 2005). Dessa maneira, a partir da Equação 2.15:

$$I(T) = a_0 + \frac{a_1}{T} + a_2 \ln(T) \quad (2.16)$$

em que a_0 , a_1 e a_2 são constantes determinadas experimentalmente por meio de ajustes por mínimos quadrados usando a técnica de regressão não linear; e T é a temperatura expressa em Kelvin.

A Equação 2.16 tem sido usada para expressar a relação $I = f(T)$ por muitos pesquisadores (HÉBERGER *et al.*, 2002; GÖRGÉNYI e HÉBERGER, 2004; HÉBERGER *et al.*, 2004; MIJIN e ANTONOVIC, 2004; PÉREZ-PARAJÓN *et al.*, 2004). No entanto, a equação geral (Equação 2.15) também tem sido usada anteriormente até o quarto termo (VEZZANI *et al.*, 1994; 1999).

2.6.2 Obtenção de entalpia de vaporização por métodos de predição

Como visto na seção anterior, a entalpia de vaporização pode ser determinada por diversas técnicas analíticas. A principal vantagem das técnicas experimentais é que elas podem apresentar resultados tanto exatos como precisos. Contudo, a maioria tem custo elevado e requerem boa prática laboratorial e escolha correta dos compostos de referência. Quando se trata de substância de preço elevado, algumas podem ser inviáveis de serem analisadas experimentalmente. Além disso, frequentemente, muitos produtos químicos são misturas de várias substâncias; a determinação de dados termodinâmicos de cada substância na mistura é difícil, quando eles não estão disponíveis na forma pura. Assim, os modelos preditivos surgem como uma alternativa interessante para a obtenção de grandezas termodinâmicas.

Os métodos usuais de estimação de entalpia de vaporização baseiam-se no princípio dos estados correspondentes (PEC). Os métodos baseados nessa teoria necessitam de dados críticos (como pressão e temperatura) que geralmente não estão disponíveis para muitos compostos, principalmente novas substâncias descobertas. Por outro lado, existem os métodos preditivos baseados no princípio de contribuição de grupos, cujos modelos necessitam apenas da estrutura química da molécula e, por isso, são frequentemente utilizados pela comunidade

científica. A seguir, são apresentados alguns modelos encontrados na literatura usados para predição de entalpias de vaporização de substâncias orgânicas.

2.6.2.1 Modelos baseados no princípio dos estados correspondentes

Método de Basarová e Svoboda (1995)

Basarová e Svoboda (1995) desenvolveram um método de contribuição de grupos baseado na seguinte equação

$$\Delta H_{vap} = A(1 - T_r)^\alpha e^{(-\alpha T_r)} \quad (2.17)$$

em que T_r é a temperatura reduzida utilizada em Kelvin; A e α são os parâmetros ajustáveis, expressos como a soma das contribuições dos grupos:

$$A = \sum_j \nu_j P_j(A) \quad (2.18)$$

$$\alpha = \sum_j \nu_j P_j(\alpha) \quad (2.19)$$

onde o subscrito j denota o tipo de contribuição (grupo); ν_j é o número de contribuições e $P_j(A)$ e $P_j(\alpha)$ são dos valores dessas contribuições.

O modelo utilizou 307 substâncias, distribuídas em hidrocarbonetos (44) e derivados de hidrocarbonetos (263). Este último incluiu: derivados de halogênios (76), éteres (13), álcoois (31), fenóis (4), aldeídos (4), cetonas (20), oxigênio heterocíclicos (2), além de compostos sulfurosos, tiocompostos e derivados de nitrogênio. Segundo os autores, o modelo foi testado na predição em substâncias com dados de ΔH_{vap} a 298,15 K e na temperatura normal de ebulição, encontrando desvio relativo médio de 1,4 % e 1,6 %, respectivamente. Além disso, reportaram desvio relativo médio de 1,9 % em um conjunto de dados de 113 dados de ΔH_{vap} (próximo à 298,15 K) medidos por calorimetria.

Método de Tu e Liu (1996)

Tu e Liu (1996) estudaram quatorze relações que correlacionam a dependência da temperatura da entalpia de vaporização e desenvolveram um método baseado na equação

$$\Delta H_{vap} = \sum_i N_i [a_i(1-T_r)^{1/3} + b_i(1-Tr)^{2/3} + c_i(1-T_r)] \quad (2.20)$$

em que N_i é o número de vezes que um grupo i se repete na molécula e a_i , b_i e c_i são as contribuições do grupo i ; T_r é a temperatura reduzida em Kelvin. O modelo é aplicável para predições até uma temperatura de 750 K e entalpia de vaporização de até 100 kJ mol⁻¹.

O modelo consistiu de 509 compostos orgânicos com 3.396 pontos de dados, dos quais: 1.743 foram hidrocarbonetos, 1.186 foram derivados de hidrocarbonetos, sendo: aromáticos (543), álcoois (232), aldeídos (20), éteres (73), cetonas (81), ácidos (49), ésteres (188). Os demais foram divididos entre compostos nitrogenados, derivados de halogênios, sulfetos, etc.

O desvio relativo médio encontrado pelos autores foi de 2,5 %. No entanto, os autores testaram modelos de outros pesquisadores, incluindo o de Basarová e Svoboda (1995), com o mesmo número de pontos de dados e tiveram melhor desempenho.

2.6.2.2 Modelos baseados somente na estrutura química

Modelo de Joback e Reid (1987)

O modelo de Joback e Reid (1987) é amplamente utilizado devido à sua praticidade. O modelo é capaz de estimar 11 tipos de propriedades termodinâmicas diferentes via contribuição de grupos. É possível distinguir os grupos que fazem parte de anel (aromático ou cíclico) e os que não fazem parte de anel, além de grupos oxigenados, nitrogenados, halogenados e sulfurados. A predição da entalpia de vaporização na T_{nb} de substâncias orgânicas é feita pela seguinte equação

$$\Delta H_{vap}(T_{nb}, K) = 15,30 + \sum_i (N_i h_{v,i}) \quad (2.21)$$

em que N_i é a quantidade de repetições no grupo i na molécula; $h_{v,i}$ é a contribuição do grupo i .

Modelo de Chickos et al. (1998)

O método de Chickos *et al.* (1998) pode ser usado para determinar entalpias de vaporização (a 298,15 K) e dados entropia. O modelo é relativamente simples de ser aplicado

e requer apenas informação sobre o número de carbonos da cadeia carbônica (N_C) e o número de carbonos quaternários (N_Q). A predição é feita segundo,

$$\Delta_{vap} H_m^0(298,15 \text{ K}) = 4,69(N_C - N_Q) + 1,3N_Q + 3,0 + \sum_i n_i F_i b_i + C \quad (2.22)$$

em que b é a contribuição do grupo funcional; C é um parâmetro de correção para compostos de natureza cíclica, aromática e grupos alquil ramificados em carbonos sp^3 ; o termo F_i corrige a natureza da substituição (carbono primário, secundário, terciário ou quaternário).

Modelo de Kolská et al. (2005)

O modelo desenvolvido por Kolská *et al.* (2005) é baseado nos critérios de Marrero e Gani (2001), onde a estimação de propriedades é feita em três tipos de procedimentos. Os grupos são divididos em níveis de primeira, segunda e terceira ordens. Basicamente, os grupos de primeira ordem permitem descrever ampla variedade de compostos orgânicos e devem compreender completamente a molécula. Em outras palavras, a molécula deve ser fragmentada de modo que todo fragmento deve ser único, sem haver qualquer parte (ou átomo) de um fragmento em outros fragmentos.

Os grupos de segunda ordem são usados para prever interação entre grupos e, na fragmentação, pode haver sobreposição de grupos (um fragmento pode fazer parte de outro). Já os grupos de terceira ordem geralmente ocorrem em compostos policíclicos. A Figura 2.10 exemplifica como os três níveis de contribuição de grupo podem ser aplicados. Na Figura 2.10a, a molécula de benzil fenil éter é fragmentada em 10 grupos CH (aromáticos), um grupo aC-CH₂ (carbono aromático) e um grupo aC-O (carbono aromático). Na Figura 2.10b, a L-anilina compartilha o grupo CH com os grupos NH₂ e CO-OH. A Figura 2.10c, exemplifica a interação de terceira ordem para três anéis aromáticos fundidos. Também é possível haver interação de terceira ordem de anéis fundido com anéis não-aromáticos e de anéis não fundidos conectados por uma cadeia.

A predição de ΔH_{vap} a 298,15 K pelo modelo de Kolská *et al.* (2005) é baseado na seguinte equação,

$$\Delta H_{vap} (298,15 \text{ K}) = h_0 + \sum_i N_{h0,i} C_{h0,i} + \sum_j M_{h0,j} D_{h0,j} + \sum_k O_{h0,k} E_{h0,k} \quad (2.23)$$

em que $N_{h0,i}$, $M_{h0,j}$ e $O_{h0,k}$ indicam o número de vezes que o grupo de primeira, segunda e terceira ordem aparece na estrutura da molécula; $C_{h0,i}$, $D_{h0,j}$ e $E_{h0,k}$ representam as contribuições correspondentes para a entalpia de vaporização.

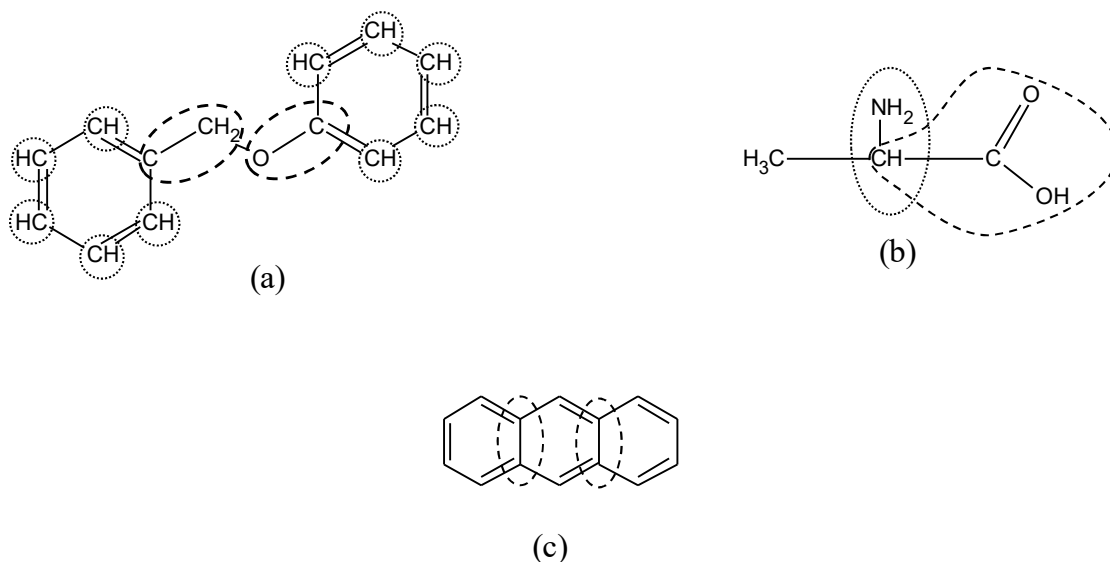


Figura 2.10 - Descrição correta do modelo de Kolská *et al.* (2005) para: **(a)** grupos de primeira ordem; **(b)** grupos de segunda ordem; **(c)** grupos de terceira ordem.

2.7 Desenvolvimento de modelos preditivos

Em etapas de desenvolvimento de modelos é fundamental a utilização de ferramentas estatísticas. As duas grandes áreas da estatística são a descritiva e a inferencial. A estatística descritiva cuida da parte de organização e descrição de um conjunto de dados, ao passo que a estatística inferencial se preocupa com os métodos que tornam possível a estimação de características (ou parâmetros) dos dados (MARTINS, 2005). Uma vez que os conceitos de estatística descritiva estão amplamente apresentados em muitos livros textos, não serão comentados no presente trabalho. Entretanto, faz-se necessário apresentar algumas definições sobre métodos de estimação de parâmetros, que neste trabalho serão abordados em termos de aprendizagem de máquina.

2.7.1 Aprendizagem de máquina

O termo aprendizagem de máquina (AM) vem sendo muito utilizado para descrever processos em que um computador é utilizado para resolver problemas do mundo real. Existem várias formas de abordagem sobre a aprendizagem de máquina. As mais comuns são a aprendizagem *supervisionada* e a *não supervisionada* (SHALEV-SHWARTZ e BEM-DAVID, 2014; MICHIE *et al.*, 1994).

A aprendizagem *supervisionada* baseia-se num conjunto de dados, a partir do qual usa-se um algoritmo para *generalizar* a resposta correta sobre todas as entradas possíveis, e geralmente envolvem métodos de regressão e de classificação. Em aprendizagem *não supervisionada*, um algoritmo tenta identificar similaridades entre as entradas de maneira que as entradas em comum sejam categorizadas (MARSLAND, 2015). O *clustering* é método mais usual em aprendizagem não supervisionada.

As principais tarefas de AM envolvem reconhecimento de padrões, diagnóstico, planejamento, controle robótico, previsão etc. Alguns pesquisadores usaram AM com sucesso para reconhecimento de fala a partir de um texto (SEJNOWSKI e ROSENBERG, 1987), controle adaptativo (NARENDRA e BALAKRISHNAN, 1987) e estimação de preços na indústria de construção (HEGAZY e MOSELHI, 1994).

No presente trabalho é usado o método de regressão em aprendizagem supervisionada e, portanto, será o foco para a sequência desta seção.

2.7.1.1 Regressão

A técnica de regressão é um dos métodos mais utilizados em aprendizagem de máquina. Esta técnica avalia a dependência entre as variáveis estudadas e responde quais delas afetam um certo valor. Por exemplo, o preço de um carro usado pode ser alterado conforme seu ano de fabricação, marca, capacidade do motor, quilometragem, etc. Para prever o preço de um determinado carro, pode-se usar a regressão com base em vendas realizadas anteriormente usando procedimentos de regressão (ALPAYDIN, 2010).

De forma simplificada, a tarefa de regressão pode entendida conforme ilustração da Figura 2.11, onde a coleção de *dados* é inspecionada e são retiradas as entradas X (atributos); os atributos são incluídos no modelo para *regressão* por meio da qual será feita a aprendizagem

da relação das entradas com o valor predito Y ; no bloco do *conhecimento* (Figura 2.11) tem-se o valor predito, o qual pode ser um valor contínuo ou uma resposta às entradas.

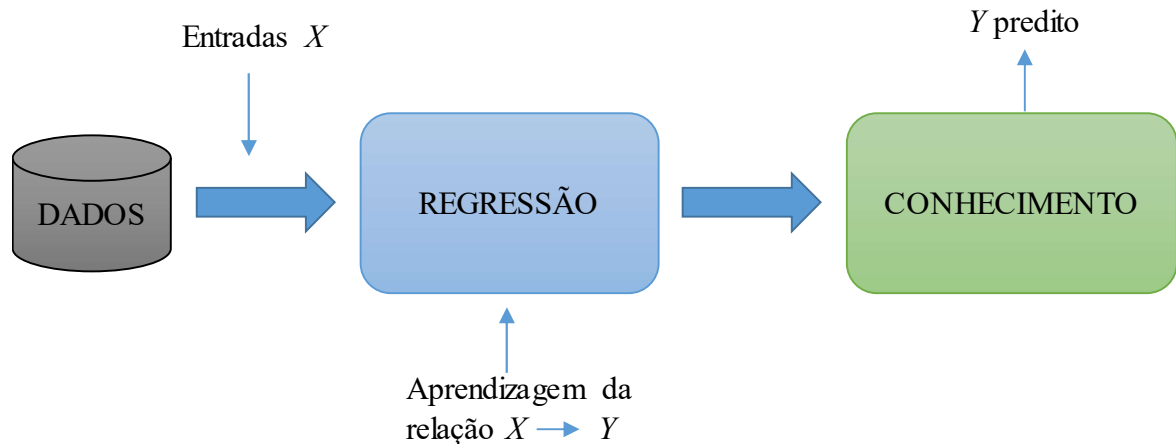


Figura 2.11 - Esquema simples de uma tarefa de regressão.

2.7.1.2 Regressão linear simples

Os modelos de regressão linear simples envolvem apenas uma variável independente e afirma que a média verdadeira da variável dependente muda a uma taxa constante quando a variável independente aumenta ou diminui (RAWLINGS *et al.*, 1998; WEISBERG, 2005). Frequentemente, esses modelos fornecem uma descrição interpretável e adequada de como as entradas (variáveis independentes) afetam a saída (variável dependente, resultado). Para fins de predição, os modelos lineares podem muitas vezes superar os modelos não lineares, além de estarem sujeitos a transformações, como logaritmos, raiz quadrada, expansões polinomiais e interações entre as entradas (HASTIE *et al.*, 2008). Em tarefas de regressão de aplicação real, o processo é feito como ilustrado na Figura 2.12.

Em modelos lineares, a equação da reta relaciona a média verdadeira da variável dependente Y_i , denominado por $\varepsilon(Y_i)$, e a variável dependente X_i

$$\varepsilon(Y_i) = w_0 + w_1 X_i \quad (2.24)$$

em que w_0 é o intercepto, ou o valor de $\varepsilon(Y_i)$ quando $X = 0$, e w_1 é o coeficiente angular da reta, que representa a taxa de mudança em $\varepsilon(Y_i)$ por unidade de mudança em X .

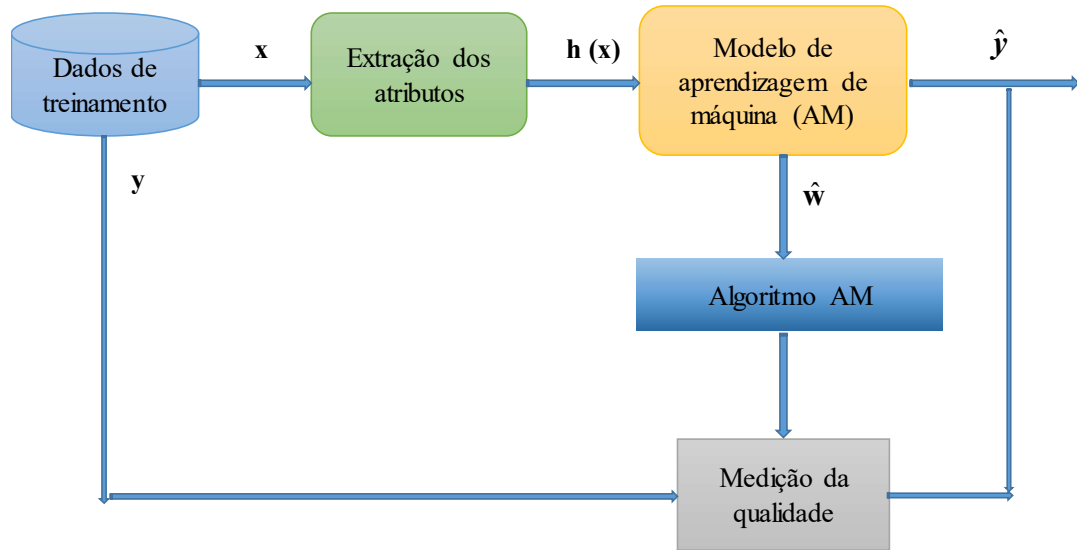


Figura 2.12 - Fluxograma de um processo de aprendizagem de máquina por regressão.

As observações sobre a variável dependente Y_i são assumidas como observações aleatórias de populações de variáveis aleatórias com a média de cada população, $\varepsilon(Y_i)$. Desse modo, o desvio de uma dada observação Y_i da sua média populacional $\varepsilon(Y_i)$ é levado em conta adicionando-se um erro (ou desvio) aleatório e_i

$$Y_i = w_0 + w_1 X_i + e_i \quad (2.25)$$

onde o subscrito i indica um observação particular.

2.7.1.3 Regressão ordinária por mínimos quadrados

Nos métodos de estimação por mínimos quadrados, o principal critério é que a solução deve dar a menor soma possível de desvios quadrados do Y_i observado em relação às estimativas de suas médias verdadeiras. Sendo \hat{w}_0 e \hat{w}_1 valores estimados dos parâmetros w_0 e w_1 . A média estimada de Y para cada X_i , com $i = 1, 2, \dots, n$, fica

$$\hat{Y}_i = \hat{w}_0 + \hat{w}_1 X_i \quad (2.26)$$

O princípio dos mínimos quadrados escolhe \hat{w}_0 e \hat{w}_1 que minimiza a soma dos quadrados residuais, SQR :

$$SQR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.27)$$

a qual também é conhecida como função objetivo.

2.7.1.4 Regressão linear múltipla

Muitos problemas reais não podem ser resolvidos por um modelo de regressão simples. O exemplo do carro (ilustrado no início do item 2.7.1.1) mostra que a estimação depende de muitos fatores (entradas, atributos ou *features*, termo usado no inglês). Nesse caso, o modelo de regressão linear múltipla pode conter quantas variáveis independentes forem necessárias para explicar uma estimação (RENCHEER, 2002). Assim, o modelo linear simples é estendido, adicionando-se outras variáveis independentes ou, usando termos da AM, aumentando a complexidade do modelo, tornando-se

$$Y_i = w_0 + w_1 X_{i1} + w_2 X_{i2} + \dots + w_p X_{ip} + e_i \quad (2.28)$$

A aplicação do método de estimação por mínimos quadrados na Equação 2.28 resulta na seguinte função objetivo

$$\begin{aligned} SQR &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\ &= \sum_{i=1}^n (Y_i - [\hat{w}_0 + \hat{w}_1 X_{i1} + \hat{w}_2 X_{i2} + \dots + \hat{w}_p X_{ip}])^2 \end{aligned} \quad (2.29)$$

a qual é usualmente utilizada em algoritmos de otimização como critério de minimização.

2.7.1.5 Regressão regularizada por mínimos quadrados

Como acontece nos métodos dos mínimos quadrados, a regressão regularizada minimiza uma função objetivo que inclui a soma dos quadrados residuais. No entanto, a função objetivo também inclui um termo constituído de um parâmetro de regularização positivo λ_2 multiplicado pela complexidade do modelo (VLAMING e GROENEN, 2015). Este parâmetro de regularização (ou penalidade, de *penalty*, do inglês) tem como objetivo evitar superajuste (*overfitting*) sobre os dados de treinamento e, assim, melhorar a capacidade de generalização (OKSER *et al.*, 2014).

As abordagens clássicas de regularização dependem do λ aplicado sobre os pesos do modelo. Os tipos mais comuns de regressão regularizada são a *lasso regression* (λ_1) e a *ridge regression* (λ_2). Esses dois termos têm sido usados em conjunto com a função objetivo em vários tipos de problemas de predição (WHITTAKER *et al.* 2000),

$$SQR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda_1 \sum_{j=1}^p \|w_j\| + \lambda_2 \sum_{j=1}^p \|w_j^2\| \quad (2.30)$$

Quando se utiliza apenas o parâmetro de regularização λ_2 (definindo $\lambda_1 = 0$), a regressão recebe o nome de *ridge regression* e o parâmetro λ_2 é usado para “encolher” o valor numérico (ou peso) das variáveis para prevenir que alguma variável tenha um valor muito grande e um efeito maior sobre o modelo. O parâmetro λ_2 tende a favorecer que o modelo dependa de todas as variáveis.

Em *lasso regression* (definindo $\lambda_2 = 0$), o parâmetro λ_1 é geralmente usado para seleção de variáveis (ou coeficientes) no próprio processo de ajuste modelo. Há outro caso em regressão regularizada que é a regressão denominada de *elastic net* (ZOU e HASTIE, 2003), a qual utiliza ambos os parâmetros de regularização λ_1 e λ_2 e é empregada para selecionar variáveis correlacionadas. Tanto *lasso* quanto *elastic net* são normalmente usadas em problemas cujo número de variáveis é extremamente alto em relação ao número de observações. No caso do presente trabalho de Tese, *lasso* e *elastic net* não têm influência significativa. Portanto, deu-se maior atenção ao termo do parâmetro de regularização λ_2 na função objetivo,

$$SQR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda_2 \sum_{j=1}^p w_j^2 \quad (2.31)$$

2.7.1.6 Algoritmos de otimização

Os algoritmos de otimização são partes fundamentais de uma tarefa de aprendizagem de máquina. O algoritmo é escolhido de acordo com o problema (NILSSON, 2005). No caso de predição, eles são os responsáveis por encontrar os valores mínimos ou minimizar a função objetivo. São usados para avaliar vantagens e desvantagens de projetos, avaliar sistemas de controle, encontrar padrões em dados, entre outras aplicações.

Otimização por gradiente descendente

O gradiente descende clássico minimiza uma função convexa e diferencial $f(w)$. Uma função é dita convexa se, para quaisquer vetores \mathbf{a} e \mathbf{b} , o gráfico da função encontra-se

abaixo do seguimento de linha que une os pontos $f(\mathbf{a})$ e $f(\mathbf{b})$, como mostra o exemplo de uma função convexa, apresentado na Figura 2.13.

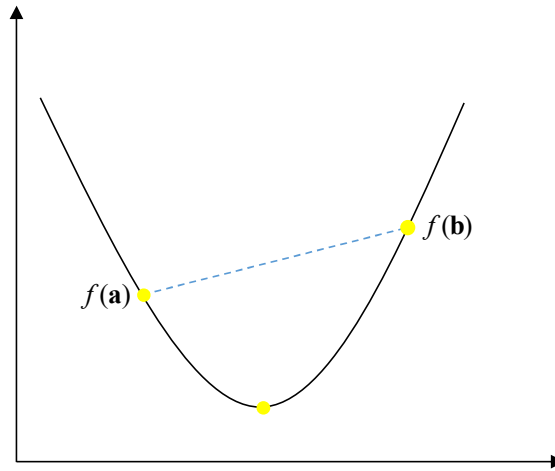


Figura 2.13 - Ilustração geral de uma função convexa.

O gradiente de uma função diferencial $f: \mathbb{R}^d \rightarrow \mathbb{R}$ em \mathbf{w} , denominado $\nabla f(\mathbf{w})$, é o vetor das derivadas parciais de f ,

$$\nabla f(\mathbf{w}) = \left(\frac{\partial f(\mathbf{w})}{\partial w[1]}, \dots, \frac{\partial f(\mathbf{w})}{\partial w[d]} \right) \quad (2.32)$$

O gradiente descendente é um algoritmo iterativo, onde começa-se com um valor inicial para os parâmetros \mathbf{w} e, à cada iteração, dá-se um passo na direção negativa do gradiente no ponto atual. O passo de atualização é dado por

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} - \eta \Delta f(\mathbf{w}^{(t)}) \quad (2.33)$$

onde t representa as iterações e η é o tamanho do passo em que o gradiente avança.

2.8 Avaliação de desempenho

Uma finalidade crucial em aprendizagem de máquina é a de avaliar o grau em que um modelo de predição é robusto e replicável. Em problemas de regressão, o desempenho de um modelo normalmente é avaliado por meio de uma *taxa de erro* (ou desvio, denominado

neste trabalho). Esta taxa de desvio é simplesmente a proporção de desvios feitos ao longo de um conjunto de dados, e mede o desempenho global da regressão.

2.8.1 Treinamento, teste e validação

Como visto anteriormente, em predição tenta-se minimizar os desvios sobre um determinado conjunto de dados. Esta etapa é chamada (usando termos da AM) de *treinamento* e os desvios advindo desta etapa é denominado de *desvio de treinamento*. O desvio de treinamento por si só não é uma boa medida de desempenho porque tende a diminuir com o aumento de número de dados (observações), como ilustrado na Figura 2.14a. Nesse caso, utiliza-se um novo conjunto de dados para medir o desempenho, denominado *dados de teste*, que são totalmente independentes dos dados de treinamento. A ideia, portanto, é fazer o treinamento e avaliar a capacidade preditiva nos dados de teste, o qual deve apresentar a menor taxa de desvio possível (NILSSON, 2005; ALPAYDIN, 2010).

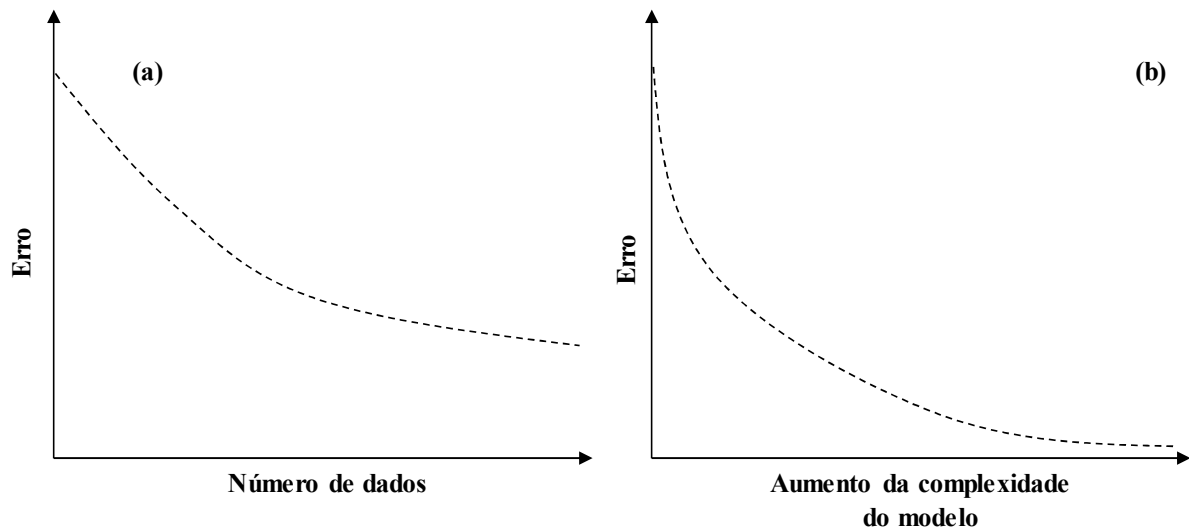


Figura 2.14 - Variação do desvio de treinamento em função; (a) do número de pontos de dados e; (b) aumento da complexidade.

O processo de desenvolvimento de um modelo requer várias etapas. Em regressão linear múltipla, por exemplo, os parâmetros precisam ser definidos antes da etapa de treinamento. Para que esta etapa seja realizada corretamente, usa-se um novo conjunto de dados

chamado de *validação*. O conjunto de validação deve conter dados independentes dos dados de treinamento e de teste.

A adição de um parâmetro em um modelo linear é chamada de aumento da *complexidade* do modelo. O desvio de treinamento tende a zero com o aumento da complexidade do modelo (Figura 2.14b). Assim, o conjunto de validação pode ser usado para a escolha dos parâmetros de modelo e para avaliar o desempenho final.

2.8.2 Validação cruzada

Em muitas aplicações que envolvem aprendizagem de máquina, dividir o banco de dados em três conjuntos diferentes pode não ser problema. Porém, em algumas situações os dados são escassos e reservar uma quantidade de dados para validação diminuiria a quantidade de dados da etapa de treinamento do modelo, o que poderia deixá-lo menos representativo. Em outras palavras, reservar uma quantidade de dados para a etapa de validação pode diminuir a capacidade de generalização do modelo. Uma forma alternativa para suprir essa necessidade é uma técnica estatística chamada de *validação cruzada* (NILSSON, 2005).

A técnica de validação cruzada consiste em dividir o banco de dados apenas em dados de treinamento (conjunto de treinamento ou *training set*) e dados de teste (conjunto de teste ou *test set*). O conjunto de treinamento é, por sua vez, dividido em um número fixo de subconjuntos ou pastas (ou *k-fold*, no inglês), em partições aproximadamente iguais. Cada uma das pastas é usada para testar o desempenho, enquanto os dados remanescentes são usados para treinar o modelo.

Supondo que conjunto de treinamento seja dividido em cinco subconjuntos (*5-fold*), na primeira iteração uma das pastas é definida como dados de validação (conjunto de validação ou *validation set*) e as demais formam o *conjunto de treinamento*. Na segunda iteração, uma pasta diferente é usada como conjunto de validação, enquanto as pastas restantes são usadas para treinamento de modelo, assim por diante, como ilustrado na Figura 2.15. Finalmente, a média de todos os desvios é o desvio verdadeiro (SHALEV-SHWARTZ e BEM-DAVID, 2014). A validação cruzada é muito utilizada na seleção de parâmetros de modelos e de outros termos utilizados no desenvolvimento do modelo final, como o parâmetro de regularização λ_2 . Recentemente, Shrivastava *et al.* (2016) aplicaram a técnica de validação cruzada para seleção

de *features* em modelos de predição de doenças de pele e Nazarenko *et al.* (2016) usaram validação cruzada em tarefas de reconhecimento de plantas medicinais.

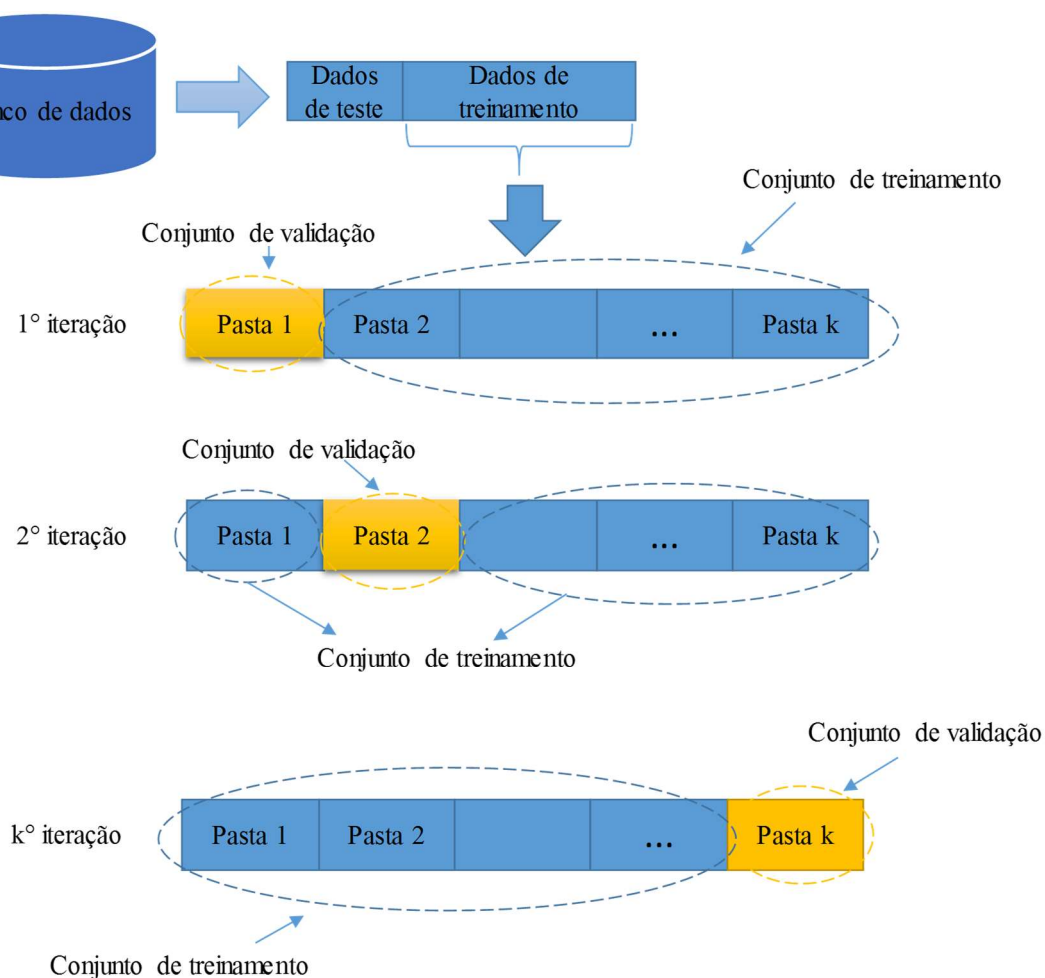


Figura 2.15 - Ilustração de um processo de validação cruzada.

A partir da revisão bibliográfica apresentada, é importante enfatizar os conteúdos importantes para a obtenção e compreensão dos resultados. Assim, sempre que for oportuno, os dados termodinâmicos obtidos pela técnica de cromatografia gasosa serão discutidos com base na revisão sobre a natureza das substâncias orgânicas (item 2.3). Para efeito de comparação com o modelo de predição de entalpia de vaporização proposto, foram utilizados apenas os modelos de Joback e Reid (1987), Chickos *et al.* (1998) e Kolská *et al.* (2005), pois são baseados

apenas na estrutura químicas dos compostos (item 2.6.2.2). O modelo proposto, por sua vez, terá sua base teórica fundamentada nos conceitos de aprendizagem de máquina (item 2.7.1), especialmente sobre as ideias de regressão ordinária por mínimos quadrados e regressão linear múltipla (itens 2.7.1.3 e 2.7.1.4). Além disso, visando a obtenção do melhor modelo possível, será empregada a técnica de regressão regularizada (item 2.7.1.5). Por fim, a técnica de validação cruzada (item 2.8.2) será fundamental para a seleção dos parâmetros do modelo.

Capítulo 3

METODOLOGIA

Neste capítulo são descritas as etapas necessárias para a obtenção do objetivo central deste trabalho de Tese. A metodologia foi dividida em duas seções principais: parte experimental e proposição do modelo de predição. A parte experimental apresenta os detalhes para a obtenção da pressão de vapor, entalpia de vaporização e temperatura normal de ebulição por cromatografia gasosa. A seção de proposição do modelo mostra primeiramente os procedimentos realizados para predição de entalpias de vaporização com modelos da literatura, bem como os caminhos feitos para o desenvolvimento do modelo preditivo que é proposto.

3.1 PARTE EXPERIMENTAL – GC/FID

Este trabalho foi conduzido no Laboratório de Processos em Meios Porosos (LPMP), do Departamento de Engenharia de Processos (DEPro), da Faculdade de Engenharia Química (FEQ), da Universidade Estadual de Campinas (Unicamp).

3.1.1 Material

3.1.1.1 Compostos terpênicos analisados

Ao todo, neste trabalho, foram avaliados 25 compostos terpênicos, apresentados Tabela 3.1. Os compostos foram adquiridos junto a fontes diferentes e estão organizados por fórmula química, pureza, massa específica, número CAS e marca do fabricante/distribuidor. Os compostos (+)- α -pineno, (+)- β -pineno, (+)-3-careno, (+)-limoneno, (-)- α -terpineol e linalol foram gentilmente cedido pelo Prof. Dr. Edson Tomaz, do Laboratório de Pesquisa e Desenvolvimento em Tecnologias Ambientais (LPDTA/DEPro/FEQ/Unicamp). As estruturas químicas de todos os terpenos analisados podem ser vistas no Apêndice A.

Tabela 3.1 - Compostos terpênicos analisados neste trabalho.

Composto	Fórmula Química	Pureza (g/g), %	ρ, g cm⁻³	CAS	Marca
(+)- α -pineno	C ₁₀ H ₁₆	98,5	0,858	7785-70-8	SIAL
(+)- β -pineno	C ₁₀ H ₁₆	98,5	0,858	19902-08-0	SIAL
(+)-3-careno	C ₁₀ H ₁₆	98,5	0,857	498-15-7	SIAL
(+)-limoneno	C ₁₀ H ₁₆	99,0	0,842	5989-27-5	SIAL
(+)- α -terpineol	C ₁₀ H ₁₈ O	97,0	0,930	7785-53-7	SIAL
linalol	C ₁₀ H ₁₈ O	99,0	0,870	78-70-6	SIAL
safrol	C ₁₀ H ₁₀ O ₂	97,0	1,095	94-59-7	SIAL
isosafrol	C ₁₀ H ₁₀ O ₂	99,5	1,122	120-58-1	CS
timol metil éter	C ₁₁ H ₁₆ O	98,0	0,940	1076-56-8	SIAL
3-BHA	C ₁₁ H ₁₆ O ₂	98,0	1,059	88-32-4	SIAL
timol	C ₁₀ H ₁₄ O	99,5	0,965	89-83-8	SIAL
carvacrol	C ₁₀ H ₁₄ O	99,9	0,976	499-75-2	MPB
terpinoleno	C ₁₀ H ₁₆	85,0	0,861	586-62-9	SIAL
γ -terpineno	C ₁₀ H ₁₆	98,0	0,847	99-85-4	AO
α -humuleno	C ₁₅ H ₁₄	96,0	0,886	6753-98-6	SIAL
<i>p</i> -cimeno	C ₁₀ H ₁₄	99,0	0,857	99-87-6	AO
(-)-terpinen-4-ol	C ₁₀ H ₁₈ O	95,0	0,934	20126-76-5	SIAL
(-)- <i>trans</i> -cariofileno	C ₁₅ H ₁₄	98,5	0,902	87-44-5	SIAL
biciclo[2.2.1]hepta-2,5-dieno	C ₇ H ₁₀	98,0	0,906	121-46-0	Aldrich
5-vinil-2-norborneno	C ₉ H ₁₂	95,0	0,841	3048-64-4	Aldrich
5-etilideno-2-norborneno	C ₉ H ₁₂	99,0	0,893	16219-75-3	Aldrich
álcool fenchil	C ₁₀ H ₁₀ O	96,0	-	1632-73-1	SIAL
(-)-borneol	C ₁₀ H ₁₀ O	99,0	-	464-45-9	SIAL
biciclo[2.2.2]octano-2,5-diona	C ₈ H ₁₀ O ₂	97,0	-	57346-05-1	Aldrich
biciclo[4.1.0]heptano-7-ácido carboxílico	C ₈ H ₁₂ O ₂	-	-	-	Aldrich

ρ – massa específica a 25 °C (dados do fabricante); SIAL – Sigma-Aldrich; CS - Chem Service; MPB - MP Biomedicals; AO - Acros Organics

3.1.1.2 Hidrocarbonetos alifáticos

Os hidrocarbonetos alifáticos (saturados, de cadeia cíclica e acíclica), incluídos para validar o método e avaliar o uso da técnica frente à várias volatilidades, foram gentilmente

cedidos pelo LPDTA/DEPro/FEQ/Unicamp. Os detalhes sobre fórmula química, pureza, massa específica e marca do fabricante/distribuidor estão apresentados na Tabela 3.2.

Tabela 3.2 - Hidrocarbonetos alifáticos analisados neste trabalho.

Composto	Fórmula Química	Pureza (g/g), %	ρ , g cm ⁻³	Marca
2,2-dimetilbutano	C ₆ H ₁₄	98,0	0,791	AO
2,3-dimetilbutano	C ₆ H ₁₄	98,0	0,660	AO
3-metilpentano	C ₆ H ₁₄	99,0	0,664	AO
2,2,4-trimetilpentano	C ₈ H ₁₈	99,0	0,996	SIAL
2-metilhexano	C ₇ H ₁₆	99,0	0,679	AO
3-metilhexano	C ₇ H ₁₆	99,0	0,687	AO
ciclohexano	C ₆ H ₁₂	90,0	0,779	Sinth
metilciclopentano	C ₆ H ₁₂	95,0	0,749	AO
metilciclohexano	C ₇ H ₁₄	99,0	0,770	AO

ρ – massa específica a 25 °C (dados do fabricante); AO - Acros Organics; SIAL - Sigma-Aldrich

3.1.1.3 *n*-Alcanos

Para a obtenção do índice de retenção de Kóvats, foi necessário utilizar alguns *n*-alcanos como padrão de referência. Os *n*-alcanos utilizados neste trabalho estão apresentados na Tabela 3.3, com suas respectivas informações fórmula química, pureza, densidade e marca. Os compostos *n*-pentano, *n*-hexano, *n*-heptano, *n*-octano, *n*-nonano, *n*-decano, foram gentilmente cedidos pelo LPDTA/FEQ/DEPro/UNICAMP.

Tabela 3.3 - *n*-Alcanos usados como padrão de referência neste trabalho.

<i>n</i> -Alcano	Fórmula Química	Pureza (g/g), %	ρ , g cm ⁻³	Marca
pentano	C ₅ H ₁₂	99,0	0,630	Merck
hexano	C ₆ H ₁₄	96,0	0,660	Merck
heptano	C ₇ H ₁₆	99,8	0,680	CE
octano	C ₈ H ₁₈	99,0	0,700	SIAL
nonano	C ₉ H ₂₀	99,0	0,720	Merck
decano	C ₁₀ H ₂₂	99,0	0,730	Merck
undecano	C ₁₁ H ₂₄	99,9	0,740	Supelco
dodecano	C ₁₂ H ₂₆	99,0	0,750	SIAL
tridecano	C ₁₃ H ₂₈	99,9	0,756	Supelco
tetradecano	C ₁₄ H ₃₀	99,9	0,764	Supelco
pentadecano	C ₁₅ H ₃₂	99,9	0,769	Supelco
hexadecano	C ₁₆ H ₃₄	99,9	0,770	Supelco

ρ – massa específica a 25 °C (dados do fabricante); CE - Carlos Erba; SIAL - Sigma-Aldrich

3.1.1.4 Sistema cromatográfico a gás

Os tempos de retenção dos compostos examinados foram obtidos usando um cromatógrafo a gás (AutoSystem XL GC Perkin Elmer, EUA), equipado com detector por ionização de chamas (GC/FID) e injetor *split/splitless* à temperatura de 230 °C (Figura 3.1). Uma coluna capilar não polar (NST05, 5 % difenil + 95 % dimetipolisiloxano, diâmetro interno de 0,25 mm e espessura 0,25 µm) de 60 m de comprimento foi usada para a separação dos compostos e foi configurada no modo *split* com razão 1:50. A temperatura do forno foi mantida a 250 °C. Os experimentos foram realizados isotermicamente usando o hélio como gás de arraste com fluxo de 2,0 mL min⁻¹. As análises foram feitas em duplicata e realizadas no LPDTA/DEPro/FEQ/Unicamp.

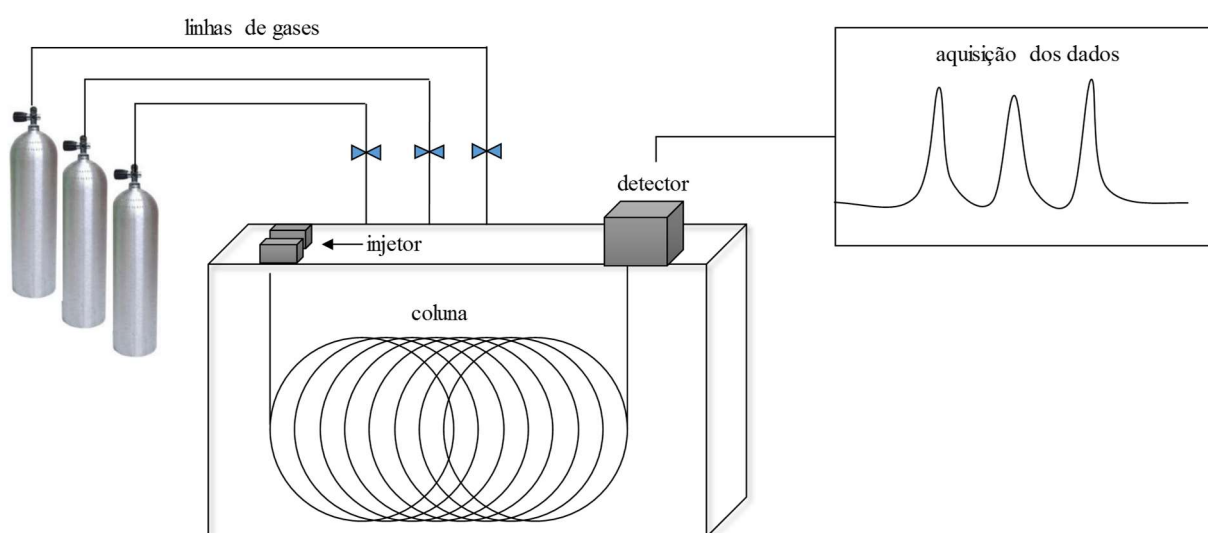


Figura 3.1 - Esquema de um sistema cromatográfico a gás usado na análise dos compostos examinados.

3.1.2 Métodos

Para a realização das análises cromatográficas foi necessário preparar soluções dos compostos terpênicos, hidrocarbonetos alifáticos e dos *n*-alcanos de referência. Um estudo preliminar demonstrou haver sobreposição de picos, quando feita apenas uma solução com todos os terpenos e hidrocarbonetos alifáticos. Além disso, o tempo de análise demonstrou-se

longo. Desse modo, os experimentos foram organizados da seguinte maneira: (i) experimentos com os terpenos de menor massa molar ou de menores forças intermoleculares (terpenos leves); (ii) experimentos com os terpenos de massa molar elevada ou de maiores forças intermoleculares (terpenos pesados); (iii) experimentos com os hidrocarbonetos alifáticos; (iv) experimentos com os *n*-alcanos de menor cadeia carbônica (*n*-alcanos leves) e; (v) experimentos com os *n*-alcanos de maior cadeia carbônica (*n*-alcanos pesados).

Foram considerados terpenos leves: (+)- α -pineno, (+)- β -pineno, (+)-3-careno, (+)-limoneno, (-)- α -terpineol, linalol, γ -terpineno, *p*-cimeno (*mistura 1*). Os terpenos safrol, isosafrol, timol metil éter, 3-BHA, timol, carvacrol, terpinoleno, α -humuleno, terpinen-4-ol e (-)-*trans*-cariofileno foram considerados terpenos pesados (*mistura 2*). Com relação aos *n*-alcanos, foram considerados leves: pentano, hexano, heptano, octano, nonano e decano (*mistura C₅-C₁₀*). O undecano, dodecano, tridecano, tetradecano, pentadecano e hexadecano (*mistura C₁₁-C₁₆*) foram considerados alcanos pesados. Sete terpenos bicíclicos foram analisados separadamente: biciclo[2.2.1]hepta-2,5-dieno, biciclo[2.2.2]octano-2,5-diona, álcool fenchil, biciclo[4.1.0]heptano-7-ácido carboxílico, 5-etilideno-2-norborneno, 5-vinil-2-norborneno e (-)-borneol, formando uma terceira mistura de terpenos (*mistura 3*).

3.1.2.1 Preparo das soluções

Tentou-se obter soluções nas condições mais semelhantes possível. As soluções foram preparadas diluindo cerca de 7,0 mg de cada analito em 10,0 mL de metanol. Para as substâncias líquidas, a massa foi obtida por meio da massa específica (temperatura ambiente de, aproximadamente, 25 °C). A amostra real foi uma solução de 0,5 mg de analito por 1,0 mL de metanol.

3.1.2.2 Análises cromatográficas

Os ensaios experimentais foram realizados no LPDTA/DEPro/FEQ/Unicamp. O experimento consistiu em injetar manualmente 1,0 μ L da amostra no cromatógrafo, com o auxílio de uma seringa de vidro de 1,0 μ L (Modelo Microliter™ 7001, Hamilton Co., EUA). A faixa de temperatura foi escolhida de acordo com a natureza dos compostos analisados. Para os hidrocarbonetos alifáticos, a faixa de temperatura foi de 30 a 100 °C. Para os terpenos da

mistura 1, a temperatura variou de 40 a 140 °C. Alguns terpenos da mistura 1 apresentaram alta retenção na coluna cromatográfica em temperatura abaixo de 70 °C. Assim, para diminuir o tempo de análise, alguns terpenos foram submetidos à uma faixa de temperatura de 70 a 140 °C. Para os terpenos da mistura 2, a faixa foi de 100 a 200 °C. Os terpenos da mistura 3 foram analisados foram avaliados, no geral, numa faixa de temperatura de 50 a 200 °C. Os cromatogramas apresentados no Apêndice I.

