

UNIVERSIDADE ESTADUAL PAULISTA – UNESP
Instituto de Química - Câmpus de Araraquara

IGOR LENO DE SOUZA FERNANDES

Comparação entre métodos de contribuição de grupos e métodos semiempíricos para o cálculo de entalpias padrão de formação em sistemas relevantes à produção de biodiesel

Araraquara - SP

2026

IGOR LENO DE SOUZA FERNANDES

Comparação entre métodos de contribuição de grupos e métodos semiempíricos para o cálculo de entalpias padrão de formação em sistemas relevantes à produção de biodiesel

Trabalho de Conclusão de Curso apresentado ao Instituto de Química – Câmpus de Araraquara, da Universidade Estadual Paulista "Júlio de Mesquita Filho", para obtenção do grau de Bacharel em Engenharia Química.

Orientador: Prof. Dr. Elias de Souza Monteiro Filho

Araraquara - SP

2026

F363c Fernandes, Igor Leno de Souza

Comparação entre métodos de contribuição de grupos e métodos semiempíricos para o cálculo de entalpias padrão de formação em sistemas relevantes à produção de biodiesel / Igor Leno de Souza Fernandes. -- Araraquara, 2026

58 p. : il., tabs.

Trabalho de conclusão de curso (Bacharelado - Engenharia Química) - Universidade Estadual Paulista (UNESP), Instituto de Química, Araraquara

Orientador: Elias de Souza Monteiro Filho

1. Engenharia química. 2. Biodiesel. 3. Termodinâmica. 4. Entalpia. 5. Química Quântica. I. Título.

IGOR LENO DE SOUZA FERNANDES

Comparação entre métodos de contribuição de grupos e métodos semiempíricos para o cálculo de entalpias padrão de formação em sistemas relevantes à produção de biodiesel

Trabalho de Conclusão de Curso apresentado ao Instituto de Química – Câmpus de Araraquara, da Universidade Estadual Paulista "Júlio de Mesquita Filho", para obtenção do grau de Bacharel em Engenharia Química.

Data da defesa: 30/06/2026

Banca Examinadora:

Prof. Dr. Elias de Souza Monteiro Filho

Profa. Dra. Marcela Marques de Freitas Lima

Profa. Dra. Érica Regina Filletti Nascimento

AGRADECIMENTOS

As conquistas só têm brilho quando compartilhadas. Até hoje, esta é minha maior e é com muita gratidão e alegria que gostaria de compartilhá-la com vocês: minha família, meu TIME. É sempre por vocês.

Agradeço à minha irmã, Monique, por ser minha parceira número um. O primeiro pedaço do meu bolo é sempre seu.

Aos meus pais, Telma e Edney, por serem meu porto seguro e minha catapulta. É por vocês que tenho confiança para voar e a tranquilidade de saber que sempre tenho para quem voltar.

Ao meu orientador, Elias, por ter dado forma e o campo para que eu pudesse expressar minha engenharia.

À minha melhor amiga, Mariana, por ser minha melhor amiga.

À minha madrinha, Luciana, por me ensinar que a vida em movimento é brilhante.

Ao meu irmão, Adriano, por ser o quinto Beatle da minha família. Seu suporte, ímpeto e coragem são admiráveis e me motivam a ser uma pessoa melhor a cada dia.

Aos meus amigos de infância, Juan e Catarine, por me proporcionarem por tanto tempo tantas histórias boas para contar. É um prazer continuar compartilhando a vida boa com vocês.

Se não fosse por todos vocês, talvez eu nunca conseguisse dizer isto. E, mesmo que conseguisse, não teria o mesmo brilho. Por essa e pelas que estão por vir: deu certo!