3.1.2.3 Ajuste dos tempos de retenção

Os tempos de retenção foram ajustados utilizando os parâmetros da Equação 3.1, obtidos por regressão não linear através dos tempos de retenção dos *n*-alcanos utilizados neste trabalho (QUINTANILLA-LÓPEZ *et al.*, 2000; MILLER e BRUNO, 2003):

$$t_R(z) = A + \exp(B + Cz^D) \quad (3.1)$$

Na Equação 3.1, o tempo de retenção dos *n*-alcanos, $t_R(z)$, é regredido em função do seu número de carbono (z). Os termos $A + \exp(B)$ equivalem ao tempo morto da coluna (t_M) (LEBRÓN-AGUILAR *et al.*, 2002). Dessa forma, o tempo de retenção dos compostos foram ajustados por $t'_R = t_R - t_M$. Os ensaios foram realizados em duplicatas; para os índices de retenção, foram reportados apenas os valores médios. Já para os cálculos realizados neste trabalho, foram considerados os valores em duplicata.

3.1.2.3 Dados termodinâmicos disponíveis na literatura

No presente trabalho, a acurácia dos valores das grandezas termodinâmicas calculadas para os compostos analisados foi medida com base nos dados existentes na literatura (Tabela 3.4). Os dados foram retirados, em sua maioria, da plataforma online gratuita do *National Institute of Standards and Technology* (NIST), da base de dados do *software Aspen®* versão 8.4, denominada *National Institute of Standards and Technology ThermoData Engine* (NIST/TDE) e complementados com alguns dados da literatura, conforme legenda da Tabela 3.4. Para os hidrocarbonetos alifáticos foi encontrada uma gama de dados de entalpia de vaporização e de temperatura normal de ebulição, com valores muito próximos (Apêndice G). Desse modo, para esses compostos foram considerados os valores médios. Para os compostos

isosafrol e α -humuleno, cujos dados de pressão de vapor não foram encontrados na literatura, foram usados os valores preditos pelo NIST/TDE, mostrados na Tabela 3.5.

Tabela 3.4 - Dados termodinâmicos retirados da literatura para os compostos analisados.

Composto	P_{vap} , kPa	ΔH_{vap} , kJ mol ⁻¹	T_{nb} , K
2,2-dimetilbutano	42,548 ^a ; 42,551 ^b	27,77 ^a ; 27,71 ^{b,5}	322,90 ^a ; 322,88 ^b
2,3-dimetilbutano	31,34 ^a ; 31,304 ^{b,1}	29,18 ^a ; 29,16 ^{b,1}	331,20 ^a ; 331,12 ^b
3-metilpentano	25,306 ^a ; 25,372 ^{b,1}	30,31 ^a ; 30,27 ^{b,1}	336,55 ^a ; 336,40 ^b
metilciclopentano	18,331 ^a ; 18,331 ^b	31,67 ^a ; 31,64 ^{b,1}	345,00 ^a ; 344,94 ^b
ciclohexano	13,010 ^a ; 12,999 ^{b,1}	33,07 ^a ; 33,00 ^{b,5}	353,90 ^a ; 353,85 ^b
2-metilhexano	8,788 ^a ; 8,781 ^b	34,87 ^a ; 34,84 ^b	363,20 ^a ; 363,10 ^b
3-metilhexano	8,213 ^a ; 7,958 ^{b,2}	35,12 ^a ; 35,04 ^b	365,00 ^a ; 364,89 ^b
2,2,4-trimetilpentano	6,578 ^a ; 6,601 ^{b,3}	35,11 ^a ; 35,15 ^{b,1}	372,40 ^a ; 372,36 ^b
metilciclohexano	6,180 ^a ; 6,183 ^{b,4}	35,32 ^a ; 35,35 ^{b,4}	377,00 ^a ; 374,04 ^b
(+)- α -pineno	0,529 ^c ; 0,544 ^d ; 0,515 ^e	44,85 ^e ; 44,57 ^{b,1} ; 45,40 ^l	429,11 ^b ; 429,35 ^j
(+)- β -pineno	0,341 ^e	46,19 ^e ; 45,80 ^{b,1}	438,93 ^b ; 439,15 ^j
safrol	0,009 ^f	54,6 ^{m,1}	508,56 ^b ; 507,65 ^j
isosafrol	-	54,9 ^{m,2}	521,5 ^b
(+)-3-careno	0,231 ^e ; 0,221 ^{b,1}	48,30 ^a ; 48,51 ^e	445,05 ^b ; 444,15 ^j
(+)-limoneno	0,213 ^c ; 0,202 ^d ; 0,183 ^e ; 0,200 ^g	48,92 ^{b,1} ; 49,60 ^e ; 49,5 ^l	450,79 ^b ; 451,15 ^j
linalol	0,035 ^e ; 0,027 ^g ; 0,039 ^{b,6}	55,25 ^e	471,15 ^j
(-)- α -terpineol	0,015 ^a ; 0,012 ^h	54,0 ^{m,3}	492,72 ^b ; 493,15 ^j
carvacrol	0,0035 ⁱ	68,20 ⁱ	512,19 ^b ; 510,85 ^j
timol	0,004 ^e ; 0,0036 ⁱ	70,85 ^e	505,79 ^b ; 505,65 ^j
γ -terpineno	0,103 ^d ; 0,137 ^e ; 0,124 ^h	51,38 ^e	456,15 ^b ; 456,15 ^j
<i>p</i> -cimeno	0,193 ^e ; 0,180 ^h	49,24 ^e	450,27 ^b ; 450,25 ^j
terpinen-4-ol	0,0195 ^c ; 0,0173 ^h	55,52 ^e	495,60 ^j
terpinoleno	0,099 ^d	50,80 ^a	459,61 ^b ; 459,15 ^j
biciclo[2.2.1]hepta- 2,5-dieno	-	36,6 ^a ; 33,8 ^b ; 34,8 ⁿ	369,2 ^a ; 363,7 ^b
5-vinil-2-norborneno	-	42,29 ^a ; 42,3 ^a	414,2 ^a ; 413,89 ^b
5-etilidene-2- norborneno	-	44,14 ^b ; 42,3 ⁿ	419,2 ^a ; 421,02 ^b

a) base de dados NIST online – dados a 298,15 K; **b)** base de dados NIST/TDE do Aspen® - dados a 298,15 K, exceto: **b,1** (298,14 K); **b,2** (297,48 K); **b,3** (298,20 K); **b,4** (298,13 K); **b,5** (298,12 K), **b,6** (297,03 K); **c)** Fichan *et al.* (1999) – dados a 298,15 K; **d)** Li *et al.* (1998) – dados a 296,65 K; **e)** Hoskovec *et al.* (2005) – dados a 298,15 K; **f)** Klauck *et al.* (2014) – dados a 298,15 K; **g)** Espinosa Díaz *et al.* (1999); *dado do *d*-linalol – dados a 298,15 K; **h)** Lucia *et al.* (2009) – dados a 298,15 K; **i)** van Roon *et al.* (2002) – dados a 298,15 K; **j)** Lide (2004); **l)** Clará *et al.* (2009) – dados a 298,15 K; **m)** Stephenson e Malanowski (1897) – **m,1** refere-se a dado na faixa de 336-506 K, **m,2** (393-531 K), **m,3** (325-491 K); **n)** Chickos e Acree (2010) – dados a 298,15 K.

Tabela 3.5 - Dados termodinâmicos preditos pelo NIST/TDE para os compostos analisados com dados não disponíveis na literatura.

Composto	*P_{vap}, kPa	*ΔH_{vap}, kJ mol⁻¹
isosafrol	0,004	-
α -humuleno	-	66,11

* Dados preditos a 298,15 K pelo NIST/TDE do Aspen®, Versão 8.4.

3.2 PROPOSIÇÃO DE MODELO DE PREDIÇÃO

3.2.1 Seleção de modelos de predição de ΔH_{vap} da literatura

A fim de avaliar o modelo proposto, selecionou-se alguns modelos preditivos baseados na teoria de contribuição de grupos. Os modelos foram usados para estimar a entalpia de vaporização dos compostos determinadas experimentalmente por cromatografia gasosa.

Conforme apresentado na Seção 2.6.2, podem ser encontrados na literatura uma gama de modelos para prever a entalpia de vaporização de compostos orgânicos. Alguns modelos carecem de informações adicionais para a estimação, como o método de Basarová e Svoboda (1995) e Tu e Liu (1996) que precisam da temperatura crítica. A principal desvantagem desses modelos é que além de serem mais trabalhosos, para muitos compostos orgânicos há escassez de dados de propriedades críticas na literatura, que os tornam menos utilizados.

Em alternativa, os métodos baseados no conceito de contribuição de grupos são bem aceitos porque requerem apenas valores de contribuição de grupos tabelados e os cálculos geralmente são feitos de forma simples e rápida. A propriedade é, portanto, apenas em função das contribuições dos grupos necessários para representar a estrutura química de uma molécula. Com base nisso, no presente trabalho, foram selecionados os métodos de Joback e Reid (1987), Chickos *et al.* (1998) e Kolská *et al.* (2005) que necessitam apenas da estrutura química dos compostos, conforme as equações apresentadas no item 2.6.2.2.

3.2.2 Proposição de modelo de predição de ΔH_{vap}

3.2.2.1 Seleção dos compostos

Posto que os óleos essenciais são constituídos basicamente por hidrocarbonetos monoterpenos, sesquiterpenos e derivados majoritariamente oxigenados (Seção 2.2), não foram

incluídos no modelo compostos orgânicos sulfurados, nitrogenados, bem como hidrocarbonetos alcinos, pois ocorrem em baixíssima frequência em óleos essenciais. Os compostos selecionados para compor o banco de dados, portanto, seguiu-se da seguinte forma:

- Hidrocarbonetos (saturados e insaturados), de cadeia aberta e ramificada, com quantidade de carbono de C_7 a C_{15} ;
- Hidrocarbonetos monoterpênicos, sesquiterpênicos e derivados contendo oxigênio, com número de carbonos de C_7 a C_{15} ;
- Hidrocarbonetos bicíclicos e derivados contendo oxigênio.

3.2.2.2 Fragmentação dos grupos

O conceito de contribuição de grupos baseia-se no pressuposto de que as moléculas podem ser divididas em grupos, que podem ser utilizados para descrever o comportamento da molécula. A fragmentação foi feita com base em apenas um nível de estimação. O método proposto foi baseado na natureza da molécula, fazendo distinção em grupos de cadeia aberta, cíclicos, aromáticos e compostos bicíclicos. Foram atribuídas informações diferentes para grupos presentes em anéis (aromáticos ou cíclicos) ou ligados à carbonos aromáticos ou cíclicos. A Figura 3.2 apresenta uma molécula hipotética que ilustra os grupos selecionados da Tabela 3.6, num total de 26 grupos.

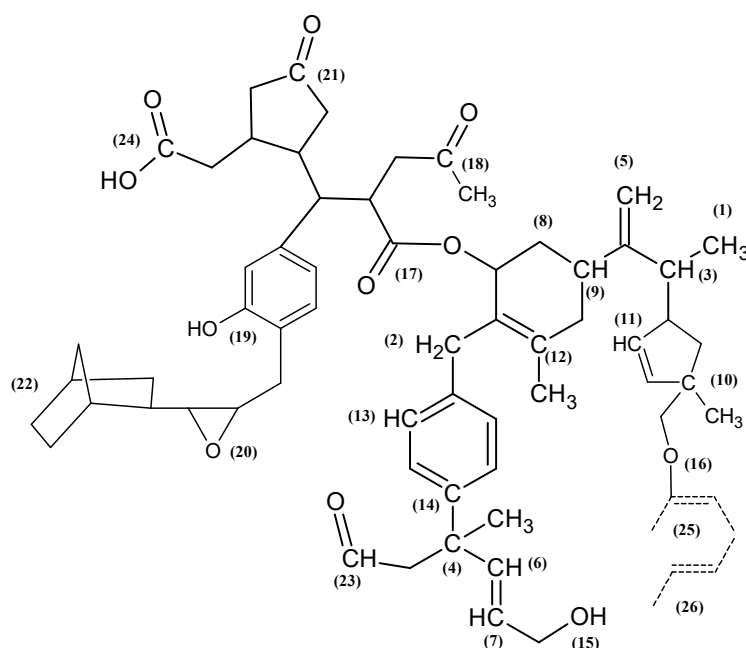


Figura 3.2 - Molécula hipotética ilustrando os grupos utilizados para o desenvolvimento do modelo proposto.

Tabela 3.6 - Grupos usados para o desenvolvimento do modelo proposto.

Nº	Grupo	Ocorrência
1	-CH ₃	hidrocarbonetos acíclicos ^a
2	-CH ₂ -	hidrocarbonetos acíclicos ^a
3	>CH-	hidrocarbonetos acíclicos ^a
4	>C<	hidrocarbonetos acíclicos ^a
5	=CH ₂	hidrocarbonetos acíclicos ^b
6	=CH-	hidrocarbonetos acíclicos ^b
7	=C<	hidrocarbonetos acíclicos ^b
8	-CH ₂ - (cíclico)	hidrocarbonetos cíclicos ^a
9	>CH- (cíclico)	hidrocarbonetos cíclicos ^a
10	>C< (cíclico)	hidrocarbonetos cíclicos ^a
11	=CH- (cíclico)	hidrocarbonetos cíclicos ^b
12	=C< (cíclico)	hidrocarbonetos cíclicos ^b
13	=CH (aromático)	hidrocarbonetos aromáticos
14	=C< (aromático)	hidrocarbonetos aromáticos
15	OH (álcool)	álcoois
16	-O-	éteres
17	O-C=O	éster
18	C=O	cetona
19	OH (fenol)	fenóis
20	-O- (cíclico)	éter em anel
21	C=O (cíclico)	cetona em anel
22	<i>bicíclico</i> (contribuição para estrutura bicíclica)	terpenos
23	HC=O	aldeídos
24	OHC=O	ácidos carboxílicos
25	<i>cis</i> (contribuição para estrutura <i>cis</i>)	terpenos e hidrocarbonetos
26	<i>trans</i> (contribuição para estrutura <i>trans</i>)	terpenos e hidrocarbonetos

a) alifáticos saturados; b) alifáticos insaturados

3.2.2.3 Aquisição dos dados de entalpia de vaporização

Os dados de entalpia de vaporização foram retirados da compilação crítica sobre dados de entalpia de vaporização, determinados de 1880 a 2010, publicada por Chickos e Acree (2010). Para organização e formação do conjunto de dados, desenvolveu-se uma macro em VBA Excel, conforme a Figura 3.3. Por meio da macro, é possível carregar a imagem da estrutura do composto, atribuir-lhe um nome e outras informações, como o número CAS. As caixas de textos ao lado dos grupos representativos são preenchidas conforme as informações do composto que está sendo adicionado no conjunto de dados. As informações salvas são enviadas diretamente para uma planilha do Excel. Uma segunda macro foi construída para visualizar as informações adicionadas (Figura 3.4), na qual pode-se carregar a estrutura química de um composto juntamente com os dados adicionados a ele, possibilitando revisar a

fragmentação da molécula. É importante mencionar que esta parte do trabalho é essencial para posterior atualização do conjunto de dados, seja acionando-se novos compostos ou atualizando os grupos.

Vizualizador de Figuras

Ver Fotos Nova Foto

Adicionar Novo Composto

Nome do Composto

Número CAS

GRUPOS ACÍCLICOS

GRUPOS CÍCLICOS E AROMÁTICOS

GRUPOS FUNCIONAIS

OUTROS

-CH3- -CH2- OH_Alcool BiCyclic

-CH2- >CH- -O- Cis

>CH- >C< O-C=O Trans

>C< =CH- C=O

=CH2 =C< OH_Fenol

=CH- =CH_Arom -O-

=C< =C<_Arom C=O_Cyclic

CH=O

(OH)C=O

Salvar Novos Detalhes

Figura 3.3 - Macro para aquisição de dados.

Vizualizador de Figuras

Ver Fotos Nova Foto

Carregar Informações do Composto

Nome do Composto

Número CAS

GRUPOS ACÍCLICOS

GRUPOS CÍCLICOS E AROMÁTICOS

GRUPOS FUNCIONAIS

OUTROS

-CH3- -CH2- OH_Alcool BiCyclic

-CH2- >CH- -O- Cis

>CH- >C< O-C=O Trans

>C< =CH- C=O

=CH2 =C< OH_Fenol

=CH- =CH_Arom -O-

=C< =C<_Arom C=O_Cyclic

CH=O

(OH)C=O

Composto Anterior

Próximo Composto

Figura 3.4 - Macro para visualizar as informações adicionadas.

3.2.2.4 Formação do conjunto de dados

O conjunto de dados utilizado para o desenvolvimento do modelo proposto na presente Tese contém substâncias da família dos hidrocarbonetos e derivados contendo oxigênio. Até o presente momento, consiste de 1.719 substâncias orgânicas diferentes e um total de 3.643 dados de entalpia de vaporização, com temperatura variando de 288 – 426 K e cadeia carbônica da faixa de C₇ a C₁₅.

Os hidrocarbonetos (incluindo saturados e insaturados, cíclicos, aromáticos e bicíclicos) possuem dados mais abundantes na literatura e, por isso, estão em maior número no banco de dados (371 compostos). A Tabela 3.7 mostra a divisão do número de compostos por classe orgânica. Os compostos multifuncionais referem-se àqueles com mais de um grupo funcional. O Apêndice C apresenta informações mais detalhadas sobre os compostos do banco de dados.

Tabela 3.7 - Classe dos compostos inseridos no banco de dados.

Classe	Número de compostos
hidrocarbonetos	623
álcoois	101
aldeídos	22
ácidos carboxílicos	90
cetonas	90
éteres	206
ésteres	235
fenóis	103
multifuncionais	249

3.2.2.5 Procedimento para estimação dos parâmetros do modelo

A tarefa de regressão dos parâmetros do modelo foi feita por meio da linguagem Python, que é uma linguagem de programação interativa e orientada a objeto, comparada muitas vezes a linguagens como C# e *Visual Basic*. É uma ferramenta livre e utilizada no ramo da ciência de dados, motivo pela qual foi utilizada neste trabalho. Com o Python, é possível trabalhar com manipulação de tabelas (conjunto de dados) mais facilmente que outras linguagens, tipo o Matlab.

Os resultados da implementação serão comparados com os advindos do pacote GraphLab Create, Versão 2.1, (Turi Create Intelligence™), cuja licença (1 ano), em caso de fins acadêmicos, foi adquirida gratuitamente.

3.2.2.6 Divisão do conjunto de dados

Uma vez o conjunto de dados contém compostos com mais de um dado de entalpia de vaporização, a divisão foi feita com base no número de compostos. A fim de obter maior capacidade de generalização do modelo proposto, o treinamento foi realizado de duas maneiras distintas. A primeira usando três conjuntos de dados: treinamento, teste e validação, sendo reservado 20 % dos compostos para teste e 20 % para validação. A segunda usando a técnica de validação cruzada em que 20 % dos compostos foram destinados para teste e os dados remanescentes para treinamento e validação cruzada.

3.2.2.7 Regressão ordinária por mínimos quadrados

A otimização foi realizada usando o método do gradiente descendente, implementado em Python 2 (Versão 2.7.11). A implementação foi realizada conforme o algoritmo 1, apresentado no Apêndice H, onde inicia-se a primeira iteração com $t = 1$ e pesos $\mathbf{w}^{(1)} = 0$ (ou valor otimizado). O algoritmo otimiza os parâmetros até que o gradiente da Soma dos Quadrados Residuais em função dos parâmetros $\nabla SQR(\mathbf{w}^{(t)})$ seja menor que uma tolerância especificada, ε . Os parâmetros são atualizados subtraindo o parâmetro antigo da derivada parcial da função objetivo parcial[j] vezes um passo especificado, η .

3.2.2.8 Regressão regularizada por mínimos quadrados

O efeito da regularização também foi avaliado. Para isso, o algoritmo 1 foi modificado para receber o termo do parâmetro de regularização λ_2 . A implementação foi feita conforme o algoritmo 2 (Apêndice H).

3.2.2.9 Medida da qualidade da regressão

A qualidade da regressão foi avaliada com base no monitoramento dos seguintes desvios:

Soma dos quadrados residuais (*SQR*)

$$SQR = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.2)$$

Raiz do desvio quadrado médio (*RDQM*)

$$RDQM = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (3.3)$$

Desvio médio absoluto (*DA*)

$$DA = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (3.4)$$

Desvio médio relativo (*DM*)

$$DM = \frac{1}{n} \sum_{i=1}^n \left(\frac{|y_i - \hat{y}_i|}{y_i} \times 100 \right) \quad (3.5)$$

em que y_i é o valor real; \hat{y}_i é o valor predito e; n é o número de elementos do conjunto de dados.

Capítulo 4

RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados obtidos experimentalmente e os resultados advindos da proposição do modelo de predição de entalpia de vaporização. Primeiramente, são apresentados os resultados referentes aos dados de pressão de vapor, entalpia de vaporização e temperatura normal de ebulição, obtidos pela técnica de cromatografia gasosa. Logo após, são apresentados os resultados de predição obtidos para os modelos analisados da literatura. Finalmente, são apresentados os resultados de predição encontrados para o modelo proposto. Além disso, é apresentada uma comparação dos resultados obtidos pelo modelo de predição proposto na presente Tese com os modelos de predição da literatura que foram estudados.

4.1 RESULTADOS EXPERIMENTAIS DE GC/FID

4.1.1 Obtenção dos índices de retenção

Com os tempos de retenção (t_R) obtidos para os n -alcanos e para os compostos estudados, por meio dos experimentos de cromatografia gasosa, e após a correção de t_R usando a Equação 3.1 (obtenção de t'_R), os índices de retenção de Kováts (I_R) foram calculados com o auxílio da Equação 2.7a. Tais índices de retenção foram obtidos em função da temperatura (T), conforme descrito na Seção 3.1.2.2. Os dados foram organizados em duas tabelas. A Tabela 4.1 apresenta os I_R para os hidrocarbonetos alifáticos e a Tabela 4.2 apresenta os dados de I_R para os terpenos.

Após breve inspeção nas Tabelas 4.1 e 4.2, foi possível notar que os compostos apresentaram comportamentos de I_R distintos quanto ao aumento da temperatura. As Figuras 4.1, 4.2, 4.3 e 4.4 mostram a curva I_R vs. T e ilustram os principais perfis encontrados para os

compostos analisados (ver Apêndice D). Por exemplo, entre os hidrocarbonetos alifáticos, o 2,3-dimetilbutano e o 3-metilpentano apresentaram comportamento muito próximo ao do composto 2,2-dimetilbutano (Figura 4.1). Os compostos carvacrol e timol apresentaram comportamento bastante diferente, quando comparados aos dos demais compostos analisados, como mostram as Figuras 4.2 e 4.3.

Tabela 4.1 - Índice de retenção obtidos por meio da Equação 2.7a para os hidrocarbonetos estudados.

Composto	Índice de retenção							
	Temperatura, °C							
	30	40	50	60	70	80	90	100
2,2-dimetilbutano	534,23	533,91	535,22	536,04	536,09	536,60	537,71	
2,3- dimetilbutano	564,19	564,23	565,37	566,25	566,41	567,11	568,10	
3-metilpentano	582,51	582,44	583,06	583,63	583,33	583,46	584,02	
metilciclopentano	628,38	629,68	631,39	633,19	634,28	636,18	637,61	641,64
ciclohexano	668,09	669,13	670,17	671,03	672,09	673,60	677,11	678,78
2-metilciclohexano	666,40	666,34	666,76	666,79	667,56	666,58	675,44	678,78
3-metilhexano	675,09	675,26	675,70	676,20	676,34	677,82	677,83	678,93
2,2,4-trimetilpentano	686,93	687,88	688,80	689,89	690,56	691,38	691,98	693,95
metilciclohexano	720,20	722,66	725,08	727,73	730,25	732,82	735,50	739,60

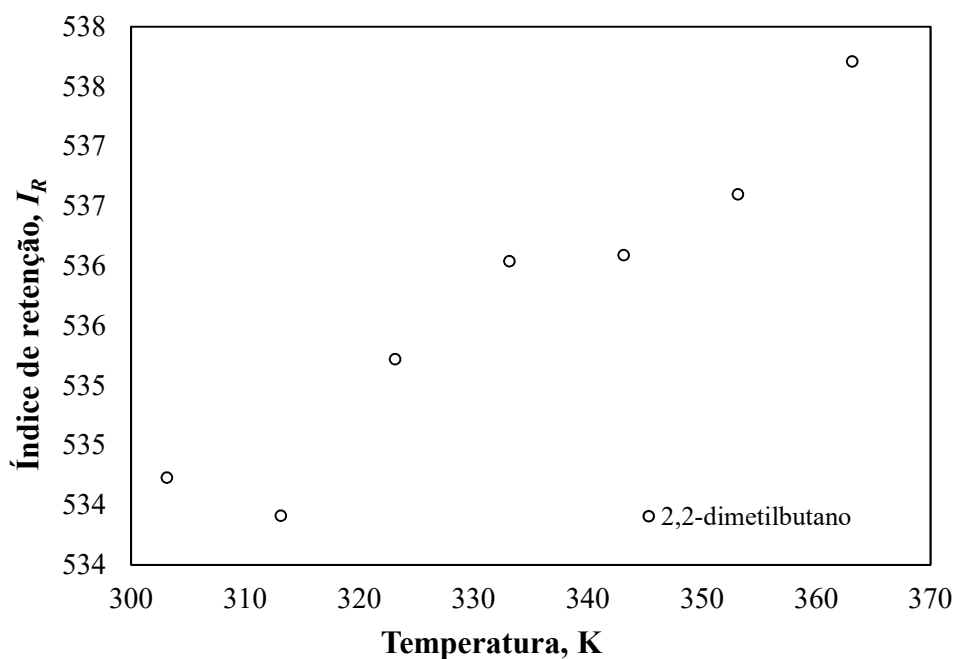


Figura 4.1 - Ilustração da dependência do índice de retenção de Kováts em função da temperatura para o 2,2-dimetilbutano.

Tabela 4.2 - Índice de retenção para os terpenos estudados.

Composto*	Índice de retenção										
	Temperatura, °C										
	40	50	60	70	80	90	100	110	120	130	140
(+)- α -pineno	926,35	929,58	932,52	936,44	939,51	943,49	947,91	951,31	954,41	958,69	963,81
(+)- β -pineno	965,20	969,20	972,80	978,02	982,15	986,63	992,24	994,56	998,16	1002,91	1.007,63
B[2.2.1]		650,23	651,35	651,63	651,93	652,16	652,23				
VNB			878,20	880,76	883,20	885,24	886,95	891,24	891,70	893,70	894,45
ENB			909,84	911,97	913,99	915,47	917,07	918,9	918,52	920,52	921
	Temperatura, °C										
	50	60	70	80	90	100	110	120	130	140	150
	(+)-3-careno	1.002,92	1.005,70	1.010,29	1.013,56	1.016,95	1.021,34	1.022,91	1.025,80	1.029,79	1.033,94
(+)-limoneno	1.021,19	1.023,72	1.027,83	1.030,81	1.033,93	1.038,35	1.039,34	1.041,91	1.045,54	1.049,26	1.052,04
linalol		1.096,66	1.097,63	1.098,36	1.099,00	1.101,27	1.100,38	1.100,52	1.101,77	1.102,74	1.103,22
	Temperatura, °C										
	70	80	90	100	110	120	130	140	150	160	170
	α -terpineol	1.182,00	1.185,28	1.188,73	1.193,15	1.195,30	1.198,32	1.202,22	1.206,25	1.209,48	1.214,44
γ -terpineno	1.055,95	1.058,06	1.061,21	1.066,12	1.065,39	1.068,03	1.072,73	1.073,89	1.076,29	1.080,62	1.082,16
<i>p</i> -cimeno	1.021,81	1.024,16	1.027,72	1.033,00	1.032,39	1.035,39	1.040,90	1.042,26	1.045,11	1.050,28	1.052,53
terpinen-4-ol	1.168,16	1.171,74	1.176,09	1.181,80	1.183,33	1.187,42	1.193,46	1.196,20	1.201,46	1.205,41	1.213,64
terpinoleno	1.084,32	1.086,89	1.090,38	1.095,54	1.095,34	1.098,35	1.103,17	1.104,48	1.107,11	1.111,04	1.112,56
B[4.1.0]			246,58	256,58	266,58	276,58	286,58	296,58	246,58		
	Temperatura, °C										
	80	90	100	110	120	130	140	150	160	170	180
	timol metil éter	1.228,54	1.230,54	1.233,65	1.233,11	1.234,87	1.238,14	1.239,03	1.241,75	1.243,69	1.245,59
	Temperatura, °C										
	70	80	90	100	110	120	130	140	150	160	170
	carvacrol	1.182,00	1.185,28	1.188,73	1.193,15	1.195,30	1.198,32	1.202,22	1.206,25	1.209,48	1.214,44
timol	1.055,95	1.058,06	1.061,21	1.066,12	1.065,39	1.068,03	1.072,73	1.073,89	1.076,29	1.080,62	1.082,16
álcool fenchil	1.021,81	1.024,16	1.027,72	1.033,00	1.032,39	1.035,39	1.040,90	1.042,26	1.045,11	1.050,28	1.052,53
(-)-borneol	1.168,16	1.171,74	1.176,09	1.181,80	1.183,33	1.187,42	1.193,46	1.196,20	1.201,46	1.205,41	1.213,64
B[2.2.2]	1.084,32	1.086,89	1.090,38	1.095,54	1.095,34	1.098,35	1.103,17	1.104,48	1.107,11	1.111,04	1.112,56

**trans*-CAR = (-)-*trans*-cariofileno; B[2.2.1] = biciclo[2.2.1] hepta-2,5-dieno; VNB = 5-vinil-2-norborneno; ENB = 5-etilideno-2-norborneno; B[2.2.2] = biciclo[2.2.2]octano-2,5-diona; B[4.1.0] = biciclo[4.1.0]heptano-7-ácido carboxílico.

Tabela 4.2 - continuação

Composto*	Índice de retenção										
	Temperatura, °C										
	100	110	120	130	140	150	160	170	180	190	200
safrol	1.283,85	1.288,55	1.292,31	1.296,49	1.302,03	1.304,52	1.311,06	1.313,23	1.316,41	1.322,87	1.327,59
isosafrol	1.366,00	1.371,13	1.375,50	1.380,21	1.386,25	1.389,55	1.396,49	1.399,67	1.403,93	1.410,80	1.416,84
3-BHA			1.482,38	1.483,21	1.483,90	1.484,89	1.486,72	1.489,69	1.495,82	1.505,38	1.512,55
α -humeleno		1.414,90	1.421,76	1.429,44	1.436,36	1.443,62	1.451,26	1.458,72	1.464,78	1.474,71	1.482,13
<i>trans</i> -CAR			1.452,88	1.460,54	1.467,64	1.474,88	1.482,53	1.489,69	1.488,98	1.493,68	1.495,86

**trans*-CAR = (-)-*trans*-cariofileno; B[2.2.1] = biciclo[2.2.1] hepta-2,5-dieno; VNB = 5-vinil-2-norborneno; ENB = 5-etilideno-2-norborneno; B[2.2.2] = biciclo[2.2.2]octano-2,5-diona; B[4.1.0] = biciclo[4.1.0]heptano-7-ácido carboxílico.

Os outros hidrocarbonetos e terpenos analisados tiveram curva I_R vs. T quase linear, como a apresentada pelo metilciclohexano (Figura 4.4). As curvas estão apresentadas no Apêndice D.

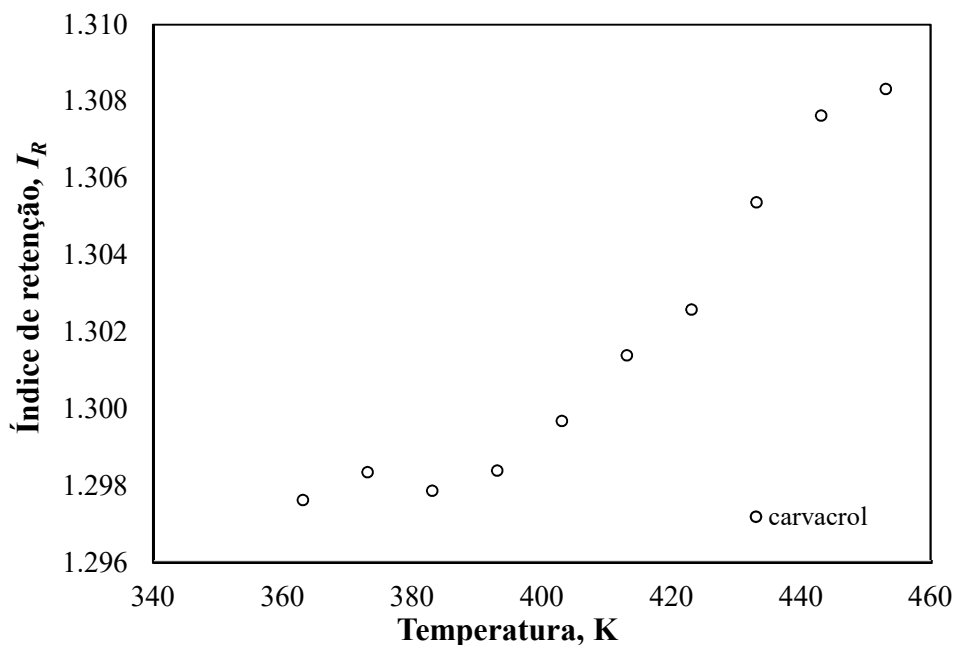


Figura 4.2 - Ilustração da dependência do índice de retenção de Kováts em função da temperatura para o carvacrol.

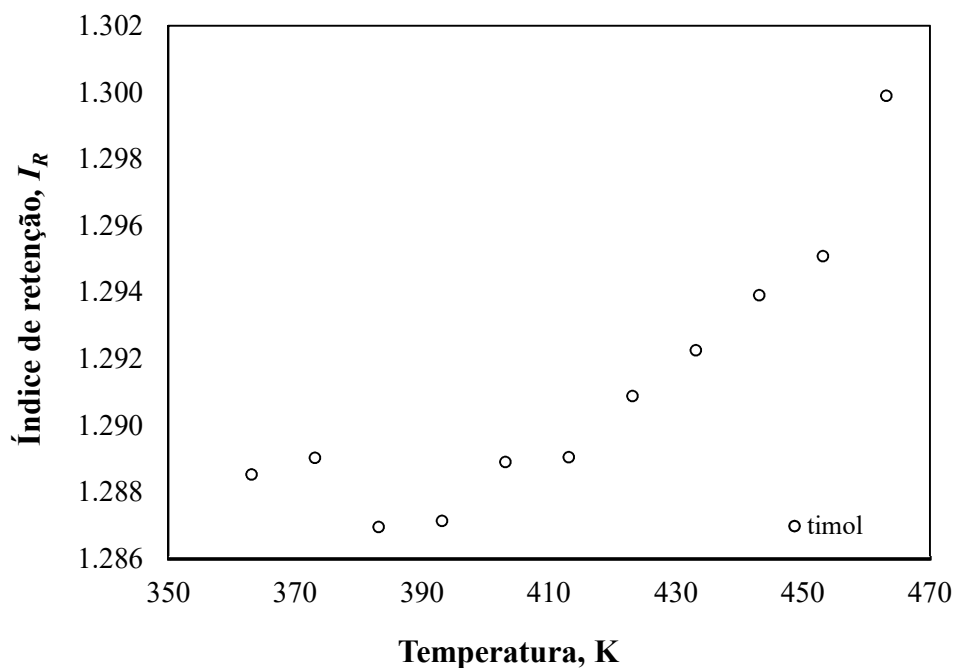


Figura 4.3 - Ilustração da dependência do índice de retenção de Kováts em função da temperatura para o timol.

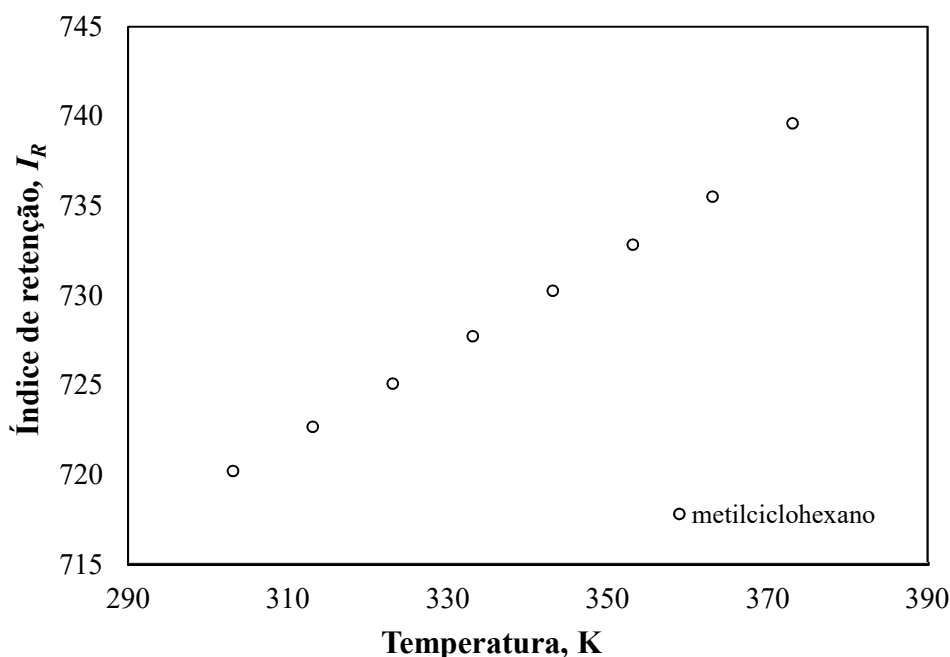


Figura 4.4 - Ilustração da dependência do índice de retenção de Kováts em função da temperatura para o metilciclohexano.

Ajustes preliminares da série de Taylor na forma da Equação 2.16 (item 2.6.1.1) apresentaram desvios bastante elevados para os coeficientes a_0 , a_1 e a_2 , principalmente para compostos com comportamento $I = f(T)$ similar ao do 2,2-dimetilbutano (Figura 4.1). Com o intuito de encontrar a melhor relação do índice de retenção em função da temperatura, a equação geral da série de Taylor (Equação 2.15) foi estudada na forma da Equação 2.16 e nas formas com dois e quatro parâmetros, respectivamente

$$I = a_0 + \frac{a_1}{T} \quad (4.1)$$

$$I = a_0 + \frac{a_1}{T} + a_2 \ln T + a_3 T \quad (4.2)$$

Como ilustração, as Figuras 4.5, 4.6 e 4.7 mostram os ajustes I_R vs. T para o 2,2-dimetilbutano para as três formas da série de Taylor analisadas. Os coeficientes foram determinados por meio da técnica de regressão não linear usando o programa *Statistica* (Versão 7.0) e uma rotina baseada no algoritmo de Levenberg-Marquardt.

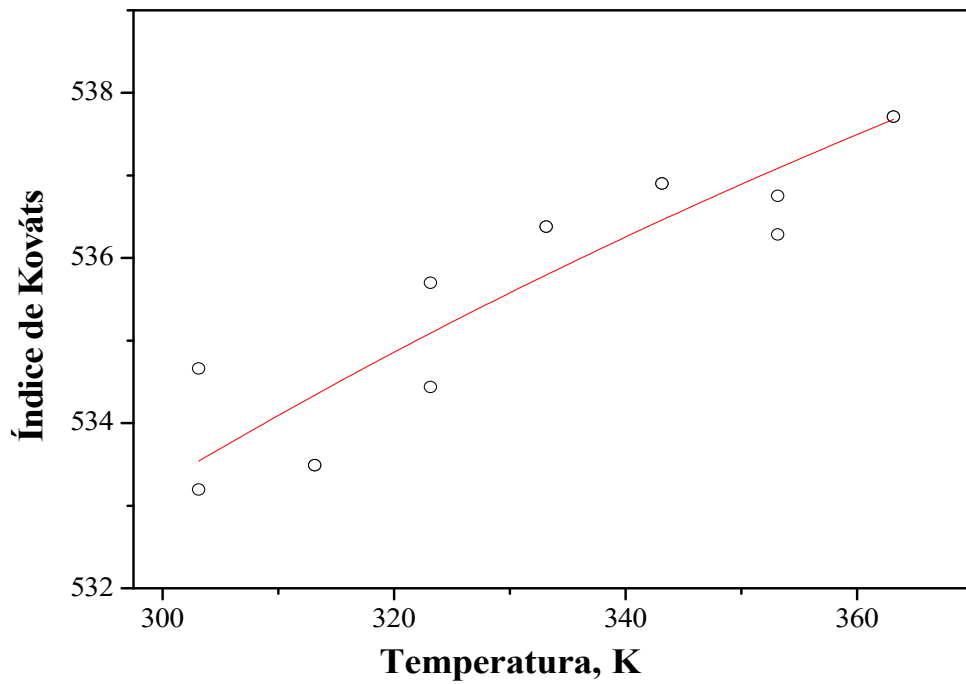


Figura 4.5 - Ajuste com a série de Taylor para o 2,2-dimetilbutano usando a Equação 4.1.

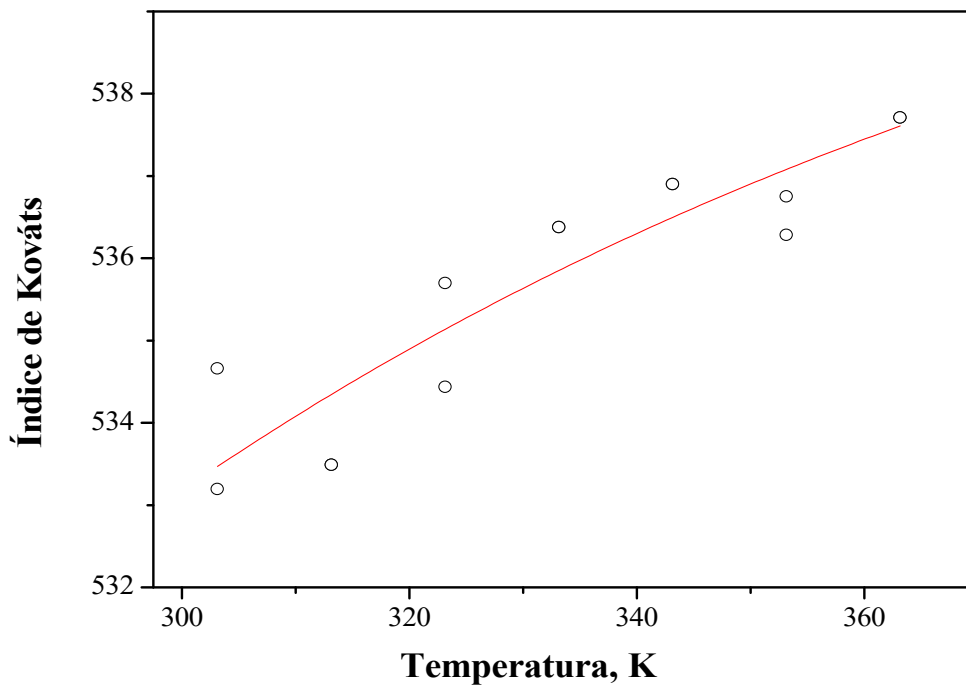


Figura 4.6 - Ajuste com a série de Taylor para o 2,2-dimetilbutano usando a Equação 2.16.

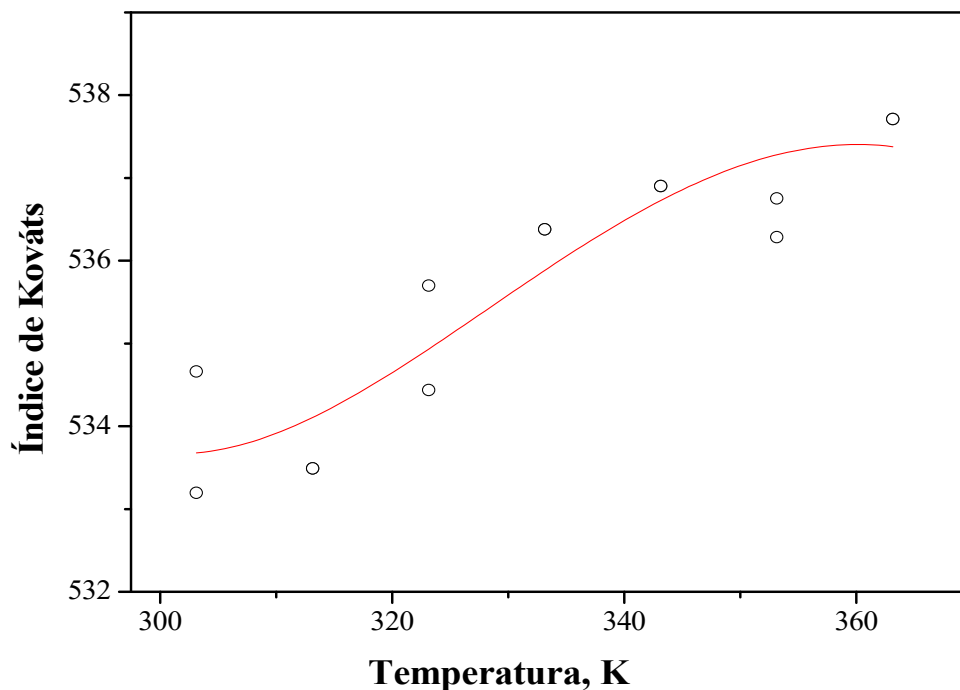


Figura 4.7 - Ajuste com a série de Taylor para o 2,2-dimetilbutano usando a Equação 4.2.

Os resultados de regressão para os ajustes das Figuras 4.5, 4.6 e 4.7 estão apresentados na Tabela 4.3. Conforme pode ser notado em tal tabela, o aumento do número de parâmetros não proporcionou melhor ajuste da série de Taylor aos dados. Baseado no valor coeficiente de determinação r^2 , pode-se notar que o melhor ajuste foi encontrado para a série de Taylor com dois parâmetros, mostrando r^2 mais próximo da unidade. Além disso, ao avaliar a relação do desvio sobre o valor do parâmetro (relação *desvio/parâmetro*), é possível verificar que o maior número de parâmetros produziu aumento da relação desvio/parâmetro. Por exemplo, o ajuste da Equação 4.1 (Figura 4.5) apresentou $r^2 = 0,81599$, com $a_0 = 558,6 \pm 3$ e $a_1 = -7.593,3 \pm 991,5$, e relação *desvio/parâmetro*, respectivamente, de 0,5 % e 13,1 %.

Os ajustes das Equações 2.16 e 4.2 (Figura 4.6 e 4.7, respectivamente) apresentaram valores da relação desvio/parâmetro bastante elevados, conforme mostra a Tabela 4.3. Desse modo, ajustes com número de parâmetros acima de três foram descartados para o estudo da relação $I = f(T)$. Os mesmos procedimentos foram feitos para os dados de todos os outros compostos estudados e a equação foi escolhida com base nesta análise.

Tabela 4.3 - Resultados de ajustes da série de Taylor com as Equações 2.16, 4.1 e 4.3 para o 2,2-dimetilbutano.

Parâmetro	Equação 4.1 ($r^2=0,81599$)		Equação 2.16 ($r^2=0,80050$)		Equação 4.2 ($r^2=0,80393$)	
	Valor	rel*, %	Valor	rel*, %	Valor	rel*, %
a_0	558,6 (± 3)	0,5	770,7 ($\pm 813,2$)	105,5	-45.949,4 ($\pm 42.787,6$)	93,1
a_1	-7.593,3 ($\pm 991,5$)	13,1	-17.919 ($\pm 39.604,5$)	221,0	1.315.710 ($\pm 1.221.790$)	92,7
a_2			-31,2 ($\pm 119,5$)	383,4	8.018,1 ($\pm 7.371,4$)	91,9
a_3					-12,1 ($\pm 11,1$)	91,5

*Relação desvio/parâmetro

4.1.2 Obtenção da pressão de vapor

Após definida a equação de Taylor como a mais adequada para o ajuste e a obtenção dos coeficientes, a sua forma geral foi usada com dois ou com três parâmetros, de acordo com a curva I_R vs. T apresentada para cada composto examinado.

Os coeficientes a_0 , a_1 e a_2 estão apresentados nas Tabelas 4.4 e 4.5. Como pode ser visto em tais tabelas, os compostos submetidos à ajuste da série de Taylor com dois parâmetros apresentaram coeficiente de determinação, r^2 , próximo de 0,99, enquanto os hidrocarbonetos alifáticos apresentaram r^2 em torno de 0,90.

Com base na ideia geral do comportamento de I_R em função da temperatura e da qualidade da regressão, a pressão de vapor dos compostos analisados foi calculada utilizando-se a Equação 2.14 (item 2.6.1.1). Para tanto, foi necessário calcular o valor da pressão de vapor para os n -alcanos de referência, P_z e P_{z+1} , usando a equação de Cox

$$\ln\left(\frac{P}{p_0}\right) = \left(1 - \frac{T_0}{T}\right) \exp(A_0 + A_1T + A_2T^2) \quad (4.3)$$

em que os coeficientes A_0 , A_1 e A_2 foram retirados de Růžička e Majer (1994); T_0 é temperatura de referência no ponto normal de ebulição e p_0 é a pressão de referência igual a 101,325 kPa.

Tabela 4.4 - Coeficientes da Equações 2.16 (a_0 , a_1 e a_2) e 4.1 (a_0 e a_1) para os hidrocarbonetos alifáticos estudados.

Composto	Coeficientes			r^2
	a_0	a_1	a_2	
2,2-dimetilbutano	558,6 ± 3,0	-7.593,3 ± 991,5	-	0,8160
2,3-dimetilbutano	590,1 ± 2,4	-8.019,0 ± 785,7	-	0,9469
3-metilpentano	593,2 ± 2,1	-3.311,6 ± 706,1	-	0,8044
metilciclopentano	686,3 ± 1,2	-17.690,7 ± 408,2	-	0,9968
ciclohexano	743,5±9,5	-25.383,3±3.199,8	-	0,9044
2-metilhexano	-7.697,3±1.589,6	397.309,2±78.302,8	1.234,5±233,1	0,9294
3-metilhexano	694,9±1,6	-6.156,1±523,8	-	0,9529
2,2,4-trimetilpentano	721,3±1,3	-10.473,4±419,5	-	0,9828
metilciclohexano	819,2±2,5	-30.277,6±829,9	-	0,9918

Uma vez que a Equação 4.3 permite sua extrapolação para outras temperaturas, P_z e P_{z+1} puderam ser calculadas de acordo com a faixa de temperatura especificada para cada composto estudado (ver Tabelas 4.1 e 4.2). Desse modo, foi possível usar a Equação 2.18 para avaliar a dependência da pressão de vapor em função da temperatura ($\ln P_X^{vap}$ (Pa) vs. $1/T$) para ambos os hidrocarbonetos alifáticos e terpenos analisados no presente trabalho, as quais estão apresentadas nas Figuras 4.8, 4.9, 4.10 e 4.11. A Figura 4.8 mostra a curva $\ln P_X^{vap}$ (Pa) vs. $1/T$ para os hidrocarbonetos alifáticos, enquanto a Figura 4.9 mostra a curva $\ln P_X^{vap}$ (Pa) vs. $1/T$ para os compostos terpênicos da *mistura 1*, a Figura 4.10 a curva para os compostos terpênicos da *mistura 2*, e a Figura 4.11, por sua vez, mostra a curva para os compostos terpênicos bicíclicos da *mistura 3*.

É possível notar nessas figuras o efeito da cadeia carbônica, do formato molecular e das forças de atração sobre a pressão de vapor dos compostos examinados. Os hidrocarbonetos alifáticos (Figura 4.8), por exemplo, são todos apolares e suas estruturas se diferem em cadeia carbônica (de C_6 a C_8) e em cadeias lineares ramificadas e cíclicas. Pose-se verifica o deslocamento para baixo das curvas de pressão de vapor com o aumento da cadeia carbônica, indo do 2,2-dimetilbutano (C_6H_{14}) para o 2,2,4-trimetilpentano (C_8H_{18}).

Tabela 4.5 - Coeficientes da Equações 2.16 (a_0 , a_1 e a_2) e 4.1 (a_0 e a_1) para compostos terpênicos estudados.

Composto	Coeficientes			r^2
	a_0	a_1	a_2	
(+)- α -pineno	1.076,5 \pm 3,8	-47.760,2 \pm 1363,1	-	0,9920
(+)- β -pineno	1.138,6 \pm 2,8	-54.866,7 \pm 1.022,9	-	0,9965
safrol	1.484,8 \pm 5,9	-75.505,3 \pm 2.442,4	-	0,9953
isosafrol	1.599,0 \pm 7,1	-87.690,9 \pm 2.968,5	-	0,9949
(+)-3-careno	1.144,7 \pm 2,4	-46.179,4 \pm 895,8	-	0,9965
(+)-limoneno	1.145,0 \pm 2,3	-41.952,6 \pm 856,4	-	0,9956
linalol	1.126,5 \pm 1,8	-9.917,5 \pm 661,7	-	0,9622
(-)- α -terpineol	1.337,2 \pm 3,6	-53.857,9 \pm 1.405,4	-	0,9933
timol metil éter	1.312,2 \pm 3,1	-29.796,8 \pm 1.246,0	-	0,9830
3-BHA	-18.471,0 \pm 1.557,1	1.160.531,0 \pm 94.936,0	2.846,0 \pm 220,3	0,9914
α -humuleno	1.763,6 \pm 6,4	-134.605,8 \pm 2.715,8	-	0,9964
carvacrol	-2.706,6 \pm 446,7	213.782,9 \pm 25.834,2	579,4 \pm 63,8	0,9889
timol	-3.804,8 \pm 575,9	282.784,4 \pm 33.608,6	732,0 \pm 82,1	0,9695
γ -terpineno	1.171,5 \pm 3,8	-39.988,5 \pm 1.465,8	-	0,9868
<i>p</i> -cimeno	1.156,0 \pm 4,6	-46.524,4 \pm 1.791,5	-	0,9855
(-)- <i>trans</i> -cariofileno	1.717,1 \pm 10,7	-102.983,5 \pm 4.615,0	-	0,9843
terpinen-4-ol	1.357,1 \pm 6,8	-65.676,0 \pm 2.663,4	-	0,9840
terpinoleno	1.210,2 \pm 3,2	-43.443,0 \pm 1.259,6	-	0,9917
biciclo[2.2.1]hepta-2,5-dieno	664,4 \pm 1,2	-4455 \pm 416,3	-	0,8972
5-vinil-2-norborneno	966,0 \pm 2,5	-29.247,8 \pm 915,7	-	0,9836
5-etilideno-2-norborneno	967,8 \pm 2,3	-19.094,7 \pm 870,2	-	0,9658
álcool fenchil	1.301,6 \pm 8,8	-71.874,5 \pm 3.697,8	-	0,9472
(-)-borneol	1.366,3 \pm 7,1	-76.571,9 \pm 2.946,3	-	0,9726
biciclo[2.2.2]octano-2,5-dione	1.486 \pm 4,4	-86.906,0 \pm 1.883,1	-	0,9912
biciclo[4.1.0]heptano-7-ácido carboxílico	-10.441 \pm 1740	719.122 \pm 101.032	1658,2 \pm 248,2	0,9830

O efeito da forma molecular é evidência analisando as curvas de pressão de vapor do 2,2,4-trimetilpentano (C_8H_{18}) e a do metilciclohexano (C_7H_{14}). O 2,2,4-trimetilpentano tem linear e ramificado, que diminuem a atração intermolecular. Por isso, embora tenha cadeia carbônica maior que o metilciclohexano, apresenta pressão de vapor menor pois a forma cíclica do metilciclohexano favorece maior contato molecular e maior força de atração.

O mesmo efeito foi verificado para os terpenos analisados. O α -terpineol e o linalol (terpenos da mistura 1, Figura 4.9), que fazem ligação de hidrogênio, devido a presença dos grupos hidroxilas, mostraram curvas de pressão de vapor menores que as dos hidrocarbonetos

terpênicos, que possuem forças de atração mais fracas. De maneira semelhante, entre os terpenos da mistura 2 (Figura 4.10), o terpinoleno ($C_{10}H_{16}$) teve a maior curva de pressão de vapor, devido ser hidrocarboneto e possuir menores forças intermoleculares. Também fica evidente o efeito do tamanho da cadeia carbônica do α -humuleno e do (-)-*trans*-cariofileno, ambos com fórmula $C_{15}H_{14}$, e da aditividade de grupos mostrada pelo 3-BHA, que faz duas ligações de hidrogênio pelos grupos -OH e -O- presentes em sua molécula. Por isso esses compostos apresentaram curvas de pressão de vapor mais baixas.

Por fim, foram apresentadas menores de curvas de pressão de vapor os compostos bicíclicos da mistura 3 (Figura 4.11) com maiores forças intermoleculares, como é o caso do biciclo[4.1.0]heptano-7-ácido carboxílico e do biciclo[2.2.2]octano-2,5-diona.

Em posse das curvas de pressão de vapor em função da temperatura, foi possível obter as constantes da equação de Antoine,

$$\ln P_{vap}(\text{Pa}) = A - \frac{B}{T(\text{K}) + C} \quad (4.4)$$

por meio do ajuste não linear (*Statistica*, Versão 7.0). Os resultados estão apresentados na Tabela 4.6. Todos os compostos analisados apresentaram r^2 acima de 0,999 e os ajustes são os mostrados nas Figuras 4.8, 4.9, 4.10 e 4.11.

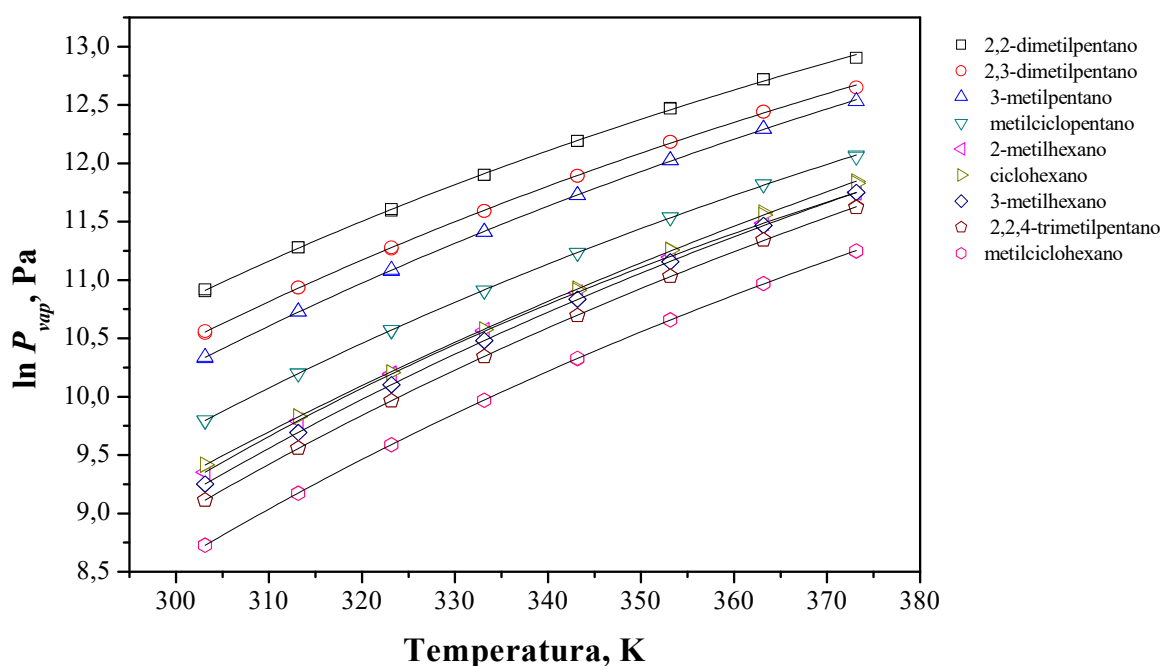


Figura 4.8 - Curvas de pressão de vapor em função da temperatura obtidas pela Equação 2.14 e ajuste da equação de Antoine (Equação 4.4) para os hidrocarbonetos alifáticos analisados.

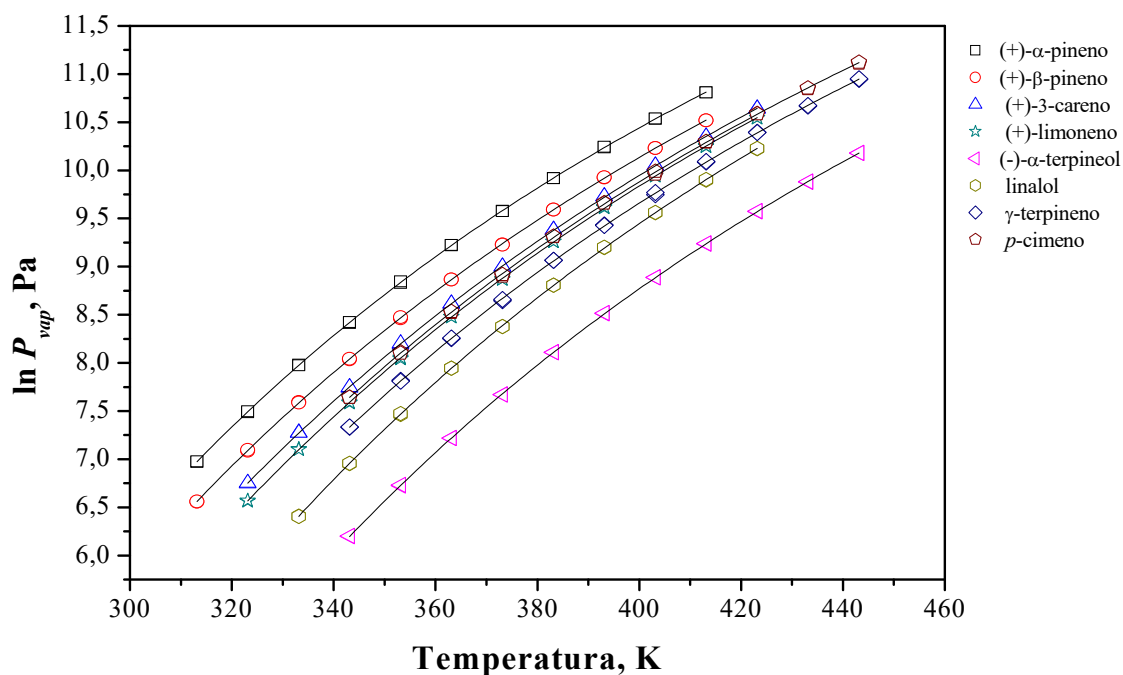


Figura 4.9 - Curvas de pressão de vapor em função da temperatura obtidas pela Equação 2.14 e ajuste da equação de Antoine (Equação 4.4) para os terpenos da *mistura 1*.

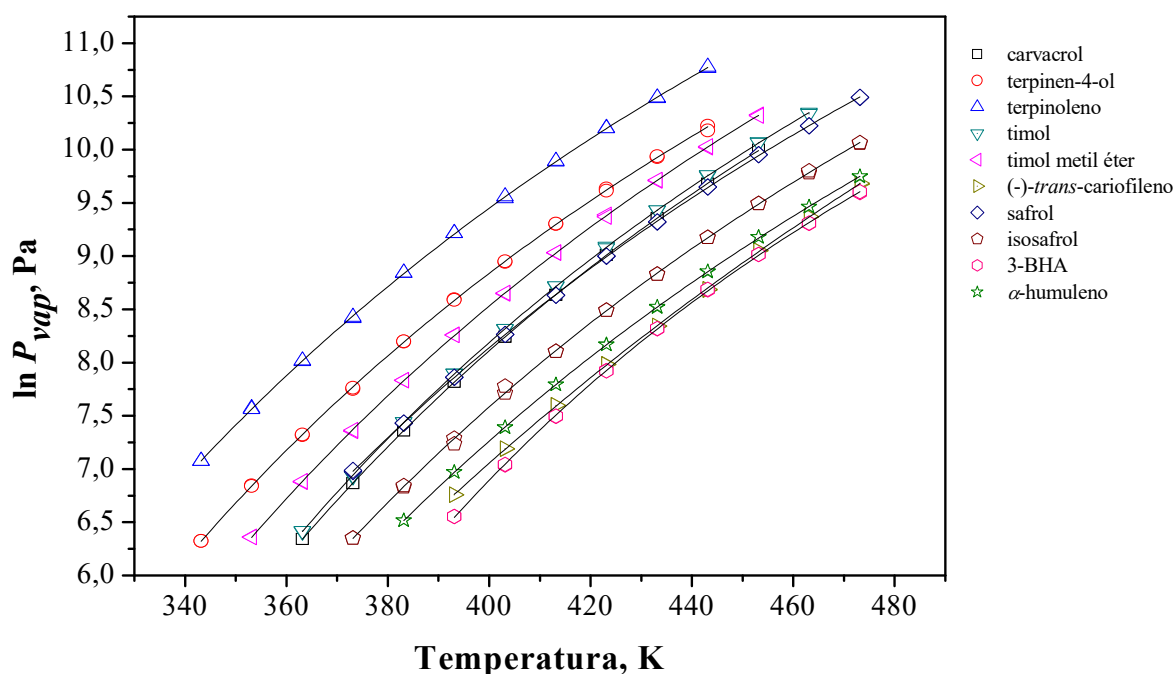


Figura 4.10 - Curvas de pressão de vapor em função da temperatura obtidas pela Equação 2.14 e ajuste da equação de Antoine (Equação 4.4) para os terpenos da *mistura 2*.

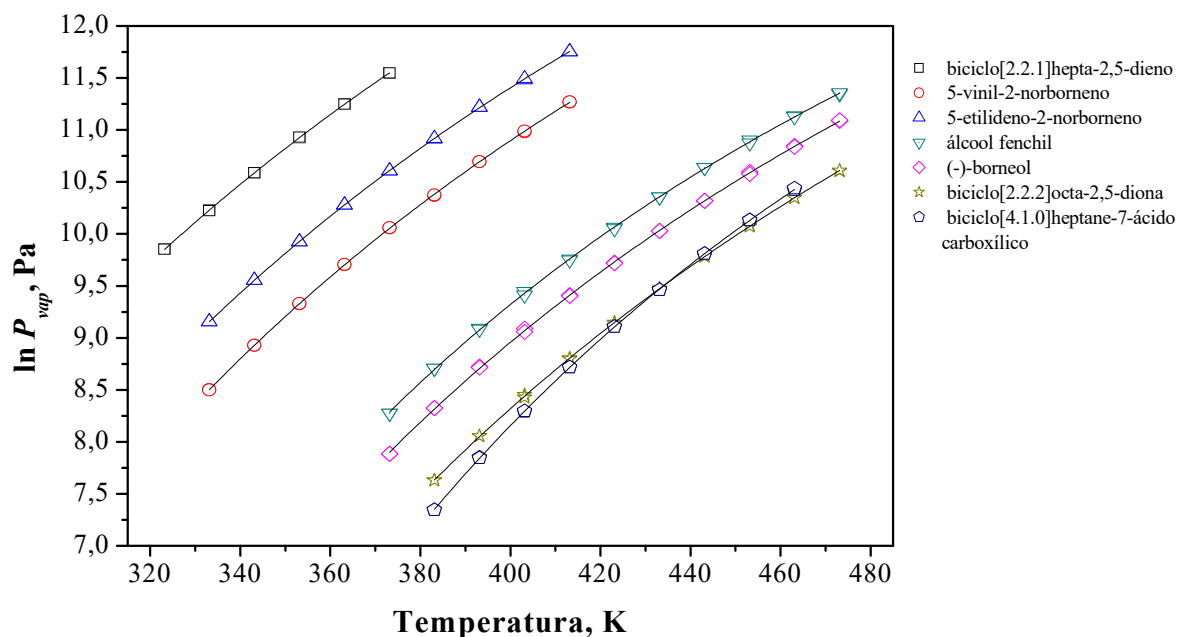


Figura 4.11 - Curvas de pressão de vapor em função da temperatura obtidas pela Equação 2.14 e ajuste da equação de Antoine (Equação 4.4) para os terpenos da *mistura 3*.

Tabela 4.6 - Constantes de Antoine para todos os compostos analisados obtidas por meio de ajustes das curvas de pressão de vapor das Figuras 4.8, 4.9, 4.10 e 4.11.

Composto*	Constantes de Antoine			r^2
	<i>A</i>	<i>B</i>	<i>C</i>	
2,2-dimetilbutano	18,9416 ± 0,5428	1684,76 ± 262,57	-93,42 ± 18,75	0,99940
2,3-dimetilbutano	19,4175 ± 0,4298	1985,68 ± 220,32	-79,16 ± 14,15	0,99970
3-metilpentano	20,0900 ± 0,3526	2341,16 ± 192,09	-63,15 ± 11,13	0,99987
metilciclopentano	19,8910 ± 0,2384	2428,34 ± 130,17	-62,61 ± 7,28	0,99994
ciclohexano	22,7360 ± 0,6881	4174,42 ± 476,50	10,22 ± 19,69	0,99980
2-metilhexano	18,5508 ± 0,1430	1827,59 ± 65,98	-104,45 ± 4,14	0,99996
3-metilhexano	20,3056 ± 0,1132	2648,46 ± 61,60	-63,57 ± 3,15	0,99999
2,2,4-trimetilpentano	20,5987 ± 0,1451	2870,14 ± 81,95	-53,23 ± 4,01	0,99998
metilciclohexano	19,7761 ± 0,1330	2608,63 ± 71,40	-67,08 ± 3,66	0,99998
(+)- α -pineno	20,1190 ± 0,0975	3185,64 ± 56,06	-70,82 ± 2,51	0,99999
(+)- β -pineno	20,6422 ± 0,1360	3597,34 ± 81,82	-57,67 ± 3,39	0,99998
safrol	20,6833 ± 0,2299	3966,4 ± 154,13	-83,83 ± 6,47	0,99997
isosafrol	20,5880 ± 0,2017	4032,42 ± 132,71	-89,98 ± 5,38	0,99998
(+)-3-careno	20,8529 ± 0,1378	3699,88 ± 84,81	-60,77 ± 3,5	0,99998
(+)-limoneno	20,9228 ± 0,1385	3741,14 ± 84,79	-62,53 ± 3,44	0,99998
linalol	21,2750 ± 0,1395	3868,22 ± 84,12	-72,98 ± 3,26	0,99998
(-)- α -terpineol	20,5712 ± 0,1041	3749,37 ± 63,75	-82,31 ± 2,59	0,99999
timol metil éter	20,8794 ± 0,1598	3866,04 ± 99,65	-86,93 ± 3,99	0,99997
3-BHA	17,8664 ± 0,2204	2445,71 ± 111,41	-177,09 ± 5,73	0,99993

Tabela 4.6 - continuação

Composto*	Constantes da equação de Antoine			r^2
	A	B	C	
α -humuleno	20,2776 \pm 0,1371	4032,42 \pm 91,79	-90,16 \pm 3,79	0,99999
carvacrol	20,2357 \pm 0,1013	3507,85 \pm 59,48	-110,70 \pm 2,48	0,99999
timol	19,9899 \pm 0,1286	3325,53 \pm 74,66	-118,23 \pm 3,23	0,99998
γ -terpineno	20,7221 \pm 0,2343	3621,19 \pm 148,10	-72,69 \pm 6,42	0,99993
<i>p</i> -cimeno	20,4691 \pm 0,2614	3449,46 \pm 164,43	-74,25 \pm 7,44	0,99991
(-)- <i>trans</i> -cariofileno	25,1631 \pm 0,7221	7809,29 \pm 667,75	31,24 \pm 19,73	0,99993
terpinen-4-ol	20,2207 \pm 0,2664	3572,22 \pm 161,11	-86,16 \pm 6,77	0,99991
terpinoleno	20,9594 \pm 0,2207	3828,42 \pm 141,91	-67,37 \pm 5,92	0,99995
B[2.2.1]	22,6586 \pm 0,4911	4183,57 \pm 344,22	3,43 \pm 14,39	0,99997
VNB	21,1942 \pm 0,237	3648,92 \pm 153,9	-45,65 \pm 6,82	0,99996
ENB	21,4477 \pm 0,1868	3667,36 \pm 125,39	-34,8 \pm 5,71	0,99998
álcool fenchil	19,3271 \pm 0,2757	2869,94 \pm 168,46	-113,19 \pm 8,91	0,99980
(-)-borneol	19,8373 \pm 0,3566	3289,4 \pm 226,09	-97,64 \pm 10,84	0,99986
B[2.2.2]	20,6535 \pm 0,2496	3954,9 \pm 172,46	-79,38 \pm 7,5	0,99995
B[4.1.0]	18,8018 \pm 0,2952	2528,18 \pm 144,33	-162,43 \pm 6,97	0,99997

*B[221] = biciclo[2.2.1]hepta-2,5-dieno; VNB = 5-vinil-2-norborneno; ENB = 5-etilideno-2-norborneno; B[222] = biciclo[2.2.2]octano-2,5-diona; B[4.1.0] = biciclo[4.1.0]heptano-7-ácido carboxílico.

4.1.2.1 Extrapolação da Equação 2.14 para a obtenção de dados de pressão de vapor a 298,15 K

Considerando que a temperatura de 298,15 K é tida como referência em muitas aplicações, a Equação 2.14 (item 2.6.1.1) foi extrapolada com o intuito de obter a pressão de vapor dos compostos analisados nesta temperatura ($P_{vap}^{298,15\text{ K}}$). Assim, a Equação 2.14 foi reescrita como

$$\ln P_{vap}^{298,15\text{ K}} = \ln P_z + \frac{(100 \cdot z - I_X^{298,15\text{ K}}) \cdot \ln(P_z / P_{z+1})}{100} \quad (4.5)$$

em que $I_X^{298,15\text{ K}}$ é o índice de retenção a 298,15 K para uma substância X, extrapolado usando os coeficientes das Tabelas 4.4 e 4.5. As pressões de vapor dos *n*-alcanos também foram usadas a 298,15 K por extrapolação feita pela equação de Cox (Equação 4.3).

Os dados de $P_{vap}^{298,15 K}$ encontrados para todos os compostos analisados, e uma comparação com os dados da literatura, estão apresentados na Tabela 4.7. A acurácia dos resultados foi baseada no desvio relativo e no desvio relativo médio, conforme, respectivamente

Tabela 4.7 - Dados de pressão de vapor para os compostos analisados calculados através da Equação 4.5 e uma comparação com os dados da literatura.

Composto*	$I_X^{298,15 K}$	* $P_{vap}^{lit.}$, kPa	$P_{vap}^{calc.}$, kPa	d_R , %	d_M , %
2,2-dimetilbutano	534,91	42,548 ^a ; 42,551 ^b	44,648	-4,94 ^a ; -4,93 ^b	4,94
2,3-dimetilbutano	564,90	31,34 ^a ; 31,304 ^{b,1}	30,967	1,19 ^a ; 1,08 ^{b,1}	1,14
3-metilpentano	583,08	25,306 ^a ; 25,372 ^{b,1}	24,807	1,97 ^a ; 2,23 ^{b,1}	2,1
metilciclopentano	626,49	18,331 ^a ; 18,331 ^b	14,700	19,81 ^a ; 19,81 ^b	19,81
ciclohexano	658,36	13,010 ^a ; 12,999 ^{b,1}	9,980	23,29 ^a ; 23,22 ^{b,1}	23,26
2-metilhexano	668,96	8,788 ^a ; 8,781 ^b	8,805	-0,2 ^a ; -0,28 ^b	0,24
3-metilhexano	674,25	8,213 ^a ; 7,958 ^{b,2}	8,272	-0,71 ^a ; -3,94 ^{b,2}	2,33
2,2,4-TMP	686,17	6,578 ^a ; 6,601 ^{b,3}	7,185	-9,22 ^a ; -8,85 ^{b,3}	9,04
metilciclohexano	717,65	6,180 ^a ; 6,183 ^{b,4}	4,907	20,6 ^a ; 20,63 ^{b,4}	20,62
(+)- α -pineno	916,31	0,529 ^c ; 0,544 ^d ; 0,515 ^e	0,481	9,15 ^c ; 11,66 ^d ; 6,68 ^e	9,16
(+)- β -pineno	954,58	0,341 ^e	0,308	9,6 ^e	9,6
safrol	1231,55	0,009 ^f	0,0125	-32,95 ^f	32,95
isosafrol*	1304,88	0,004	0,0054	-25,93	25,93
(+)-3-careno	989,81	0,231 ^e ; 0,221 ^{b,1}	0,2049	11,3 ^e ; 7,29 ^{,1}	9,3
(+)-limoneno	1004,29	0,213 ^c ; 0,202 ^d ; 0,183 ^e ; 0,200 ^g	0,1731	18,71 ^c ; 14,29 ^d ; 5,59 ^e ; 13,43 ^g	13,01
linalol	1093,24	0,035 ^e ; 0,027 ^g ; 0,039 ^{b,5}	0,0615	-75,84 ^e ; -127,94 ^g ; -57,8 ^{b,5}	87,19
(-)- α -terpineol	1156,56	0,015 ^a ; 0,012 ^h	0,0297	-97,95 ^a ; -147,44 ^h	122,7
timol metil éter	1212,26	-	0,0156	-	-
3-BHA	1636,80	-	0,000123	-	-
α -humuleno	1312,13	-	0,00487	-	-
carvacrol	1311,62	0,0035 ⁱ	0,00497	-41,51 ⁱ	41,51
timol	1314,30	0,004 ^e ; 0,0036 ⁱ	0,00482	-20,39 ^e ; -33,77 ⁱ	27,08
γ -terpineno	1037,38	0,103 ^d ; 0,137 ^e ; 0,124 ^h	0,11784	-14,41 ^d ; 14,11 ^e ; 4,96 ^h	11,16
<i>p</i> -cimeno	999,96	0,193 ^e ; 0,180 ^h	0,18208	5,85 ^e ; -1,17 ^h	3,51
<i>trans</i> -CAR	1371,69	-	0,00248	-	-
terpinen-4-ol	1136,82	0,0195 ^e ; 0,0173 ^h	0,0373	-91,06 ^e ; -114,98 ^h	103,02
terpinoleno	1064,49	0,099 ^d	0,086	13,16 ^d	13,16
B[2.2.1]	649,47	-	10,724	-	-
VNB	867,93	-	0,8454	-	-
ENB	903,73	-	0,5575	-	-
álcool fenchil	1060,56	-	0,090	-	-
(-)-borneol	1109,50	-	0,0509	-	-
B[2.2.2]	1194,52	-	0,0190	-	-
B[4.1.0]	1418,94	-	0,00139	-	-

a) NIST online – dados a 298,15 K; b) NIST/TDE do Aspen® - dados a 298,15 K, exceto: **b,1** (298,14 K); **b,2** (297,48 K); **b,3** (298,20 K); **b,4** (298,13 K); **b,5** (297,03 K); c) Fichan *et al.* (1999) – dados a 298,15 K; d) Li *et al.* (1998) – dados a 296,65 K; e) Hoskovec *et al.* (2005) – dados a 298,15 K; f) Klauck *et al.* (2014) – dados a 298,15 K; g) Espinosa Díaz *et al.* (1999) – dados a 298,15 K (*d*-linalol); h) Lucia *et al.* (2009) – dados a 298,15 K; i) van Roon *et al.* (2002) – dados a 298,15 K. *2,2,4-TMP = 2,2,4-trimetilpentano; *trans*-CAR = (-)-*trans*-cariofileno; B[2.2.1] = biciclo[2.2.1] hepta-2,5-dieno; VNB = 5-vinil-2-norborneno; ENB = 5-etilideno-2-norborneno; B[2.2.2] = biciclo[2.2.2]octano-2,5-diona; B[4.1.0] = biciclo[4.1.0]heptano-7-ácido carboxílico; * valor predito pelo TDE do Aspen®

$$d_R = \left(\frac{P_{vap}^{calc.} - P_{vap}^{lit.}}{P_{vap}^{lit.}} \right) \times 100 \quad (4.6)$$

$$d_M = \frac{1}{n} \sum_{i=1}^n \left(\frac{|P_{vap}^{calc.} - P_{vap}^{lit.}|}{P_{vap}^{lit.}} \times 100 \right) \quad (4.7)$$

nas quais $P_{vap}^{calc.}$ é a pressão de vapor (apresentada em kPa) a 298,15 K calculada no presente trabalho de Tese e $P_{vap}^{lit.}$ é a pressão de vapor (kPa) da literatura.

Com uma breve inspeção do desvio relativo médio (d_M) da Tabela 4.7, é possível notar que entre os hidrocarbonetos alifáticos, apenas o 2,2-dimetilbutano (4,94 %), 2,3-dimetilbutano (1,14 %), 3-metilpentano (2,10 %), 2-metilhexano (0,24 %) e o 3-metilhexano (2,33 %), apresentaram desvio relativo abaixo de 5,0 %. Quanto aos terpenos, exceto o *p*-cimeno (3,51 %), todos apresentaram desvio relativo extremamente elevado, como é o caso do linalol (87,19 %), do (-)- α -terpineol (127,7 %) e do terpinen-4-ol (103,02 %).

A Equação 4.5, portanto, demonstrou-se não adequada para a obtenção dos dados de pressão de vapor a 298,15 K para a maioria dos compostos estudados. Contudo, utilizou-se procedimentos baseado em Hoskovec *et al.* (2005) para melhorar a acurácia da Equação 4.5

Conforme apresentaram Hoskovec *et al.* (2005), a diferença nos logaritmos dos coeficientes de atividade de duas substâncias (*i* e *j*) pode ser aproximada pela diferença correspondente na solubilidade do gás ideal, X_i^g , advinda da equação de van't Hoff,

$$\begin{aligned} \ln \gamma_i^\infty - \ln \gamma_j^\infty &\approx \ln X_i^g - \ln X_j^g \\ &\cong \frac{\Delta S_{vap}^{b,i}(T_{nb}^i - T)}{RT} - \frac{\Delta S_{vap}^{b,j}(T_{nb}^j - T)}{RT} \end{aligned} \quad (4.8)$$

em que ΔS_{vap}^b é a entropia na temperatura normal de ebulição T_{nb} .

A Equação 4.8 mostra que a razão dos coeficientes de atividades pode ser representada pelos parâmetros não cromatográficos ΔS_{vap}^b e T_{nb} , que inseridos na Equação 4.5, obtém-se

$$\ln P_{vap,X}^{298,15K} = \ln P_z + (\ln X_z^g - \ln X_X^g) + \frac{(100.z - I_X^{298,15K}) [\ln(P_z / P_{z+1}) + (\ln X_z^g - \ln X_{z+1}^g)]}{100} \quad (4.9)$$

em que o subscrito X representa o componente analisado e os subscritos z e $z+1$ referem-se aos n -alcanos de referência.

A temperatura normal de ebulição para muitos compostos está disponível na literatura (como na NIST, por exemplo). Já ΔS_{vap}^b pode ser estimada a partir da estrutura química, de acordo com os procedimentos encontrados em Myrdal *et al.* (1996), que envolve o número de ligação torsional (estimativa da flexibilidade molecular)

$$\tau = \sum (\text{SP3} + 0,5 \text{ SP2} + 0,5 \text{ RING}) - 1 \quad (4.10)$$

e o número de grupos polares capazes de fazer ligação de hidrogênio

$$N_{LH} = \frac{\sqrt{OH + COOH}}{MM} \quad (4.11)$$

em que, na Equação 4.10, SP3 e SP2 são respectivamente o número de átomos sp^3 e sp^2 em moléculas acíclicas; RING indica o número de anéis (cíclico ou aromático) no composto. Na Equação 4.11, OH refere-se à álcoois e fenóis, $COOH$ à ácidos carboxílicos e MM , à massa molar.

As informações das Equações 4.10 e 4.11 são, portanto, inseridas em

$$\Delta S_{vap}^{b,X} = 86 + 0,4 \tau_{efet.} + 1421 N_{LH} \quad (4.12)$$

e a entropia de vaporização na temperatura normal de ebulição de uma substância X ($\Delta S_{vap}^{b,X}$) pode ser estimada. Em caso de τ apresentar valor negativo, atribui-se a ele o valor “zero” (MYRDAL *et al.*, 1999), chamado neste trabalho de $\tau_{efet.}$. Baseado na estrutura química dos compostos analisados (Apêndice A), realizou-se os cálculos de $\tau_{efet.}$ e de N_{LH} , bem como a estimativa de $\Delta S_{vap}^{b,X}$. Os resultados estão apresentados na Tabela 4.8.

O mesmo procedimento foi realizado para obtenção da entropia de vaporização no ponto normal de ebulição dos n -alcanos de referência ($\Delta S_{vap}^{b,z}$ e $\Delta S_{vap}^{b,z+1}$) cujos valores podem ser vistos no Apêndice B. Com os dados de ΔS_{vap}^b dos compostos analisados e dos n -alcanos de

referência utilizados, e com o auxílio da Equação 4.9, foi possível calcular o novo valor da pressão de vapor a 298,15 K. Os resultados estão apresentados na Tabela 4.9, que também mostra uma comparação com os dados da literatura e com os resultados obtidos anteriormente pela Equação 4.5.

Tabela 4.8 - Número de ligação torsional efetivo ($\tau_{efet.}$), número de ligação de hidrogênio (N_{LH}) e estimação de $\Delta S_{b,X}^{vap}$ para os compostos analisados.

Composto	$\tau_{efet.}$	N_{LH}	$\Delta S_{b,X}^{vap}$ (J K ⁻¹ mol ⁻¹)
2,2-dimetilbutano	1,0	0,000	86,40
2,3-dimetilbutano	1,0	0,000	86,40
3-metilpentano	2,0	0,000	86,80
metilciclopentano	0,0	0,000	86,00
ciclohexano	0,0	0,000	86,00
2-metilhexano	3,0	0,000	87,20
3-metilhexano	3,0	0,000	87,20
2,2,4-trimetilpentano	2,0	0,000	86,80
metilciclohexano	0,0	0,000	86,00
(+)- α -pineno	0,0	0,000	86,00
(+)- β -pineno	0,0	0,000	86,00
safrol	1,5	0,000	86,60
isosafrol	1,0	0,000	86,40
(+)-3-careno	0,0	0,000	86,00
(+)-limoneno	0,0	0,000	86,00
linalol	3,5	0,006	96,61
(-)- α -terpineol	0,5	0,006	95,41
timol metil éter	1,5	0,000	86,60
3-BHA	1,5	0,006	94,48
α -humuleno	0,0	0,000	86,00
carvacrol	0,5	0,007	95,66
timol	0,5	0,007	95,66
γ -terpineno	0,5	0,000	86,20
<i>p</i> -cimeno	0,5	0,000	86,20
(-)- <i>trans</i> -cariofileno	0,0	0,000	86,00
terpinen-4-ol	0,5	0,006	95,41
terpinoleno	0,0	0,000	86,00
biciclo[2.2.1]hepta-2,5-dieno	0,0	0,000	86,00
5-vinil-2-norborneno	0,5	0,000	86,20
5-etilideno-2-norborneno	0,5	0,000	86,20
álcool fenchil	0,0	0,006	95,21
(-)-borneol	0,0	0,006	95,21

biciclo[2.2.2]octano-2,5-diona	0,0	0,000	86,00
biciclo[4.1.0]heptano-7-ácido carboxílico	0,0	0,007	96,14

Tabela 4.9 - Dados de pressão de vapor a 298,15 K para os compostos analisados obtidos pela Equação 4.9 e comparação com dados da literatura e com dados advindos da Equação 4.5.

Composto*	* P_{vap}^{lit} , kPa	$P_{vap}^{calc.}$, kPa	d_R , %	d_M , %	
				Eq. (4.5)	Eq. (4.9)
2,2-dimetilbutano	42,548 ^a ; 42,551 ^b	41,482	2,5 ^a ; 2,51 ^b	4,94	2,51
2,3-dimetilbutano	31,34 ^a ; 31,304 ^{b,1}	30,435	2,89 ^a ; 2,77 ^{b,1}	1,14	2,83
3-metilpentano	25,306 ^a ; 25,372 ^{b,1}	24,924	1,51 ^a ; 1,76 ^{b,1}	2,1	1,64
metilciclopentano	18,331 ^a ; 18,331 ^b	17,822	2,78 ^a ; 2,78 ^b	19,81	2,78
ciclohexano	13,010 ^a ; 12,999 ^{b,1}	12,827	1,41 ^a ; 1,32 ^{b,1}	23,26	1,37
2-metilhexano	8,788 ^a ; 8,781 ^b	8,809	-0,24 ^a ; -0,32 ^b	0,24	0,28
3-metilhexano	8,213 ^a ; 7,958 ^{b,2}	8,180	0,41 ^a ; -2,78 ^{b,2}	2,33	1,6
2,2,4-TMP	6,578 ^a ; 6,601 ^{b,3}	6,228	5,33 ^a ; 5,66 ^{b,3}	9,04	5,5
metilciclohexano	6,180 ^a ; 6,183 ^{b,4}	5,928	4,08 ^a ; 4,12 ^{b,3}	20,62	4,1
(+)- α -pineno	0,529 ^c ; 0,544 ^d ; 0,515 ^e	0,515	2,65 ^c ; 5,33 ^d ; 0,004 ^e	9,16	2,66
(+)- β -pineno	0,341 ^e	0,338	0,85 ^e	9,6	0,85
safrol	0,009 ^f	0,00971	-3,12 ^f	32,95	3,12
Isosafrol**	0,004	0,0046	-13,04	-	13,04
(+)-3-careno	0,231 ^e ; 0,221 ^{b,1}	0,232	-0,61 ^e ; -5,17 ^{b,1}	9,3	2,89
(+)-limoneno	0,213 ^c ; 0,202 ^d ; 0,183 ^e ; 0,200 ^g	0,185	13,3 ^c ; 8,58 ^d ; -0,69 ^e ; 7,67 ^g	13,01	7,56
linalol	0,035 ^e ; 0,027 ^g ; 0,039 ^{b,5}	0,0311	11,17 ^e ; -15,15 ^g ; 20,28 ^{b,5}	87,19	15,53
(-)- α -terpineol	0,015 ^a ; 0,012 ^h	0,0115	23,33 ^a ; 4,17 ^h	122,7	13,75
timol metil éter	-	0,0137	-	-	-
3-BHA	-	0,000045	-	-	-
α -humuleno	-	0,0014	-	-	-
carvacrol	0,0035 ⁱ	0,0030	15,38 ⁱ	41,51	15,38
timol	0,004 ^e ; 0,0036 ⁱ	0,0035	13,25 ^e ; 3,61 ⁱ	27,08	8,43
γ -terpineno	0,103 ^d ; 0,137 ^e ; 0,124 ^h	0,1345	-30,56 ^d ; 1,98 ^e ; -8,46 ^h	11,16	16,27
<i>p</i> -cimeno	0,193 ^e ; 0,180 ^h	0,1921	0,68 ^e ; -6,72 ^h	3,51	3,7
<i>trans</i> -CAR	-	0,0016	-	-	-
terpinen-4-ol	0,0195 ^e ; 0,0173 ^h	0,0189	3,28 ^e ; -8,83 ^h	103,02	6,06
terpinoleno	0,099 ^d	0,1106	-11,73 ^d	13,16	11,73
B[2.2.1]	-	2,209	-	-	-
VNB	-	0,389	-	-	-
ENB	-	0,5329	-	-	-
álcool fenchil	-	0,0261	-	-	-
(-)-borneol	-	0,0135	-	-	-
B[2.2.2]	-	0,0105	-	-	-
B[4.1.0]	-	0,0013	-	-	-

a) NIST online – dados a 298,15 K; b) NIST/TDE do Aspen® - dados a 298,15 K, exceto: b,1 (298,14 K); b,2 (297,48 K); b,3 (298,20 K); b,4 (298,13 K); b,5 (297,03 K); c) Fichan *et al.* (1999) – dados a 298,15 K; d) Li *et al.* (1998) – dados a 296,65 K; e) Hoskovec *et al.* (2005) – dados a 298,15 K; f) Klauck *et al.* (2014) – dados a 298,15 K; g) Espinosa Díaz *et al.* (1999); – dados a 298,15 K (*d*-linalol); h) Lucia *et al.* (2009) – dados a 298,15 K; i) van Roon *et al.* (2002) – dados a 298,15 K. *2,2,4-TMP = 2,2,4-trimetilpentano; *trans*-CAR = (-)-*trans*-cariofileno; B[2.2.1] = biciclo[2.2.1] hepta-2,5-dieno;

VNB = 5-vinil-2-norborneno; ENB = 5-etilideno-2-norborneno; B[2.2.2] = biciclo[2.2.2]octano-2,5-diona; B[4.1.0] = biciclo[4.1.0]heptano-7-ácido carboxílico; ** valor predito pelo TDE do Aspen®

Pode-se notar na Tabela 4.9, por meio do desvio relativo médio (d_M), que os dados de $P_{vap}^{298,15\text{ K}}$ calculados pela Equação 4.9 apresentaram-se em concordância com os dados da literatura. Verifica-se também que o termo de correção incluído na Equação 4.9 provocou diminuição no d_M para quase todos os compostos estudados, o que sugere o uso direto da Equação 4.9 em futuras análises de cromatografia gasosa.

Com exceção do 2,2,4-trimetilpentano, que apresentou $d_M = 5,5\%$, e do metilciclohexano (4,1%), todos os hidrocarbonetos alifáticos analisados apresentaram d_M abaixo de 3,0%. A respeito dos terpenos, que antes haviam apresentado desvios bastante elevados, somente o linalol (15,53%), (-)- α -terpineol (13,75%), carvacrol (15,38%), γ -terpineno (16,27%), terpinoleno (11,73%) apresentaram desvios acima de 10%. Esses dois últimos compostos, nos entanto, tiveram seus dados de P_{vap} publicados a 296,15 K (Li *et al.*, 1998), que explicado a menor pressão de vapor e o aumento dos desvios relativos.

Por outro lado, percebe-se que os compostos (+)- α -pineno (2,66%), (+)- α -pineno (0,85%), safrol (3,12%), (+)-3-careno (2,84%) e *p*-cimeno (3,7%) tiveram desvio relativo médio abaixo de 4,0%. É importante mencionar que as purezas dos compostos analisados (ver Tabela 3.1), em geral, tiveram pouca influência nos resultados de pressão de vapor obtidos. O terpinoleno, por exemplo, que apresentou $d_M = 11,73\%$, tem pureza de 85,0% (m/m), mas esse desvio pode estar mais relacionado com a temperatura, conforme já comentado. Isso reforça mais uma vantagem da técnica de cromatografia gasosa utilizada neste trabalho.

Considerando-se apenas os valores médios de P_{vap} dos compostos que têm dados disponíveis na literatura, pode-se traçar o gráfico do logaritmo da pressão de vapor dos dados da literatura ($\ln P_{vap}^{lit.}$) contra os dados obtidos pelas Equações 4.5 e 4.9 ($\ln P_{vap}^{calc.}$), para verificar a acurácia de obtenção de P_{vap} por tais equações. As Figuras 4.12 e 4.13 mostram respectivamente a comparação para os dados obtidos pela Equação 4.5 e 4.9.

Deve-se ressaltar que não foram encontrados na literatura dados de pressão de vapor a 298,15 K para os compostos timol metil éter, 3-BHA, (-)-*trans*-cariofileno, α -humuleno, isosafrol, biciclo[2.2.1]hepta-2,5-dieno, biciclo[4.1.0]heptano-7-ácido carboxílico, (-)-borneol, biciclo[2.2.2]octano-2,5-diona, 5-etilideno-2-norborneno, 5-vinil-2-norborneno e álcool fenchil. Portanto, além das curvas de pressão de vapor, apresentadas no item anterior, esses compostos tiveram seus respectivos dados de P_{vap} obtidos a 298,15 K pela primeira vez.

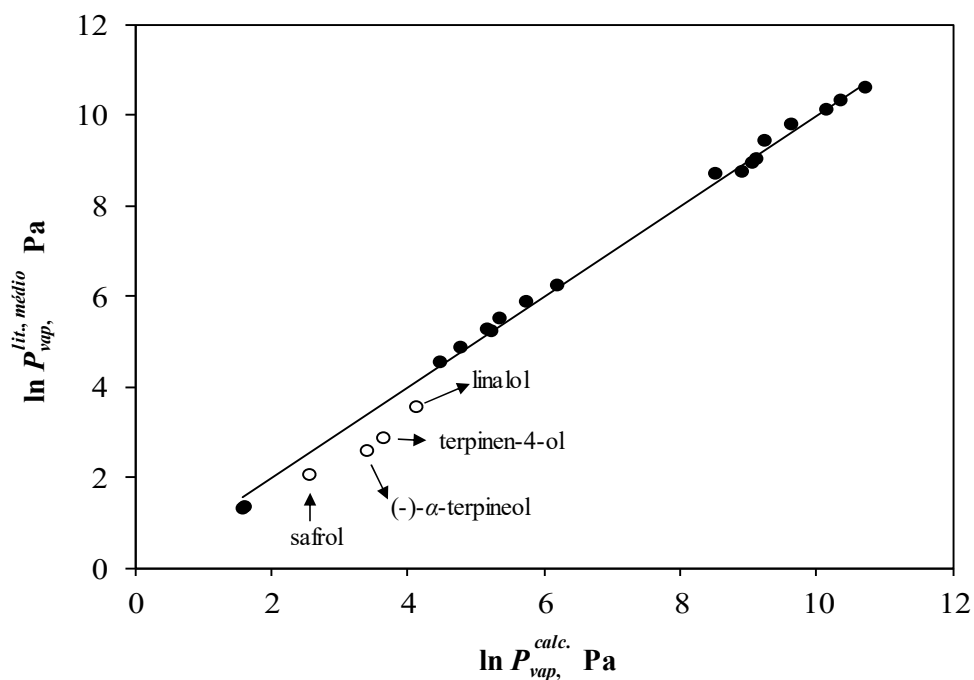


Figura 4.12 - Valores de pressão de vapor calculados a 298,15 K versus valores médios da literatura (Equação 4.5).

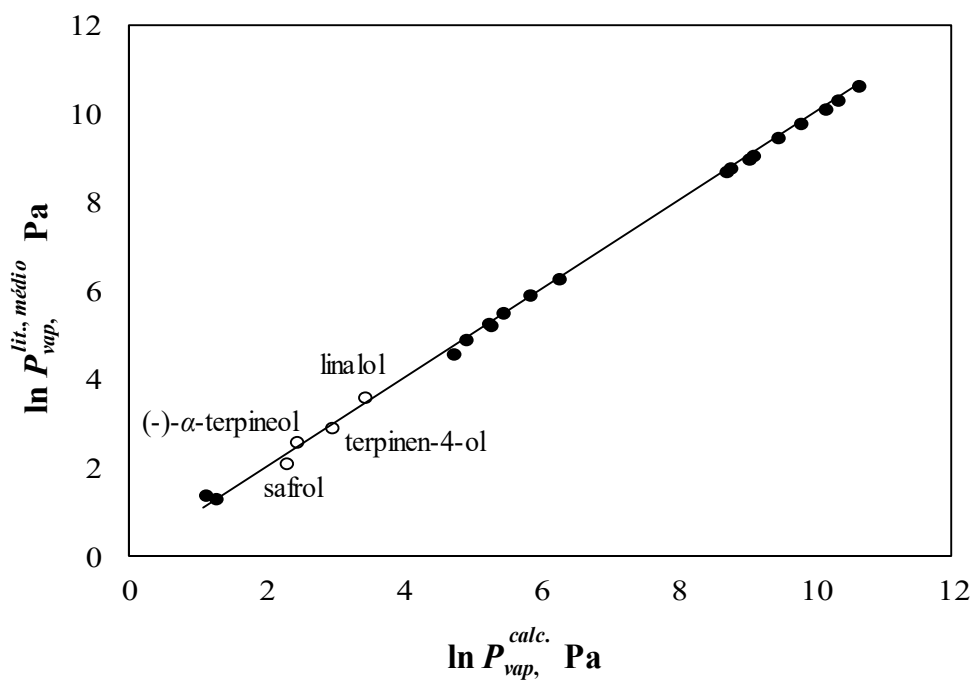


Figura 4.13 - Valores de pressão de vapor calculados a 298,15 K vs. valores médios da literatura (Equação 4.8).

4.1.3 Obtenção de dados de entalpias de vaporização e temperatura normal de ebulição

Sabendo-se que a entalpia de vaporização (ΔH_{vap}) tem relação direta com a pressão de vapor (ver item 2.5.1), os dados de ΔH_{vap} puderam ser derivados a partir da dependência da pressão de vapor com a temperatura. No presente trabalho de Tese, em vez de usar a clássica equação de Clausius-Clapeyron, foi usada a equação do tipo Kirchhoff-Rankin (KULIKOV *et al.*, 2001; HOSKOVEC *et al.*, 2005)

$$\ln P = \frac{1}{R} b_0 + \frac{1}{(RT)} b_1 + \frac{1}{R} b_2 \ln(T/T_0) \quad (4.13)$$

em que R é constante universal dos gases ($8,3145 \text{ kJ mol}^{-1}$); T_0 é uma temperatura de referência, escolhida arbitrariamente como $298,15 \text{ K}$.

Os coeficientes b_1 e b_2 da Equação 4.13 podem ser usados para obter as mudanças na entalpia molar de vaporização da fase líquida para a fase gasosa ($\Delta_l^g H_m^0$) em dada temperatura por

$$\Delta_l^g H_m^0 = -b_1 + b_2 T \quad (4.14)$$

Desse modo, é possível aplicar os dados de pressão de vapor em função da temperatura obtidos no item anterior (Figuras 4.8, 4.9, 4.10 e 4.11) na Equação 4.13 para obter os coeficientes b_0 , b_1 e b_2 . Os coeficientes foram obtidos pela técnica de regressão não linear, baseada no algoritmo de Levenberg-Marquardt, usando o *Statistica*, Versão 7.0).

Para todos os compostos analisados, o ajuste $\ln P$ vs. T da Equação 4.13 foi bastante adequado. Para ilustrar, estão apresentados na Figura 4.14 os ajustes do 2,2-dimetilbutano, 3-metilpentano, 3-BHA e (-)-*trans*-cariofileno (ver os demais ajustes no Apêndice E). Todos os ajustes apresentaram coeficiente de determinação (r^2) acima de 0,9999, o que explica a boa relação $P = f(T)$. Os coeficientes da Equação 4.13 para ambos os hidrocarbonetos alifáticos e terpenos estudados estão mostrados na Tabela 4.10.