RESUMO

A produção de biodiesel envolve reações de esterificação e transesterificação cuja modelagem termodinâmica depende de valores confiáveis de entalpia padrão de formação (ΔH_f°), frequentemente escassos ou inexistentes para triglicerídeos, ácidos graxos de cadeia longa e seus ésteres correspondentes. Este trabalho compara três métodos de contribuição de grupos (Joback e Reid, Constantinou e Gani, e Marrero e Gani) e três métodos semiempíricos (AM1, PM3 e PM7) na estimativa de entalpias de formação e de reação em sistemas relevantes à produção de biodiesel. A avaliação foi organizada em dois níveis: comparação frente a dados experimentais de 31 moléculas representativas e extensão preditiva a quatro compostos de cadeia longa (ácido oleico, ácido linoleico, metil oleato e metil linoleato), cujas referências foram obtidas por DFT segundo Silva, Custódio e Reis (2015). No primeiro nível, os métodos de contribuição de grupos e o PM7 apresentaram desvios médios próximos, situados entre 1,5% e 2,2%, enquanto PM3, com aproximadamente 3,0%, e sobretudo AM1, com aproximadamente 8,9%, apresentaram desempenho global inferior. No segundo nível, restrito aos métodos mais promissores, os desvios ficaram entre 3,5% e 4,2%: PM7 foi o mais preciso para os ácidos graxos (1,61%) e Constantinou e Gani apresentou o melhor desempenho global. Nas entalpias de reação, os métodos previram valores exotérmicos frente a referências gasosas endotérmicas ou próximas da termoneutralidade; apenas o PM7 forneceu valores distintos para as duas reações e reproduziu parcialmente a separação associada à diferença de insaturação, recuperando cerca de 67% da diferença esperada (16,18 de 24,20 kJ/mol), enquanto os três métodos de contribuição de grupos forneceram valores iguais ou praticamente idênticos para as duas reações, devido ao cancelamento das contribuições dos grupos insaturados. Os resultados indicam que o protocolo CREST+PM7 tem potencial como alternativa computacionalmente acessível para estimativas de ΔH_f° e análise qualitativa de diferenças estruturais em sistemas lipídicos. Entretanto, suas limitações quantitativas em ésteres metílicos e em entalpias de reação indicam que seu uso deve ser acompanhado de validação e de critérios de incerteza.

Palavras-chave: entalpia de formação; entalpia de reação; PM7; contribuição de grupos; biodiesel; ácidos graxos insaturados.

ABSTRACT

Biodiesel production involves esterification and transesterification reactions whose thermodynamic modeling depends on reliable values of standard enthalpy of formation (ΔH_f°), which are often scarce or unavailable for triglycerides, long-chain fatty acids, and their corresponding esters. This work compares three group contribution methods (Joback and Reid, Constantinou and Gani, and Marrero and Gani) and three semiempirical methods (AM1, PM3, and PM7) for estimating formation and reaction enthalpies in systems relevant to biodiesel production. The evaluation was organized into two levels: comparison against experimental data for 31 representative molecules and a predictive extension to four long-chain compounds (oleic acid, linoleic acid, methyl oleate, and methyl linoleate), whose reference values were obtained by DFT following Silva, Custódio, and Reis (2015). At the first level, the group contribution methods and PM7 showed similar mean deviations, ranging from 1.5% to 2.2%, whereas PM3, at approximately 3.0%, and especially AM1, at approximately 8.9%, performed worse overall. At the second level, restricted to the most promising methods, deviations ranged from 3.5% to 4.2%: PM7 was the most accurate for fatty acids (1.61%), and Constantinou and Gani showed the best overall performance. For reaction enthalpies, the methods predicted exothermic values against gas-phase references that were endothermic or close to thermoneutrality; only PM7 yielded distinct values for the two reactions and partially reproduced the separation associated with the difference in unsaturation, recovering about 67% of the expected difference (16.18 of 24.20 kJ/mol), whereas the three group contribution methods yielded equal or nearly identical values for the two reactions because the contributions of the unsaturated groups canceled out. The results indicate that the CREST+PM7 protocol has potential as a computationally accessible alternative for ΔH_f° estimation and qualitative analysis of structural differences in lipid systems. However, its quantitative limitations for methyl esters and reaction enthalpies indicate that its use should be accompanied by validation and uncertainty criteria.

Keywords: enthalpy of formation; reaction enthalpy; PM7; group contribution; biodiesel; unsaturated fatty acids.