A partir dos coeficientes b_1 e b_2 apresentados na Tabela 4.10 e definindo uma temperatura T , pode-se obter os dados de entalpia de vaporização. No presente trabalho, foi definido $T = 298,15 \text{ K}$ e, dessa forma, os dados de entalpia de vaporização foram obtidos na temperatura de referência de $298,15 \text{ K}$, $\Delta H_{vap}^{298,15 \text{ K}}$.

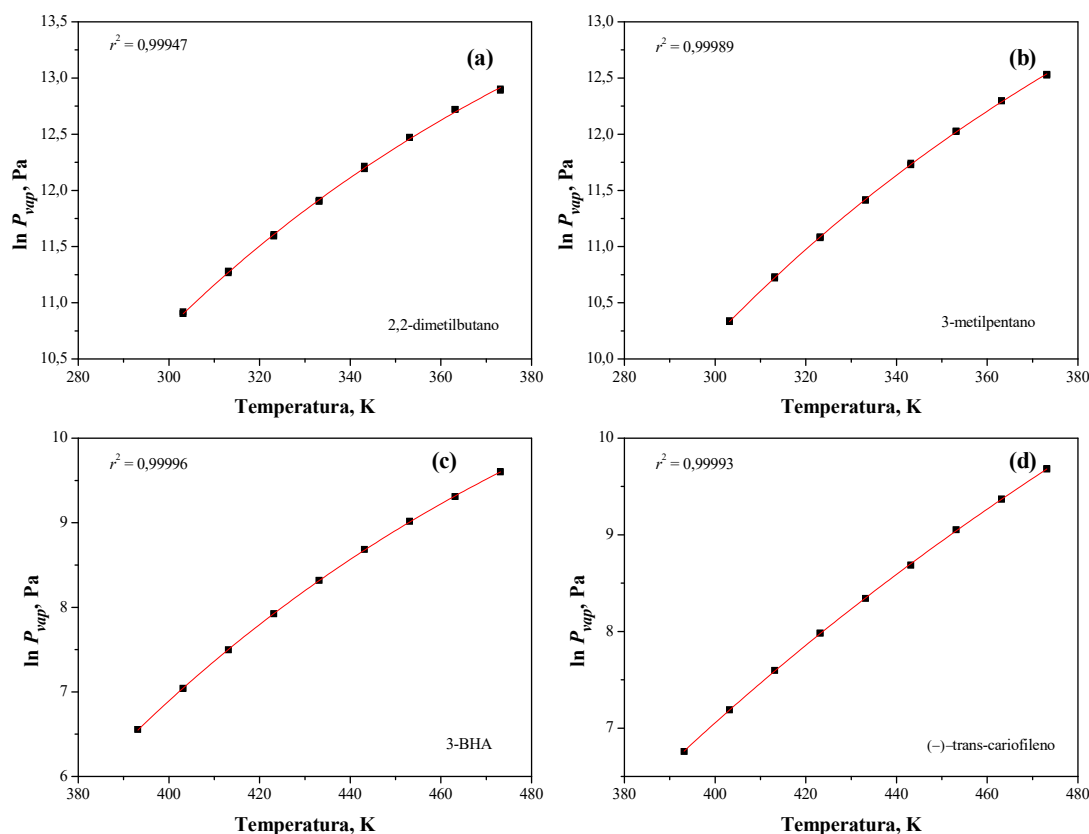


Figura 4.14 - Ajuste $\ln P$ vs. T da Equação 4.13 para o: (a) 2,2-dimetilbutano; (b) 3-metilpentano; (c) 3-BHA e; (d) (-)-trans-cariofileno.

Adicionalmente, sabendo-se que a temperatura normal de ebulição (T_{nb}) de uma substância na fase líquida é aquela cuja pressão de vapor é igual a pressão atmosférica (atm), é possível definir P na Equação 4.13 como 1 atm para obter os dados de T_{nb} dos compostos analisados. Para isso, foi usada uma rotina iterativa no *Maple software* (Versão 16).

Os valores de ΔH_{vap} e T_{nb} obtidos para ambos os compostos terpênicos e hidrocarbonetos analisados encontram-se na Tabela 4.11 e 4.12, respectivamente. Comparando os dados encontrados no presente trabalho com os dados existentes na literatura e baseando-se no desvio relativo médio, calculado respectivamente por

$$d_{\Delta H} = \frac{1}{n} \sum_{i=1}^n \left(\frac{|\Delta H_{vap}^{calc.} - \Delta H_{vap}^{lit.}|}{\Delta H_{vap}^{lit.}} \times 100 \right) \quad (4.15)$$

$$d_{T_{nb}} = \frac{1}{n} \sum_{i=1}^n \left(\frac{|T_{nb}^{calc.} - T_{nb}^{lit.}|}{T_{nb}^{lit.}} \times 100 \right) \quad (4.16)$$

nas quais $\Delta H_{vap}^{calc.}$ e $T_{nb}^{calc.}$ são a entalpia de vaporização e temperatura normal de ebulição calculados neste trabalho e; $\Delta H_{vap}^{lit.}$ e $T_{nb}^{lit.}$, os dados da literatura.

Tabela 4.10 - Coeficientes b_0 , b_1 e b_2 da Equação 4.13 para todos os compostos analisados.

Composto	Coeficiente			r^2
	b_0	b_1	b_2	
2,2-dimetilbutano	262,3 ± 14,6	-51665 ± 4368	-73,3 ± 13,0	0,99950
2,3- dimetilbutano	248,5 ± 10,2	-48444 ± 3065	-59,7 ± 9,1	0,99980
3-metilpentano	235,7 ± 7,3	-45184 ± 2191	-46,3 ± 6,5	0,99990
metilciclopentano	233,7 ± 5,3	-45935 ± 1596	-45,7 ± 4,8	0,99994
ciclohexano	180,4 ± 12,1	-30989 ± 3622	5,0 ± 10,8	0,99984
2-metilhexano	280,4 ± 6,6	-60981 ± 1975	-85,9 ± 5,9	0,99995
3-metilhexano	240,1 ± 3,3	-49240 ± 1001	-46,7 ± 3,0	0,99999
2,2,4-trimetilpentano	232,8 ± 3,0	-47392 ± 898	-40,5 ± 2,7	0,99999
metilciclohexano	241,9 ± 3,3	-51090 ± 995	-51,1 ± 3,0	0,99999
(+)- α -pineno	257,5 ± 3,1	-61608 ± 922	-56,8 ± 2,6	0,99998
(+)- β -pineno	244,7 ± 4,0	-58834 ± 1205	-45,4 ± 3,4	0,99997
safrol	280,0 ± 7,6	-77649 ± 2365	-62,0 ± 5,6	0,99998
isosafrol	295,4 ± 6,6	-84520 ± 2062	-71,6 ± 4,9	0,99998
(+)-3-careno	251,2 ± 3,8	-61789 ± 1143	-47,7 ± 3,1	0,99998
(+)-limoneno	256,2 ± 3,9	-63822 ± 1168	-50,5 ± 3,2	0,99998
linalol	282,4 ± 4,1	-73936 ± 1263	-64,6 ± 3,4	0,99999
(-)- α -terpineol	287,4 ± 3,0	-77550 ± 913	-70,1 ± 2,3	0,99999
timol metil éter	298,1 ± 5,5	-82189 ± 1705	-73,9 ± 4,3	0,99997
3-BHA	472,5 ± 10,9	-143110 ± 3428	-195,3 ± 8	0,99996
α -humuleno	289,3 ± 4,6	-83466 ± 1445	-69,0 ± 3,4	0,99999
carvacrol	340,5 ± 5,0	-97141 ± 1537	-103 ± 3,8	0,99999
timol	349,8 ± 5,6	-99799 ± 1742	-109,6 ± 4,3	0,99998
γ -terpineno	263,3 ± 7,3	-66804 ± 2243	-54,4 ± 5,8	0,99994
<i>p</i> -cimeno	260,1 ± 8,2	-64847 ± 2526	-53,8 ± 6,5	0,99991
(-)- <i>trans</i> -cariofileno	174,9 ± 14,2	-48613 ± 4467	18,1 ± 10,4	0,99993
terpinen-4-ol	289,0 ± 8,8	-77595 ± 2687	-73,1 ± 6,9	0,99992
terpinoleno	259,2 ± 6,7	-66323 ± 2062	-50,5 ± 5,3	0,99995
biciclo[2.2.1]hepta-2,5-dieno	185,6 ± 9,2	-33559 ± 2777	1,6 ± 0,1	0,99997
5-vinil-2-norborneno	225,4 ± 6,4	-50452 ± 1954	-29,6 ± 5,3	0,99996
5-etilideno-2-norborneno	229,5 ± 4,6	-52642 ± 1398	-30,2 ± 3,8	0,99998
álcool fenchil	295,4 ± 12,0	-77914 ± 3736	-78,6 ± 8,9	0,99980
(-)-borneol	281,9 ± 14,2	-75008 ± 4434	-67,8 ± 10,7	0,99980
biciclo[2.2.2]octano-2,5-diona	266,9 ± 8,5	-72752 ± 2648	-54,0 ± 6,2	0,99995
biciclo[4.1.0]heptano-7-ácido carboxílico	464,1 ± 17,4	-136073 ± 5412	-190,9 ± 13,3	0,99997

Tabela 4.11 - Valores de entalpia de vaporização a 298,15 K (em kJ mol⁻¹) para os hidrocarbonetos alifáticos e os terpenos estudados e comparação com os dados da literatura.

Composto*	$\Delta H_{vap}^{lit.}$	$\Delta H_{vap,calc.}^{298,15 K}$	** d_R , %	d_M , %
2,2-dimetilbutano	27,77 ^a ; 27,71 ^{b,3}	29,81	-7,35 ^a ; -7,58 ^{b,3}	7,47
2,3-dimetilbutano	29,18 ^a ; 29,16 ^{b,1}	30,64	-5,0 ^a ; -5,08 ^{b,1}	5,04
3-metilpentano	30,31 ^a ; 30,27 ^{b,1}	31,38	-3,53 ^a ; -3,67 ^{b,1}	3,6
metilciclopentano	31,67 ^a ; 31,64 ^{b,1}	32,31	-2,02 ^a ; -2,12 ^{b,1}	2,07
ciclohexano	33,07 ^a ; 33,00 ^{b,3}	32,48	1,78 ^a ; 1,58 ^{b,3}	1,68
2-metilhexano	34,87 ^a ; 34,84 ^b	35,37	-1,43 ^a ; -1,52 ^b	1,48
3-metilhexano	35,12 ^a ; 35,04 ^b	35,32	-0,57 ^a ; -0,8 ^b	0,69
2,2,4-TMP	35,11 ^a ; 35,15 ^{b,1}	35,30	-0,54 ^a ; -0,43 ^{b,1}	0,49
metilciclohexano	35,32 ^a ; 35,35 ^{b,2}	35,78	-1,3 ^a ; -1,22 ^{b,2}	1,26
(+)- α -pineno	44,85 ^c ; 44,57 ^{b,1} ; 45,40 ^e	44,67	0,4 ^c ; -0,22 ^{b,1} ; 1,61 ^e	0,74
(+)- β -pineno	46,19 ^c ; 45,80 ^{b,1}	45,30	1,93 ^c ; 1,09 ^{b,1}	1,51
safrol	54,6 ^{f,1}	59,15	-8,33 ^{f,1}	8,33
isosafrol	54,9 ^{f,2}	63,17	-15,06 ^{f,2}	15,06
(+)-3-careno	48,30 ^a ; 48,51 ^e	47,57	1,51 ^a ; 1,94 ^e	1,73
(+)-limoneno	48,92 ^{b,1} ; 49,60 ^c ; 49,5 ^e	48,77	0,31 ^{b,1} ; 1,67 ^c ; 1,47 ^e	1,15
linalol	55,25 ^c	54,68	1,03 ^c	1,03
(-)- α -terpineol	54,0 ^{f,3}	56,65	-4,91 ^{f,3}	4,91
timol metil éter	-	60,16	-	-
3-BHA	-	84,88	-	-
α -humuleno**	66,11	62,89	3,22	3,22
carvacrol	68,20 ^d	66,43	2,6 ^d	2,6
timol	70,85 ^c	67,12	5,26 ^c	5,26
γ -terpineno	51,38 ^c	50,58	1,56 ^c	1,56
<i>p</i> -cimeno	49,24 ^c	48,81	0,87 ^c	0,87
(-)- <i>trans</i> -cariofileno	-	54,01	-	-
terpinen-4-ol	55,52 ^c	55,80	-0,5 ^c	0,5
terpinoleno	50,80 ^a	51,27	-0,93 ^a	0,93
B[2.2.1]	36,6 ^a ; 33,8 ^b ; 34,8 ^g	34,03	7,03 ^a ; -0,67 ^b ; 2,22 ^g	3,31
VNB	42,29 ^a ; 42,3 ^a	41,64	1,53 ^a ; 1,56 ^a	1,55
ENB	44,14 ^b ; 42,3 ^g	43,63	1,16 ^b ; -3,14 ^g	2,15
álcool fenchil	-	54,47	-	-
(-)-borneol	-	54,80	-	-
B[2.2.2]	-	56,65	-	-
B[4.1.0]	-	79,15	-	-

a) NIST online – dados a 298,15 K; **b)** NIST/TDE do Aspen® - dados a 298,15 K, exceto: **b,1** (298,14 K), **b,2** (298,13 K) e **b,3** (298,12 K); **c)** Hoskovec *et al.* (2005) – dados a 298,15 K; **d)** van Roon *et al.* (2002) – dados a 298,15 K; **e)** Clará *et al.* (2009) – dados a 298,15 K; **f)** Stephenson e Malanowski (1897) – **f,1** refere-se a dado na faixa de 336-506 K, **f,2** (393-531 K), **f,3** (325-491 K); **g)** Chickos e Acree (2010) – dados a 298,15 K. *2,2,4-TMP = 2,2,4-trimetilpentano; B[2.2.1] = biciclo[2.2.1] hepta-2,5-dieno; VNB = 5-vinil-2-norborneno; ENB = 5-etilideno-2-norborneno; B[2.2.2] = biciclo[2.2.2]octano-2,5-diona; B[4.1.0] = biciclo[4.1.0]heptano-7-ácido carboxílico. ** desvios relativos calculados conforme a Equação 4.6; ** valor predito pelo TDE do Aspen®.

Tabela 4.12 - Valores de temperatura normal de ebulição (em K) para os hidrocarbonetos alifáticos e os terpenos estudados e comparação com os dados da literatura.

Composto*	$T_{nb}^{lit.}$	$T_{nb}^{calc.}$	** d_R , %	d_M , %
2,2-dimetilbutano	322,90 ^a ; 322,88 ^b	320,62	0,71; 0,7	0,71
2,3-dimetilbutano	331,20 ^a ; 331,12 ^b	330,73	0,14; 0,12	0,13
3-metilpentano	336,55 ^a ; 336,40 ^b	336,55	0,0; -0,04	0,02
metilciclopentano	345,00 ^a ; 344,94 ^b	352,91	-2,29; -2,31	2,30
ciclohexano	353,90 ^a ; 353,85 ^b	362,20	-2,35; -2,36	2,36
2-metilhexano	363,20 ^a ; 363,10 ^b	364,14	-0,26; -0,29	0,28
3-metilhexano	365,00 ^a ; 364,89 ^b	365,35	-0,1; -0,13	0,12
2,2,4-TMP	372,40 ^a ; 372,36 ^b	369,43	0,8; 0,79	0,80
metilciclohexano	377,00 ^a ; 374,04 ^b	383,45	-1,71; -2,52	2,12
(+)- α -pineno	429,11 ^b ; 429,35 ^c	442,44	-3,11; -3,05	3,08
(+)- β -pineno	438,93 ^b ; 439,15 ^c	453,01	-3,21; -3,16	3,19
safrol	508,56 ^b ; 507,65 ^c	517,84	-1,82; -2,01	1,92
isosafrol	521,5 ^b	537,29	-3,03	3,03
(+)-3-careno	445,05 ^b ; 444,15 ^c	458,10	-2,93; -3,14	3,04
(+)-limoneno	450,79 ^b ; 451,15 ^c	461,44	-2,36; -2,28	2,32
linalol	471,15 ^c	470,74	0,09	0,05
(-)- α -terpineol	492,72 ^b ; 493,15 ^c	498,69	-1,21; -1,12	1,17
timol metil éter	-	501,77	-	-
3-BHA	-	580,38	-	-
α -humuleno**	-	553,63	-	-
carvacrol	512,19 ^b ; 510,85 ^c	516,53	-0,85; -1,11	0,98
timol	505,79 ^b ; 505,65 ^c	513,31	-1,49; -1,51	1,50
γ -terpineno	456,15 ^b ; 456,15 ^c	466,97	-2,37; -2,26	2,32
<i>p</i> -cimeno	450,27 ^b ; 450,25 ^c	460,18	-2,2; -2,21	2,21
<i>trans</i> -CAR	-	541,04	-	-
terpinen-4-ol	495,60 ^c	498,89	-0,66	0,66
terpinoleno	459,61 ^b ; 459,15 ^c	473,83	-3,09; -3,2	3,15
B[2.2.1]	369,2 ^a ; 363,7 ^b	372,36	-0,86; -2,38	1,62
VNB	414,2 ^a ; 413,89 ^b	423,12	-2,15; -2,23	2,19
ENB	419,2 ^a ; 421,02 ^b	429,15	-2,37; -1,93	2,15
álcool fenchil	-	481,30	-	-
(-)-borneol	-	493,90	-	-
B[2.2.2]	-	513,28	-	-
B[4.1.0]	-	516,95	-	-

a) NIST online – dados a 298,15 K; b) NIST/TDE do Aspen®; c) Lide (2004); *2,2,4-TMP = 2,2,4-trimetilpentano; *trans*-CAR = (-)-*trans*-cariofileno; B[2.2.1] = biciclo[2.2.1] hepta-2,5-dieno; VNB = 5-vinil-2-norborneno; ENB = 5-etilideno-2-norborneno; B[2.2.2] = biciclo[2.2.2]octano-2,5-diona; B[4.1.0] = biciclo[4.1.0]heptano-7-ácido carboxílico. ** desvios relativos calculados conforme a Equação 4.6.

De modo geral, os dados de ΔH_{vap} e T_{nb} encontrados neste trabalho ficaram em conformidade com os dados da literatura. Analisando, os dados de ΔH_{vap} dos hidrocarbonetos alifáticos contidos na Tabela 4.11, observa-se que os compostos mais voláteis dessa classe (2,2-dimetilbutano, 2,3-dimetilbutano e 3-metilpentano) mostraram maior desvio relativo médio, d_M . O 2,2-dimetilbutano (composto mais volátil entre os compostos estudados neste trabalho) teve d_M igual a 7,47 %, que equivale à diferença de cerca de 2 kJ mol⁻¹ no valor da entalpia. Os demais hidrocarbonetos alifáticos apresentaram $d_M < 5,0$ %, sendo que o 3-metilhexano e o 2,2,4-trimetilpentano apresentaram d_M , respectivamente de 0,49 e 0,69 %. Em se tratando dos compostos terpênicos, apenas o safrol (8,33%) e o isosafrol (15,06 %) apresentaram desvio relativo médio um pouco elevado, que pode estar relacionado à não correção para 298,15 K da faixa de temperatura, respectivamente de 336 – 506 K e 393 – 531 K, utilizada pelos autores (Stephenson e Malanowski, 1897). Com exceção do timol (5,26 %), o restante dos terpenos analisados apresentou desvio relativo médio abaixo de 5,0 %.

Em relação aos valores de T_{nb} , apresentados na Tabela 4.12, foram encontrados desvios relativos médios extremamente baixos, que, para todos os compostos estudados, ficaram abaixo de 4,0 %. Por exemplo, os compostos 3-metilpentano e 2-metilhexano mostraram uma diferença na temperatura normal de ebulição de apenas 0,07 e 0,09 K, respectivamente.

Baseado nesses resultados, pode-se destacar que o método utilizado no presente trabalho de Tese, além dos dados de pressão de vapor, é capaz de determinar dados de entalpia de vaporização e temperatura normal de ebulição dentro de uma pequena margem de desvio. Desse modo, o presente trabalho pôde contribuir com dados novos de ΔH_{vap} e T_{nb} para os compostos timol metil éter, 3-BHA, α -humuleno e (-)-*trans*-cariofileno e para os compostos bicíclicos biciclo[2.2.2]octano-2,5-diona, biciclo[4.1.0]heptano-7-ácido carboxílico, álcool fenchil e (-)-borneol.

4.2 DESEMPENHO PREDITIVO DOS MODELOS DA LITERATURA

4.2.1 Modelo de Joback e Reid (1987)

As predições das entalpias de vaporização para o modelo de Joback e Reid (1987) foram realizadas por meio da Equação 2.21 (item 2.6.2.2),

$$\Delta H_{vap}(T_{nb}, K) = 15,30 + \sum_i (N_i h_{v,i})$$

Os grupos, seguidos de suas respectivas contribuições, aplicados para os compostos analisados estão organizados na Tabela 4.13 e estão distribuídos em grupos acíclicos (para compostos saturados e insaturados), grupos cíclicos e aromáticos e os grupos funcionais.

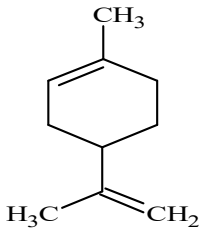
Tabela 4.13 - Principais grupos extraídos de Joback e Reid (1987) utilizados neste trabalho.

Grupo	Contribuição
<i>Grupos acíclicos</i>	
Constante	15,30
-CH ₃	2,374
-CH ₂ -	2,227
>CH-	1,691
=CH-	2,206
=C<	2,139
>C<	0,636
=CH ₂	1,725
<i>Grupos cíclicos e aromáticos</i>	
-CH ₂ - (cíclico)	2,399
>CH- (cíclico)	1,943
>C< (cíclico)	0,644
=CH- (aromático)	2,546
=C< (aromático)	3,061
<i>Grupos funcionais</i>	
-OH (fenol)	12,506
-OH (álcool)	16,835
-O-	2,412
-O- (em anel)	4,685
>C=O (ring)	8,972
-COOH (ácido)	19,537
O=CH- (aldeído)	9,093
-COO- (éster)	9,633

Uma exemplificação da aplicação do método de Joback e Reid está apresentada nas Tabelas 4.14 e 4.15. A Tabela 4.15 ilustra a estimação da entalpia de vaporização do (+)-limoneno, que contém um anel cíclico insaturado e carbonos saturados e insaturados. A soma do produto dos grupos vezes a sua contribuição apresentou entalpia de vaporização para

o (+)-limoneno de $35,64 \text{ kJ mol}^{-1}$. O desvio relativo absoluto (d_R) em relação à literatura (HOSKOVEC *et al.*, 2005) foi de 28,15 %, que é relativamente alto.

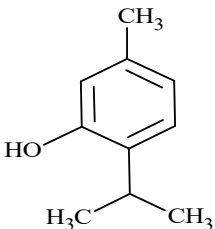
Tabela 4.14 - Estimação da entalpia de vaporização do (+)-limoneno pelo método de Joback e Reid (1987).

(+)-limoneno	Contagem dos grupos	Descrição dos grupos	Contribuição total
	2	-CH ₃	4,748
	1	=C<	2,139
	1	=CH ₂	1,725
	3	-CH ₂ - (cíclico)	7,197
	2	>CH- (cíclico)	3,886
	1	>C< (cíclico)	0,644
		$\Delta H_{\text{estimado}}$	35,64 kJ mol ⁻¹
		$\Delta H_{\text{experimental}}^*$	49,60 kJ mol ⁻¹

*Hoskovec *et al.* (2005)

Outra ilustração do uso do método de Joback e Reid (1987) foi feita para o timol, que tem natureza aromática, além de uma parte de carbonos saturados e uma hidroxila fenólica. O valor de ΔH_{vap} estimado para o timol foi de $53,44 \text{ kJ mol}^{-1}$. Comparando com o dado de Hoskovec *et al.* (2005), o d_R foi 24,57 %.

Tabela 4.15 - Estimação da entalpia de vaporização do timol pelo método de Joback e Reid (1987).

timol	Contagem dos grupos	Descrição dos grupos	Contribuição total
	3	-CH ₃	7,122
	1	>CH-	1,691
	3	=CH- (aromático)	7,638
	3	=C< (aromático)	9,183
	1	-OH (fenol)	12,506
		$\Delta H_{\text{estimado}}$	53,44 kJ mol ⁻¹
		$\Delta H_{\text{experimental}}^*$	70,85 kJ mol ⁻¹

*Hoskovec *et al.* (2005)

4.2.2 Modelo de Chickos *et al.* (1998)

Os grupos envolvidos na predição de ΔH_{vap} pelo método de Chickos *et al.* (1998) estão mostrados na Tabela 4.16. Embora seja um modelo relativamente simples de ser utilizado, deve-se ter cautela quando for aplicar os termos de correções. O modelo faz distinção de substâncias que fazem ligação de hidrogênio (como álcoois) e o tipo de hibridização dos átomos (qualquer elemento) cujos grupos funcionais fazem ligação.

Tabela 4.16 - Termos envolvidos na predição pelo modelo de Chickos *et al.* (1998).

Grupos funcionais	Valor
>C=O	10,5
-OH	29,4
>O (éter)	5,0
-C(=O)OH (ácido carboxílico)	38,8
-CHO (aldeído)	12,9
-C(=O)O (éster)	10,5
<i>Correções</i>	
C (para cetonas cíclicas)	2,9
Ramificação alquila vicinal e <i>orto</i> para carbonos sp^2 e sp^3 em anéis de 5 e 6 membros.	-2,0
Ligação de hidrogênio intramolecular para álcoois (em anéis de 5-9 membros)	-7,0
Substituição única em átomo primário sp^3	1,62
Substituição única em átomo terciário sp^2	0,69
Substituição única em átomo quaternário sp^3	0,79

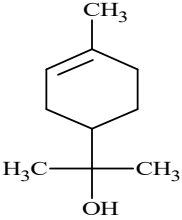
As Tabelas 4.17 e 4.18 ilustram os resultados de predição do (-)- α -terpineol e do 3-BHA feitas através do modelo de Chickos *et al.* (1998), conforme a Equação 2.22 (item 2.6.2.2),

$$\Delta_{vap} H_m^0(298,15 \text{ K}) = 4,69(N_C - N_O) + 1,3N_O + 3,0 + \sum_i n_i F_i b_i + C$$

O (-)- α -terpineol contém um anel cíclico insaturado e uma parte de carbonos saturados. É um álcool terpênico, portanto, faz ligação de hidrogênio (correção C) e sua hidroxila está ligada em um átomo de carbono quaternário sp^3 (correção F_i). Aplicando esses valores na Equação 2.26, obteve-se $\Delta H_{vap} = 62,74 \text{ kJ mol}^{-1}$ (Tabela 4.17). Comparado com o dado de Stephenson e Malanowski (1987), o d_R foi de $14,91 \text{ kJ mol}^{-1}$.

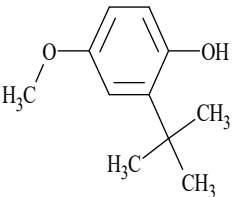
A ilustração do método de Chickos *et al.* (1998) para o 3-BHA (Tabela 4.18) necessitou de mais atenção. O 3-BHA contém dois grupos funcionais (éter e álcool) ligados ao anel aromático. O grupo *tert*-butila está na posição *orto* em relação à hidroxila (contribuição F_i), que também faz substituição em um átomo terciário sp^2 . O grupo funcional éter está ligado em dois átomos diferentes, fazendo uma substituição em átomo terciário sp^2 e outra em átomo terciário sp^3 (as duas foram consideradas na predição). O 3-BHA apresentou o valor de entalpia de 72,30 kJ mol⁻¹. A comparação com a entalpia determinada neste trabalho (84,88 kJ mol⁻¹) resultou em um desvio relativo absoluto de 14,82 %.

Tabela 4.17 - Predição da entalpia de vaporização para o α -terpineol pelo modelo de Chickos *et al.* (1998).

(-)- α -terpineol	Descrição do grupo/fator de correção	Contribuição total
	N_C (número de carbonos)	10
	N_Q (número de carbonos quaternários)	1
	-OH	29,4
	Ligação de hidrogênio intramolecular	-7
	Substituição única em átomo quaternário sp^3	0,79
	$\Delta H_{\text{estimado}}$	62,74 kJ mol ⁻¹
	$\Delta H_{\text{experimental}}^*$	54,60 kJ mol ⁻¹

*Stephenson e Malanowski (1987)

Tabela 4.18 - Predição da entalpia de vaporização para o 3-BHA pelo modelo de Chickos *et al.* (1998).

3-BHA	Descrição do grupo/fator de correção	Contribuição total
	N_C (número de carbonos)	11
	N_Q (número de carbonos quaternários)	1
	-OH	29,4
	>O (Éter)	5,0
	Ramificação alquila vicinal e <i>orto</i> em carbonos sp^2 e sp^3 em anéis de 5 e 6 membros.	-2,01
	Ligação de hidrogênio intramolecular	-7,0
	Substituição única em átomo primário sp^3	1,62
	Substituição única em átomo terciário sp^2 (duas contribuições)	0,69
	$\Delta H_{\text{estimado}}$	72,30 kJ mol ⁻¹
	$\Delta H_{\text{experimental}}^*$	84,88 kJ mol ⁻¹

*Determinada neste trabalho

4.2.3 Modelo de Kolská *et al.* (2005)

O modelo de Kolská *et al.* (2005) utiliza três níveis de contribuição de grupos. No entanto, para os compostos analisados neste trabalho, foi necessário aplicar apenas os grupos de primeira e segunda ordens. Os grupos e suas respectivas contribuições empregados nas predições dos compostos analisados estão apresentados nas Tabelas 4.19 e 4.20.

Tabela 4.19 - Grupos de primeira ordem e suas contribuições empregados na predição de entalpia de vaporização pelo modelo de Kolská *et al.* (2005).

Grupos de primeira ordem	Contribuição
x	9,672
CH ₃	2,266
CH ₂	4,724
CH	4,724
C	4,809
CH ₂ =CH	5,620
CH=CH	8,516
CH ₂ =C	7,741
CH=C	10,409
aCH	4,297
aC-CH ₃	8,121
aC-CH ₂	9,840
aC-CH	9,261
aC-C	8,843
OH	24,930
aC-OH	43,072
aC-O	-4,968
H ₂ C _{cíclico}	4,013
CH _{cíclico}	4,075
C _{cíclico}	3,667
(CH=CH) _{cíclico}	9,179
(CH=C) _{cíclico}	10,070
H ₂ C=C _{cíclico}	6,749
(C=O)	14,837
(-O-) _{cíclico}	6,171
CHO	13,027
aC-COO	21,425
COO	16,457
aC	3,910

Tabela 4.20 - Grupos de segunda ordem e suas contribuições empregados na predição de entalpia de vaporização pelo modelo de Kolská *et al.* (2005).

Grupos de segunda ordem	Contribuição
CH(CH ₃) ₂	-0,066
(CH ₃) ₃ C	-0,022
COH	-0,405
(CH=C) _{cíclico} -CH ₃	3,216
CH _{cíclico} -CH ₃	0,112
CH _{cíclico} -CH	0,736
CH _{cíclico} -C	1,565
(CH) _{cíclico} -C=CH _n (n = 0..2)	-2,416
C _{cíclico} -CH ₃	-1,355
CH _{cíclico} -OH	3,07
aC-CH _n -COO (n = 1...2)	0,758
AROMRINGs1s2	-0,832
AROMRINGs1s3	-0,073
AROMRINGs1s4	0,473
AROMRINGs1s2s4	1,245

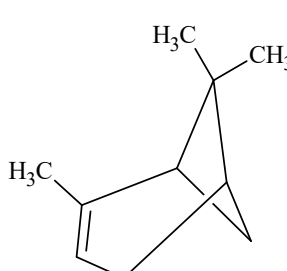
Para ilustrar o uso do modelo de Kolská *et al.* (2005), utilizou-se os exemplos da Tabela 4.21, usando o (+)- α -pineno, e da Tabela 4.23, usando o (+)- β -pineno. O (+)- α -pineno possui molécula com anel bicíclico insaturado e grupos metilas saturados. O modelo não diferencia moléculas cíclicas de moléculas bicíclicas. Assim, a molécula foi dividida em grupos metila (não cíclicos) e os demais como grupos cíclicos. O (+)- α -pineno ainda contém duas informações de segunda ordem: a primeira é referente ao carbono saturado cíclico dissustituído por dois grupos metilas (C_{cíclico}-CH₃), e conta como duas interações; e a segunda refere-se ao grupo insaturado cíclico (CH=C)_{cíclico} ligado ao grupo metila.

Conforme pode ser visto na Tabela 4.21, o ΔH_{vap} predito para o (+)- α -pineno foi 46,90 kJ mol⁻¹. O valor apresentado por Hoskovec *et al.* (2005) foi de 44,84 kJ mol⁻¹. Isso representa desvio relativo absoluto de 4,59 %, que aponta uma boa aproximação.

Uma vez que o modelo de Kolská *et al.* (2005) contém vários níveis de contribuição, pode diferenciar uma variedade de compostos isômeros. Desse modo, o (+)- β -pineno (isômero do (+)- α -pineno) foi escolhido para ilustrar uma segunda aplicação do modelo de Kolská *et al.* (2005) cujos resultados estão apresentados na Tabela 4.22.

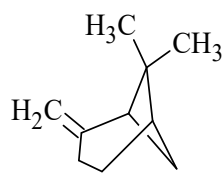
O (+)- β -pineno, assim como o seu isômero (+)- α -pineno, também possui um anel bicíclico, porém com a instauração fora do anel (grupo metileno, H₂C=). O modelo considera o grupo metileno ligado à um carbono cíclico como um único grupo (H₂C=C_{cíclico}). As demais fragmentações foram feitas como para o (+)- α -pineno, incluindo as interações de segunda ordem no carbono cíclico saturado com os grupos metilas (C_{cíclico}-CH₃). O ΔH_{vap} estimado para o (+)- β -pineno foi de 42,10 kJ mol⁻¹. Comparado com o dado da literatura (46,19 kJ mol⁻¹), o d_R apresentado foi de 8,85 %.

Tabela 4.21 - Predição da entalpia de vaporização para o (+)- α -pineno pelo modelo de Kolská *et al.* (2005).

(+)- α -pineno	Contagem dos grupos	Descrição do grupo	Contribuição total	
	1	x	9,672	
	3	CH ₃	6,798	
	2	H ₂ C _{cíclico}	8,026	
	2	CH _{cíclico}	8,15	
	1	C _{cíclico}	3,667	
	1	(CH=C) _{cíclico}	10,07	
	<i>Segunda ordem</i>			
	1	(CH=C) _{cíclico} -CH ₃	3,216	
	2	C _{cíclico} -CH ₃	-2,710	
			$\Delta H_{estimado}$	46,90 kJ mol ⁻¹
		$\Delta H_{experimental*}$	44,84 kJ mol ⁻¹	

*Hoskovec *et al.* (2005)

Tabela 4.22 - Predição da entalpia de vaporização para o (+)- β -pineno pelo modelo de Kolská *et al.* (2005).

(+)- β -pineno	Contagem dos grupos	Descrição do grupo	Contribuição total	
	1	x	9,672	
	2	CH ₃	4,532	
	3	H ₂ C _{cíclico}	12,039	
	2	CH _{cíclico}	8,15	
	1	C _{cíclico}	3,667	
	1	CH ₂ =C _{cíclico}	6,749	
	<i>Segunda ordem</i>			
	2	C _{cíclico} -CH ₃	-2,71	
			$\Delta H_{estimado}$	42,10 kJ mol ⁻¹
			$\Delta H_{experimental*}$	46,19 kJ mol ⁻¹

*Hoskovec *et al.* (2005)

4.2.4 – Desempenho geral dos modelos da literatura na determinação de ΔH_{vap} dos terpenos analisados por CG

Para avaliar a capacidade preditiva de ΔH_{vap} para compostos presentes em óleos essenciais, os modelos da literatura foram usados para estimar as entalpias de vaporização dos compostos terpênicos estudados. Para facilitar a discussão, os valores de ΔH_{vap} preditos pelos modelos selecionados foram comparadas apenas com os dados de ΔH_{vap} obtidos neste trabalho, tendo em vista que os resultados encontrados se apresentaram de acordo com os dados da literatura. Os resultados desta comparação estão apresentados na Tabela 4.23.

Analisando a Tabela 4.23, nota-se que houve grande divergência entre as predições feitas pelos modelos analisados da literatura e os dados de ΔH_{vap} obtidos neste trabalho. De forma geral, pode-se notar que o modelo de Joback e Reid (1987) teve pior desempenho de predição para os compostos selecionados, apresentando desvios relativos bastante elevados, no geral, em torno de 20 % a 30 %. Apenas o terpinen-4-ol, o linalol e o (-)- α -terpineol, álcool fenchil e (-)-borneol tiveram d_R abaixo de 10 %.

O modelo de Kolská *et al.* (2005) apresentou ótimo desempenho para as predições do *p*-cimeno e (+)-limoneno ($d_R < 1,0$ %). No entanto, para a maioria dos terpenos, o desvio relativo ficou acima de 30 %. O modelo de Chickos *et al.* (1998), por sua vez, apresentou melhor desempenho (comparado aos modelos de Joback e Reid (1987) e Kolská *et al.* (2005)). Dos 25 compostos terpênicos analisados, 10 compostos demonstraram $d_R < 5$ %. Chama a atenção, porém, a predição para o (-)-*trans*-cariofileno, que apresentou d_R próximo de 30 % e que mostra a baixa capacidade preditiva desse modelo.

A Tabela 4.23 também evidencia que os três modelos da literatura analisados se mostraram inadequados para predição de ΔH_{vap} para a maioria dos compostos bicíclicos estudados neste trabalho (destacados em negrito na Tabela 4.23). O modelo de Chickos *et al.* (1998), por exemplo, apresentou desvio relativos em torno de 50 % para o álcool fenchil e (-)-borneol. Para esses mesmos compostos, o modelo de Kolská *et al.* (2005) apresentou d_R próximo de 20 %. Por fim, o modelo de Joback e Reid apresentou d_R em torno de 16 % para o 5-vinil-2-norborneno, 5-etilideno-2-norborneno e biciclo[2.2.2]octano-2,5-diona e desvio relativo de 36,5 % para o biciclo[4.1.0]heptano-7-ácido carboxílico.

Tabela 4.23 - Desempenho dos modelos de predição testados neste trabalho. Em parênteses encontra-se o desvio relativo (em %) em relação aos dados obtidos por cromatografia gasosa.

Composto*	ΔH_{vap} (obtidos	ΔH_{vap} (estimado), kJ mol ⁻¹		
	neste trabalho)	Joback e Reid	Chickos <i>et al.</i>	Kolská <i>et al.</i>
<i>p</i> -cimeno	48,81	40,42 (17,2)	49,9 (2,2)	48,71 (0,2)
terpinoleno	51,27	40,43 (21,2)	49,9 (2,7)	45,7 (10,9)
(+)-limoneno	48,77	38,66 (20,7)	49,9 (2,3)	48,93 (0,3)
(+)-α-pineno	44,67	37,36 (16,4)	46,51 (4,1)	46,89 (5,0)
(+)-β-pineno	45,30	36,56 (19,3)	46,51 (2,7)	42,1 (7,1)
(+)-3-careno	47,57	37,36 (21,5)	46,51 (2,2)	46,89 (1,4)
γ -terpineno	50,58	40,13 (20,7)	49,9 (1,3)	52,58 (3,9)
(-)-<i>trans</i>-cariofileno	54,01	49,34 (8,6)	69,96 (29,5)	65,68 (21,6)
α -humuleno	62,89	51,34 (18,4)	69,96 (11,2)	71,5 (13,7)
carvacrol	66,43	53,44 (19,6)	77,29 (16,3)	86,65 (30,4)
timol	67,12	53,44 (20,4)	77,29 (15,2)	87,41 (30,2)
terpinen-4-ol	55,80	54,4 (2,5)	62,74 (12,4)	74,71 (33,9)
linalol	54,68	52,62 (3,8)	62,74 (14,7)	71,28 (30,4)
(-)- α -terpineol	56,65	54,64 (3,5)	62,74 (10,7)	75,2 (32,8)
safrol	59,15	50,05 (15,4)	61,28 (3,6)	32,1 (45,7)
isosafrol	63,17	50,68 (19,8)	61,28 (3,0)	41,56 (34,2)
3-BHA	84,88	57,17 (32,6)	72,3 (14,8)	76,76 (9,6)
timol metil éter	60,16	45,72 (24,0)	62,69 (4,2)	41,71 (30,7)
biciclo[2.2.1]hepta-2,5-dieno	34,03	31,77 (6,6)	35,83 (5,3)	40,19 (18,1)
5-vinil-2-norborneno	41,64	34,95 (16,1)	45,21 (8,6)	42,31 (1,6)
5-etilideno-2-norborneno	43,63	36,72 (15,8)	45,21 (3,6)	40,96 (6,1)
álcool fenchil	54,47	51,63 (5,2)	83,75 (53,8)	64,86 (19,1)
(-)-borneol	54,80	51,63 (5,8)	83,75 (52,8)	64,86 (18,4)
biciclo[2.2.2]octano-2,5-diona	56,65	46,73 (17,5)	67,32 (18,8)	63,55 (12,2)
biciclo[4.1.0]heptano-7-ácido carboxílico	79,15	50,26 (36,5)	82,42 (4,1)	66,51 (16,0)

* Destaque em negrito refere-se aos compostos bicíclicos

A seguir são apresentados os procedimentos para a obtenção do modelo de contribuição de grupos proposto neste trabalho e uma discussão comparando os resultados encontrados e os resultados obtidos pelos modelos da literatura avaliados.

4.3 PROPOSTA DE MODELO DE CONTRIBUIÇÃO DE GRUPOS PARA PREDIÇÃO DE ENTALPIAS DE VAPORIZAÇÃO

4.3.1 Comportamento dos dados

Antes de partir para os procedimentos de regressão dos parâmetros do modelo preditivo proposto, foi necessário entender o comportamento dos dados utilizados, com o auxílio de algumas ferramentas estatísticas. Conforme exposto na seção da metodologia, Capítulo 3, do presente trabalho de Tese, o conjunto de dados utilizado é constituído por 3.643 dados de entalpia de vaporização (ΔH_{vap}), com faixa de temperatura de 288 – 426 K.

A Figura 4.15 mostra como as entalpias de vaporização estão distribuídas em função do número de dados. Pode-se notar em tal figura que a maior concentração dos dados de entalpia encontra-se entre, aproximadamente, na faixa delimitada de 30 a 80 kJ mol⁻¹.

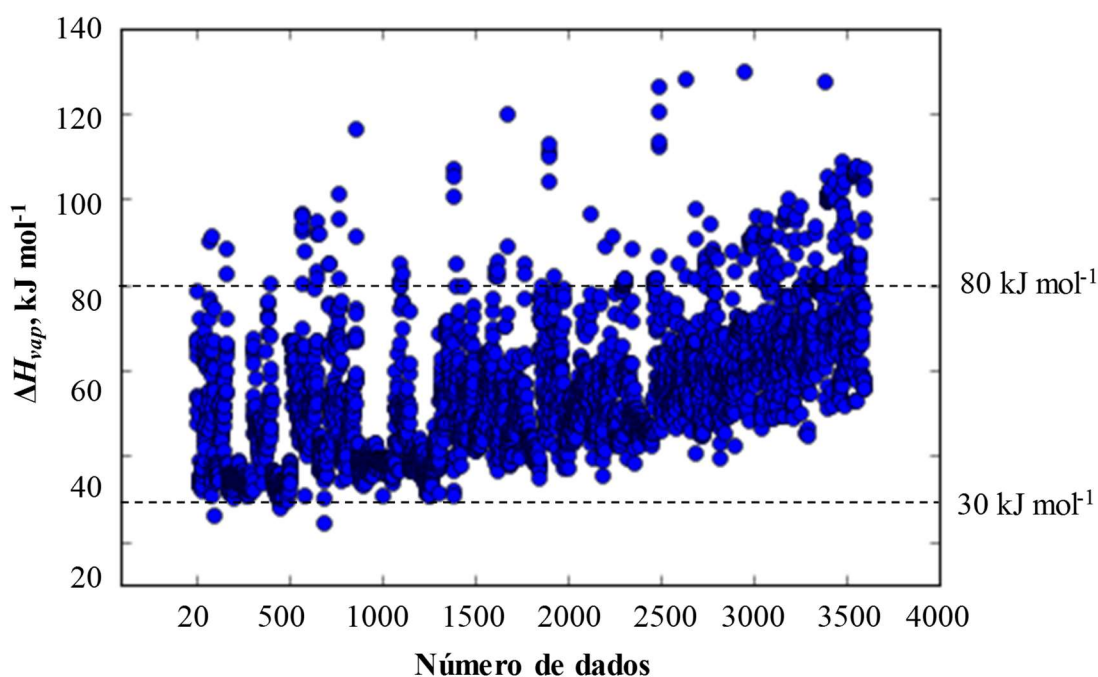


Figura 4.15 - Gráfico de dispersão dos dados de entalpias incluídos no modelo.

A distribuição dos dados pode ser melhor compreendida por meio de construção de um histograma (Figura 4.16). Para isso, foi necessário obter a tabela de distribuição de frequências, apresentada na Tabela 4.24, construída usando a equação de Sturges (MARTINS, 2005),

$$k = 1,0 + 1,33 \times \log(n) \quad (4.17)$$

$$h \cong \frac{A_T}{k} \quad (4.18)$$

em que k é o número de classes; h é o intervalo entre as classes; n é o número de elementos do conjunto de dados e; A_T é a diferença entre os elementos de maior e menor valor.

Utilizando-se o número de elementos do conjunto de dados ($n = 3.643$) na equação de Sturges, encontra-se o número aproximado de classes considerado ideal para a construção da tabela de distribuição de frequência, $k = 13$. O menor e o maior valor de ΔH_{vap} do conjunto de dados foi, respectivamente, de 24,7 e 130,0 kJ mol⁻¹. Assim, o conjunto de dados apresenta amplitude total A_T de 105,3 kJ mol⁻¹. O intervalo entre as classes (h) foi obtido dividindo-se A_T por k . Para o conjunto de dados do presente trabalho, h apresentou valor de 7,188 kJ mol⁻¹.

A Tabela 4.24 evidencia a faixa de valores de entalpia notado na Figura 4.15. Como pode ser visto nesta tabela, a maior concentração de dados de ΔH_{vap} está compreendida no intervalo de 32,9 a 82,0 kJ mol⁻¹ (classes de 2 a 7), que representa cerca de 92 % dos dados, conforme mostra a frequência relativa acumulada (f_{i_ac} , %). Os dados de ΔH_{vap} com maior frequência F_i estão contidos na classe 7, cuja faixa de ΔH_{vap} é de 49,3 – 57,5 kJ mol⁻¹ ($F_i = 802$; f_i , % = 22,01), seguida da classe 3, com intervalo de 41,1 – 49,3 kJ mol⁻¹ (f_i , % = 19,96) e da classe 2 (intervalo 32,9 – 41,1 kJ mol⁻¹; f_i , % = 19,14). Assim, com o uso da Tabela 4.24, construiu-se o histograma das frequências absolutas, mostrado na Figura 4.16.

Após a realização da distribuição de frequências, foi necessário construir um gráfico de caixa (*boxplot*) com o intuito de identificar possíveis *outliers* no conjunto de dados (Figura 4.17). Para tanto, foi necessário aplicar a regra de Tukey (MARTINS, 2005; DÜMBGEN e RIEDWYL, 2007), determinando os quartis inferior e superior (Q_1 e Q_3), bem como os limites inferior e superior (f_L e f_U), conforme,

$$Q_1 = l_{Q_1} + \frac{((n/4) - \sum f) \times h}{F_{Q_1}} \quad (4.19)$$

$$Q_3 = l_{Q_3} + \frac{((3n/4) - \sum f) \times h}{F_{Q_3}} \quad (4.20)$$

$$f_L = Q_1 - 1,5 \times I_Q \quad (4.21)$$

$$f_U = Q_3 + 1,5 \times I_Q \quad (4.22)$$

nas quais l_{Q_1} é o limite inferior da classe que contém o elemento que representa Q_1 (classe Q_1); l_{Q_3} é o limite inferior da classe que contém o elemento que representa Q_3 (classe Q_3); Σf é a soma das frequência das classes que antecedem as classes Q_1 e Q_3 ; F_{Q_1} e F_{Q_3} são as frequências absolutas de suas respectivas classes e; h é o intervalo das classes.

Tabela 4.24 - Distribuição de frequência para os 3.643 dados de entalpia de vaporização.

Classes (k)	Intervalos (h)	F_i^a	$F_{i_ac}^b$	$f_i(\%)^c$	$f_{i_ac}(\%)^d$	x_i^e
1	24,7 - 32,9	76	76	2,09	2,09	28,79
2	32,9 - 41,1	661	737	18,14	20,23	36,98
3	41,1 - 49,3	727	1.464	19,96	40,19	45,17
4	49,3 - 57,5	802	2.266	22,01	62,2	53,36
5	57,5 - 65,6	586	2.852	16,09	78,29	61,55
6	65,6 - 73,8	384	3.236	10,54	88,83	69,74
7	73,8 - 82,0	194	3.430	5,33	94,16	77,92
8	82,0 - 90,2	87	3.517	2,39	96,55	86,11
9	90,2 - 98,4	62	3.579	1,7	98,25	94,30
10	98,4 - 106,6	39	3.618	1,07	99,32	102,49
11	106,6 - 114,8	18	3.636	0,49	99,81	110,68
12	114,8 - 123,0	3	3.639	0,08	99,89	118,87
13	123,0 - 130,0	4	3.643	0,11	100,00	126,48

a) frequência; b) frequência acumulada (ΣF_i); c) frequência relativa (F_i/n); d) frequência relativa acumulada (Σf_i); e) ponto médio entre as classes ((limite inferior + limite superior)/2)

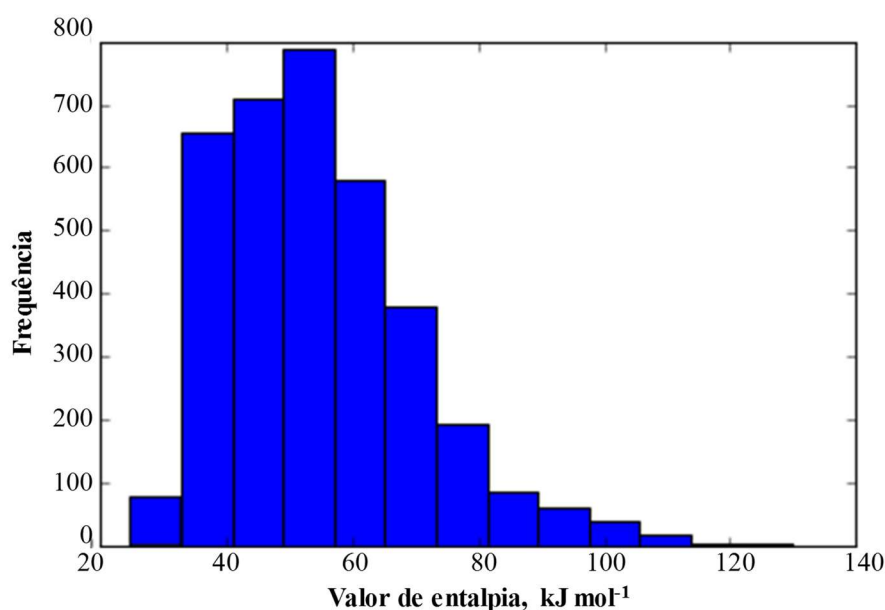


Figura 4.16 - Histograma das frequências absolutas dos 3.643 dados de entalpia do conjunto de dados analisado.

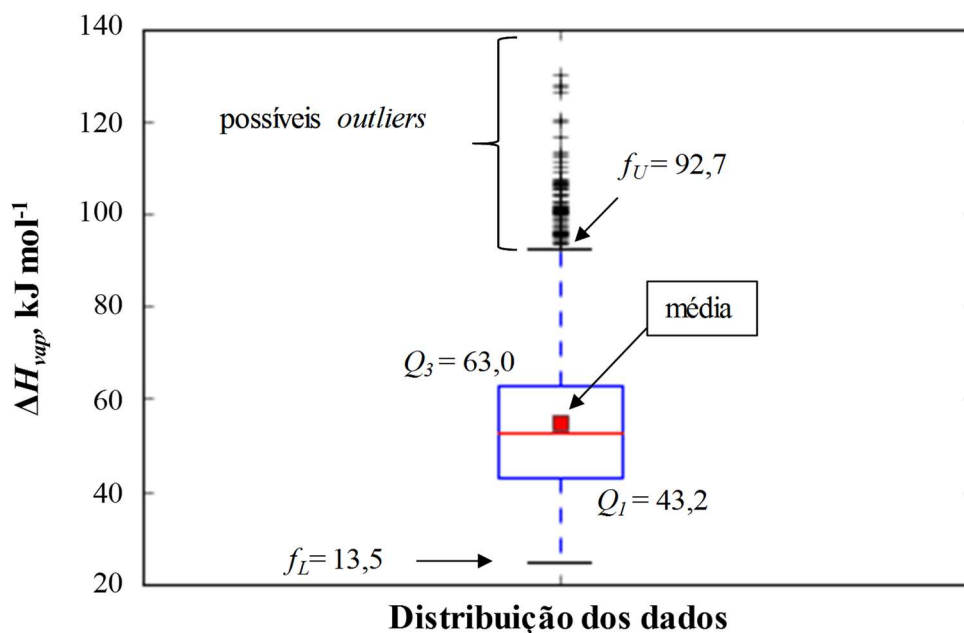


Figura 4.17 - Gráfico de caixa para o conjunto de 3.643 dados de entalpia de vaporização.

A partir dos valores do primeiro (Q_1) e terceiro (Q_3) quartis, determinou-se os valores dos limites extremos inferior (f_L) e superior (f_U), apresentados no gráfico de caixa (Figura 4.17). O limite extremo inferior de 13,5 sugere que o conjunto de dados sob consideração pode conter valores menores de entalpia de até 13,5 kJ mol⁻¹. Como o valor mínimo de ΔH_{vap} é de 24,7 kJ mol⁻¹, o conjunto de dados analisado não apresenta *outliers* com relação à f_L . No que se refere à f_U , que apresentou valor igual a 92,7, há existência de possíveis *outliers* (representados por “+”, na Figura 4.17) a partir de aproximadamente 92,7 kJ mol⁻¹, o que justifica a existência de distribuição assimétrica à direita apresentado pelo histograma mostrado na Figura 4.16.

4.3.2 Procedimentos de regressão para obtenção do modelo proposto

Antes de iniciar esta etapa do presente trabalho de Tese, foi realizada uma inspeção sobre conjunto de dados com o intuito de identificar e retirar os dados de ΔH_{vap} referentes aos compostos analisados por CG, que posteriormente são utilizados para avaliar o desempenho preditivo do modelo proposto. Após este procedimento, o conjunto de dados passou a conter

3.598 dados de entalpia de vaporização, que são referentes a 1.719 compostos orgânicos diferentes.

As predições das entalpias de vaporização foram feitas conforme a equação

$$\Delta H_{vap} \text{ (kJ mol}^{-1}\text{)} = h_0 + \sum_{i=1}^n N_i C_i \quad (4.23)$$

em que h_0 é um valor constante; N_i é o número de grupos e C_i é sua respectiva contribuição.

4.3.2.1 Regressão ordinária por mínimos quadrados usando o GL

Seleção dos parâmetros do modelo

Para avaliar a correta implementação do algoritmo, os resultados de regressão foram comparados com os resultados obtidos pelo pacote GraphLab Create. A etapa de seleção dos parâmetros do modelo foi feita de duas maneiras distintas. Na primeira, o banco de dados foi dividido em três conjuntos de dados diferentes (*treinamento*, *validação* e *teste*), em que 60 % dos dados foram reservados para treinamento, 20 % para validação e 20 % para teste do modelo (denominado neste trabalho de procedimento *tradicional*). Na segunda parte, a seleção dos parâmetros foi conduzida por validação cruzada, na qual 80 % foram dedicados para treinamento e validação cruzada, e 20 % para teste.

Os resultados sobre os primeiros procedimentos estão apresentados na Figura 4.18, que mostra a variação da soma dos quadrados residuais (SQR) em função da complexidade do modelo. Na Figura 4.18 pode-se notar que tanto o desvio de treinamento (linha contínua) como o desvio de validação (linha contínua com marcação) diminuem à medida que a complexidade aumenta (aumento do número de atributos). Isso sugere que todos os parâmetros usados na obtenção do modelo são importantes para aumentar a sua capacidade de predição. Desse modo, foram considerados todos os grupos contidos na Tabela 3.6 (item 3.2.2.2) para a sequência do desenvolvimento do modelo final. O resultado do desempenho da regressão com todos os atributos selecionados sobre o banco de teste apresentou $SQR = 38.608,30$, raiz do desvio quadrado médio ($RDQM$) igual a 7,204, desvio médio absoluto (DA) de 5,032 e desvio relativo médio (DM) de 8,686 (Tabela 4.25).

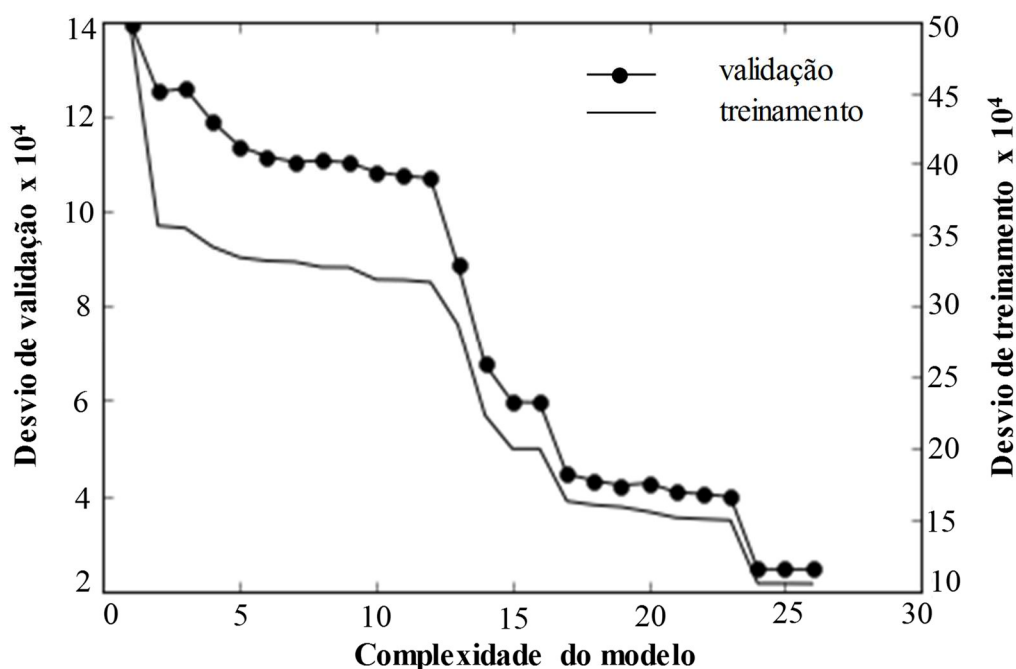


Figura 4.18 - Desvios de treinamento e validação vs. complexidade do modelo.

Tabela 4.25 - Desempenho de regressão: validação tradicional vs. validação cruzada.

Validação	Tipo de desvio			
	<i>SQR</i>	<i>RDQM</i>	<i>DA</i>	<i>DM</i>
Tradicional	38.608,30	7,204	5,032	8,686
Cruzada	37.925,12	7,140	4,921	8,442

Com relação aos resultados advindos do procedimento de validação cruzada, aplicado com o intuito de diminuir os desvios de regressão e de aumentar a generalização do modelo, verifica-se, por meio da Figura 4.19, que apresentou comportamento semelhante ao encontrado para o desvio de treinamento. No entanto, o desempenho de regressão usando validação cruzada apresentou-se melhor que o primeiro procedimento, conforme comparação dos resultados apresentada na Tabela 4.25. É possível notar uma diminuição nos valores dos desvios, embora sutil, do método da validação cruzada em relação à regressão com conjunto de validação tradicional. Portanto, foi escolhido o método da validação cruzada para a sequência deste trabalho.

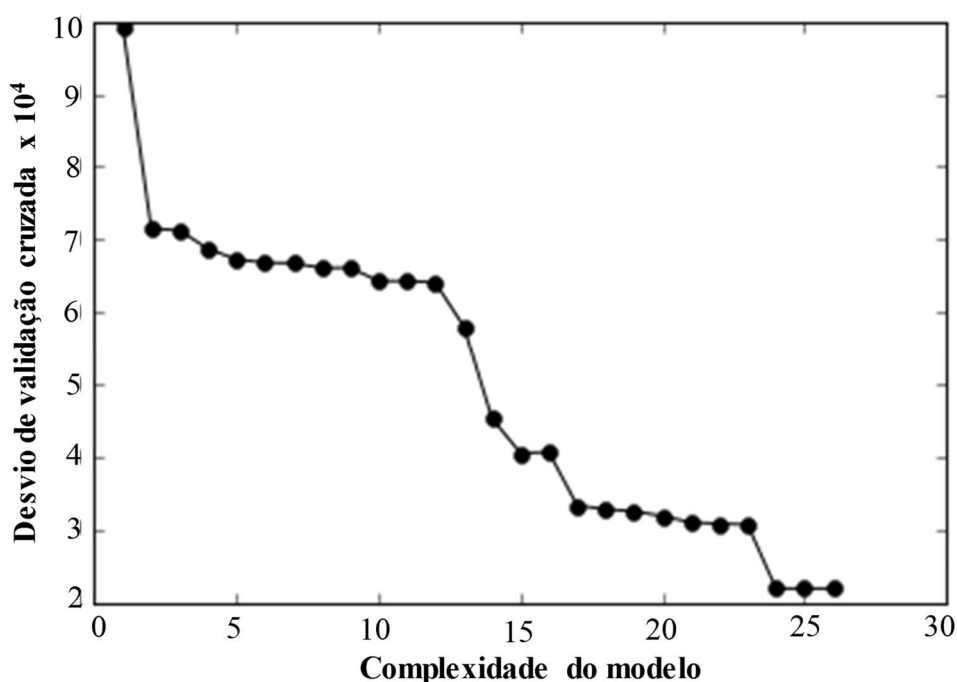


Figura 4.19 - Desvio de validação cruzada vs. complexidade do modelo.

4.3.2.3 Regressão ordinária por mínimos quadrados usando o GD

Seleção dos parâmetros de entrada

Os parâmetros de entrada são fundamentais para o processo de otimização via gradiente descendente. Os parâmetros de entrada: *pesos iniciais* (\mathbf{w}), *tamanho do passo* (η) e *tolerância* foram escolhidos através do método de validação cruzada, utilizando $k\text{-fold} = 5$. O melhor valor de cada parâmetro de entrada foi escolhido com base no menor valor da soma dos quadrados residuais médio (SQR_m) obtido em cada iteração. Os resultados da aplicação de validação cruzada para a escolha dos parâmetros de entrada supracitados estão apresentados nas Figuras 4.20, 4.21 e 4.22.

Para a escolha do valor inicial dos pesos do modelo, foram estudados valores de 0,0, 0,5, 1,0 e 2,0. Embora a Figura 4.20 mostre claramente melhor resultado para o valor inicial dos pesos definido como 1,0, de forma geral não houve variação significativa no valor de SQR_m (Tabela 4.26). Por exemplo, ao mudar os valores de \mathbf{w} de 1,0 para 2,0, a diminuição no total de SQR_m foi apenas de 36,2 (27.674,04 contra 27.710,25). No caso da escolha do valor da tolerância (Figura 4.21), foram estudados valores na ordem de $1,0 \times 10^{-5}$, $1,0 \times 10^{-4}$, $1,0 \times 10^{-3}$

e $1,0 \times 10^{-2}$. O melhor desempenho foi apresentado usando tolerância $1,0 \times 10^{-4}$. No entanto, para os três primeiros valores, não houve diferença expressiva no valor de SQR_m , sugerindo a utilização de qualquer um desses valores.

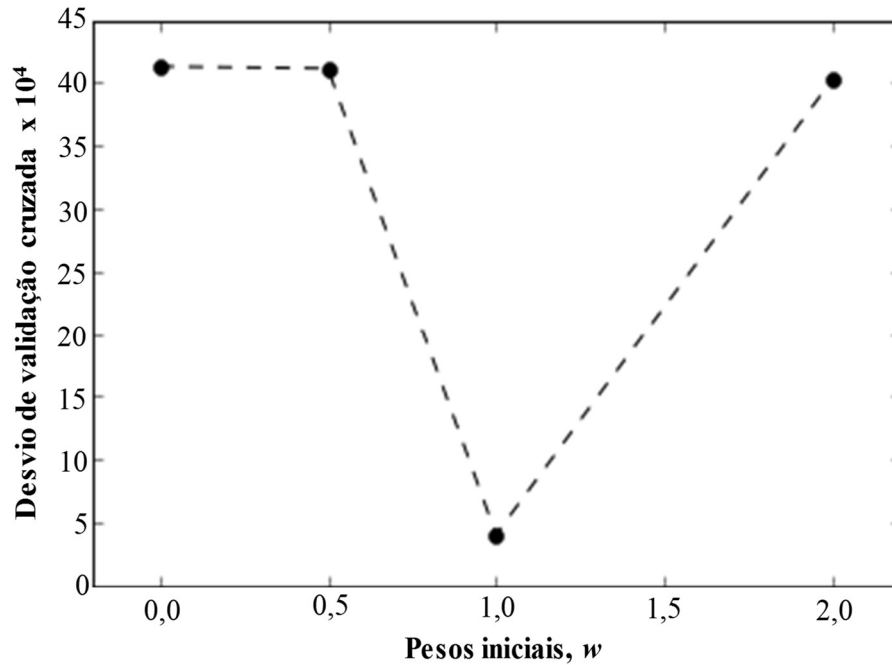


Figura 4.20 - Variação do SQR_m em função do valor inicial dos pesos.

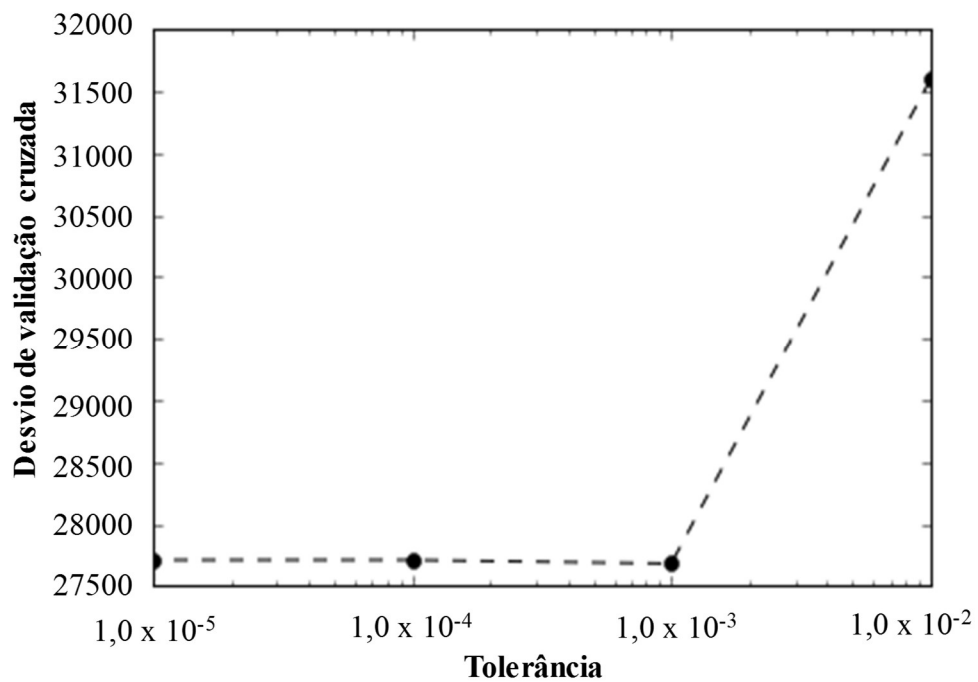


Figura 4.21 - Variação do SQR_m em função do valor da tolerância.

Outro fator importante em desenvolvimento de modelos e em procedimentos de otimização é o tamanho do passo (η). Definir o valor de η muito grande pode levar a *loop* infinito, podendo não alcançar o critério de convergência. Em contrapartida, valor de η muito pequeno pode demandar tempo de convergência demasiadamente longo, podendo levar à desvios elevados devido o critério de número máximo de iterações. Para a escolha do melhor η , a validação cruzada foi usada para os valores de η na ordem de grandeza de $1,0 \times 10^{-7}$, $1,0 \times 10^{-6}$, $1,0 \times 10^{-5}$ e $1,0 \times 10^{-4}$ (Figura 4.22). No caso deste parâmetro, a diferença no valor de SQR_m foi bastante relevante. Conforme pode ser notado na Figura 4.22, ao variar a tolerância de $1,0 \times 10^{-5}$ para $1,0 \times 10^{-4}$, SQR_m saiu da ordem de grandeza de $1,0 \times 10^5$ para $1,0 \times 10^8$, indicando que a escolha aleatória desse valor poderia diminuir a capacidade preditiva do modelo.

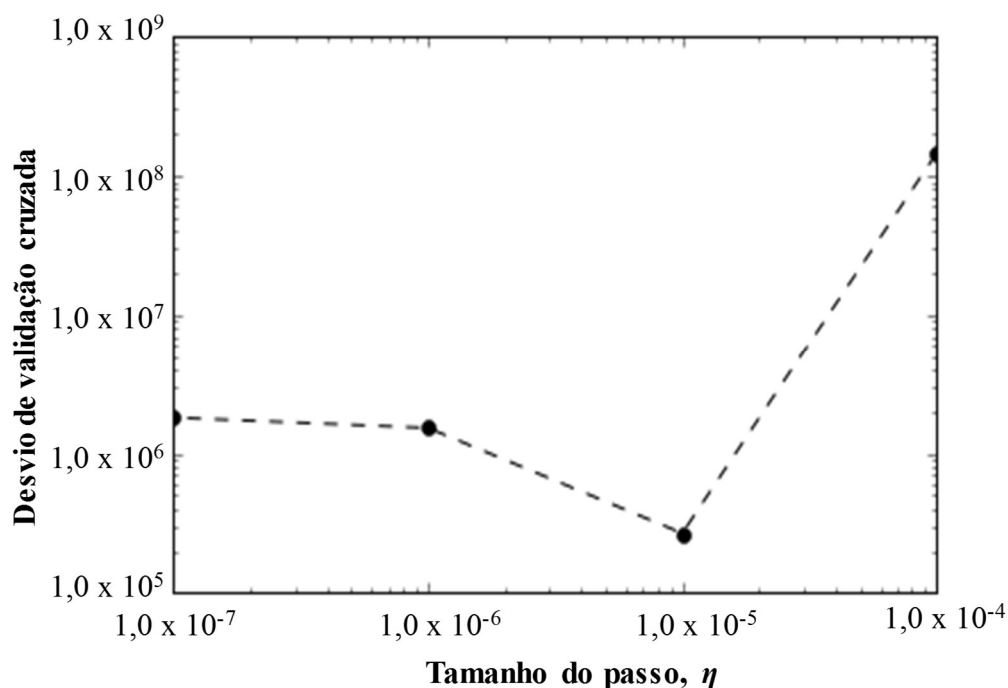


Figura 4.22 - Variação do SQR_m em função do tamanho do passo η .

Com base nesses resultados, para sequência do desenvolvimento do modelo proposto, portanto, foi considerado os valores de w iguais a 1,0, tolerância igual a $1,0 \times 10^{-4}$ e $\eta = 1,0 \times 10^{-5}$, pois apresentaram menor SQR_m no processo de validação cruzada.

Um resumo do desempenho dos parâmetros de entrada em função do SQR_m está mostrado na Tabela 4.26. Vale ressaltar que nesta etapa não foi avaliada a ordem de grandeza

do valor de SQR_m , e sim como parâmetro comparativo para a seleção dos parâmetros de entrada. Embora a diferença no valor de SQR_m tenha se mostrado pequena, quando os parâmetros foram variados, este procedimento foi preponderante para melhorar o desempenho final do modelo preditivo desenvolvido neste trabalho.

Treinamento do modelo com regressão ordinária por mínimos quadrados

Após a seleção dos valores dos parâmetros de entrada, o modelo foi novamente treinado por gradiente descendente (GD). Os resultados foram comparados com os obtidos pelo pacote GraphLab (GL) (Tabela 4.27). Conforme a análise dos desvios, pode-se notar que o modelo desenvolvido por GD apresentou-se de acordo com o obtido por GraphLab. Houve leve diminuição em todos os tipos de desvios, indicando que o algoritmo foi implementado corretamente.

Tabela 4.26 - Comparação da soma dos quadrados residuais médio (SQR_m) para escolha dos melhores parâmetros de entrada.

Parâmetro de entrada/Desvio médio	Iteração			
	0,0	0,5	1,0	2,0
Pesos iniciais SQR_m	27.711,30	27.711,14	27.674,04	27.710,25
Tolerância SQR_m	$1,0 \times 10^{-2}$ 31.593,13	$1,0 \times 10^{-3}$ 27.682,25	$1,0 \times 10^{-4}$ 27711,30	$1,0 \times 10^{-5}$ 27.710,97
Tamanho do passo SQR_m	$1,0 \times 10^{-4}$ 145.707.480,99	$1,0 \times 10^{-5}$ 268.516,11	$1,0 \times 10^{-6}$ 1.526.155,04	$1,0 \times 10^{-7}$ 1.809.934,29

Tabela 4.27 - Desempenho de regressão por gradiente descendente (GD) vs. GraphLab.

Validação	Tipo de desvio			
	SQR	$RDOM$	DA	DM (%)
GD	37.926,87	7,140	4,921	8,443
GL	37.951,73	7,143	4,926	9,099

Treinamento do modelo com regressão regularizada por mínimos quadrados

Com a intenção melhorar a qualidade do ajuste e controlar a magnitude dos coeficientes, nesta parte do trabalho foi explorada a implementação do parâmetro de regularização λ_2 . Conforme exposto no item 2.8.5, este termo afeta diretamente os valores dos coeficientes, podendo evitar possível ocorrência de *overfitting*. Porém, deve ser escolhido de forma criteriosa.

A Figura 4.23 exemplifica o efeito de λ_2 sobre os coeficientes do modelo. Apenas como exemplo, o modelo foi treinado somente em função da massa molar dos compostos (regressão linear simples). Para isso, foi aplicado valor parâmetro de regularização igual a zero ($\lambda_2 = 0,0$) e outro muito elevado ($\lambda_2 = 1,0 \times 10^8$). Observa-se na Figura 4.23 que sem o uso de λ_2 a linha de regressão seguiu a tendência normal, ou seja, o aumento do valor de entalpia de vaporização com o aumento da massa molar. Por outro lado, ao aumentar a magnitude de λ_2 a valores muito elevados, verifica-se o deslocamento da linha de regressão, tendendo os valores dos coeficientes a zero.

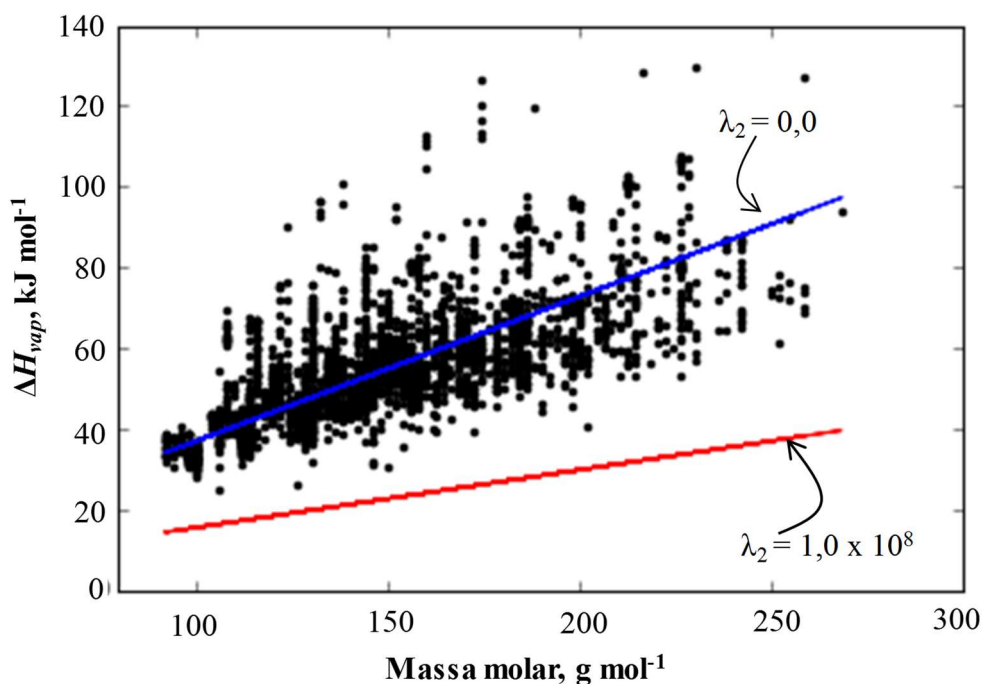


Figura 4.23 - Ilustração do efeito do parâmetro de regularização λ_2 para o modelo treinado apenas com um atributo (-CH₃).

O efeito do parâmetro de regularização λ_2 pode ser melhor visualizado por meio da Figura 4.24. Neste processo o modelo foi treinado usando todos os atributos para a faixa de λ_2 variando 20 valores de $1,0 \times 10^{-5}$ a $1,0 \times 10^5$, igualmente espaçados em escala logarítmica. Pode-se notar, na Figura 4.24, que com λ_2 muito baixo (entre $1,0 \times 10^{-5}$ e $1,0 \times 10^0$), os coeficientes não sofreram influência relevante nos seus valores. Quando λ_2 começa a aumentar de forma mais efetiva, a partir de $1,0 \times 10^1$, o valor dos coeficientes passa a ser afetado e começa a diminuir, indo a zero quando o parâmetro de regularização λ_2 recebe o valor máximo de $1,0 \times 10^5$.

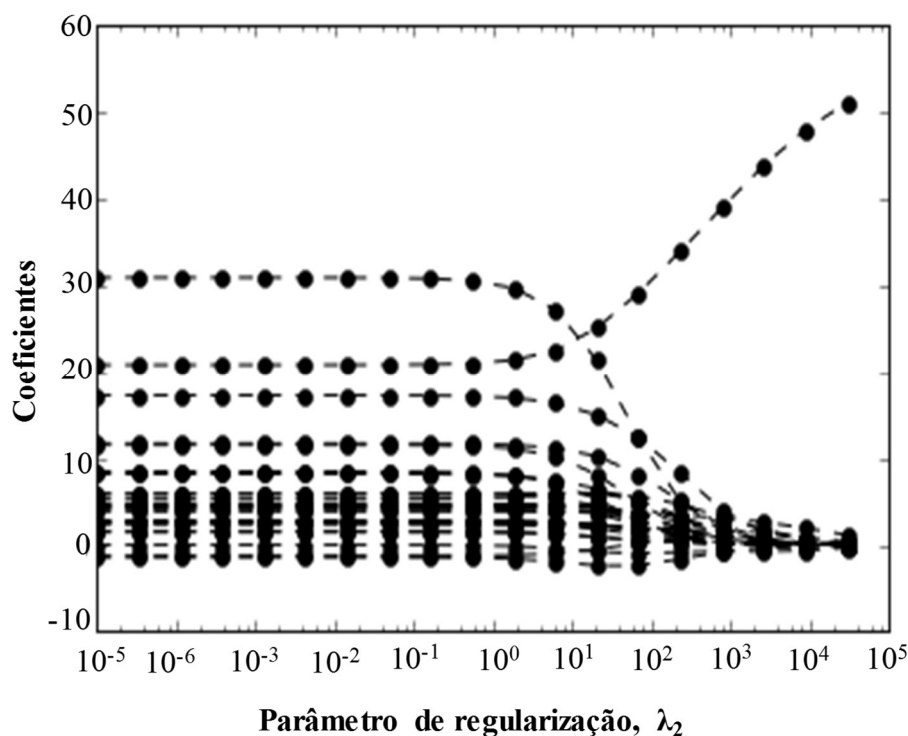


Figura 4.24 - Efeito do aumento do parâmetro de regularização λ_2 sobre o valor dos coeficientes.

Seleção do parâmetro de regularização λ_2

Para a seleção do o melhor valor de λ_2 , recorreu-se novamente à técnica de validação cruzada. Uma vez que a região onde possivelmente pode-se encontrar o valor ótimo de λ_2 já é conhecida, usou-se a validação cruzada para a faixa de 0 a 7,0, linearmente espaçados entre 20

valores. O resultado deste procedimento está apresentado na Figura 4.25. Conforme pode ser notado, λ_2 que apresentou menor SQR_m está entre 0,0 e 2,0, ou mais exatamente $\lambda_2 = 0,7368$. Isso sugere que os parâmetros do modelo são sensíveis ao valor de λ_2 , indicando que um pequeno valor de λ_2 é suficiente para deslocar os valores dos coeficientes.

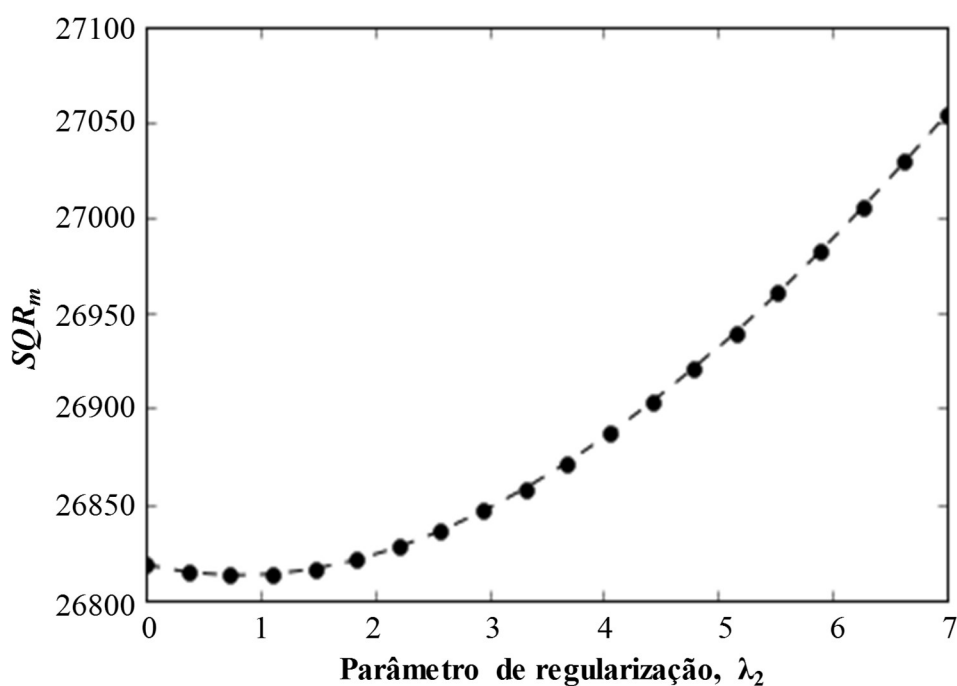


Figura 4.25 - Seleção do parâmetro de regularização λ_2 por validação cruzada.

4.3.2.4 Desempenho de regressão regularizada por mínimos quadrados usando GD

O valor de λ_2 igual a 0,7368 encontrado na seção anterior foi usado para avaliar o desempenho de regressão com a inclusão deste termo. Os resultados foram comparados com os obtidos pelo pacote GL e com os resultados encontrados na regressão sem λ_2 (Tabela 4.28).

Conforme pode ser conferido na Tabela 4.28, a técnica de regressão regularizada aplicada aos parâmetros do modelo proposto não provocou efeito positivo. Após os procedimentos de regularização houve pequena piora no desempenho do modelo, explicado pelo aumento dos desvios analisados quando comparados com a regressão sem regularização (coluna *regressão (GD)* da Tabela 4.28) e com os resultados do GraphLab Create.

Tabela 4.28 - Desempenho da regressão regularizada e sua comparação com a regressão regularizada do pacote GraphLab Create e com a regressão sem regularização.

Tipo de desvio	Regressão (GD)*	Regressão regularizada (GD)	Regressão regularizada (GL)
<i>SQR</i>	37.926,87	37.957,66	37.951,73
<i>RDQM</i>	7,140	7,143	7,142
<i>DA</i>	4,921	4,926	4,926
<i>DM (%)</i>	8,443	8,453	8,450

Legenda: GD – gradiente descendente; GL - GraphLab Create; * regressão sem λ_2

4.3.2.5 Desempenho do modelo sem os dados considerados *outliers*

De acordo com o gráfico de caixa (Figura 4.17) houve a existência de possíveis *outliers* no conjunto de dados. Conforme a regra de Tukey, os dados de entalpia de vaporização acima de aproximadamente $93,0 \text{ kJ mol}^{-1}$ são considerados *outliers*. Esses dados representam 66 compostos (87 dados de ΔH_{vap}), apresentados no Apêndice F. Inspeccionando-se as estruturas químicas desses compostos (Apêndice F), nota-se que a maioria apresenta mais de um grupo funcional, ou contém cadeia carbônica alongada, como por exemplo, pentadecan-1-ol, ácido nonanodióico, heptano-1,7-diol. A simples exclusão desses dados poderia diminuir o desempenho preditivo do modelo, além de ir contra o princípio da aditividade de grupos, pois a influência de dois ou mais grupos funcionais em uma mesma molécula deixaria de ser avaliada.

Assim, foram considerados *outliers* apenas os valores extremamente altos e que apareceram com pouca frequência no conjunto de dados. Para este caso, foram excluídos os valores de ΔH_{vap} acima de 115 kJ mol^{-1} . Após a eliminação dos *outliers*, o conjunto de dados passou a conter 3.591 dados de entalpia de vaporização.

Após isso, o conjunto de dados foi dividido novamente em *dados de treinamento* e *dados de teste*, conforme a Tabela 4.29. Os resultados do treinamento após a exclusão dos *outliers* estão apresentados na Tabela 4.30.

Os resultados obtidos demonstram que a eliminação dos *outliers* com dados de ΔH_{vap} acima de 115 kJ mol^{-1} melhorou o desempenho do modelo. Houve diminuição no valor

dos quatro parâmetros de medida de qualidade utilizados (*SQR*, *RDQM*, *DA* e *DM*). Nota-se, também, que a regressão ordinária (sem λ_2) apresentou melhor desempenho quando comparada com a regressão regularizada obtida por gradiente descendente e com a regressão regularizada obtida pelo GraphLab.

Tabela 4.29 - Organização do conjunto de dados após exclusão dos outliers acima de 115 kJ mol⁻¹.

Dados	Número de compostos	Número de dados
Teste	356	718
Treinamento	1.360	2.873

Tabela 4.30 - Desempenho do modelo após a exclusão dos outliers acima de 115 kJ mol⁻¹.

Desvio	Regressão sem pontos discrepantes		
	Regressão (GD)*	Regressão regularizada (GD)	Regressão regularizada (GL)
<i>SQR</i>	33.538,46	37.957,66	37.951,73
<i>RDQM</i>	6,835	7,143	7,142
<i>DA</i>	4,805	4,926	4,926
<i>DM</i> (%)	8,831	8,453	8,450

Legenda: GD – gradiente descendente; GL - GraphLab Create; * Regressão sem λ_2

4.3.2.9 Construção do modelo final

Com base nos resultados encontrados até aqui, o modelo de contribuição de grupos para predição de entalpias de vaporização proposto foi desenvolvido por meio da otimização por gradiente descendente sem a inclusão do parâmetro de regularização λ_2 e de dados ΔH_{vap} acima de 115 kJ mol⁻¹.

Após a etapa de treinamento, o desempenho do modelo foi avaliado sobre os dados de teste, apresentados nas Figuras 4.26 e 4.27. Na Figura 4.26, em que os dados encontram-se

misturados, pode-se notar que a predição se mostrou adequada para muitos compostos dos dados de teste. Nota-se, no entanto, que alguns dados, como os valores de entalpia acima de 80 kJ mol⁻¹, apresentaram desvios elevados. Isso pode ser atribuído a lacuna de descritores moleculares ou pelos chamados ruídos inerentes aos dados, como a temperatura de referência e a técnica em que foram obtidos, apresentados nos dados retirados da literatura.

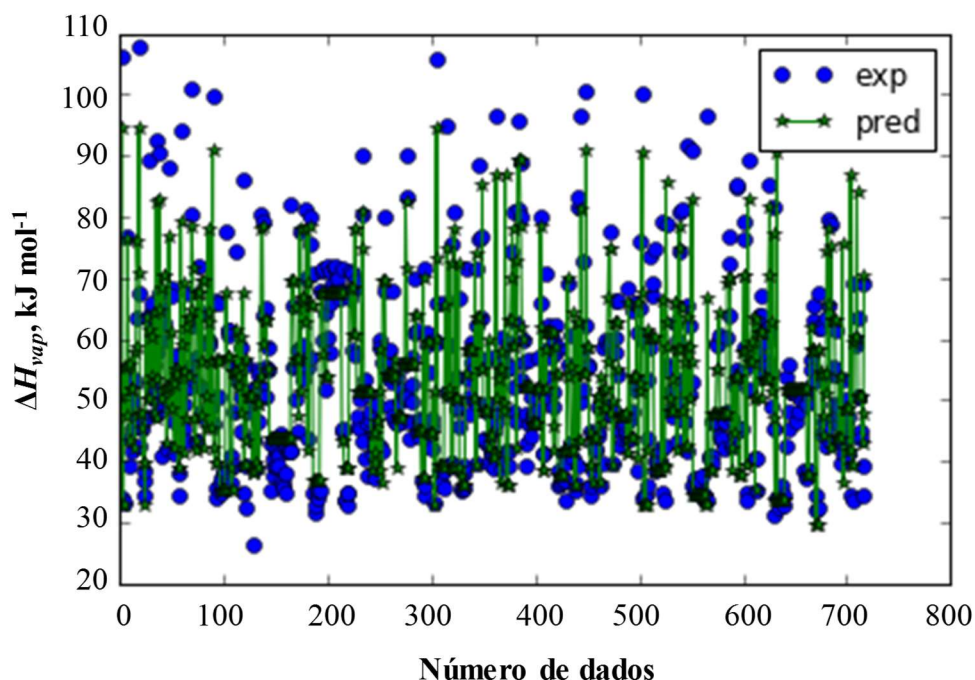


Figura 4.26 - Valores preditos vs. literatura sobre os dados de teste.

Outra maneira de visualizar o desempenho de predição sobre os dados de teste é organizando-os em ordem crescente do valor de ΔH_{vap} , como mostra a Figura 4.27. Por meio desta figura é possível notar mais claramente que os valores de entalpia de vaporização que sofreram maiores desvios referem-se aos valores entre 70 - 110 kJ mol⁻¹. No entanto, esta é apenas uma parte minoritária. Traçando-se uma linha diagonal entre os valores preditos e os valores experimentais (Figura 4.28), pode-se verificar grande concentração dos dados em torno da reta e existência de poucos dados mais dispersos. É importante frisar que a eliminação desses dados experimentais não necessariamente melhoraria o desempenho de predição, pois deixaria o banco de dados com menos número de dados e diminuiria a frequência de descritores moleculares.

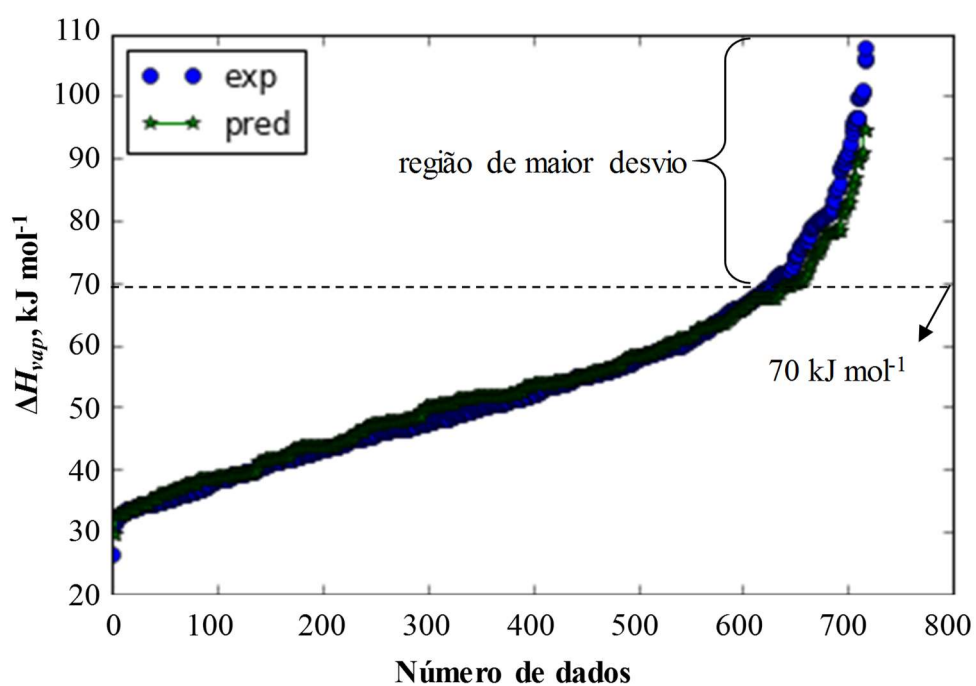


Figura 4.27 - Valores preditos vs. literatura sobre os dados de teste organizados em ordem crescente do valor de entalpia de vaporização.

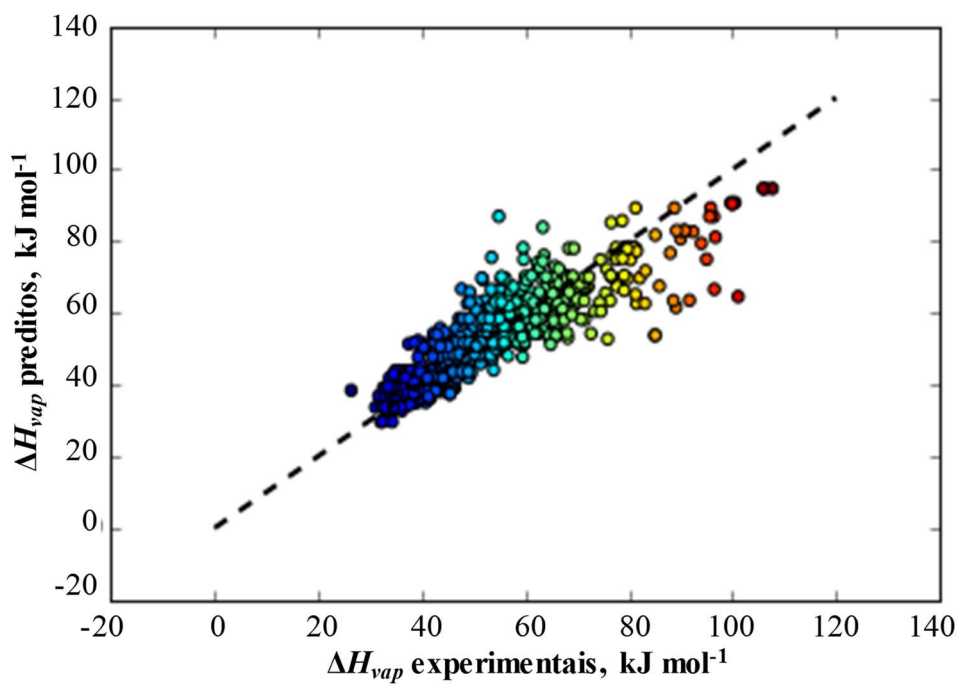


Figura 4.28 - Comparação entre os resultados preditos vs. experimentais.

A Tabela 4.31 mostra todos os grupos utilizados na regressão e suas respectivas contribuições para predição de entalpia de vaporização. Um teste de sensibilidade não apresentou diferença significativa no valor predito. Assim, todos os valores das contribuições foram mantidos em três casas decimais.

Tabela 4.31 - Grupos selecionados e as suas respectivas contribuições.

Nº	Grupo	Contribuição
0	h_0	19,548
1	-CH ₃	0,475
2	-CH ₂ -	3,900
3	>CH-	4,172
4	>C<	4,380
5	=CH ₂	-1,196
6	=CH-	4,525
7	=C<	5,008
8	-CH ₂ - (cíclico)	2,793
9	>CH- (cíclico)	3,197
10	>C< (cíclico)	4,535
11	=CH- (cíclico)	3,048
12	=C< (cíclico)	4,862
13	=CH (aromático)	2,808
14	=C< (aromático)	5,610
15	OH (álcool)	17,421
16	-O-	2,255
17	O-C=O	12,162
18	C=O	6,301
19	OH (fenol)	7,979
20	-O- (cíclico)	4,307
21	C=O (cíclico)	10,589
22	<i>bicíclico</i> (correção para estrutura bicíclica)	-2,053
23	HC=O	8,912
24	OHC=O	30,201
25	<i>cis</i> (correção para estrutura cis)	0,983
26	<i>trans</i> (correção para estrutura trans)	1,338

A respeito dos valores das contribuições, são necessárias algumas considerações:

- Os grupos contendo carbonila (C=O) e hidroxila (-OH) apresentaram maior contribuição. Isso é extremamente coerente pois os grupos que contêm carbonila são polares e possuem

forças intermoleculares fortes do tipo dipolo-dipolo; os contendo hidroxila, como álcoois, fenóis e ácidos carboxílicos, apresentam ligação de hidrogênio, que são forças extremamente fortes;

- As contribuições para os compostos bicíclicos e com isomeria cis-trans mostraram valores negativos; esses valores são coerentes devido a posição estereoquímica em compostos desse tipo;
- Os grupos referentes à compostos alifáticos ou às partes alifáticas de moléculas, como por exemplo o -CH₃, -CH₂-, apresentaram menor contribuição, que pode estar relacionado as ramificações que diminuem a atração entre as moléculas e influenciam suas propriedades termodinâmicas, como temperatura normal de ebulição e entalpia de vaporização.

4.3.3 – Avaliação de desempenho do modelo de predição proposto

A predição de entalpias de vaporização a partir do modelo proposto neste trabalho pode ser feita conforme a Equação 4.23, apresentada no item 4.3.2,

$$\Delta H_{vap} \text{ (kJ mol}^{-1}\text{)} = h_0 + \sum_{i=1}^n N_i C_i$$

a qual é válida para predição de ΔH_{vap} na faixa de 288 – 426 K. Novamente, h_0 é um valor constante; N_i é o número de grupos e C_i é sua respectiva contribuição.

O modelo proposto difere dos modelos da literatura analisados, trazendo contribuição de grupos para compostos com isomeria *cis*, *trans* e contribuição para compostos de natureza *bicíclica*. É importante destacar que o modelo proposto é específico para substâncias presentes em óleos essenciais e contempla apenas compostos que contêm oxigênio em sua estrutura, enquanto os modelos da literatura analisados são modelos gerais e são aplicáveis à várias substâncias, inclusive inorgânicas.

O resultado do desempenho do modelo proposto no presente trabalho de Tese foi realizado com base na predição feita pela equação desenvolvida, com a contribuição de cada grupo da Tabela 4.31 e através da comparação com os dados experimentais obtidos para os compostos terpênicos analisados por meio da técnica de cromatografia gasosa (item 4.1.3). Os resultados obtidos pelo modelo proposto também foram comparados com os resultados

advindos dos modelos de predição da literatura, cujas informações estão apresentadas na Tabela 4.32.

Tabela 4.32 - Desempenho do modelo desenvolvido neste trabalho em comparação com os dados experimentais obtidos por cromatografia gasosa e com os modelos preditivos analisados da literatura. Em parênteses encontra-se o desvio relativo (%) em relação aos dados experimentais.

Composto*	^a ΔH_{vap} , kJ mol ⁻¹	^b ΔH_{vap} , kJ mol ⁻¹			
		Neste trabalho	Joback e Reid	Chickos <i>et al.</i>	Kolská <i>et al.</i>
<i>p</i> -cimeno	48,81	50,4 (3,3)	40,42 (17,2)	49,9 (2,2)	48,71 (0,2)
terpinoleno	51,27	47,13 (8,1)	40,43 (21,2)	49,9 (2,7)	45,7 (10,9)
(+)-limoneno	48,77	43,8 (10,2)	38,66 (20,7)	49,9 (2,3)	48,93 (0,3)
(+)-α-pineno	44,67	43,35 (3,0)	37,36 (16,4)	46,51 (4,1)	46,89 (5,0)
(+)-β-pineno	45,30	41,42 (8,6)	36,56 (19,3)	46,51 (2,7)	42,1 (7,1)
(+)-3-careno	47,57	43,35 (8,9)	37,36 (21,5)	46,51 (2,2)	46,89 (1,4)
γ -terpineno	50,58	46,55 (8,0)	40,13 (20,7)	49,9 (1,3)	52,58 (3,9)
(-)-<i>trans</i>-cariofileno	54,01	56,73 (5,0)	49,34 (8,6)	69,96 (29,5)	65,68 (21,6)
α -humuleno	62,89	59,07 (6,1)	51,34 (18,4)	69,96 (11,2)	71,5 (13,7)
carvacrol	66,43	58,38 (12,1)	53,44 (19,6)	77,29 (16,3)	86,65 (30,4)
timol	67,12	58,38 (13,0)	53,44 (20,4)	77,29 (15,2)	87,41 (30,2)
terpinen-4-ol	55,80	63,39 (13,6)	54,4 (2,5)	62,74 (12,4)	74,71 (33,9)
linalol	54,68	63,44 (16,0)	52,62 (3,8)	62,74 (14,7)	71,28 (30,4)
(-)- α -terpineol	56,65	62,26 (9,9)	54,64 (3,5)	62,74 (10,7)	75,2 (32,8)
safrol	59,15	63,44 (7,2)	50,05 (15,4)	61,28 (3,6)	32,1 (45,7)
isosafrol	63,17	65,73 (4,1)	50,68 (19,8)	61,28 (3,0)	41,56 (34,2)
3-BHA	84,88	61,32 (27,8)	57,17 (32,6)	72,3 (14,8)	76,76 (9,6)
timol metil éter	60,16	53,13 (11,7)	45,72 (24)	62,69 (4,2)	41,71 (30,7)
B[2.2.1]**	34,03	38,87 (14,2)	31,77 (6,6)	35,83 (5,3)	40,19 (18,1)
VNB**	41,64	42,1 (1,1)	34,95 (16,1)	45,21 (8,6)	42,31 (1,6)
ENB**	43,63	45,43 (4,1)	36,72 (15,8)	45,21 (3,6)	40,96 (6,1)
álcool fenchil**	54,47	60,18 (10,5)	51,63 (5,2)	83,75 (53,8)	64,86 (19,1)
(-)-borneol**	54,80	60,18 (9,8)	51,63 (5,8)	83,75 (52,8)	64,86 (18,4)
B[2.2.2]**	56,65	56,24 (0,7)	46,73 (17,5)	67,32 (18,8)	63,55 (12,2)
B[4.1.0]**	79,15	68,46 (13,5)	50,26 (36,5)	82,42 (4,1)	66,51 (16,0)

a) valores obtidos neste trabalho usando a cromatografia gasosa; b) valores estimados; *os dados em negrito referem-se aos compostos bicíclicos; ** siglas referentes à: B[2.2.1] = biciclo[2.2.1]hepta-2,5-dieno; ENB = 5-vinil-2-norborneno; VNB = 5-etilideno-2-norborneno; B[2.2.2] = biciclo[2.2.2]octano-2,5-diona; B[4.1.0] = biciclo[4.1.0]heptano-7-ácido carboxílico.