SUMÁRIO

1 INTRODUÇÃO.....	8
1.1 PROBLEMA DE PESQUISA.....	9
1.2 OBJETIVO GERAL	9
1.3 OBJETIVOS ESPECÍFICOS.....	10
1.4 JUSTIFICATIVA.....	11
1.5 ESTRUTURA DO TRABALHO.....	12
2 FUNDAMENTAÇÃO TEÓRICA.....	14
2.1 TERMODINÂMICA E ENTALPIA.....	14
2.2 ENTALPIA PADRÃO DE FORMAÇÃO.....	14
2.3 ENTALPIA DE REAÇÃO E APLICAÇÕES EM PROCESSOS.....	15
2.4 MÉTODOS DE CONTRIBUIÇÃO DE GRUPOS.....	16
2.4.1 Conceitos gerais.....	16
2.4.2 Método de Joback e Reid.....	16
2.4.3 Método de Constantinou e Gani.....	17
2.4.4 Método de Marrero e Gani	18
2.4.5 Limitações dos métodos de contribuição de grupos.....	18
2.5 MÉTODOS QUÂNTICOS SEMIEMPÍRICOS.....	20
2.5.1 Características gerais.....	20
2.5.2 Métodos AM1 e PM3.....	20
2.5.3 Método PM7 e relação entre custo e precisão.....	20
2.5.4 Busca conformacional com CREST e otimização no MOPAC.....	22
2.6 REFERÊNCIA DFT COM CORREÇÃO EMPÍRICA.....	22
2.7 COMPOSTOS E REAÇÕES ASSOCIADOS À PRODUÇÃO DE BIODIESEL....	23
3 MATERIAIS E MÉTODOS.....	25
3.1 ORGANIZAÇÃO DOS COMPOSTOS.....	25
3.2 CURADORIA E ORGANIZAÇÃO DA BASE DE DADOS	26
3.3 ΔH_f° POR CONTRIBUIÇÃO DE GRUPOS	27
3.4 ΔH_f° POR MÉTODOS SEMIEMPÍRICOS	28
3.5 ΔH_r° DAS ESTERIFICAÇÕES	29
3.6 MÉTRICAS ESTATÍSTICAS DE AVALIAÇÃO.....	30
3.7 FERRAMENTAS E TRATAMENTO DOS DADOS	31

4 RESULTADOS E DISCUSSÃO.....	33
4.1 VISÃO GERAL E ESCOLHA DAS MÉTRICAS.....	33
4.2 ΔH_f° DOS COMPOSTOS — NÍVEL 1.....	33
4.2.1 Desempenho global.....	33
4.2.2 Desempenho por grupo molecular.....	35
4.3 ΔH_f° DOS COMPOSTOS C18 — NÍVEL 2.....	39
4.3.1 Resultados e desvios frente à referência DFT.....	39
4.3.2 Análise por subgrupo e padrões de desvio.....	41
4.4 ΔH_r° DAS ESTERIFICAÇÕES — NÍVEL 2.....	42
4.4.1 Resultados das entalpias de reação.....	42
4.4.2 Análise dos desvios e fenômeno de sinal.....	44
4.4.3 Diferenciação por grau de insaturação.....	46
4.5 SÍNTESE CRUZADA E IMPLICAÇÕES.....	47
5 CONCLUSÕES.....	49
5.1 SÍNTESE GERAL.....	49
5.2 SÍNTESE DO NÍVEL 1.....	49
5.3 SÍNTESE DO NÍVEL 2 — ΔH_f°	50
5.4 SÍNTESE DO NÍVEL 2 — ΔH_r°	50
5.5 TRABALHOS FUTUROS.....	52
REFERÊNCIAS.....	55
APÊNDICE A — MÉTRICAS POR CLASSE E MÉTODO (NÍVEL 1).....	58

1 INTRODUÇÃO

A previsão confiável de propriedades termodinâmicas é um elemento central no projeto, na análise e na otimização de processos da engenharia química, uma vez que influencia diretamente balanços de massa e energia, dimensionamento de equipamentos e avaliação de viabilidade técnico-econômica (Green e Perry, 2008). Entre essas propriedades, as entalpias padrão de formação e de reação ocupam papel de destaque, pois são parâmetros fundamentais para o cálculo de variações de entalpia associadas a operações unitárias e reações químicas industriais.

A obtenção experimental de entalpias pode ser trabalhosa, cara e, em alguns casos, inviável, seja pela disponibilidade limitada de substâncias puras, seja por questões de segurança e tempo de medição. Por esse motivo, métodos preditivos têm sido amplamente desenvolvidos e empregados como alternativas ou complementos aos dados experimentais, permitindo estimar propriedades com base em estruturas moleculares e modelos teóricos (Joback e Reid, 1987; Constantinou e Gani, 1994). Nesse contexto, destacam-se, por um lado, os métodos de contribuição de grupos, que utilizam decomposições estruturais relativamente simples e apresentam baixo custo computacional (Marrero e Gani, 2001; Hukkerikar et al., 2012), e, por outro, métodos de química quântica semiempíricos, que incorporam de forma mais explícita aspectos eletrônicos das moléculas e permitem tratar sistemas de maior complexidade estrutural (Jensen, 2016; Rozanska et al., 2015).

A cadeia produtiva do biodiesel é um exemplo emblemático em que a disponibilidade de dados termodinâmicos confiáveis é crucial para a engenharia de processos. A síntese de biodiesel envolve reações de esterificação e transesterificação entre ácidos graxos, álcoois de cadeia curta, ésteres metílicos e triglicerídeos, formando sistemas com moléculas de diferentes tamanhos e funções químicas. A avaliação energética dessas etapas depende de entalpias de formação e de reação, as quais nem sempre estão plenamente disponíveis em tabelas clássicas de propriedades (Glišić e Skala, 2009; Joshipura e Patel, 2015).

Em particular, dados experimentais diretos de ΔH_f° para ácidos graxos insaturados de cadeia longa, como o ácido oleico (C18:1) e o ácido linoleico (C18:2), e para seus ésteres metílicos correspondentes são escassos ou inexistentes na literatura. Essa lacuna limita a capacidade de modelagem termodinâmica rigorosa das etapas de esterificação que envolvem matérias-primas insaturadas, as quais constituem a

maioria dos ácidos graxos presentes em óleos vegetais de interesse industrial, como soja, girassol e canola.

1.1 PROBLEMA DE PESQUISA

Embora a literatura apresente diversas abordagens para o cálculo preditivo de propriedades termodinâmicas, ainda persiste a necessidade de avaliar, de forma sistemática, a precisão e a aplicabilidade desses métodos quando empregados em sistemas de interesse da engenharia química. No contexto da produção de biodiesel, a disponibilidade de entalpias padrão de formação confiáveis é essencial para a modelagem de reações de esterificação e transesterificação, mas dados experimentais nem sempre estão disponíveis para todas as espécies envolvidas, especialmente em compostos orgânicos de maior cadeia ou de maior complexidade estrutural. Métodos de contribuição de grupos, como Joback e Reid, Constantinou e Gani e Marrero e Gani, são amplamente utilizados por sua simplicidade, baixo custo computacional e boa aplicabilidade em estimativas preliminares. Entretanto, por se basearem em parametrizações empíricas e na decomposição da molécula em grupos estruturais, sua capacidade de generalização pode ser limitada quando aplicados a moléculas menos representadas nas bases experimentais usadas em seu desenvolvimento. Por outro lado, métodos semiempíricos de química quântica, quando associados a uma busca conformacional adequada, podem oferecer uma alternativa intermediária entre os métodos puramente empíricos e os métodos de maior custo computacional, incorporando informações de geometria molecular e estrutura eletrônica sem exigir o custo de abordagens *ab initio* ou DFT.

Surge, assim, a seguinte questão orientadora: um protocolo semiempírico de baixo custo computacional, baseado em busca conformacional por CREST/GFN2-xTB e cálculo de entalpias de formação no MOPAC, é capaz de produzir estimativas de ΔH_f° com desempenho comparável ao de métodos consolidados de contribuição de grupos para moléculas relevantes à produção de biodiesel? Além disso, quando aplicado a compostos representativos de um processo real dessa cadeia produtiva, como ácidos graxos de cadeia longa e seus ésteres metílicos, esse protocolo mantém desempenho adequado para estimativas de entalpias de formação e de reação?