Com base na Tabela 4.32, dos 25 compostos analisados por CG, 7 compostos apresentaram desvio relativo (d_R) entre 0 a 5,0 %: *p*-cimeno (3,3 %), (+)- α -pineno (3,0 %),

(-)-*trans*-cariofileno (1,8 %), 5-etilideno-2-norborneno (4,1 %), 5-vinil-2-norborneno (1,1 %), isosafrol (4,1 %) e biciclo[2.2.2]octano-2,5-diona. Outros 8 compostos apresentaram $5,0 < d_R \leq 10,0$ %, são eles: terpinoleno (8,1 %), (+)-3-careno (8,9 %), γ -terpineno (8,0 %), safrol (7,2 %), α -humuleno (6,1 %), (-)- α -terpineol (9,9 %) e (-)-borneol (9,8 %). Por fim, 10 compostos tiveram d_R maior que 10,0 %: timol (13,0 %), carvacrol (12,1 %), linalol (16,0 %), timol metil éter (11,7 %), (+)-limoneno (10,2 %), biciclo[2.2.2]hepta-2,5-dieno (14,2 %), biciclo[4.1.0]heptano-7-ácido carboxílico (12,5 %), álcool fenchil (10,5 %), 3-BHA (27,8 %) e terpinen-4-ol (13,6 %).

Em relação aos modelos analisados na literatura, o modelo de Joback e Reid (1987) apresentou apenas os compostos linalol (3,8 %), terpinen-4-ol (2,5 %) e (-)- α -terpineol (3,5 %) com $0 < d_R \leq 5,0$ %; e os compostos biciclo[2.2.1]hepta-2,5-dieno (6,6 %), (-)-borneol (5,8 %), *trans*-cariofileno (8,6 %) e álcool fenchil (5,2 %) com $5,0 < d_R \leq 10,0$ %. Para os dados restantes, (18 dos 25 compostos comparados) apresentaram d_R acima de 10,0 %.

O modelo de Chickos *et al.* (1998) foi o que mais apresentou predição (12 compostos) com desvio relativo entre 0 e 5,0 %. O *p*-cimeno (2,2 %), terpinoleno (2,7 %), (+)-limoneno (2,3 %), (+)- α -pineno (4,1 %), biciclo[4.1.0]heptano-7-ácido carboxílico (4,1 %), (+)- β -pineno (2,7 %), (+)-3-careno (2,2 %), γ -terpineno (1,3 %), safrol (3,6 %), 5-etilideno-2-norborneno (3,6 %), isosafrol (3,0 %) e o timol metil éter (4,2 %) foram os compostos que apresentaram menores d_R . O modelo de Chickos *et al.* (1998), no entanto, apresentou apenas o 5-vinil-2-norborneno (8,6 %) com $5,0 < d_R \leq 10,0$ % e desvios elevados para o (-)-*trans*-cariofileno (29,5 %), álcool fenchil (53,8 %) e (-)-borneol (52,8 %).

Já o modelo de Kolská *et al.* (2005), que proporciona estimativa em três níveis, mostrou d_R entre 0 e 5,0 % para o *p*-cimeno (0,2 %), (+)-limoneno (0,3 %), (+)- α -pineno (5,0 %), (+)-3-careno (1,4 %), γ -terpineno (3,9 %) e 5-vinil-2-norborneno (1,6 %); e desvios entre $5,0 < d_R \leq 10,0$ % apenas para o (+)- β -pineno (7,1 %), 5-etilideno-2-norborneno (6,1 %) e 3-BHA (9,6 %). É importante frisar que o modelo de Kolská *et al.* (2005), apesar dos três níveis de predição, foi o que apresentou maiores desvios para a maioria dos compostos estudados. O carvacrol (30,4 %), timol (30,2 %), (-)- α -terpineol (33,9 %), linalol (30,7 %), γ -terpineol (32,7 %), safrol (45 %), isosafrol (34,2 %) e timol metil éter (30,7 %) apresentaram desvios relativos acima de 30 %.

Os dados em destacados em negrito na Tabela 4.32 são referentes aos compostos bicíclicos, que se incluem no objetivo central deste trabalho. Ao todo foram analisados 11

compostos bicíclicos. De modo geral, o modelo proposto neste trabalho apresentou melhor desempenho de predição para compostos bicíclicos frente aos modelos de Jaback e Reid (1987) e de Kolská *et al.* (2005). O modelo de Chickos *et al.* (1998) apresentou melhor predição para 4 dos compostos bicíclicos analisados (biciclo[2.2.1]hepta-2,5-dieno, biciclo[4.1.0]heptano-7-ácido carboxílico, (+)- β -pineno, (+)-3-careno e 5-etilideno-2-norborneno). Por outro lado, o modelo de Chickos *et al.* (1998) apresentou desvios bastante elevados para os compostos (-)-*trans*-cariofileno (29,5 %), álcool fenchil (53,8 %) e (-)-borneol (52,8 %).

A Tabela 4.33 mostra uma comparação geral do modelo de contribuição de grupos proposto com os modelos na literatura, baseando-se no desvio relativo médio (d_M),

$$d_M = \frac{1}{n} \sum_{i=1}^n \left(\frac{|\Delta H_{vap}^{pred.} - \Delta H_{vap}^{exp.}|}{\Delta H_{vap}^{exp.}} \times 100 \right) \quad (4.23)$$

em que os sobrescritos *pred.* e *exp.* referem-se aos dados de ΔH_{vap} predito e experimental, respectivamente.

Pode-se perceber que a Tabela 4.33 confirma a análise feita acima. O modelo de Kolská *et al.* (2005) apresentou $d_M = 20,5$ %, evidenciando não ter desempenho satisfatório para os compostos analisados. Os modelos de Jaback e Reid (1987) e de Chickos *et al.* (2005) apresentaram d_M em torno de 15 %. O modelo de Chickos *et al.* (1998), embora tenha apresentado menores d_R para quase metade dos compostos, teve seu desempenho geral influenciado por (-)-*trans*-cariofileno, álcool fenchil, e (-)-borneol, que tiveram desvios muito altos, conforme exposto anteriormente. A Tabela 4.33 também confirma que o modelo proposto no presente trabalho teve desempenho geral melhor do que os modelos da literatura analisados, com desvio global em cerca de 10 %.

Tabela 4.33 - Desempenho geral do modelo desenvolvido e uma comparação com os modelos preditivos da literatura analisados.

	Modelo de predição			
	Proposto neste trabalho	Joback e Reid	Chickos <i>et al.</i>	Kolská <i>et al.</i>
d_M , %	9,87	15,72	14,31	20,50

4.3.4 Avaliação com dados da literatura

Além dos compostos analisados por CG, foi importante verificar a capacidade de generalização do modelo proposto em substâncias com dados de ΔH_{vap} apresentados por outros autores na literatura. Para isso, foram selecionados alguns compostos de existência em óleos essenciais, que apresentem características similares às quais estão sendo avaliadas neste trabalho, como por exemplo, compostos bicíclicos e isomeria *cis-trans*. Ao todo, foram feitas previsões para 30 compostos e os resultados encontrados, bem como uma comparação com os modelos da literatura avaliados, estão apresentados na Tabela 4.34.

De modo geral, o modelo proposto apresentou bom desempenho de predição para a maioria dos compostos analisados da literatura. O maior desvio relativo foi encontrado para o (-)-*cis*-verbenol, de 22,31 %; e o menor d_R foi encontrado para o α -tujeno (1,07 %). Em síntese, o modelo proposto mostrou 6 compostos com d_R até 5 % e 12 compostos com $5 < d_R \leq 10$ %. A penas (-)-*cis*-verbenol mostrou d_R acima de 20 %.

Com relação aos modelos da literatura, os modelos de Kolská *et al.* (2005) e de Chickos *et al.* (1998) apresentaram desempenho semelhante com os compostos analisados. O modelo de Kolská *et al.* (2005) mostrou 8 compostos com d_R até 5 % e 8 compostos entre 5 e 10 %. Porém, 11 compostos acusaram d_R acima de 20 %. O modelo de Chickos *et al.* (1998), por sua vez, teve 10 compostos com d_R até 10 % e 5 deles com d_R entre 5 e 10 %. O Pior desempenho foi apresentado pelo modelo de Joback e Reid (1987), que teve apenas 5 compostos com d_R até 5 % e 16 com $d_R > 20$ % (Tabela 4.34). Esses resultados são corroborados pelo desvio relativo médio (d_M). O modelo proposto neste trabalho apresentou $d_M = 9,51$ %, ao passo que os modelos de Joback e Reid(1987), Chickos *et al.* (1998) e Kolská *et al.* (2005) apresentaram, respectivamente, 19,24 %, 13,85 % e 15,79 % (Tabela 4.35).

Com respeito aos compostos bicíclicos, as previsões de ΔH_{vap} feitas pelo modelo proposto, tanto para os compostos analisados por CG neste trabalho quanto com os dados da literatura, foram melhores apenas que o modelo de Joback e Reid (1987), e apresentou resultados em concordância com os modelos de Chickos *et al.* (1998) e Kolská *et al.* (2005). No entanto, esses dois últimos modelos apresentaram d_R elevados para alguns bicíclicos, como por exemplo $d_R = 53,8$ % para o álcool fenchil, apresentando pelo modelo de Chickos *et al.* (1998), e $d_R = 32,07$ % para (-)-mirtenal, mostrado pelo modelo de Kolská *et al.* (2005). Nesse sentido, o maior d_R mostrado pelo modelo proposto foi de 22,31 %, para o (-)-*cis*-verbenol, que

a princípio está dentro da lacuna deixada neste trabalho, dado que o histograma (Figura 4.16) indicou baixa probabilidade de predição para dados acima de 80 kJ mol⁻¹.

Tabela 4.34 - Predição de ΔH_{vap} para compostos terpênicos com dados experimentais da literatura. Em parênteses encontram-se os desvios relativos.

a) N ^o	Composto ^b	ΔH_{vap} , kJ mol ⁻¹ (Literatura)	Predição de ΔH_{vap} , kJ mol ⁻¹			
			Neste trabalho	Joback e Reid	Chickos <i>et al.</i>	Kolská <i>et al.</i>
1	canfeno	44,65	41,42 (7,24)	36,56 (18,12)	43,12 (3,43)	42,1 (5,71)
2	(+)-2-careno	48,45	43,35 (10,54)	37,36 (22,9)	43,12 (11)	46,89 (3,22)
3	sabineno	46,91	42,39 (9,63)	36,31 (22,6)	43,12 (8,08)	45,39 (3,24)
4	α -tujeno	44,80	44,32 (1,07)	37,11 (17,18)	43,12 (3,75)	50,18 (12,01)
5	(-)- α -felandreno	48,32	45,14 (6,58)	39,15 (18,97)	49,9 (3,27)	52,48 (8,62)
6	1,8-cineol	48,95	46,67 (4,67)	39,93 (18,42)	51,02 (4,23)	39,87 (18,56)
7	(-)-mirtenal	55,05	51,78 (5,94)	44,08 (19,93)	43,12 (21,67)	37,39 (32,07)
8	(+)-carvona	58,24	51,59 (11,41)	45,23 (22,34)	42,63 (26,8)	54,32 (6,73)
9	(+)-pulegona	57,93	53,01 (8,5)	45,73 (21,05)	42,63 (26,41)	56,01 (3,32)
10	geranial	62,54	56,75 (9,26)	44,66 (28,59)	56,02 (10,43)	46,74 (25,27)
11	neral	60,18	56,75 (5,7)	44,66 (25,79)	56,02 (6,91)	46,74 (22,34)
12	(+)- <i>cis</i> -verbenol	54,94	62,15 (13,13)	53,74 (2,19)	67,87 (23,54)	68,67 (24,98)
13	(+)- <i>trans</i> -verbenol	55,00	62,51 (13,65)	53,74 (2,3)	67,87 (23,4)	68,67 (24,85)
14	citronelol	63,50	67,7 (6,61)	54,2 (14,64)	87,14 (37,23)	75,43 (18,79)
15	(-)-mentol	56,56	60,26 (6,55)	53,97 (4,57)	74,65 (31,99)	73,39 (29,76)
16	(<i>cis</i>)- β -farneseno	72,50	62,78 (13,41)	47,82 (34,05)	63,18 (12,86)	69,55 (4,08)
17	β -cariofileno	65,54	57,44 (12,35)	49,34 (24,72)	63,18 (3,6)	62,46 (4,7)
18	anisol	44,31	41,93 (5,38)	35,88 (19,03)	40,83 (7,85)	28,46 (35,78)
19	metil salicilato	56,17	62,62 (11,48)	56,12 (0,09)	73,42 (30,71)	92,79 (65,2)
20	metil fenil acetato	57,36	55,74 (2,83)	45,33 (20,98)	55,71 (2,88)	60,48 (5,44)
21	etil fenil acetato	60,66	59,64 (1,69)	47,55 (21,61)	60,4 (0,43)	65,2 (7,49)
22	2-fenil etanol	54,55	64,42 (18,09)	52,38 (3,98)	81,15 (48,76)	70,65 (29,52)
23	vanilina	66,92	64,42 (3,73)	58,51 (12,57)	75,82 (13,3)	81,12 (21,21)
24	(-)- <i>cis</i> -verbenol	80,00	62,15 (22,31)	53,74 (32,83)	65,52 (18,1)	73,09 (8,64)
25	(-)- β -pineno	45,90	41,42 (9,76)	36,56 (20,35)	43,12 (6,06)	43,45 (5,33)
26	(-)- α -pineno	44,60	43,35 (2,81)	37,36 (16,24)	43,12 (3,32)	45,03 (0,96)
27	(-)-verbenona	58,90	51,14 (13,17)	43,93 (25,42)	56,52 (4,04)	55,85 (5,17)
28	(DL)-mentol	74,80	60,54 (19,07)	53,97 (27,84)	72,3 (3,34)	73,39 (1,88)
29	(-)- α -bisabolol	96,60	78,15 (19,10)	64,55 (33,18)	88,97 (7,9)	92,53 (4,21)
30	etoxibenzeno	50,70	45,83 (9,61)	38,1 (24,84)	45,52 (10,22)	33,18 (34,56)

a) compostos n^o 1 a 23 de Hoskovec *et al.* (2005); n^o 24 a 27 de Štejfá *et al.* (2013); n^o 28 e 29 de Keating *et al.* (2017); e n^o 30 de Gobble e Chickos (2016); b) Destaque em negrito refere-se aos compostos bicíclicos.

Tabela 4.35 - Distribuição dos desvios relativos (d_R) obtidos para todos compostos analisados da literatura.

d_R	Neste trabalho	Joback e Reid	Chickos <i>et al.</i>	Kolská <i>et al.</i>
até 5 %	6	5	10	8
5 - 10 %	12	0	5	8
10 - 20 %	11	9	6	3
acima 20 %	1	16	9	11
d_M	9,51	19,24	13,85	15,79

Capítulo 5

CONCLUSÃO E SUGESTÕES

5.1 Conclusões finais

Os resultados obtidos neste trabalho demonstraram que as grandezas termodinâmicas pressão de vapor (P_{vap}), entalpia de vaporização (ΔH_{vap}) e temperatura normal de ebulição (T_{nb}) de compostos orgânicos contidos em óleos essenciais podem ser determinados satisfatoriamente pela técnica de cromatografia gasosa (CG). Além disso, no que se refere à compostos hidrocarbonetos terpênicos e derivados oxigenados, o modelo de predição proposto demonstrou-se, de forma geral, melhor do que os três modelos de predição analisados da literatura.

Por meio dos índices de retenção de Kováts, calculados para várias temperaturas, e das equações do tipo Kirchhoff-Rankin, que combinam os índices de retenção com quantidades termodinâmicas, foram obtidos valores confiáveis das grandezas termodinâmicas examinadas. Ao todo foram avaliados 25 compostos terpênicos e mais 9 hidrocarbonetos alifáticos, usados como parâmetros de aplicabilidade do método a diferentes volatilidades. Os dados de P_{vap} foram determinados a 298,15 K e apresentaram-se em conformidade com os dados da literatura. Dos 34 compostos examinados, apenas os terpenos linalol (15,53 %), (-)- α -terpineol (13,75 %), carvacrol (15,38 %), γ -terpineno (16,27 %) e terpinoleno (11,73 %) tiveram desvio relativo médio, d_M , superior a 10 %, cujas diferenças podem ser atribuídas, principalmente, às temperaturas de referência dos dados da literatura.

Em relação aos dados de ΔH_{vap} , os dados para os hidrocarbonetos alifáticos ficaram de acordo com os encontrados nos bancos de dados da NIST (online) e TDE/NIST (Aspen® Versão 8.4), com faixa de desvio relativo médio de 0,49 – 7,47 %. Com relação aos compostos terpênicos analisados, a comparação com valores da literatura mostrou, em geral, boa concordância entre os dados. Onze terpenos apresentaram $d_R < 2,0$ %, são eles: (+)- α -pineno

(0,74 %), (+)- β -pineno (1,51 %), (+)-3-carene (1,73 %), (+)-limoneno (1,15 %), linalol (1,03 %), γ -terpineno (1,56 %), *p*-cimeno (0,87 %), terpinen-4-ol (0,5 %), terpinoleno (0,93 %) e 5-vinil-2-norborneno. Além disso, a técnica CG também se mostrou adequada para a obtenção de T_{nb} para os compostos analisados, que também se apresentaram de acordo com a literatura. Para os terpenos, os d_M ficaram entre 0,05 - 3,19 %. No caso dos hidrocarbonetos alifáticos, os d_M foram na faixa de 0,02 - 2,36 %. Quanto aos resultados de predição, o modelo de contribuição de grupos proposto demonstrou melhor desempenho em relação aos modelos de Kolská *et al.* (2005), Chickos *et al.* (1998) e Joback e Reid (1987), quando usados para predição de ΔH_{vap} para os 25 compostos terpenos analisados neste trabalho por CG e para 30 compostos da literatura.

Considerando o conjunto de dados de 55 compostos usado para análise de desempenho, o modelo proposto apresentou desvio relativo médio de 9,38 %, ao passo que os métodos Joback e Reid (1987), Chickos *et al.* (1998) e Kolská *et al.* (2005) apresentaram 17,93 %, 13,01 % e 16,49 %, respectivamente. Com respeito aos compostos bicíclicos, as predições de ΔH_{vap} feitas pelo modelo proposto, tanto para os compostos analisados por CG neste trabalho quanto com os dados da literatura, foram melhores que o modelo de Joback e Reid (1987), e apresentou resultados em concordância com os modelos de Chickos *et al.* (1998) e Kolská *et al.* (2005). No entanto, esses dois últimos modelos apresentaram d_R elevados para alguns bicíclicos, como por exemplo $d_R = 53,8$ % para o álcool fenchil, apresentando pelo modelo de Chickos *et al.* (1998), e $d_R = 32,07$ % para (-)-mirtenal, mostrado pelo modelo de Kolská *et al.* (2005). Nesse sentido, o maior d_R mostrado pelo modelo proposto foi de 22,31 %, para o (-)-*cis*-verbenol, que a princípio está dentro da lacuna deixada neste trabalho, uma vez que o histograma (Figura 4.16) indicou baixa probabilidade de predição para dados acima de 80 kJ mol⁻¹.

Reitera-se que este trabalho contribuiu com a determinação experimental de pressão de vapor, entalpias de vaporização e de temperatura normal de ebulição, que são cruciais em vários ramos da indústria química. Além disso, tendo em vista a lacuna deixada por alguns modelos da literatura, foi proposto um modelo de predição de entalpia de vaporização voltado para compostos advindos de óleos essenciais e enfatizando os compostos bicíclicos. Adicionalmente, foram apresentados dados de P_{vap} a 298,15 K inéditos para os compostos terpênicos 5-vinil-2-norborneno, 5-etilideno-2-norborneno, biciclo[2.2.1]hepta-2,5-dieno, isosafrol, 3-BHA, timol metil éter, (-)-*trans*-cariofileno, biciclo[2.2.2]octano-2,5-diona, álcool

fenchil, (-)-borneol e biciclo[4.1.0]hepta-7-ácido carboxílico, além de dados inéditos de ΔH_{vap} a 298,15 K e T_{nb} para os sete últimos compostos.

5.2 Sugestões para trabalhos futuros

Diante do que foi concluído, algumas sugestões para dar seguimento a este trabalho são propostas:

- ✚ Verificar adaptações da técnica utilizada neste trabalho em outras classes orgânicas, principalmente de alto custo, como por exemplo os compostos farmacêuticos;
- ✚ Determinar experimentalmente dados termodinâmicos de classes específicas de componentes de óleos essenciais, como por exemplo as cetonas terpênicas, que são muito utilizadas no tratamento do mal de Alzheimer.
- ✚ Propor a utilização de coeficientes de atividade na diluição infinita (γ^∞) experimentais na Equação 2.13, em vez de fazer a simplificação apresentada.
- ✚ Examinar o efeito de fases estacionárias (colunas) diferentes no cálculo do índice de retenção de Kováts.
- ✚ Propor um modelo de contribuição de grupos para predição de pressão de vapor e/ou ponto de ebulição de compostos terpênicos, tomando como ponto de partida os descritores moleculares apresentados no presente trabalho de Tese.
- ✚ Aplicar teste estatísticos como o de Kolmogorov-Smirnov, Anderson-Darling e Qui-Quadrado para verificar qual função densidade de probabilidade melhor se adequa à distribuição de frequência dos dados apresentados.
- ✚ Examinar a existência de outros descritores moleculares como parâmetros do modelo, tais como a interação entre grupos e os diferentes tipos de anéis bicíclicos.
- ✚ Estudar o efeito dos diferentes tipos de estruturas bicíclicas sobre as propriedades termodinâmicas de compostos terpênicos.
- ✚ Promover adequação do modelo proposto para predição de entalpias de vaporização a 298,15 K
- ✚ Propor a utilização de outros métodos de seleção de parâmetros, a exemplo do *lasso regression*.
- ✚ Promover a regressão de parâmetros por diferentes algoritmos, como o gradiente descendente estocástico.

Referências

ABDULAGATOV, I.M.; BAZAEVA, A.R.; BAZAEVA, E.A.; DZHAPPAROVA, T.A. PVT properties of 1-propanol in the critical and supercritical regions. *Journal of Supercritical Fluids*, p.117, 172-193, 2016.

ALPAYDIN, E. *Introduction to machine learning*. London, England: MIT Press Cambridge, 2.ed, 2010. 537p.

AMORIM, J.L.; SIMAS, D.L.R.; PINHEIRO, M.M.G.; MORENO, S.A.; ALVIANO, C.S SILVA, A.J.R.; FERNANDES, P.D. Anti-inflammatory properties and chemical characterization of the essential oils of four citrus species. *PLoS ONE*, v.11, p.1-18, 2016.

ARAUJO, A.F.O.; RIBEIRO-PAES, J.T.; DE DEUS, J.T.; CAVALCANTI, S.C.H.; NUNES, R.S.; ALVES, P.B.; MACORIS, M.L.G. Larvicidal activity of *Syzygium aromaticum* (L.) Merr and *Citrus sinensis* (L.) Osbeck essential oils and their antagonistic effects with temephos in resistant populations of *Aedes aegypti*. *Memórias do Instituto Oswaldo Cruz*, v.111, p.443-449, 2016.

BARROS, G.C.; TRESVENZOL, L.M.F.; CUNHA, L.C.; FERRI, P.H.; PAULA, J.R.; BARA, M.T.F. Chemical Composition, antibacterial activity and evaluation of acute toxicity of *Vetiveria zizanoides* L. Nash (Poaceae). *Latin American Journal of Pharmacy*, v.28, p.531-537, 2009.

BASAROVÁ, P.; SVOBODA, V. Prediction of the enthalpy of vaporization by the group contribution method. *Fluid Phase Equilibria*. v.105, p. 27-47, 1995.

BIZZO, H.R.; HOVELL, A.M.C.; REZENDE, C.M. Óleos essenciais no Brasil: aspectos gerais, desenvolvimento e perspectivas. *Química Nova*, v.32, p.588-594, 2009.

BURANT, A.; THOMPSON, C.; LOWRY, G.V.; KARAMALIDIS, A.K. New linear partitioning models based on experimental water: supercritical CO₂ partitioning data of selected organic compounds. *Environmental Science & Technology*, v.50, p.5135-5142, 2016.

CALM, J.M.; HOURAHAN, G.C. Physical safety and environmental data for current and alternative refrigerants. *Refrigeration for Sustainable Development* (proceedings of the 23rd International Congress of Refrigeration (ICR 2011, Prague, Czech Republic, 2011.08.21-26).

CARRASCO, A.; MARTINEZ-GUTIERREZ, R.; TOMAS, V.; TUDELA, J. *Lavandula angustifolia* and *Lavandula latifolia* essential oils from Spain: aromatic profile and bioactivities. *Planta Medica*, v.82, p.163-170, 2016.

CHAGAS, A.C.S.; BARROS, L.D.; COTINGUIBA, F.; FURLAN, M.; GIGLIOTI, R.; OLIVEIRA, M.C.S.; BIZZO, H.R. In vitro efficacy of plant extracts and synthesized substances on *Rhipicephalus (Boophilus) Microplus* (Acari: Ixodidae). *Parasitology Research*, v.110, p.295-303, 2012.

CHATTERJEE, K.; DOLLIMORE, D.; ALEXANDER, K. A new application for the Antoine equation in formulation development. *International journal of pharmaceuticals*, v. 213, p.31-44, 2001.

CHICKOS, J. S.; ACREE, W. E. Phase transition enthalpy measurements of organic and organometallic compounds. sublimation, vaporization and fusion enthalpies from 1880 to 2010. *Journal of Physical and Chemical Reference Data*, v. 39, p. 1-942, 2010.

CHICKOS, J. S.; ACREE, W. E.; LIEBMAN, J. F. Estimating phase-change enthalpies and entropies. In: *Computational thermochemistry (Chapter 4)*. American Chemical Society, 1998.

CHICKOS, J.S; HOSSEINI, S.; HESSE, D.G. Determination of vaporization enthalpies of simple organic-molecules by correlations of changes in gas-chromatographic net retention times. *Thermochimica Acta*, v.249, p.41-62, 1995.

CIAZYNSKA-HALAREWICZ, K.; KOWALSKA, T. A study of the dependence of the Kováts retention index on the temperature of analysis on stationary phases of different polarity. *Acta Chromatographica*, v.13, p.69-80, 2003.

CLARKE, E.C.W.; GLEW, D.N. Evaluation of thermodynamic functions from equilibrium constants. *Transactions of the Faraday Society*, v.62, p.539-547, 1965.

COLLINS, C.H.; BONATO, P.S.; BRAGA, G.L. *Fundamentos de cromatografia*. Campinas: Editora da UNICAMP, 2006.

CONCEIÇÃO, G. M.; CASTRO, A. A. J. F. Fitossociologia de uma área de cerrado marginal, Parque Estadual do Mirador, Mirador, Maranhão. *Scientia Plena*, v.5, p.1-16, 2009.

CRAVEIRO, A.A.; e QUEIROZ, D.C. Óleos essenciais e a química fina. *Química Nova*, v.16, p.224-228, 1993.

CLARÁ, R.A.; MARIGLIANO, A.C.G.; SÓLIMO, H.N. Density, viscosity, and refractive index in the range (283.15 to 353.15) K and vapor pressure of α -pinene, *d*-limonene, (\pm)-linalool, and citral over the pressure range 1.0 kPa atmospheric pressure. *Journal of Chemical and Engineering Data*, v.54, p.1087-1090, 2009.

CREMASCO, M. A.; BRAGA, N. P. Análise termogravimétrica do óleo essencial de pimenta longa (*Piper hispidinervium* C. DC). *Acta Amazonica*, v.41, p.275-278, 2011.

DENG, S.; SU, W.; ZHAO, L. A neural network for predicting normal boiling point of pure refrigerants using molecular groups and a topological index. *International Journal of Refrigeration*, v.63, p.63-71, 2016.

DIAS, F.R.F.; SILVA, W.A.; CAMPOS, V.R.; CUNHA, A.C.; FERREIRA, V.F.; SOUZA, N.A. Resgatando um método eficiente para determinação do ponto de ebulição de substâncias orgânicas: percolador versus Siwoloboff. *Química Nova*, v. 37, p.915-918, 2014.

DOBBS, A.J.; HART, G.F.; PARSONS, A.H. The determination of vapour pressures from relative volatilization rates. *Chemosphere*, v.13, p.687-692, 1984.

DOLLIMORE, D.; EVANS T.A.; LEE Y.F.; WILBUM F.W. Correlation between the shape of a TG/DTG curve and the form of the kinetic mechanism which is applying. *Thermochimica Acta*, v.198, p.249-257, 1992.

DÜMBGEN, L.; RIEDWYL, H. On fences and asymmetry in box-and-whiskers plots. *The American Statistician*, v.61, p.356-359, 2007.

ESPINOSA DÍAZ, M.A.; GUETACHEW, T.; LANDY, P.; JOSE, J.; VOILLEY, A. Experimental and estimated saturated vapour pressures of aroma compounds. *Fluid Phase Equilibria*, v.157, p.257-270, 1999.

ETTRE, L.S.; BILLEB, K. Considerations of the retention index concept: Retention index and column temperature. *Journal of Chromatography A*, v.30, p.1-11, 1967.

ETZLER, F.M.; CONNERS, J.J. A DSC/TGA method for determination of the heat of vaporization. *Thermochimica Acta*, v.189, p.185-182, 1991.

FARRITOR, R.E.; TAO, L.C. An improved method of measurement of vaporization heat of volatile liquids with a differential scanning calorimeter. *Thermochimica acta*, v.1, p.297-304, 1970.

FICHAN, I.; LARROCHE, C.; GROS, J.B. Water solubility, vapor pressure, and activity coefficients of terpenes and terpenoids. *Journal of Chemical & Engineering Data*, v.44, p.56-62, 1999.

FISCHER R.G.; BALLSCHMITER, K. Prediction of the environmental distribution of alkyl dinitrates - Chromatographic determination of vapor pressure p^0 , water solubility S_{H_2O} , gas-water partition coefficient K_{GW} (Henry's law constant) and octanol-water partition coefficient K_{OW} . *Fresenius Journal of Analytical Chemistry*, v.360, p.769-776, 1998.

FISCHER, R.C.; WITTLINGER, R.; BALLSCHMITER, K. Retention-index based vapor pressure estimation for polychlorobiphenyl (PCB) by gas chromatography. *Fresenius Journal of Analytical Chemistry*, v.342, p.421-425, 1992.

GOBBLE, C.; CHICKOS, J.S. The vapor pressure and vaporization enthalpy of R-(+)-menthofuran, a hepatotoxin metabolically derived from the abortifacient terpene, (R)-(+)-pulegone by correlation gas chromatography. *Journal of Chemical Thermodynamics*, v.98, p.135-139, 2016.

GOBBLE, C.; CHICKOS, J.S. The vapor pressure and vaporization enthalpy of R-(+)-menthofuran, a hepatotoxin metabolically derived from the abortifacient terpene, (R)-(+)-pulegone by correlation gas chromatography. *The Journal of Chemical Thermodynamics*, v.98, p.135-139, 2016.

GOBBLE, C.; WALKER, B.; CHICKOS, J.S. The vaporization enthalpy and vapor pressure of fenpropidin and phencyclidine (PCP) at $T/K=298.15$ by correlation gas chromatography. *Journal of Chemical And Engineering Data*, v.61, p.896-902, 2016.

GOLDFARB, J.L. Review of Sublimation Thermodynamics of Polycyclic Aromatic Compounds and Heterocycles. *Journal of Heterocyclic Chemistry*, v.50, p.1243-1263, 2013.

GOODRICH, S.; HASANOVIC, J.; GOBBLE, C.; CHICKOS, J.S. Vaporization enthalpies and vapor pressures of some insect pheromones by correlation gas chromatography. *Journal of Chemical And Engineering Data*, v.61, p.1524-1530, 2016.

GÖRGÉNYI, M.; HÉBERGER, K. Minimum in the temperature dependence of the Kováts retention indices of nitroalkanes and alkanenitriles on an apolar phase. *Journal of Chromatography A*, v.985, p.11-19, 2003.

GRAYSON, B.T.; FOSBRAEY, L.A. Determination of the vapour pressure of pesticides. *Pesticide Science*, v.13, 269-278, 1982.

GROB, R.L. *Modern Practice of Gas Chromatography*. New York: John Wiley & Sons, 1977.

HASHIMOTO, G.S.O.; NETO, F.M.; RUIZ, M.L.; ACCHILE, M.; CHAGAS, E.C.; CHAVES, F.C.M.; MARTINS, M.L. Essential oils of *Lippia sidoides* and *Mentha piperita* against monogenean parasites and their influence on the hematology of Nile tilapia. *Aquaculture*, v.450, p.182-186, 2016.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The elements of statistical learning: Data mining, inference, and prediction*. 2.ed. Stanford, California: Springer, 2008. 739p.

HAZRA, A.; ALEXANDER, K.; DOLLIMORE, D.; RIGA, A. Characterization of some essential oils and their key components: Thermoanalytical techniques. *Journal of Thermal Analysis and Calorimetry*, v.75, p.317-330, 2004.

HAZRA, A.; DOLLIMORE, D.; ALEXANDER, K. Thermal analysis of the evaporation of compounds used in aromatherapy using thermogravimetry. *Thermochimica Acta*, v.392, p.221-229, 2002.

HÉBERGER, K.; GÖRGÉNYI, M.; KOWALSKA, T. Effect of film thickness on the retention index of nitro and cyano compounds. Interpretation of the logarithmic equation. *Journal of Chromatographic Science*, v.42, p.288-292, 2004.

HÉBERGER, K.; GÖRGÉNYI, M.; KOWALSKA, T. Temperature dependence of Kováts indices in gas chromatography revisited. *Journal of Chromatography A*, v.973, p.135-142, 2002.

HEGAZY, T.; MOSELHI, O. Analogy-Based Solution to Markup Estimation Problem. *Journal of Computing in Civil Engineering*, v.8, p.72-87, 1994.

HINCKLEY, D. A.; BIDLEMAN, T.F.; FOREMAN, W.T. Determination of vapor pressures for nonpolar and semipolar organic compounds from gas chromatographic retention data. *Journal of Chemical & Engineering Data*, v.35, p.232-237, 1990.

HOSKOVEC, M.; GRYGAROVÁ, D.; CVAČKA, J.; STREINZ, L.; ZIMA, J.; VEREVKIN, S.P.; KOUTEK, B. Determining the vapour pressures of plant volatiles from gas chromatographic retention data. *Journal of Chromatography A*, v.1083, p.161-172, 2005.

JOBACK K.G.; REID R.C. Estimation of pure-component properties from group-contributions. *Chemical Engineering Communications.*, v.57, p.233-243, 1987.

JUAREZ, Z.N.; BACH, H.; SANCHEZ-ARREOLA, E.; BACH, H.; HERNANDEZ, L.R. Protective antifungal activity of essential oils extracted from *Buddleja perfoliata* and *Pelargonium graveolens* against fungi isolated from stored grains. *Journal of Applied Microbiology*, v.120, p.1264-1270, 2016.

KAISER, M. A. Qualitative and quantitative analysis by gas chromatography. In: GROB, R. L. *Modern practice of gas chromatography*. John Wiley: New York, 1977.

KEATING, L.; HARRIS, H.H., CHICKOS, J.S. Vapor pressures and vaporization enthalpy of (alfa) a-bisabolol and (dl) menthol by correlation gas chromatography, *J. Chem. Thermodynamics*, v.107, p.18-25, 2017.

KLAUCK, M.; SILBERMANN, R.; METASCH, R.; JASINOWSKI, T.; KALIES, G.; SCHMELZER, J. VLE and LLE in ternary systems of two associating components (water, aniline, and cyclohexylamine) and a hydrocarbon (cyclohexane or methylcyclohexane). *Fluid Phase Equilibria*, v.369, p.95-108, 2014.

KOLSKÁ, Z.; RŮŽIČKA, V.; GANI, R. Estimation of the enthalpy of vaporization and the entropy of vaporization for pure organic compounds at 298.15 K and at normal boiling temperature by a group contribution method. *Industrial & Engineering Chemistry Research.*, v.44, p.8436-8454, 2005.

KOLSKÁ, Z.; ZÁBRANSKÝ, M.; RANDOVÁ, A. Group contribution methods for estimation of selected physico-chemical properties of organic compounds. *Thermodynamics - Fundamentals and Its Application in Science*, 2012., InTech, DOI: 10.5772/49998. Disponível em: < <http://www.intechopen.com/books/> >. Acesso em: 20 ago. 2016.

KOUTEK, B.; HOSKOVEC, M.; VRKOČOVÁ, P.; FELTL, L. Gas chromatographic determination of vapour pressures of pheromone-like compounds IV. Acetates, a reinvestigation. *Journal of Chromatography A*, v.759, p.93-109, 1997.

KRUIF, C.G.; KUIPERS, T.; VAN MILTENBURG, J.C.; SCHAAKE, R.C.F.; STEVENS, G. The vapour pressure of solid and liquid naphthalene. *Journal of Chemical Thermodynamics*, v.13, p.1081-1086, 1981.

KULIKOV, D.; VEREVKIN, S.P.; HEINTZ, A. Enthalpies of vaporization of a series of aliphatic alcohols: Experimental results and values predicted by the ERAS-model. *Fluid Phase Equilibria*, v.192, p.187-207, 2001.

LEBRÓN-AGUILAR, R.; QUINTANILLA-LÓPEZ, E.; GARCÍA-DOMÍNGUEZ, J.A. Improving the accuracy of Kováts' retention indices in isothermal gas chromatography. *Journal of Chromatography A*, v.945, p.185-194, 2002.

LETCHER, T.M.; NAICKER, P.K. Determination of vapor pressures using gas chromatography. *Journal of Chromatography A*, v.1037, p.107-114, 2004.

LI, J.; PERDUE, E.M.; PAVLOSTATHIS, S.G.; ARAUJO, R. Physicochemical properties of selected monoterpenes. *Environment International*, v.24, p.353-358, 1998.

LIAN, B.; YALKOWSKY, S.H. Unified physicochemical property estimation relationships (UPPER). *Journal of Pharmaceutical Sciences*, v.103, p.2710-2723, 2014.

LIDE, D.R. *CRC Handbook of Chemistry and Physics*, 84 ed., CRC Press, 2004.

LUCIA, A.; LICASTRO, S.; ZERBA, E.; AUDINO, P.G.; MASUH, H. Sensitivity of *Aedes aegypti* adults (Diptera: Culicidae) to the vapors of *Eucalyptus* essential oils. *Bioresource Technology*, v.100, p.6083-6087, 2009.

MAJER, V.; SVOBODA, M.; Pick, J. *Heats of Vaporizations of Fluids*. Amsterdam: Elsevier, 1989.

MARRERO, J.; GANI, R. Group-contribution based estimation of pure component properties. *Fluid Phase Equilibria*, v.183/184, p.183-208, 2001.

MARSLAND, S. *Machine learning: an algorithmic perspective*. 2 ed. Taylor and Francis, 2015.

MARTINS, G.A. *Estatística geral e aplicada*. 3 ed. São Paulo: Atlas, 2005. 424 p.

McMURRY, J. *Fundamentals of organic chemistry*. 7 ed. Belmont, USA: Brooks/Cole, 2011.

MICHIE, D.; SPIEGELHALTER, D.J.; TAYLOR, C.C. *Machine learning, neural and statistical classification*. 1994. Disponível em: <<https://www1.maths.leeds.ac.uk/~charles/statlog/whole.pdf>>. Acesso em: 19 jul. 2016.

MIJIN, D.Z.; ANTONOVIC, D.G. Temperature dependence of the Kovats retention indices for alkyl 1,3-diketones on a DB-5 capillary column. *Journal of the Serbian Chemical Society*, v.69, p.759-767, 2004.

MILLER, K.E.; BRUNO, T.J. Isothermal Kovats retention indices of sulfur compounds on a poly(5% diphenyl-95% dimethylsiloxane) stationary phase. *Journal of Chromatography A*, v.1007, p.117-125, 2003.

MINISTÉRIO DO MEIO AMBIENTE. Disponível: <<http://www.mma.gov.br>>. Acesso em: 14 set. 2016.

MKADDEM, M.; BOUJILA, J.; ENNAJAR, M.; LEBRIHI, A.; MATHIEU, F.; ROMDHANE, M. Chemical composition and antimicrobial and antioxidant activities of *Mentha (longifolia L. and viridis)* essential oils. *Journal of Food Science*, v.74, p.M358-M363, 2009.

MONTEIRO, O. S. *Caracterização do óleo essencial da pimenta dioica lindl e sua aplicação como atrativo de abelhas euglossina*. Departamento de Química, Universidade Federal da Paraíba, 2008. 148 p. Tese (Doutorado).

MORI, M.; RATH, N.; GOBBLE, C.; CHICKOS, J. Vaporization, sublimation enthalpy, and crystal structures of imidazo[1,2- α]pyrazine and phthalazine. *Journal of Chemical & Engineering Data*, v.61, p.370-379, 2016.

MOUCHREK FILHO, V. E. *Introdução à Química de Óleos Essenciais*. São Luís, 2001. 72p. Apostila.

MUN, S.P.; PREWITT, L. Antifungal activity of organic extracts from *Juniperus virginiana* heartwood against wood decay fungi. *Forest Products Journal*, v.61, p. 443-449, 2011.

MYRDAL, P.B.; KRZYZANIAK, J.F.; YALKOWSKY, S.H. Modified Trouton's Rule for Predicting the Entropy of Boiling. *Industrial & Engineering Chemistry Research*, v.35, p.1788-1792, 1996.

NARENDRA, K.S.; BALAKRISHNAN, J. Adaptive control using multiple models. *IEEE Transactions on Automatic Control*, v.42, p.171-187, 1987.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. Disponível em < <http://webbook.nist.gov/chemistry/name-ser.html> >. Acesso em: 10 set. 2016.

NIST/TDE Aspen® plus version 8.4

NAZARENKO, D.V.; KHARYUK, P.V.; OSELEDETS, I.V.; RODIN, I.A.; SHPIGUN, O.A. Machine learning for LC–MS medicinal plants identification, *Chemometrics and Intelligent Laboratory Systems*, v.156, p. 174-180, 2016.

NICHOLS, G.; KWESKIN, S.; FRERICKS, M.; REITER, S.; WANG, G.; ORF, J.; CARVALLO, B.; HILLESHEIM, D.; CHICKOS, J. Evaluation of the vaporization, fusion, and sublimation enthalpies of the 1-alkanols: The vaporization enthalpy of 1-, 6-, 7-, and 9-heptadecanol, 1-octadecanol, 1-eicosanol, 1-docosanol, 1-hexacosanol, and cholesterol at $T = 298.15$ K by correlation gas chromatography. *Journal of Chemical and Engineering Data*, v.51, p.475-482, 2006.

NILSSON, N.J. *Introduction to machine learning: an early draft of a proposed textbook*. Robotics laboratory, Department of computer Science. Stanford, 2005.

OKSER, S.; PAHIKKALA, T.; AIROLA, A.; SALAKOSKI, T.; RIPATTI, S.; AITTOKALLIO, T. Regularized Machine Learning in the Genetic Prediction of Complex Traits. *Plos Genetics*, v.10, p.1-9, 2014.

OLIVEIRA, C.E.L.; CREMASCO, M.A. Determination of the vapor pressure of *Lippia gracilis* Schum essential oil by thermogravimetric analysis. *Thermochimica Acta*, v.577, p.1-4, 2014.

OLIVEIRA, C.E.L. *Caracterização do óleo essencial de Lippia Gracilis schum e de seus principais constituintes por termogravimetria*. Faculdade de Engenharia Química, Universidade Estadual de Campinas, 2012. 115 p. Dissertação (Mestrado).

PANDEY, A.K.; RAI, M.K.; ACHARYA, D. Chemical Composition and antimycotic activity of the essential oils of corn mint (*Mentha arvensis*) and lemon grass (*Cymbopogon flexuosus*) against human pathogenic fungi. *Pharmaceutical Biology*, v.41, p.421-425, 2003.

PÉREZ-PARAJÓN, J.M.; SANTIUSTE, J.M.; TAKÁCS, J.M. Sensitivity of the methylbenzenes and chlorobenzenes retention index to column temperature, stationary phase polarity, and number and chemical nature of substituents. *Journal of Chromatography A*, v.1048, p.223-232, 2004.

POLING, B.E.; PLAUSNITZ, M.; O'CONNELL, P. *The Properties of Gases and Liquids*. 5 ed. New York: McGraw-Hill, 2001.

QUINTANILLA-LÓPEZ, J.E.; LEBRÓN-AGUILAR, R.; GARCÍA-DOMÍNGUEZ, J.A. Hold-up time in gas chromatography: V. Dependence of the retention of n-alkanes on the chromatographic variables in isothermal gas chromatography. *Journal of Chromatography A*, v.878, p.125-135, 2000.

RAWLINGS, J.O.; PANTULA, S.G.; DICKEY, D.A. *Applied regression analysis: a research tool*. 2.ed. Springer: 1998.

RENCHER, A. *Methods of multivariate analysis*. 7 ed. New York, USA: John Wiley and Sons, 2002.

RODILLA, J.M.; SILVA, L.A.; MARTINEZ, N.; LORENZO, D.; DAVYT, D.; CASTILLO, L.; GIMÉNEZ, C.; CABRERA, R.; GONZÁLEZ-COLOMA, A.; ZROSTLÍKOVÁ, J.; DELLACASSA, E. Advances in the identification and agrochemical importance of sesquiterpenoids from *Bulnesia sarmientoi* essential oil. *Industrial Crops and Products*, v.33, p.497-503, 2011.

RŮŽIČKA, K.; KOUTEK, B.; FULEM, M.; MICHAL HOSKOVEC, M. Indirect determination of vapor pressures by capillary gas-liquid chromatography: analysis of the reference vapor-pressure data and their treatment. *Journal of Chemical & Engineering Data*, v.57, p.1349-1368, 2012.

RŮŽIČKA, K.; MAJER, V. Simultaneous treatment of vapor pressures and related thermal data between the triple and normal boiling temperatures for n-alkanes C5 – C20. *Journal of Physical and Chemical Reference Data*, v.23, p.1-39, 1994.

SEJNOWSKI, T.J.; ROSENBERG, C.R. Parallel networks that learn to pronounce english text. *Complex Systems*, v.1, p.145-168, 1987.

SHALEV-SHWARTZ, S.; BEM-DAVID, S. *Understanding machine learning: from theory to algorithms*. New York, USA: Cambridge University Press, 2014. Disponível em: < <http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning> >. Acesso em: 09 jul. 2016.

SHEN, L.; ALEXANDER, K.S. A thermal analysis study of long chain fatty acids. *Thermochimica Acta*, v.340/341, p.271-278, 1999.

SHRIVASTAVA, V.M.; LONDHE, N.D.; SONAWANE, R.S.; SURI, J.S. A novel approach to multiclass psoriasis disease risk stratification: Machine learning paradigm, *Biomedical Signal Processing and Control*, v.28, p.27-40, 2016.

SILVA, M.A.V.R.; MONTE, M.J.S. The construction, testing and use of a new knudsen effusion apparatus. *Thermochimica Acta*, v.171, p.169-183, 1990.

SIMÕES, C.M.O.; SPITZER, V. Óleos voláteis. In: SIMÕES, C.M.O. *Farmacognosia: da planta ao medicamento*. 6. ed. Florianópolis, SC: 2007. 1102 p.

SISTEMA DE ANÁLISE DAS INFORMAÇÕES DE COMÉRCIO EXTERIOR (AliceWeb). Disponível: < <http://aliceweb.mdic.gov.br>>. Acesso em: 16 dez. 2016.

SMITH, J.M.; VAN NESS, H.C.; ABBOTT, M.M. *Introdução à termodinâmica da engenharia química*. 7 ed. Rio de Janeiro, Brasil: LTC, 2007.

SOUZA, S.A.M.; MEIRA, M.R.; FIGUEIREDO, L.S.; MARTINS, E.R. Óleos essenciais: aspectos econômicos e sustentáveis. *Enciclopédia Biosfera*, v.6, p.1-11, 2010.

SPEZIALI, M. G. De aromas e perfumes, o mercado da indústria do “cheiro”. *Química Nova*, v.35, p.861-864, 2012.

SPIEKSMAN, W.; LUIJK, R.; GOVERS, H.A.J. Determination of the liquid vapour pressure of low-volatility compounds from the Kováts retention index. *Journal of Chromatography A*, v.972, p.141-148, 1994.

ŠTEJFA, V.; FULEM, M.; RŮŽIČKA, K.; CERVINKA, M.; ROCHA, M.A.A.; SANTOS, L.M.N.B.F.; SCHRÖDER, B. Thermodynamic study of selected monoterpenes, *Journal of Chemical Thermodynamics*, v.60, p.117-125, 2013.

STEPHENSON, R. M.; MALANOWSKI, S. *Handbook of the Thermodynamics of Organic Compounds*. New York: Elsevier, 1987.

SVOBODA, V.C.; KOUTEK, B. Determination of saturated vapor pressures and heats of vaporization from gas chromatography retention data. *Chemické Listy*, v.96, p.10-18, 2002.

TAKÁCS, J.; ROCKENBAUER, M.; OLÁCSI, I. Determination of the relationship between retention index and column temperature in gas chromatography through the temperature-dependence of the net retention volume. *Journal of Chromatography A*, v.42, p.19-28, 1969.

TIMUNG, R.; BARIK, C.R.; PUROHIT, S.; GOUD, V.V. Composition and anti-bacterial activity analysis of citronella oil obtained by hydrodistillation: Process optimization study. *Industrial Crops and Products*, v.94, p.178-188, 2016.

TU, C.H.; LIU, C.P. Group-contribution estimation of the enthalpy of vaporization of organic compounds. *Fluid Phase Equilibria*, v.121, p.45-65, 1996.