1.2 OBJETIVO GERAL

O objetivo geral deste trabalho é comparar o desempenho de métodos de contribuição de grupos: Joback e Reid (JR), Constantinou e Gani (CG) e Marrero e Gani

(MG), e métodos semiempíricos: AM1, PM3 e PM7, na estimativa de entalpias padrão de formação de moléculas relevantes à produção de biodiesel, avaliando se um protocolo computacional de baixo custo, baseado em busca conformacional e cálculo semiempírico, pode apresentar desempenho comparável ao de métodos consolidados de contribuição de grupos. Para isso, a avaliação foi estruturada em dois níveis complementares: (i) uma comparação quantitativa frente a dados experimentais para um conjunto de 31 moléculas representativas, incluindo ácidos carboxílicos, álcoois, ésteres metílicos e triglicerídeos de cadeia curta; e (ii) uma extensão preditiva a compostos de cadeia longa diretamente associados à produção de biodiesel, utilizando valores calculados em nível B3LYP/6-311+G(d,p) com correção empírica como referência (Silva, Custódio e Reis, 2015). Adicionalmente, as entalpias de reação de esterificação foram calculadas para avaliar a propagação dos desvios de ΔH_f° em um processo químico representativo dessa cadeia produtiva.

1.3 OBJETIVOS ESPECÍFICOS

Os objetivos específicos incluem:

- a) organizar uma base de dados termoquímica composta por moléculas relevantes à produção de biodiesel, estruturando-a em dois níveis metodológicos: um conjunto principal com dados experimentais rastreáveis de ΔH_f° e um conjunto complementar com moléculas de cadeia longa associadas à produção de biodiesel, para as quais foram adotadas referências obtidas por métodos de maior nível de teoria;
- b) calcular entalpias padrão de formação pelos métodos de contribuição de grupos Joback e Reid (JR), Constantinou e Gani (CG) e Marrero e Gani (MG), de modo a estabelecer uma base comparativa com métodos consolidados de baixo custo computacional;
- c) obter entalpias padrão de formação pelos métodos semiempíricos AM1, PM3 e PM7, mediante busca conformacional com o CREST (GFN2-xTB) e otimização de geometria no MOPAC com o algoritmo *Eigenvector Following*, avaliando o efeito de um tratamento conformacional sistemático sobre o desempenho desses métodos;
- d) comparar o desempenho dos métodos de contribuição de grupos e dos métodos semiempíricos frente aos dados experimentais do conjunto principal, quantificando os desvios por método e por grupo molecular com as métricas MRD, MAE, RMSE e MBE;

- e) identificar, a partir dos resultados do conjunto principal, os métodos com desempenho mais promissor para aplicação em moléculas de maior cadeia e maior relevância direta à produção de biodiesel;
- f) aplicar os métodos selecionados na etapa anterior aos compostos de cadeia longa do conjunto complementar: ácido oleico, ácido linoleico, metil oleato e metil linoleato, comparando as estimativas obtidas com valores de referência calculados em nível B3LYP/6-311+G(d,p) com correção empírica, conforme Silva, Custódio e Reis (2015);
- g) calcular e comparar as entalpias de reação de esterificação envolvendo ácido oleico e ácido linoleico com metanol, avaliando como os desvios nas entalpias de formação se propagam para uma aplicação reacional representativa da cadeia produtiva do biodiesel;
- h) discutir as limitações, vantagens e aplicabilidade dos métodos avaliados para a geração de dados termoquímicos em sistemas associados à produção de biodiesel, considerando precisão, custo computacional e potencial de extrapolação para moléculas menos representadas em bases experimentais.

1.4 JUSTIFICATIVA

A escolha do tema deste trabalho se justifica pela crescente demanda por combustíveis alternativos e renováveis, como o biodiesel, e pela necessidade de ferramentas de cálculo que permitam estimar propriedades termodinâmicas de forma rápida e confiável em diferentes rotas de processo, condições operacionais e matérias-primas. Em ambientes industriais, desvios nas entalpias de formação e de reação podem se propagar para balanços de energia, afetando o dimensionamento de equipamentos térmicos, a análise de eficiência e a segurança operacional (Green e Perry, 2008).

Métodos de contribuição de grupos são amplamente utilizados em engenharia devido à sua simplicidade, baixo custo computacional e facilidade de aplicação em estimativas preliminares. Entretanto, esses métodos dependem de parametrizações baseadas em dados experimentais, cuja disponibilidade e representatividade variam entre diferentes classes de compostos (Joback e Reid, 1987; Constantinou e Gani, 1994). Como evolução dessa classe de abordagens, Marrero e Gani (2001) propuseram um método em três níveis hierárquicos que amplia o domínio de aplicação para moléculas de maior complexidade estrutural, incluindo compostos com cadeias carbônicas longas. Ainda assim, por se basearem na decomposição da molécula em

grupos estruturais, tais métodos podem apresentar limitações quando aplicados a sistemas menos representados nas bases experimentais de parametrização.

Nesse contexto, métodos semiempíricos de química quântica, como AM1, PM3 e PM7, representam uma alternativa intermediária entre os métodos de contribuição de grupos e abordagens de maior custo computacional, como métodos *ab initio* e DFT. Quando associados a uma busca conformacional adequada, esses métodos podem incorporar informações de geometria molecular e estrutura eletrônica com custo computacional relativamente baixo, o que os torna potencialmente úteis para a geração de dados termoquímicos em conjuntos amplos de moléculas orgânicas (Stewart, 2023; Rozanska et al., 2014).

Para sistemas associados à produção de biodiesel, essa comparação é particularmente relevante, pois envolve moléculas de diferentes classes químicas, como ácidos carboxílicos, álcoois, ésteres metílicos e triglicerídeos. Além disso, compostos de cadeia longa presentes em matérias-primas lipídicas nem sempre dispõem de dados experimentais gasosos de ΔH_f° , o que exige o uso de referências computacionais de maior nível de teoria. Nesse sentido, a metodologia de Silva, Custódio e Reis (2015) baseada em cálculos DFT com correção empírica paramétrica fornece uma referência adequada para avaliar a extensão preditiva dos métodos mais promissores a moléculas diretamente associadas à produção de biodiesel. Assim, este trabalho contribui para identificar alternativas computacionalmente acessíveis para a estimativa de dados termoquímicos em sistemas lipídicos de relevância industrial.

1.5 ESTRUTURA DO TRABALHO

Este trabalho está organizado em cinco capítulos, além dos elementos pré e pós-textuais. No Capítulo 1 apresenta-se a introdução geral ao tema, incluindo o problema de pesquisa, os objetivos, a justificativa e a estrutura do trabalho. O Capítulo 2 reúne a fundamentação teórica necessária para a compreensão da pesquisa, abordando conceitos de entalpia padrão de formação e de reação, os métodos de contribuição de grupos avaliados: Joback e Reid (JR), Constantinou e Gani (CG) e Marrero e Gani (MG), os métodos semiempíricos AM1, PM3 e PM7, o protocolo conformacional CREST/MOPAC e as principais classes de compostos associadas à produção de biodiesel.

No Capítulo 3 são descritos os materiais e métodos utilizados, incluindo a organização da base de dados em dois níveis metodológicos, a curadoria dos valores

de referência, os procedimentos de cálculo por contribuição de grupos e por métodos semiempíricos, bem como as métricas estatísticas empregadas na avaliação dos desvios. O Capítulo 4 apresenta e discute os resultados em etapas complementares: inicialmente, a avaliação comparativa do conjunto principal com dados experimentais; em seguida, a aplicação dos métodos mais promissores a moléculas de cadeia longa relevantes à produção de biodiesel; e, por fim, o cálculo das entalpias de reação de esterificação como aplicação termoquímica representativa. O Capítulo 5 sintetiza as principais conclusões do trabalho, destaca as limitações observadas e indica perspectivas para desenvolvimentos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 TERMODINÂMICA E ENTALPIA

A termodinâmica fornece a base conceitual para descrever e quantificar as transformações de energia associadas a processos físicos e químicos, constituindo um dos fundamentos da engenharia química. Em aplicações industriais, a determinação e a estimativa de propriedades termodinâmicas estão diretamente relacionadas ao dimensionamento de equipamentos, à avaliação da eficiência energética e à análise de segurança de processos.