TUDOR, E. Analysis of the equations for the temperature dependence of the retention index: I. Relation between equations. *Journal of Chromatography A*, v.858, p.65-78, 1999.

VAN HENKE, G.R. A modern vapor pressure apparatus based on the isoteniscope. *Journal of Chemical Education*, v. 69, p.681-683, 1992.

VAN ROON, A.; PARSONS, J.R.; GOVERS, H.A.J. Gas chromatographic determination of vapour pressure and related thermodynamic properties of monoterpenes and biogenically related compounds, *Journal of Chromatography A*, v.955, p.105-115, 2002.

VEREVKIN, S.P.; EMEL'YANENKO, V.N.; VARFOLOMEEV, M.A.; SOLOMONOV, B.N.; ZHERIKOVA, K.V. Vaporization enthalpies of a series of the fluoro and chloro-substituted methylbenzenes. *Fluid Phase Equilibria*, v.380, p.67-75, 2014.

VEZZANI, S.; MORETTI, P.; CASTELLO, G. Fast and accurate method for the automatic prediction of programmed-temperature retention times. *Journal of Chromatography A*, v.677, p.331-343, 1994.

VEZZANI, S.; PIERANI, D.; MORETTI, P.; CASTELLO, G. Automatic prediction of retention times in programmed-pressure isothermal gas chromatography. *Journal of Chromatography A*, v.848, p.229-238, 1999.

VLAMING, R.; GROENEN, P. J.F. The current and future use of ridge regression for prediction in quantitative genetics. *BioMed Research International*, v.2015, p.1-18, 2015.

VOGEL, A. I. *Qualitative Organic Analysis*, v.2, 2nd ed., Longman: London, 1966.

WANG, X.; HAO, Q.; CHEN, Y.; JIANG, S.; YANG, Q.; LI, Q. The effect of chemical composition and bioactivity of several essential oils on *Tenebrio molitor* (Coleoptera: Tenebrionidae). *Journal of Insect Science*, v.15, p.1-7, 2015.

WEISBERG, S. *Applied linear regression*. 3 ed. New Jersey, USA: John Wiley and Sons, 2005.

WHITTAKER, J.C., THOMPSON, R.; DENHAM, M.C. Marker-assisted selection using ridge regression. *Genetics Research Cambridge*, v.75, p.249-252, 2000.

WRIGHT, S. F.; DOLLIMORE, D.; DUNN, J.G.; ALEXANDER, K. Determination of the vapor pressure curves of adipic acid and triethanolamine using thermogravimetry analysis. *Thermochimica Acta*, v. 421, p. 25-30, 2004.

ZAITSAU, D.H.; VEREVKIN, S.P.; SAZONOVA, A.Y. Vapor pressures and vaporization enthalpies of 5-nonanone, linalool and 6-methyl-5-hepten-2-one. Data evaluation. *Fluid Phase Equilibria*, v.386, p.140-148, 2015.

ZARRAD, K.; HAMOUDA, A.B.; CHAIEB, I.; LAARIF, A.; JEMÂA, J.M.B. Chemical composition, fumigant and anti-acetylcholinesterase activity of the Tunisian *Citrus aurantium* L. essential oils. *Industrial Crops and Products*, v.76, p.121–127, 2015.

ZENG, Z.; YANG, Z.; XUE, W.; LI, X. Vapor pressure, vaporization enthalpy, standard enthalpy of formation and standard entropy of n-butyl carbamate. *Chinese Journal of Chemical Engineering*, v.22, p.1145-1152, 2014.

ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. *Journal of Royal Statistical Society B*, v.67 (parte 2), p.301-320, 2005.

Apêndice A

ESTRUTURAS QUÍMICAS

A.1 Compostos analisados neste trabalho

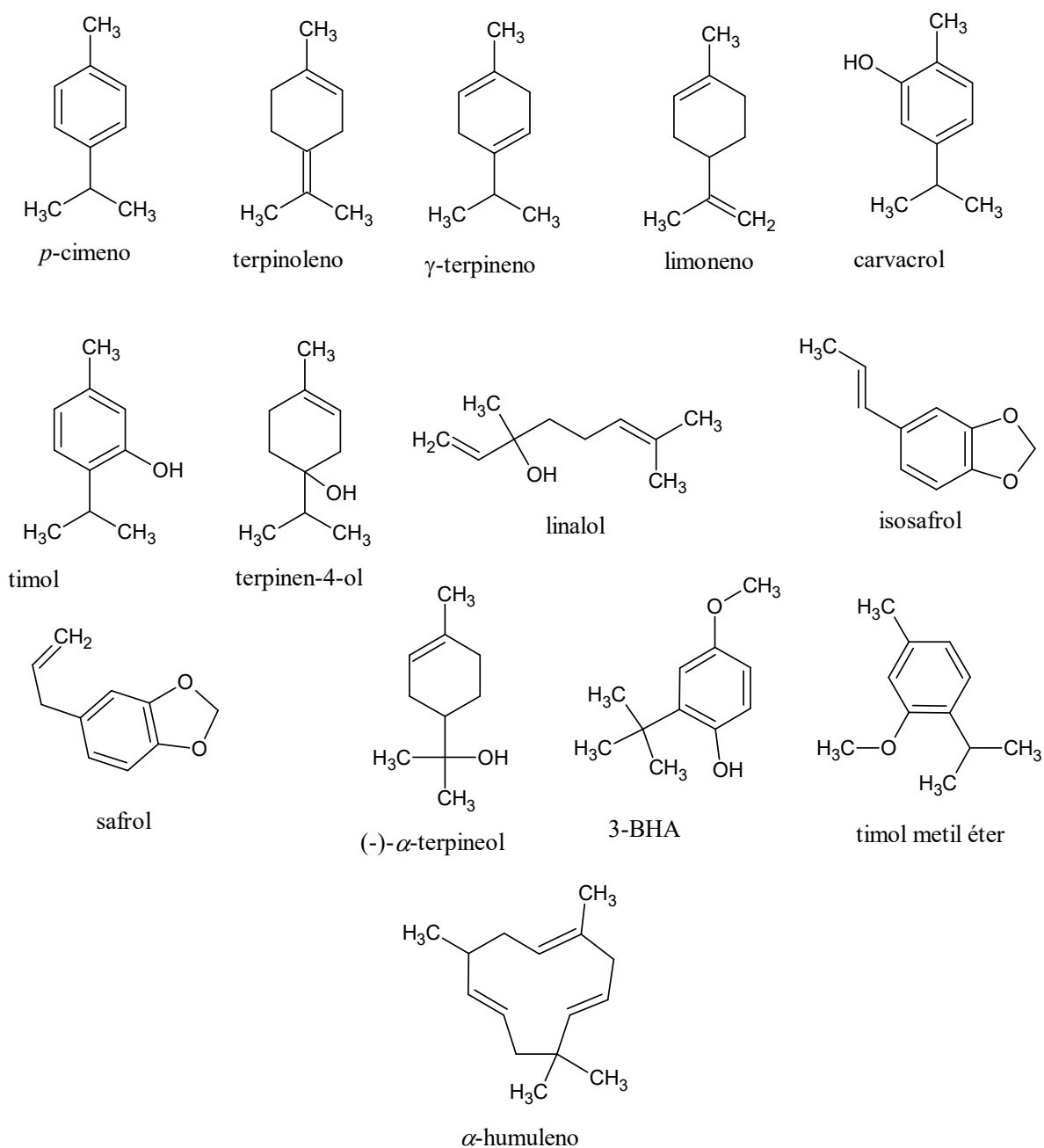


Figura A.1 – Estruturas químicas dos compostos monoterpênicos e sesquiterpênicos analisados por CG neste trabalho.

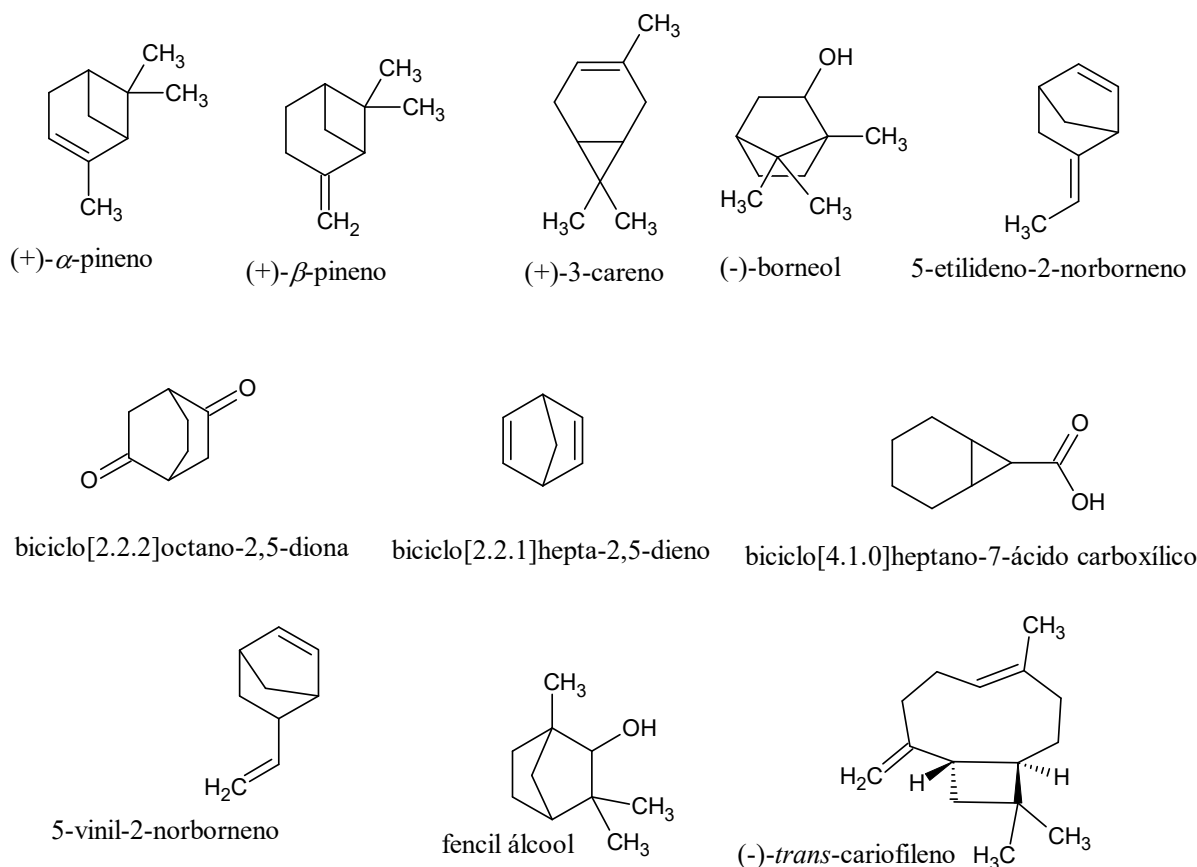


Figura A.2 – Estruturas químicas dos compostos monoterpênicos bicíclicos analisados por CG neste trabalho.

A.2 Compostos analisados da literatura

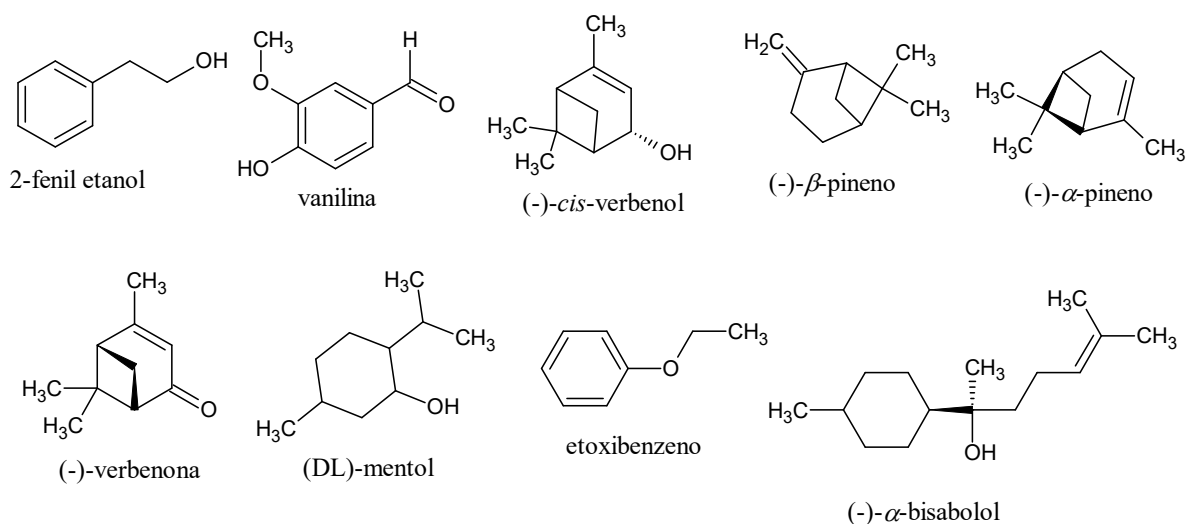


Figura A.3 – Estruturas químicas dos compostos terpênicos da literatura usados para predição de entalpias de vaporização.

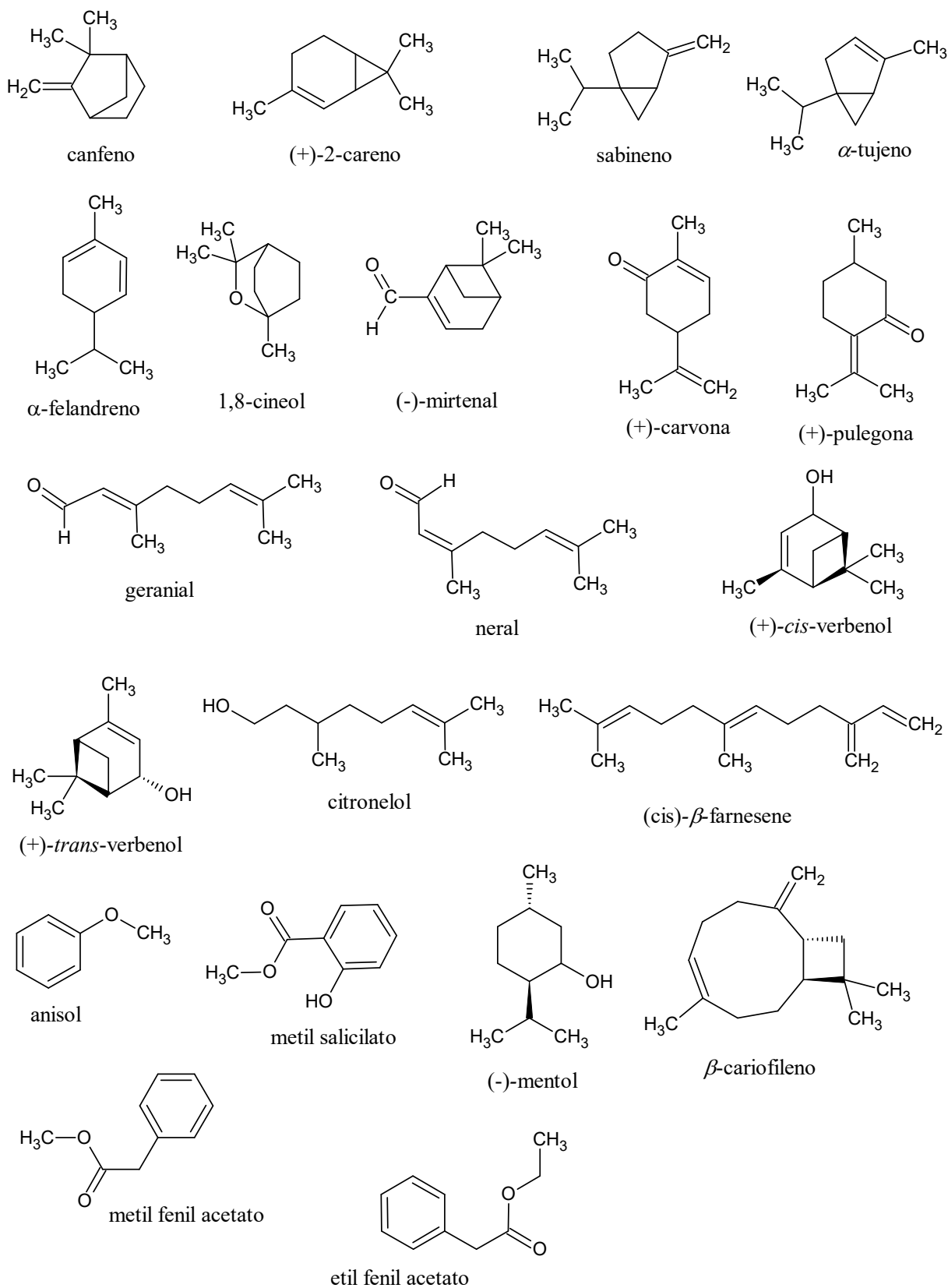


Figura A.3 – Continuação: Estruturas químicas dos compostos terpênicos da literatura usados para predição de entalpias de vaporização.

Apêndice B

PROPRIEDADES DOS *n*-ALCANOS

A Tabela B apresenta os dados dos *n*-alcanos utilizados neste trabalho, que auxiliaram nos cálculos das grandezas termodinâmicas estudadas.

Tabela B – Dados de pressão de vapor, número de ligação torsional efetivo, entropia de ebulição e temperatura normal de ebulição dos *n*-Alcanos de C₅ a C₁₆.

<i>n</i> -Alcano	$P_{vap,z}^{298,15K}$, Pa	$\ln P_{vap,z}^{298,15K}$, Pa	* $\tau_{efet.}$	ΔS_{vap}^b J K ⁻¹ mol ⁻¹	T_{nb} , K
pentano	68353,99	11,132	2	86,8	309,20
hexano	20180,29	9,912	3	87,2	341,90
heptano	6101,70	8,716	4	87,6	371,50
octano	1871,80	7,535	5	88,0	398,70
nonano	580,70	6,364	6	88,4	423,80
decano	182,00	5,204	7	88,8	447,20
undecano	56,89	4,041	8	89,2	468,00
dodecano	18,02	2,891	9	89,6	489,00
tridecano	5,68	1,737	10	90,0	507,00
tetradecano	1,80	0,588	11	90,4	523,00
pentadecano	0,58	-0,545	12	90,8	540,00
hexadecano	0,19	-1,661	13	91,2	554,00

*adimensional

Apêndice C

CLASSE DOS COMPOSTOS

A Tabela C apresenta a classe de todos os compostos incluídos no banco de dados usado para a estimação dos parâmetros do modelo de contribuição de grupos proposto neste trabalho.

Tabela C – Quantidade de compostos por classe incluídos no Banco de Dados.

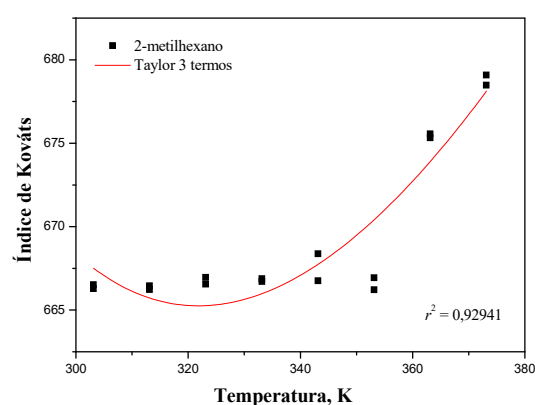
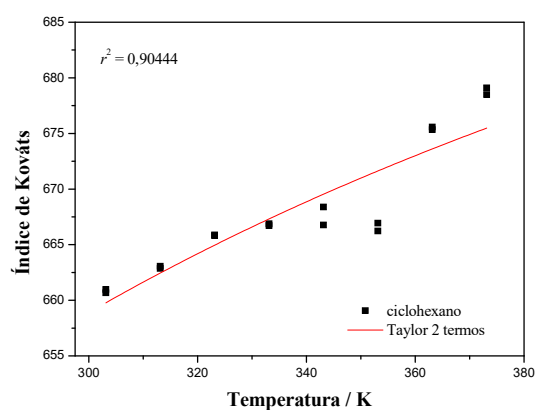
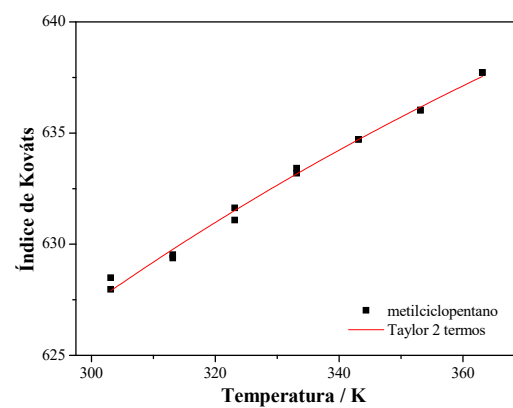
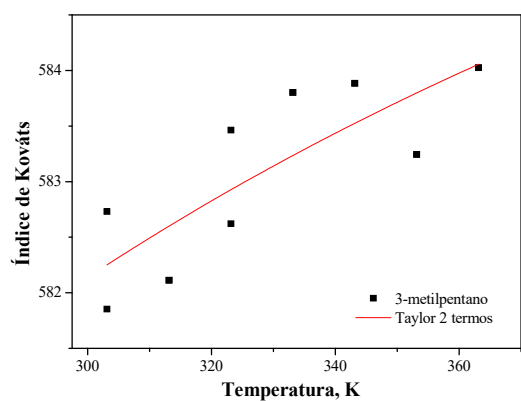
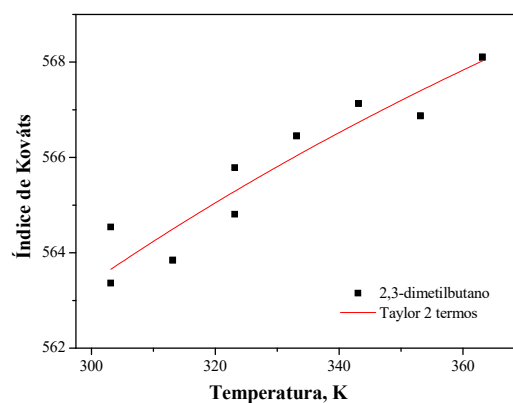
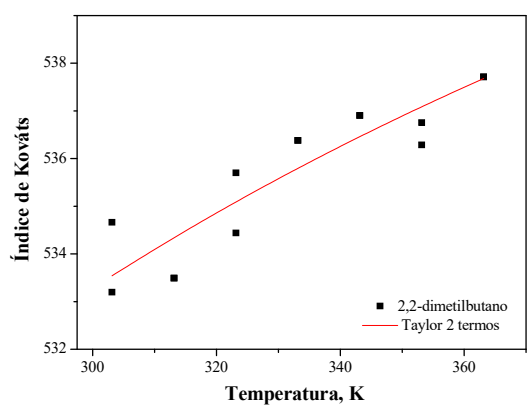
Classe	Número de compostos
<i>Hidrocarbonetos</i>	
Saturados e insaturados	344
Aromáticos	121
Cíclicos	110
Bicíclicos	48
<i>Cetonas</i>	
Saturadas e insaturadas	47
Aromáticas	16
Cíclicas	26
Bicíclicas	12
<i>Ácidos Carboxílicos</i>	
Saturados e insaturados	18
Aromáticos	3
Bicíclicos	1
<i>Aldeídos</i>	
Saturados e insaturados	81
Aromáticos	8
Aldeído c/ bicíclico	1
<i>Éteres</i>	
Saturados e insaturados	48
Aromáticos	32
Cíclicos	9
Éter c/ bicíclico	1
<i>Ésteres</i>	
Saturados e insaturados	157
Aromáticos	24
Cíclicos	25

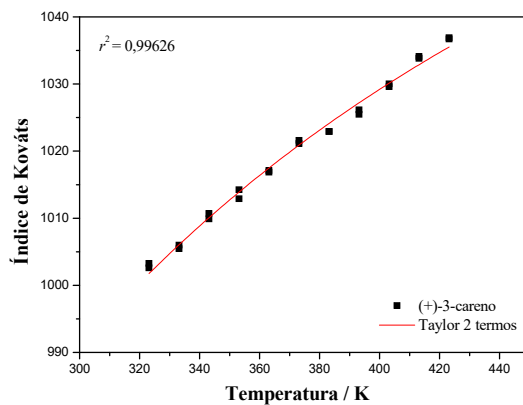
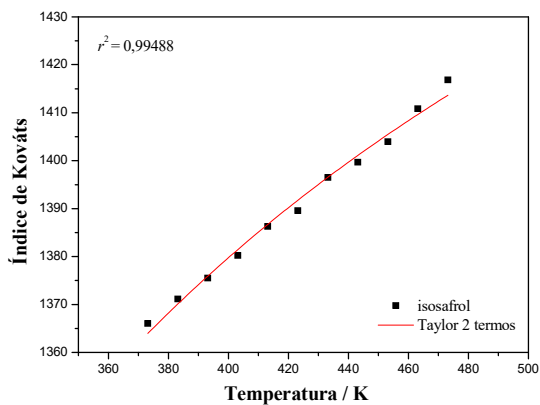
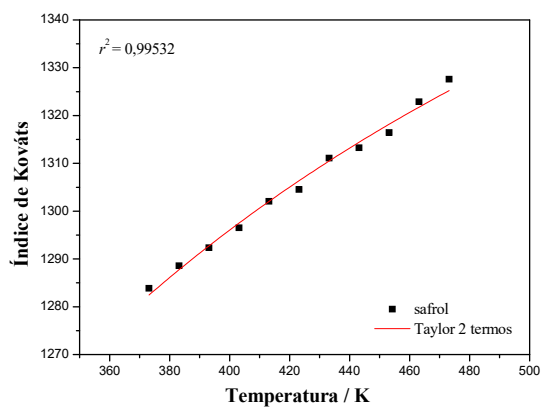
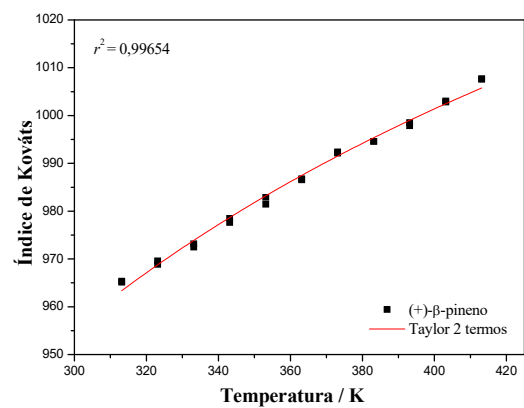
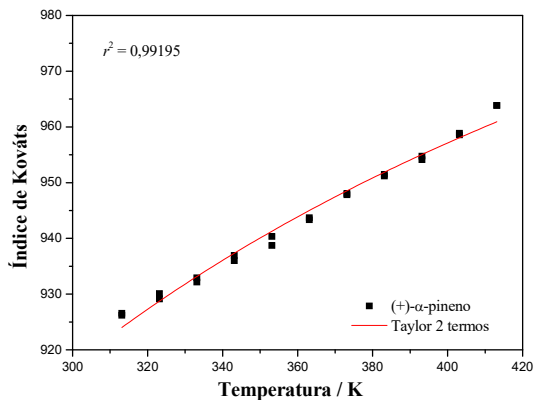
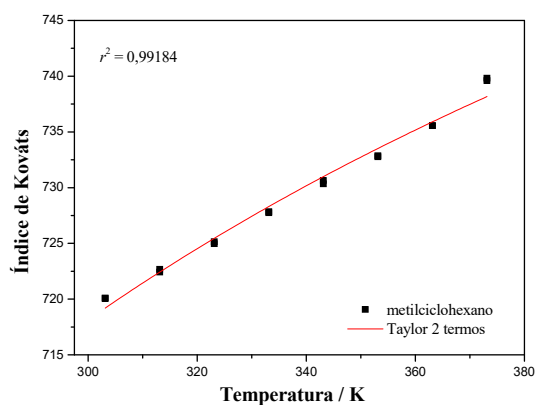
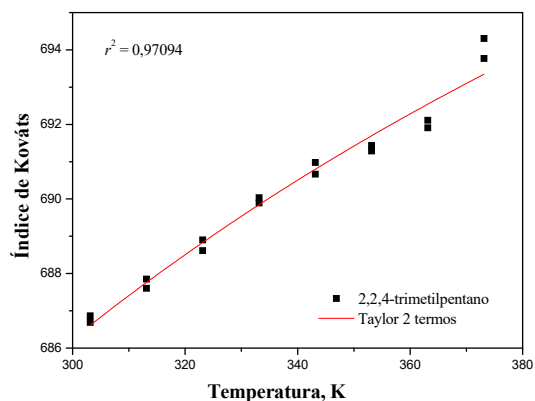
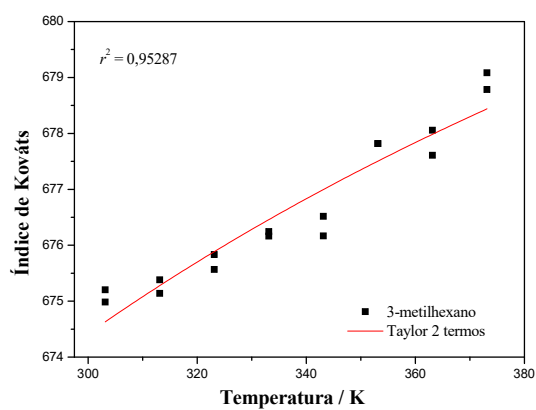
Tabela C - continuação

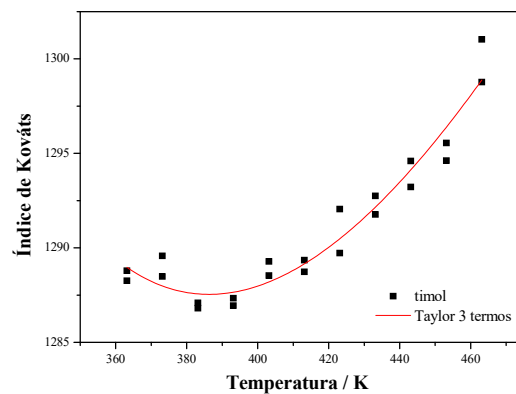
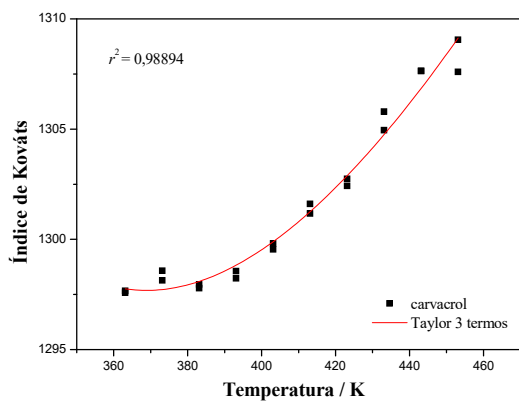
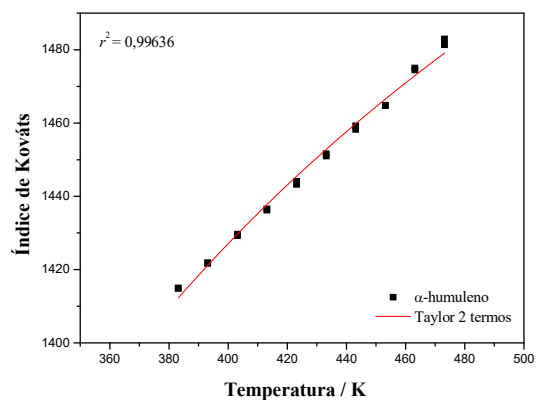
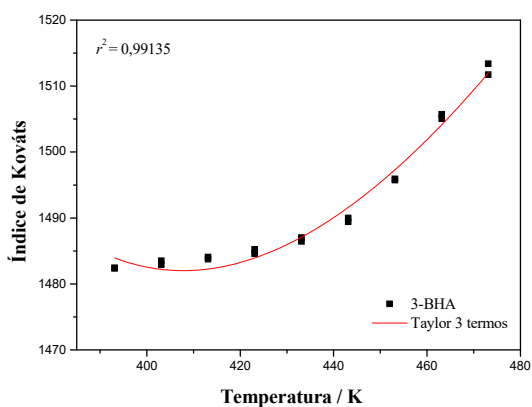
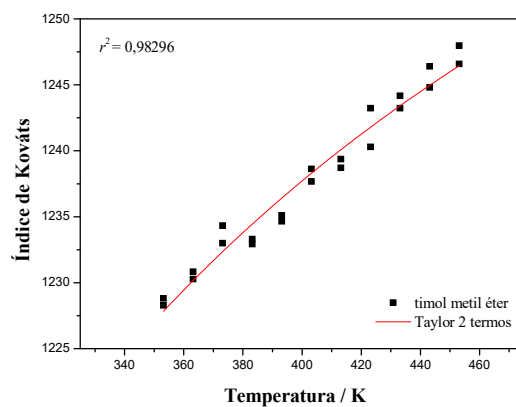
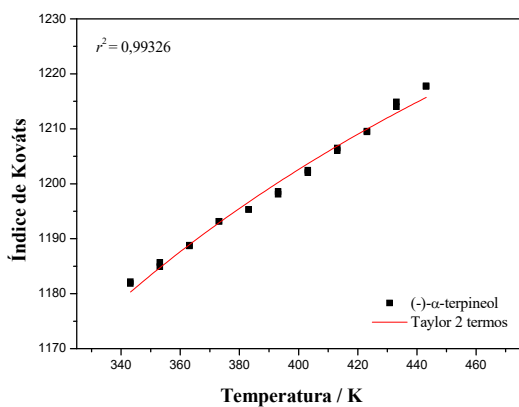
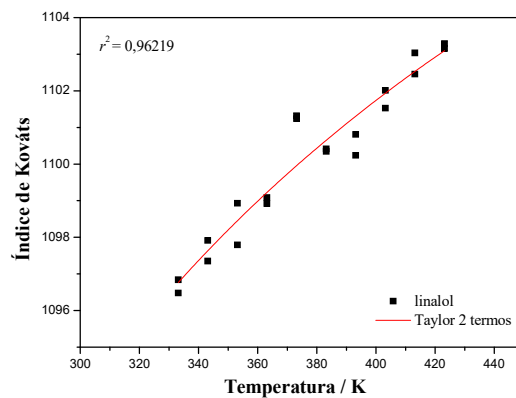
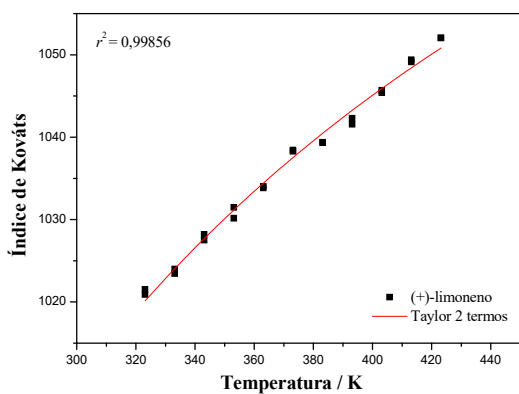
Classe	Número de compostos
<i>Álcoois</i>	
saturados e insaturados	191
cíclicos	30
diálcoois	12
álcool c/ bicíclico	2
<i>Fenóis</i>	
fenóis/di-fenóis	101
fenol c/ bicíclico	2
<i>Anel Aromático Dissubstituído</i>	
éter + álcool	7
éster + fenol	10
éter + aldeído	2
éter + ácido carboxílico	3
fenol + éter	10
fenol + cetona	1
<i>Anel Aromático Trissubstituído</i>	
difenol + éter	2
fenol + éter + aldeído	1
<i>Compostos Multifuncionais</i>	
diácido carboxílico	3
aldeído + cetona cíclica	1
aldeído + álcool	1
dialdeído	1
cetona cíclica + álcool cíclico	2
dicetona	5
éster + cetona cíclica	2
éster cíclico + cetona cíclica	1
éster + cetona	15
éster + éter	21
éster + álcool	15
éster + aldeído	1
diéster	52
triéster	6
éter + ác. carboxílico	1
éter + cetona	6
éter cíclico + cetona	10
éter cíclico + grupo peroxide (-o-o-)	1
diéter	37
diéter + álcool	5
diéter + cicíclico	4
triéter	11
outros	12

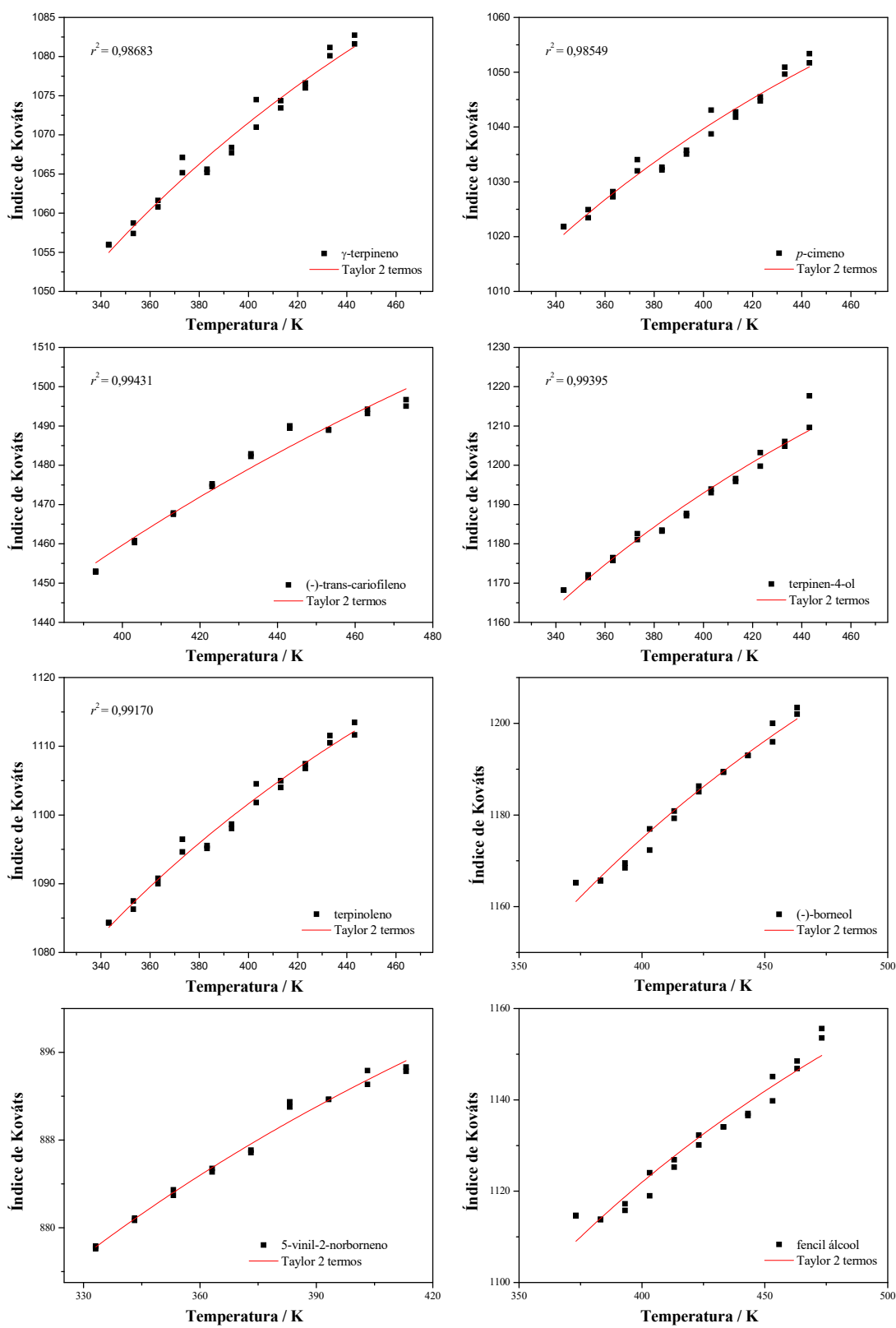
Apêndice D

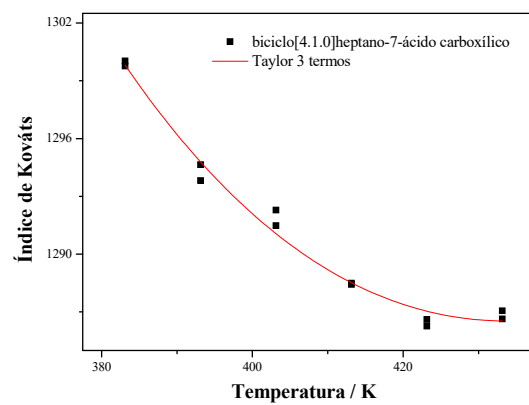
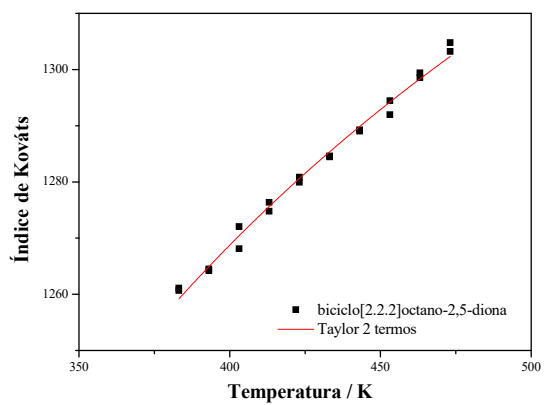
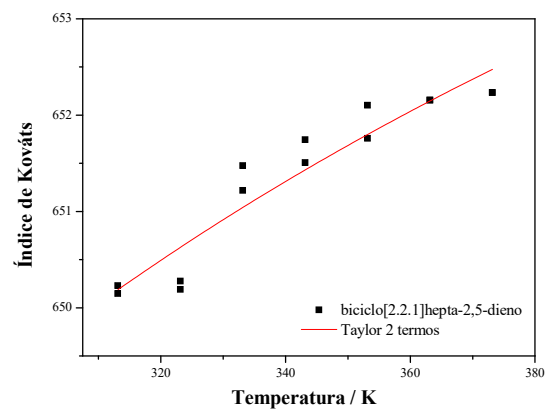
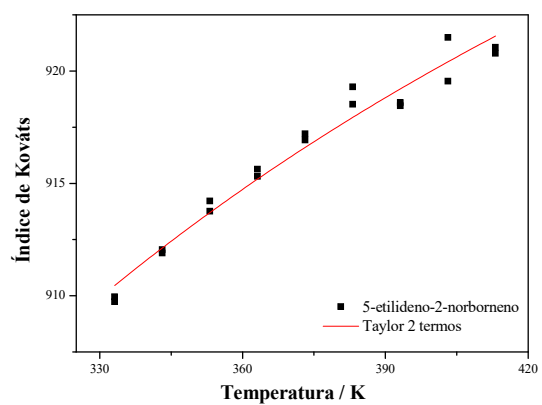
AJUSTES I vs. T





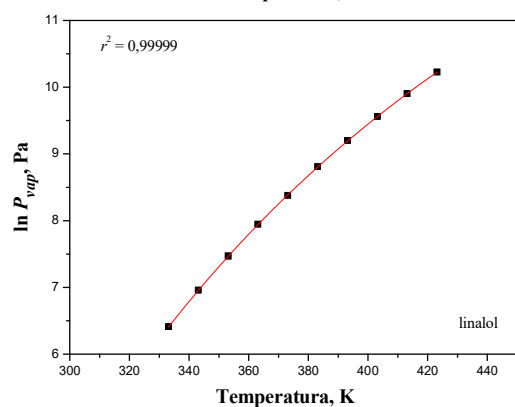
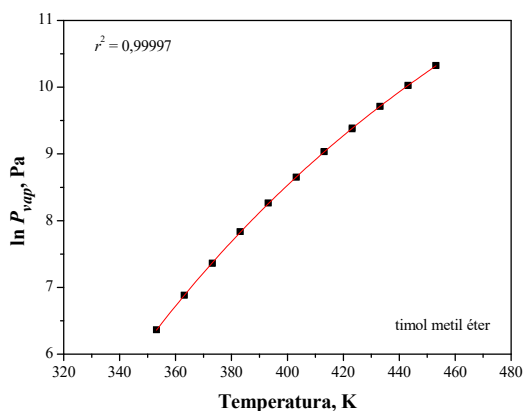
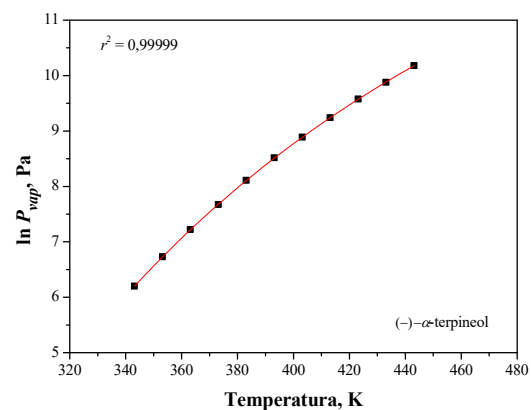
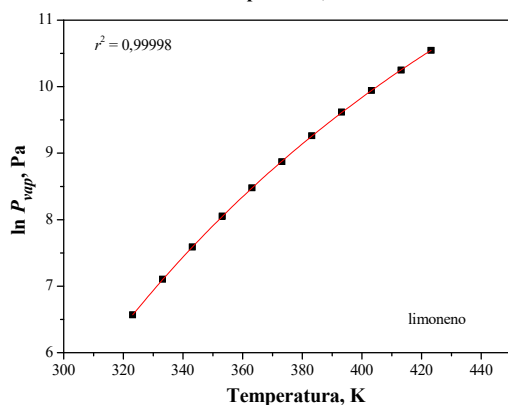
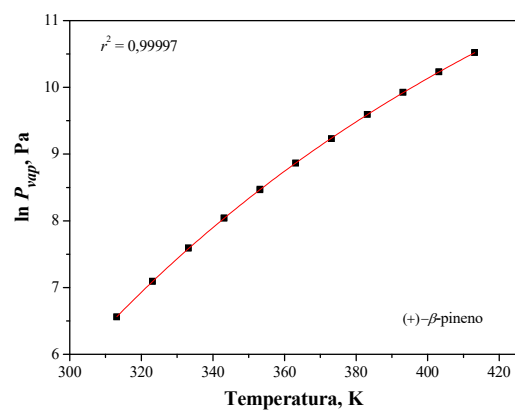
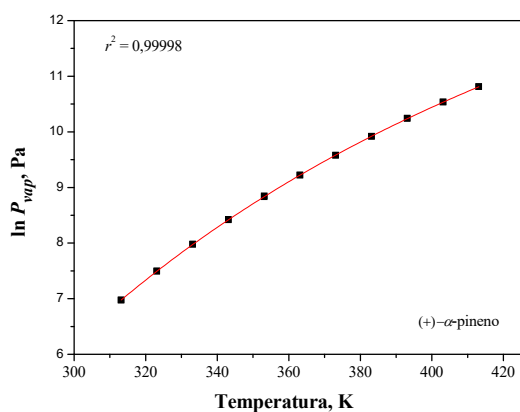


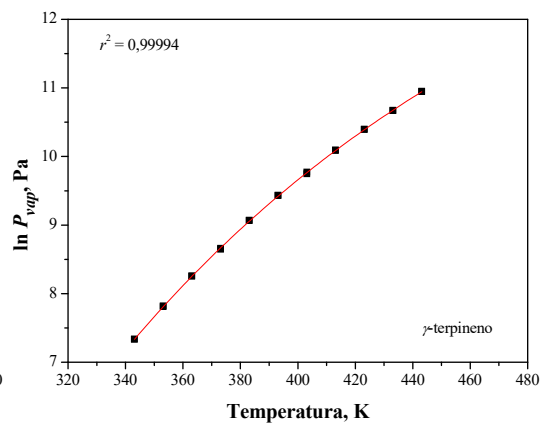
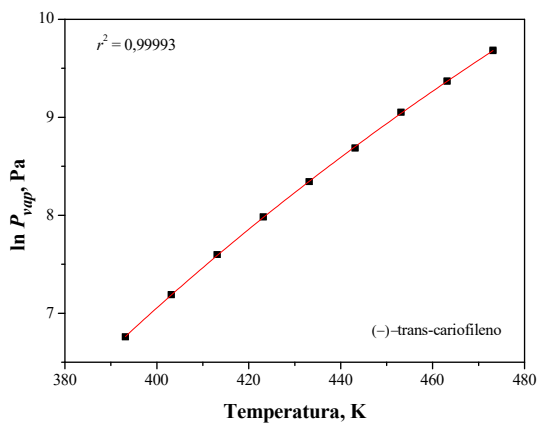
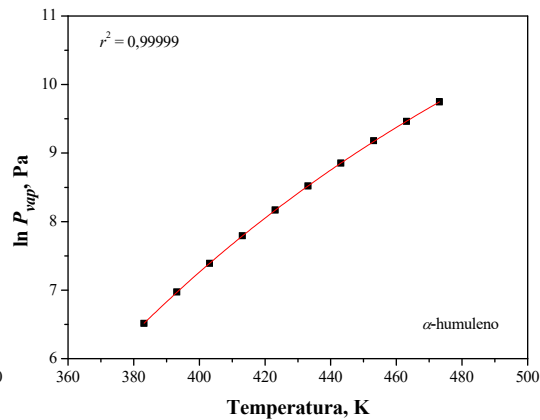
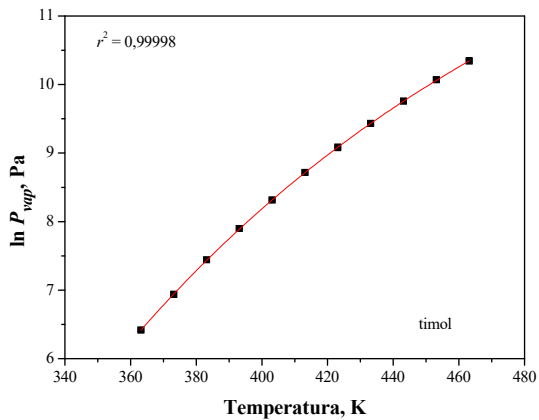
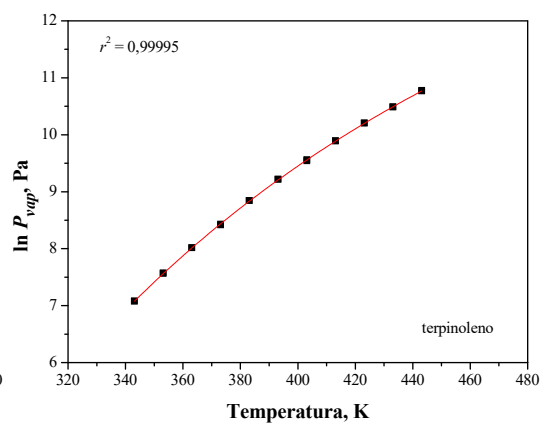
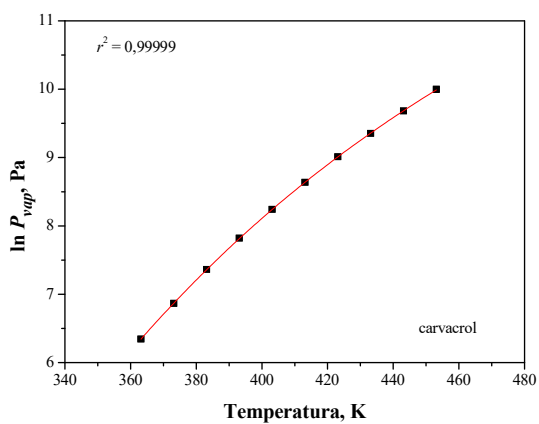
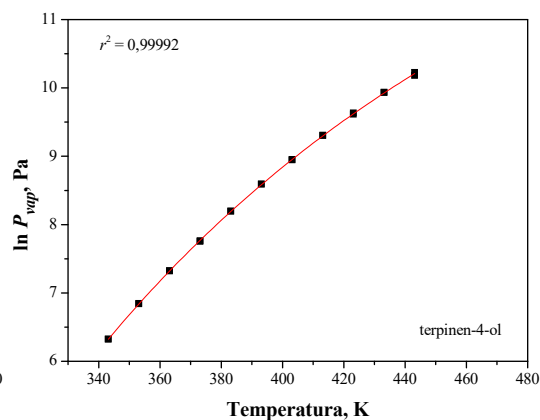
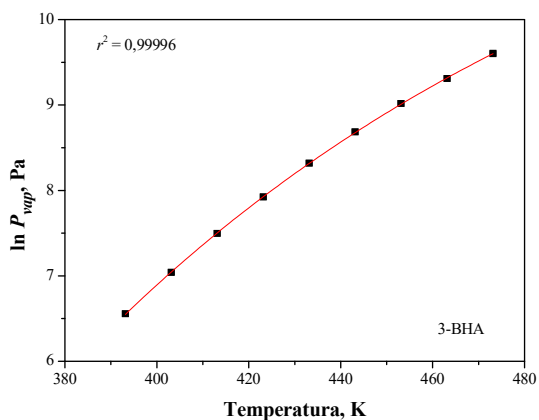