Nesse contexto, a entalpia é uma propriedade termodinâmica de estado amplamente utilizada na formulação de balanços de energia. Suas variações permitem quantificar os efeitos energéticos associados ao aquecimento, ao resfriamento, às mudanças de fase e às reações químicas. As entalpias padrão de formação e de reação constituem, portanto, dados fundamentais para o cálculo de cargas térmicas e para a avaliação energética de diferentes rotas de processo (Green e Perry, 2008).

2.2 ENTALPIA PADRÃO DE FORMAÇÃO

A entalpia padrão de formação de uma substância corresponde à variação de entalpia associada à formação de um mol do composto a partir de seus elementos constituintes em seus estados padrão mais estáveis, sob pressão padrão e em uma temperatura especificada, usualmente 298,15 K. Por convenção, a entalpia padrão de formação dos elementos em seus estados padrão de referência é igual a zero. Essa propriedade, representada por ΔH_f° , constitui a base para o cálculo de entalpias de reação por meio da Lei de Hess (Linstrom e Mallard, 2023).

Do ponto de vista da engenharia química, a disponibilidade de entalpias padrão de formação para diferentes espécies permite calcular entalpias de reação e estimar os efeitos térmicos associados a reatores e a outras unidades de processo. Entretanto, para diversos compostos orgânicos de maior tamanho ou complexidade estrutural, os dados experimentais são escassos ou não foram identificados nas principais bases de dados e compilações termoquímicas consultadas. Essa limitação é particularmente relevante para moléculas associadas à produção de biodiesel, como triglicerídeos, ácidos graxos de cadeia longa e seus ésteres correspondentes, o que motiva o emprego de métodos preditivos (Joshipura e Patel, 2015; Glišić e Skala, 2009).

2.3 ENTALPIA DE REAÇÃO E APLICAÇÕES EM PROCESSOS

A entalpia padrão de reação, representada por ΔH_r° , corresponde à variação de entalpia associada a uma reação química balanceada, considerando reagentes e produtos em seus respectivos estados padrão. De acordo com a Lei de Hess, essa propriedade pode ser calculada a partir das entalpias padrão de formação das espécies envolvidas:

$$\Delta H_r^\circ = \sum p \nu_p \Delta H_{f,p}^\circ - \sum r \nu_r \Delta H_{f,r}^\circ \quad (1)$$

em que ν_p e ν_r representam os coeficientes estequiométricos dos produtos e dos reagentes, respectivamente, e $\Delta H_{f,p}^\circ$ e $\Delta H_{f,r}^\circ$ correspondem às entalpias padrão de formação dos produtos e dos reagentes.

Para as reações de esterificação avaliadas neste trabalho, cuja estequiometria é 1:1:1:1, a Equação 1 assume a forma específica:

$$\Delta H_r^\circ = \Delta H_f^\circ(\text{éster}) + \Delta H_f^\circ(\text{H}_2\text{O}) - \Delta H_f^\circ(\text{ácido}) - \Delta H_f^\circ(\text{metanol}) \quad (2)$$

Para a aplicação consistente da Lei de Hess, todas as entalpias padrão de formação devem corresponder à mesma temperatura, à mesma pressão padrão e aos estados físicos especificados para cada espécie. Neste trabalho, as comparações entre os métodos de contribuição de grupos, os métodos semiempíricos e a referência computacional do Nível 2 são realizadas em fase gasosa. A obtenção de valores aplicáveis à fase líquida requer a inclusão de propriedades de mudança de fase, como as entalpias de vaporização.