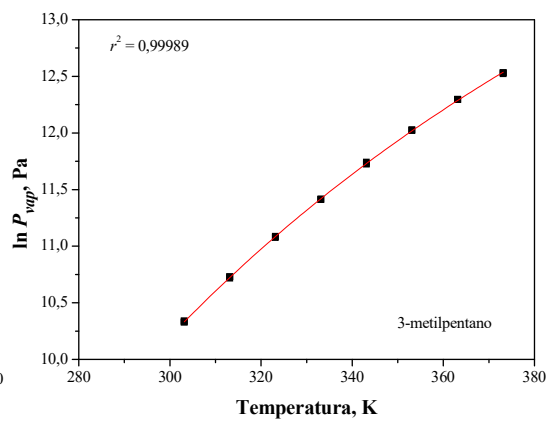
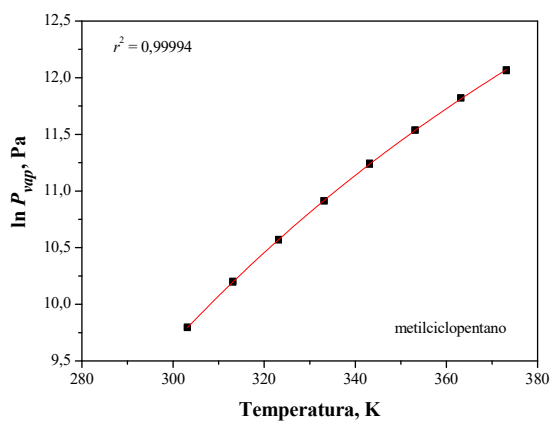
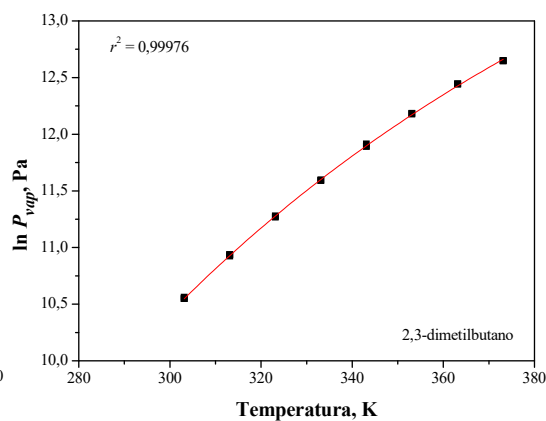
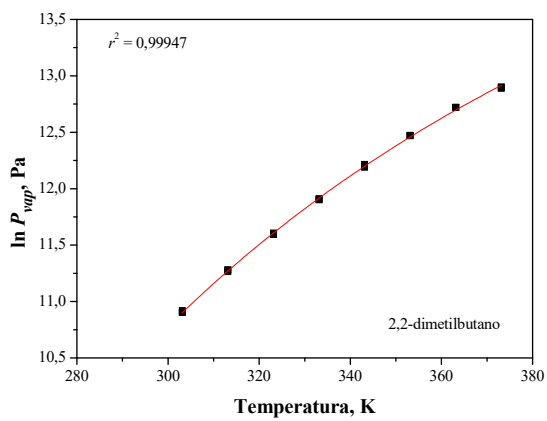
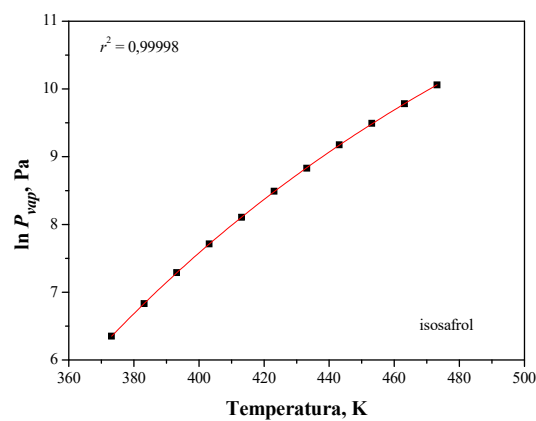
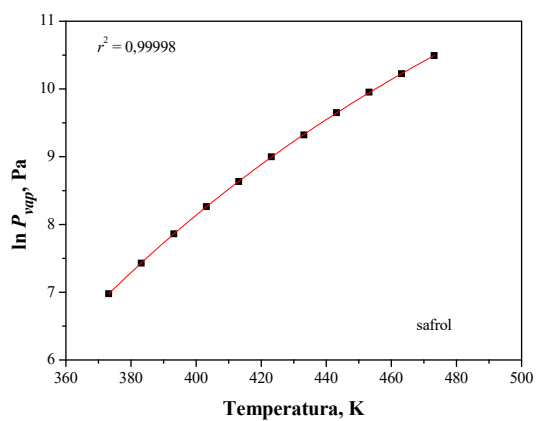
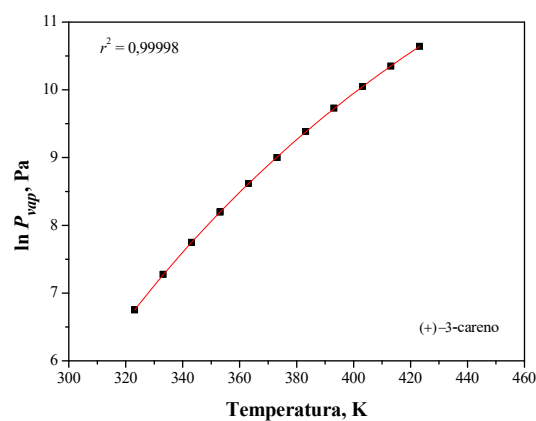
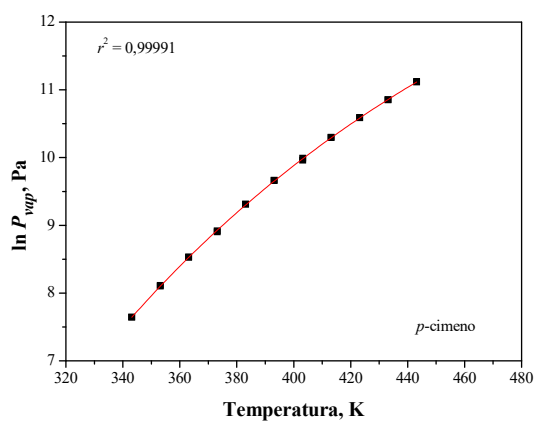


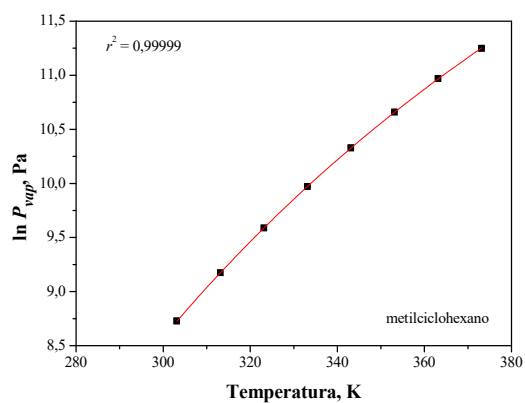
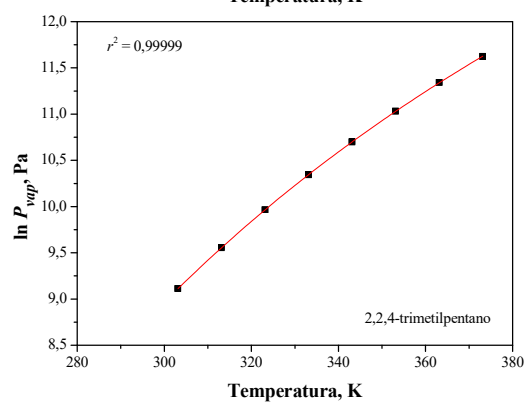
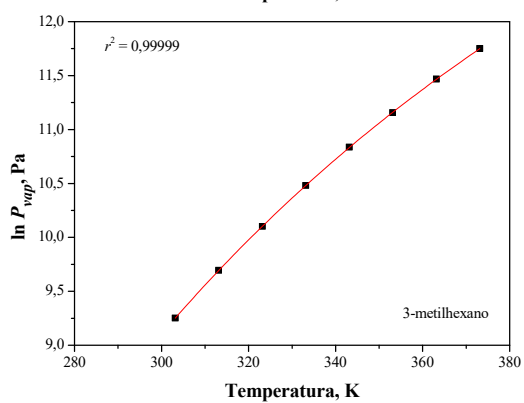
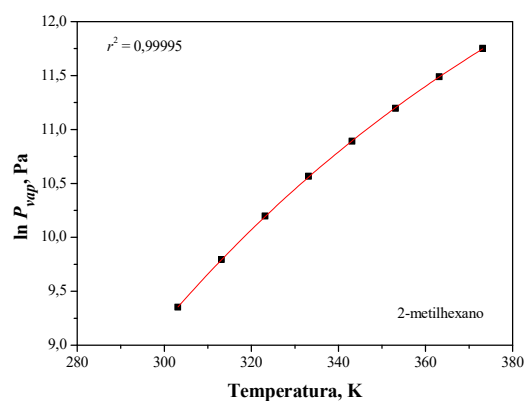
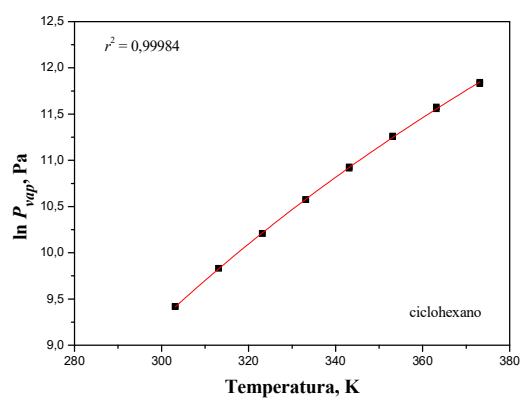
Apêndice E

AJUSTES DA EQUAÇÃO DE KIRCHHOFF-RANKIN





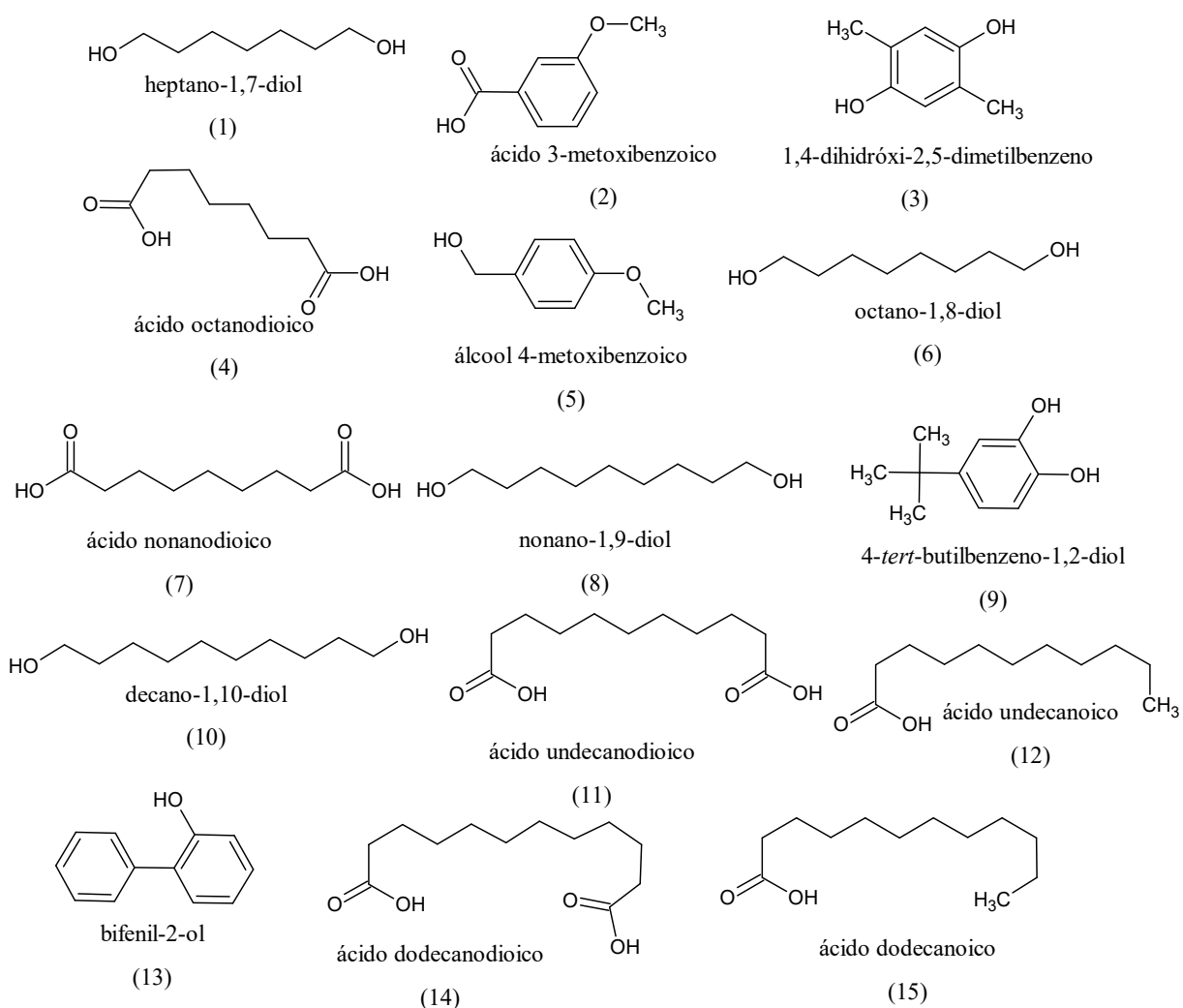


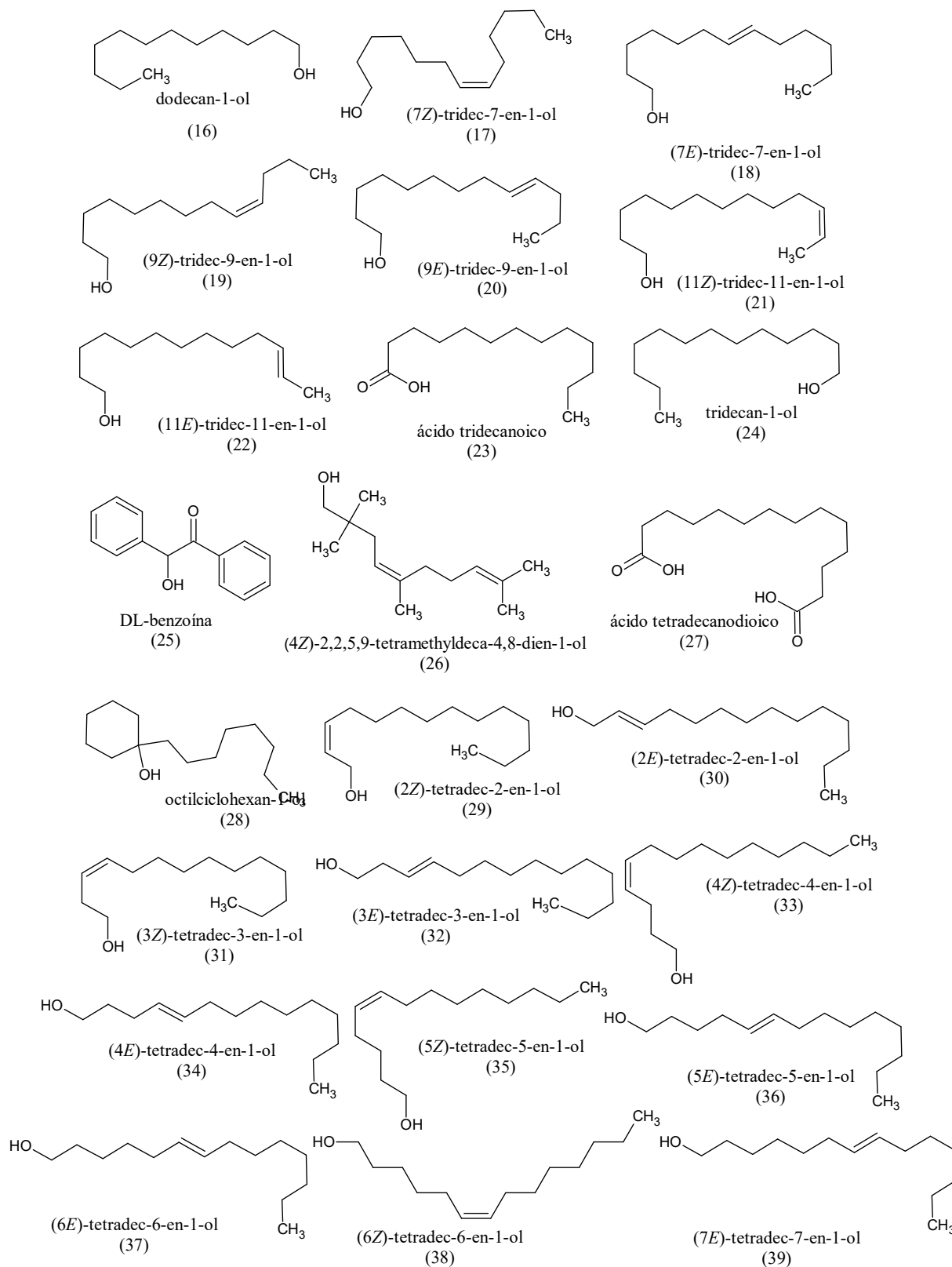


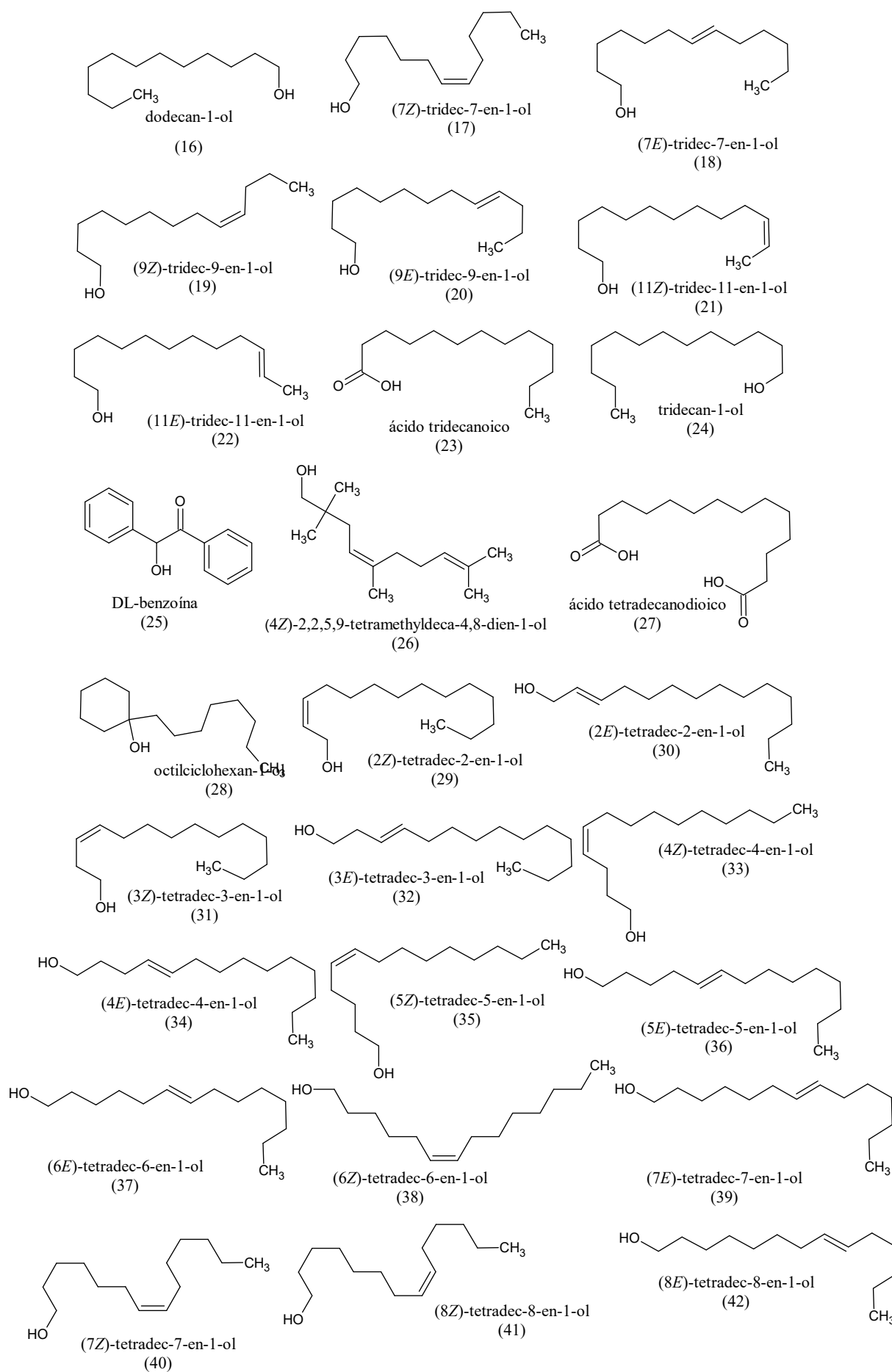
Apêndice F

COMPOSTOS *OUTLIERS*

A seguir estão apresentadas as estruturas químicas dos compostos que se apresentaram na região dos *outliers*, conforme o gráfico de caixa da Figura 4.17. Os dados de ΔH_{vap} de cada composto estão mostrados na Tabela F, conforme a ordem numérica.







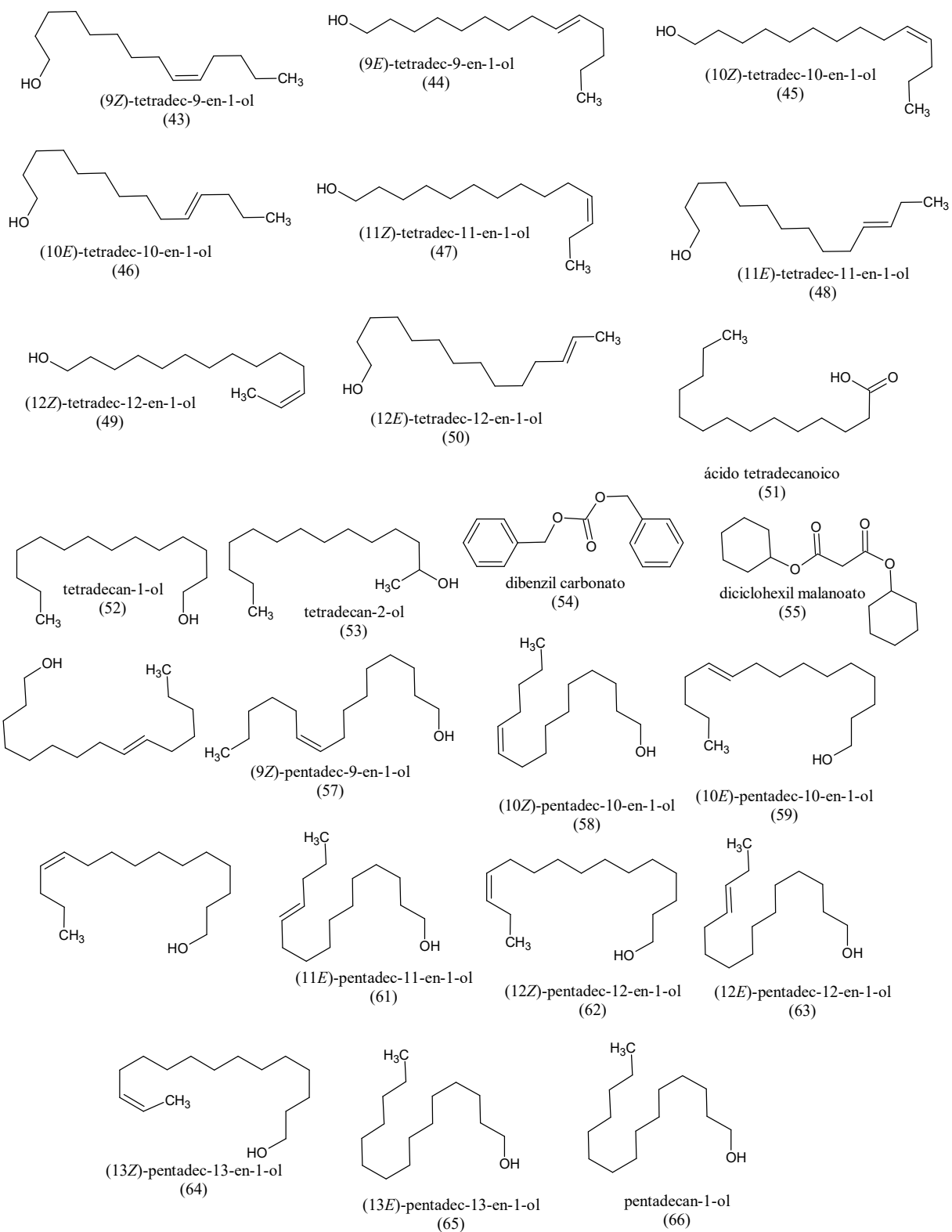


Tabela F - Compostos com dados de entalpia de vaporização na região dos *outliers*.

Nº	Composto	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K		
1	heptane-1,7-diol	96,50	298,00		
		93,80	323,00		
		96,20	298,00		
2	ácido 3-metoxibenzoico	95,00	298,00		
		3	1,4-dihidroxi-2,5-dimetilbenzeno	101,10	346,00
				116,70	298,00
4	ácido octanodioico	95,60	409,00		
		5	álcool 4-metoxibenzoico	101,00	356,00
				107,00	298,00
6	octano-1,8-diol	105,40	298,00		
		119,70	298,00		
		104,40	360,00		
7	ácido nonadioco	111,40	298,00		
		110,00	ND		
		112,80	298,00		
8	nonano-1,9-diol	96,50	298,00		
		9	4-tert-butilbenzeno-1,2-diol	126,60	298,00
				112,40	364,00
10	decano-1,10-diol	120,40	298,00		
		113,70	298,00		
		128,20	298,00		
11	ácido undecanodioico	97,90	305,00		
		12	ácido undecanoico	94,20	449,00
				130,00	298,00
13	ácido dodecanodioico	95,80	332,00		
		14	ácido dodecanoico	95,40	305,00
				95,40	305,00
15	dodecan-1-ol	95,10	298,00		
		16	(7Z)-tridec-7-en-1-ol	95,60	298,00
				(7E)-tridec-7-en-1-ol	95,60
17	(9Z)-tridec-9-en-1-ol	95,80	298,00		
		18	(9E)-tridec-9-en-1-ol	96,40	298,00
				96,40	298,00
19	(11Z)-tridec-11-en-1-ol	97,10	298,00		
		20	(11E)-tridec-11-en-1-ol	97,20	298,00
				97,20	298,00
21	ácido tridecanoico	100,40	340,00		
		22	tridecan-1-ol	94,70	298,00
				95,80	298,00
23	DL-benzoína	98,50	298,00		
		24	(4Z)-2,2,5,9-tetramethyldeca-4,8-dien-1-ol	94,00	378,00
				127,40	298,00
25	ácido tetradecanodioico	105,60	388,00		
		26	octilciclohexan-1-ol	101,10	298,00
				(2Z)-tetradec-2-en-1-ol	101,10
27	(2E)-tetradec-2-en-1-ol	101,50	298,00		
		28	(3Z)-tetradec-3-en-1-ol	99,80	298,00
				(3E)-tetradec-3-en-1-ol	99,70
29	DL-benzoína	98,50	298,00		
		30	(4Z)-2,2,5,9-tetramethyldeca-4,8-dien-1-ol	94,00	378,00
				127,40	298,00
31	ácido tetradecanodioico	105,60	388,00		
		32	octilciclohexan-1-ol	101,10	298,00
				(2Z)-tetradec-2-en-1-ol	101,10
32	(2E)-tetradec-2-en-1-ol	101,50	298,00		
		33	(3Z)-tetradec-3-en-1-ol	99,80	298,00
				(3E)-tetradec-3-en-1-ol	99,70

Tabela G – continuação

Nº	Composto	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K
33	(4Z)-tetradec-4-en-1-ol	100,00	298,00
34	(4E)-tetradec-4-en-1-ol	100,70	298,00
35	(5Z)-tetradec-5-en-1-ol	100,30	298,00
36	(5E)-tetradec-5-en-1-ol	100,80	298,00
37	6E)-tetradec-6-en-1-ol	100,00	298,00
38	(6Z)-tetradec-6-en-1-ol	100,50	298,00
39	(7E)-tetradec-7-en-1-ol	99,90	298,00
40	(7Z)-tetradec-7-en-1-ol	100,50	298,00
41	(8Z)-tetradec-8-en-1-ol	100,30	298,00
42	(8E)-tetradec-8-en-1-ol	101,40	298,00
43	(9Z)-tetradec-9-en-1-ol	100,60	298,00
44	(9E)-tetradec-9-en-1-ol	101,00	298,00
45	(10Z)-tetradec-10-en-1-ol	101,10	298,00
46	(10E)-tetradec-10-en-1-ol	101,50	298,00
47	(11Z)-tetradec-11-en-1-ol	101,70	298,00
48	(11E)-tetradec-11-en-1-ol	101,80	298,00
49	(12Z)-tetradec-12-en-1-ol	102,50	298,00
50	(12E)-tetradec-12-en-1-ol	102,50	298,00
51	ácido tetradecanoico	100,40	398,00
		104,10	349,00
52	tetradecan-1-ol	98,90	298,00
52		93,60	328,00
52		98,70	298,00
52		109,00	332,00
52		102,20	298,00
52		106,40	328,00
52		104,20	320,00
53	tetradecan-2-ol	95,70	328,00
54	dibenzil carbonato	96,70	298,00
55	diciclohexil malanoato	93,70	298,00
56	(9E)-pentadec-9-en-1-ol	105,30	298,00
57	(9Z)-pentadec-9-en-1-ol	105,90	298,00
58	(10Z)-pentadec-10-en-1-ol	105,90	298,00
59	(10E)-pentadec-10-en-1-ol	106,20	298,00
60	(11Z)-pentadec-11-en-1-ol	106,30	298,00
61	(11E)-pentadec-11-en-1-ol	106,50	298,00
62	(12Z)-pentadec-12-en-1-ol	106,70	298,00
63	(12E)-pentadec-12-en-1-ol	107,00	298,00
64	(13Z)-pentadec-13-en-1-ol	107,70	298,00
65	(13E)-pentadec-13-en-1-ol	107,70	298,00
66	pentadecan-1-ol	103,50	298,00
		95,50	339,00
		102,50	298,00
		107,20	298,00

Apêndice G

DADOS TERMODINÂMICOS PARA OS HIDROCARBONETOS ALIFÁTICOS

A seguir são apresentados os dados de pressão de vapor (P_{vap}), entalpia de vaporização (ΔH_{vap}) e temperatura normal de ebulição (T_{nb}) dos 7 hidrocarbonetos alifáticos estudados, retirados da base de dados do *National Institute of Standards and Technology* (NIST) *ThermoData Engine* (TDE), disponível no programa Aspen® V8.4. Os dados considerados no cálculo do valor médio, bem como a obtenção da incerteza associada foram definidos seguindo os procedimentos encontrados na plataforma NIST, da seguinte forma:

- 1) São calculadas estimativas da média e do desvio padrão sobre todos os dados encontrados;
- 2) São excluídos os dados superiores a três desvios padrão em relação à média;
- 3) Em havendo dados excluídos, o processo repetido para os dados retidos.

A incerteza corresponde a 1,96 vezes a estimativa final do desvio padrão.

2,2-dimetilbutano

Nº	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	42,57	298,14	27,69	298,09	322,88
2	42,56	298,14	27,69	298,09	
3	42,58	298,14	27,69	298,09	
4	42,55	298,15	27,83	298,19	
5	42,47	298,15	*27,64	298,15	
6	*42,56	298,15			
Média	42,55 ± 0,08	298,15 ± 0,03	27,71 ± 0,14	298,12 ± 0,09	

* valor predito NIST/TDE Apens® plus Versão 8.4

2,3-dimetilbutano

Nº	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	31,29	298,14	29,12	298,09	331,12
2	31,32	298,14	29,23	298,19	
3	31,28	298,14	*29,13	298,15	
4	31,33	298,15			
5	*31,30	298,15			
Média	31,30 ± 0,02	298,14 ± 0,01	29,16 ± 0,12	298,14 ± 0,10	

* valor predito NIST/TDE Apens® plus Versão 8.4

3-metilpentano

Nº	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	25,31	298,09	30,27	298,09	336,40
2	25,49	298,14	30,19	298,14	
3	25,35	298,14	30,29	298,14	
4	25,39	298,14	30,26	298,14	
5	*25,33	298,19	30,35	298,19	
6			*30,26	298,15	
Média	25,37 ± 0,14	298,14 ± 0,07	30,27 ± 0,10	298,14 ± 0,06	

* valor predito NIST/TDE Apens® plus Versão 8.4

metilciclopentano

Nº	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	18,31	298,14	31,63	298,09	344,94
2	*18,35	298,15	31,68	298,19	
3			*31,62	298,15	
Média	18,33 ± 0,05	298,15 ± 0,01	31,64 ± 0,07	298,14 ± 0,10	

* valor predito NIST/TDE Apens® plus Versão 8.4

ciclohexano

Nº	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	13,01	298,14	33,33	298,14	353,85
2	13,00	298,14	33,04	298,09	
3	13,04	298,09	33,01	298,09	
4	13,04	298,14	33,00	298,14	
5	13,05	298,14	32,97	298,09	
6	13,04	298,14	32,89	298,14	
7	13,01	298,14	32,89	298,14	
8	13,07	298,14	32,95	298,14	
9	12,96	298,14	33,02	298,19	
10	12,96	298,14	33,09	298,14	
11	13,00	298,14	32,79	298,14	
12	13,00	298,09	32,96	298,00	
13	13,00	298,14	33,04	298,15	
14	13,02	298,14	*33,33	298,14	
15	13,00	298,14			
16	13,01	298,14			
17	13,00	298,14			
18	13,02	298,14			
19	13,03	298,14			
20	13,00	298,14			
21	13,02	298,15			
22	13,01	298,14			
23	*13,01	298,14			
Média	13,00 ± 0,05	298,14 ± 0,03	33,00 ± 0,25	298,12 ± 0,09	

* valor predito NIST/TDE Apens® plus Versão 8.4

2-metilhexano

N°	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	8,78	298,14	34,85	298,14	363,10
2	*8,79	298,15	*34,83	298,15	
Média	8,78 ± 0,01	298,15 ± 0,01	34,84 ± 0,03	298,15 ± 0,01	

* valor predito NIST/TDE Apens® plus Versão 8.4

3-metilhexano

N°	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	7,70	296,80	35,05	298,14	364,89
2	*8,21	298,15	*35,04	298,15	
Média	7,96 ± 0,71	297,48 ± 1,87	35,04 ± 0,02	298,15 ± 0,01	

* valor predito NIST/TDE Apens® plus Versão 8.4

2,2,4-trimetilpentano

N°	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	6,67	298,43	35,13	298,09	372,36
2	6,58	298,14	35,13	298,14	
3	6,58	298,14	35,15	298,19	
4	6,60	298,14	35,19	298,14	
5*	6,58	298,15			
Média	6,60 ± 0,07	298,20 ± 0,25	35,15 ± 0,06	298,14 ± 0,08	

* valor predito NIST/TDE Apens® plus Versão 8.4

metilciclohexano

N°	P_{vap} , kPa	Temperatura, K	ΔH_{vap} , kJ mol ⁻¹	Temperatura, K	T_{nb} , K
1	6,20	298,14	35,36	298,09	374,04
2	6,17	298,09	35,29	298,09	
3	6,18	298,14	35,36	298,14	
4	*6,18	298,15	35,27	298,09	
5			35,40	298,14	
6			35,38	298,14	
7			35,35	298,19	
8			*35,40	298,15	
Média	6,18 ± 0,02	298,13 ± 0,05	35,35 ± 0,09	298,13 ± 0,07	

* valor predito NIST/TDE Apens® plus Versão 8.4

Apêndice H

ALGORITMOS USADO PARA ESTIMAÇÃO DOS PARÂMETROS DO MODELO

 Algoritmo 1

```

w(1) = 0
t = 1
while || ∇SQR (w(t)) || > ε
  for j = 0, 1, 2, ..., D
    parcial[j] = -2 ∑i=1N hj(xi) (yi - ŷi (w(t)))
    wj(t+1) = wj(t) - η parcial[j]
  t = t + 1
  
```

Figura I.1 – Algoritmo gradiente descendente implementado em Python usado para obter os valores ótimos dos parâmetros do modelo.

 Algoritmo 2

```

w(1) = 0
t = 1
while || ∇SQR (w(t)) || > ε
  for j = 0, 1, 2, ..., D
    parcial[j] = [ -2 ∑i=1N hj(xi) (yi - ŷi (w(t))) + 2 λ2 wj(t) ]
    wj(t+1) = wj(t) - η parcial[j]
  t = t + 1
  
```

Figura I.2 - Algoritmo gradiente descendente implementado em Python usado para regressão regularizada.

Apêndice I

CROMATOGRAMAS OBTIDOS NAS ANÁLISES CROMATOGRÁFICAS

A seguir são apresentados os cromatogramas obtidos nos experimentos de cromatográfica gasosa realizados neste trabalho. Com os terpenos não foi possível analisar as misturas numa mesma faixa de temperatura, o que impossibilitou a comparação do efeito da temperatura sobre o efeito de todos os picos. Assim são destacados apenas alguns compostos dessas misturas e mencionada a ordem de eluição dos demais compostos de cada mistura específica.

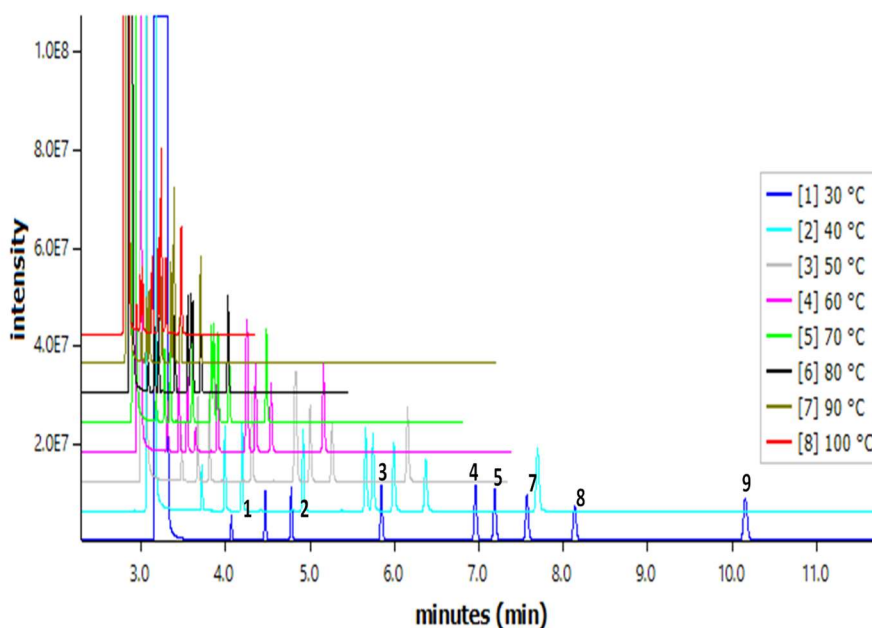


Figura I-1 – Cromatogramas obtidos para os hidrocarbonetos alifáticos. Da esquerda para a direita: (1) 2,2-dimetilbutano, (2) 2,3-dimetilbutano, (3) 3-metilpentano, (4) metilciclopentano, (5) ciclohexano, (6) 2-metilciclohexano, (7) 3-metilhexano, (8) 2,2,4-trimetilpentano e (9) metilciclohexano.

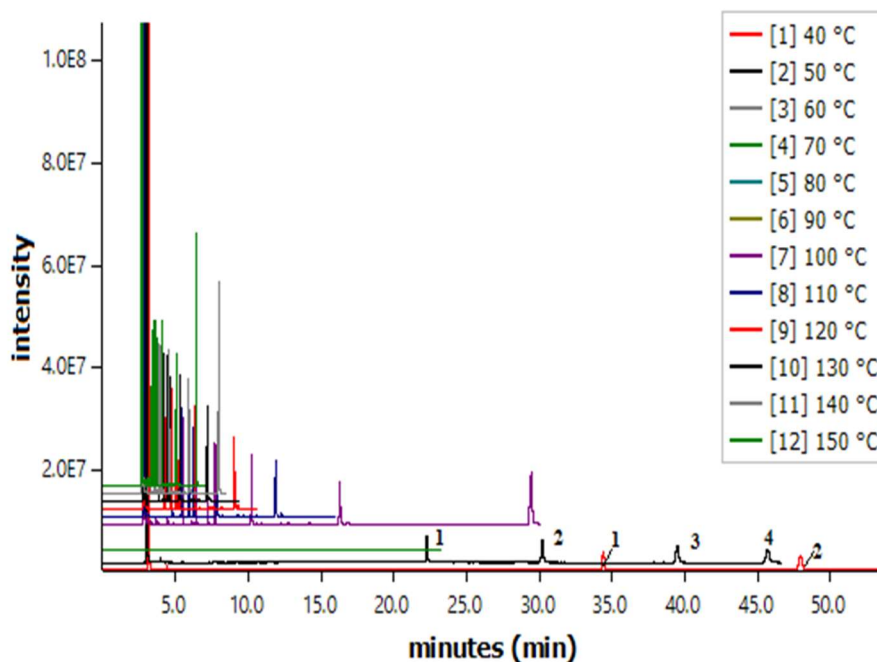


Figura I-2 – Cromatogramas obtidos para os terpenos da *mistura 1*. À temperatura de 40 °C e no intervalo de 50 minutos, apenas o (+)- β -pineno (1) e o (+)- α -pineno (2) aparecem no cromatograma. À temperatura de 40 °C aparecem mais dois picos (3 e 4), referentes ao (+)-3-careno e (+)-limoneno. Os outros compostos da *mistura 1* foram eluídos na seguinte ordem: γ -terpineno, linalol e (-)- α -terpineol.

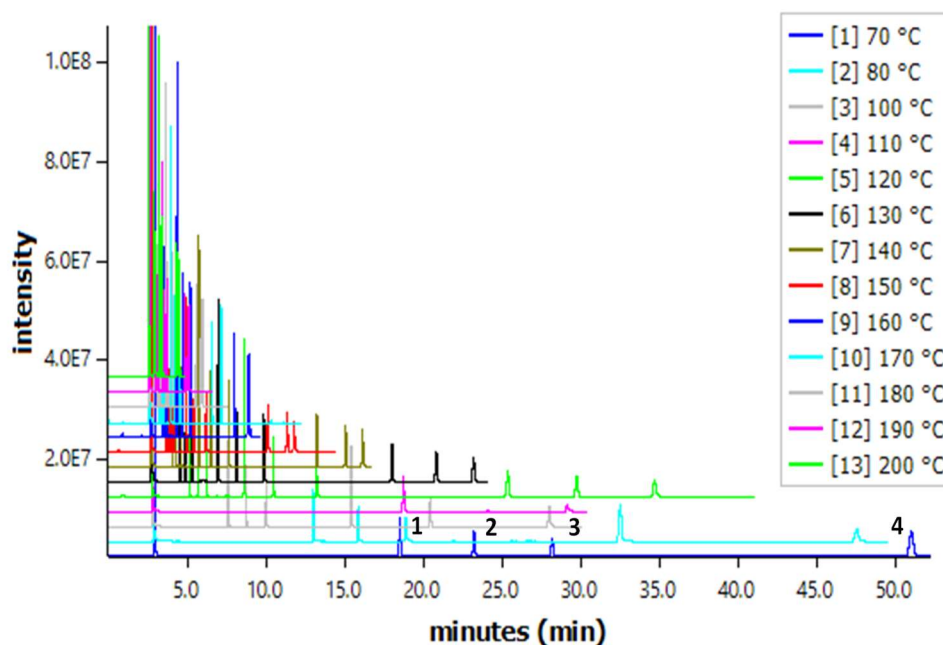


Figura I-3 – Cromatogramas obtidos para os terpenos da *mistura 2*. Da esquerda para a direita: (1) terpinoleno, (2) terpinen-4-ol, (3) timol metil éter, (4) timol. Os outros compostos

aparecem à medida que aumenta a temperatura na seguinte ordem de eluição: carvacrol, safrol, isosafrol, α -humuleno, (-)-*trans*-cariofileno e 3-BHA.

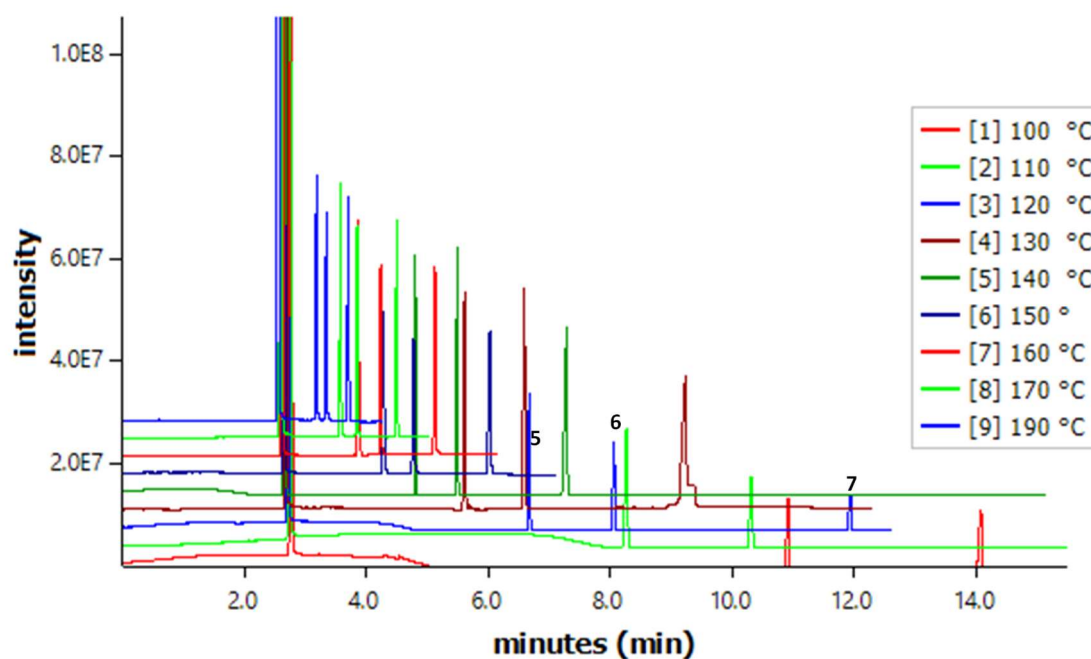
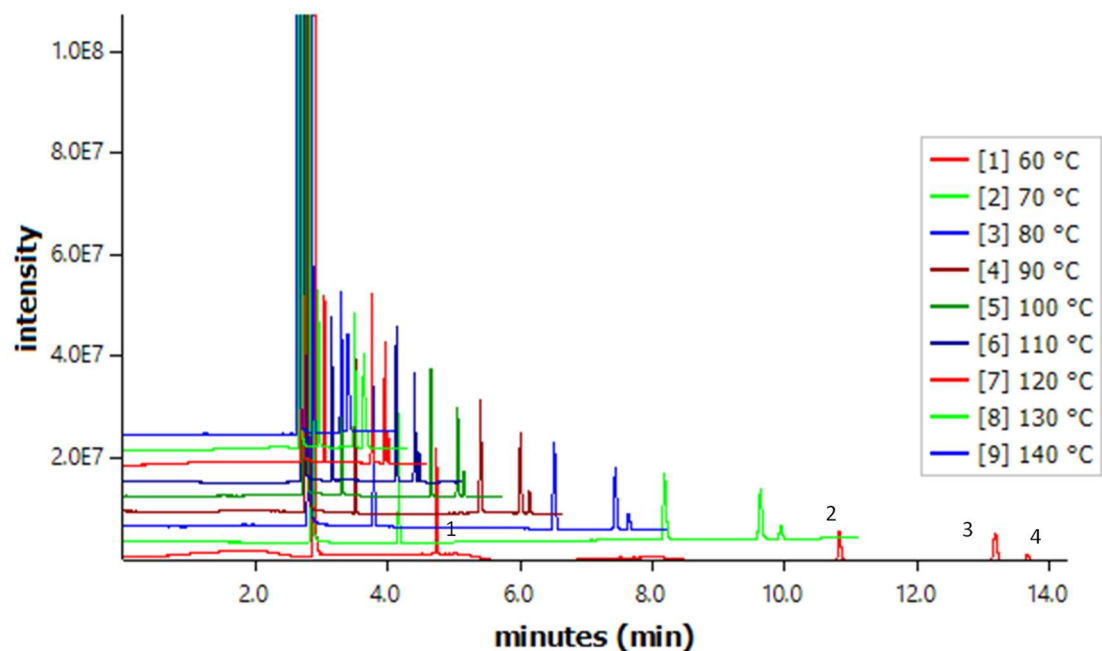


Figura I-4 – Cromatogramas obtidos para os compostos da *mistura 3*. Da esquerda para a direita: (1) biciclo[2.2.1]hepta2,5-dieno, (2) 5-etilideno-2-norborneno, (3) 5-vinil-2-norborneno, (4) álcool fenilil, (5) (-)-borneol, (6) biciclo[2.2.2]octa-2,5-diona, (7) biciclo[4.1.0]heptane-7-ácido carboxílico.