Na produção de biodiesel, as entalpias das reações de esterificação e transesterificação influenciam os balanços de energia e as necessidades de aquecimento ou resfriamento das unidades de processo. A precisão dessas estimativas é relevante para o projeto e a otimização de reatores, bem como para estudos de integração energética e avaliação do desempenho do processo (Green e Perry, 2008; Glišić e Skala, 2009).

Como a entalpia de reação é calculada pela diferença entre as entalpias de formação de produtos e reagentes, os erros presentes nas propriedades individuais podem se acumular ou se cancelar. Dessa forma, estimativas relativamente próximas dos valores de referência para ΔH_f° não resultam necessariamente em estimativas igualmente precisas para ΔH_r° , especialmente quando a entalpia de reação apresenta pequena magnitude. Esse comportamento é analisado em detalhe no Capítulo 4.

2.4 MÉTODOS DE CONTRIBUIÇÃO DE GRUPOS

2.4.1 Conceitos gerais

Métodos de contribuição de grupos (MCG) são abordagens preditivas que estimam propriedades termodinâmicas de uma molécula a partir da soma das contribuições de grupos estruturais, como fragmentos funcionais e unidades de repetição de cadeias. A premissa básica é que a propriedade alvo pode ser aproximada como uma soma ponderada das contribuições de grupos elementares, ajustadas a partir de dados experimentais.

O desenvolvimento histórico dos métodos de contribuição de grupos para propriedades termoquímicas inclui o trabalho de Yoneda (1979), que propôs o método ABWY para a estimativa de propriedades de compostos orgânicos no estado de gás ideal. O desenvolvimento posterior dos métodos de contribuição de grupos incluiu os modelos de Joback e Reid (1987) e de Constantinou e Gani (1994), além de abordagens hierárquicas mais detalhadas, como o método de Marrero e Gani (2001).

Nessa abordagem, a molécula é decomposta em grupos definidos pelo método; para cada grupo é atribuído um valor de contribuição associado à propriedade de interesse e a propriedade total é obtida pela soma das contribuições multiplicadas pelas respectivas multiplicidades. Esse tipo de método é amplamente utilizado devido ao baixo custo computacional e à rapidez de aplicação, tornando-se atrativo em contextos de projeto preliminar e de triagem de alternativas (Joback e Reid, 1987; Constantinou e Gani, 1994).

2.4.2 Método de Joback e Reid

O método de Joback e Reid é um dos modelos clássicos de contribuição de grupos para a estimação de propriedades de compostos orgânicos (Joback e Reid, 1987). No caso de entalpias padrão de formação, o método expressa a propriedade como uma soma do tipo:

$$\Delta H_f^\circ = 68,29 + \sum_k N_k H_{f,k} \quad (3)$$

em que 68,29 kJ/mol é uma constante empírica, N_k é o número de ocorrências do grupo k na molécula e $H_{f,k}$ é a contribuição de entalpia padrão de formação associada ao grupo k , tabelada em unidades de kJ/mol.

O método possui grupos para diferentes funções orgânicas, incluindo ligações duplas carbono-carbono ($=CH-$ e $=C<$), o que permite sua aplicação a compostos insaturados como ácidos graxos C18. O método foi parametrizado a partir de dados

experimentais de diferentes classes de compostos orgânicos e desenvolvido como uma abordagem simples para estimativas preliminares de propriedades (Joback e Reid, 1987). Na aplicação a sistemas de biodiesel, o método é utilizado identificando-se grupos como fragmentos de cadeia alifática, grupos carboxílicos, hidroxilas, ligações duplas C=C e demais funções presentes nos compostos de interesse.

2.4.3 Método de Constantinou e Gani

O método de Constantinou e Gani (1994) permite realizar estimativas em dois níveis: uma aproximação de primeira ordem, baseada apenas nos grupos fundamentais, e uma aproximação de segunda ordem, que acrescenta correções estruturais. Para a entalpia padrão de formação, a expressão correspondente à aproximação de segunda ordem, empregada neste trabalho, é dada por:

$$\Delta H_f^\circ = 10,835 + \sum_k N_k C_k + \sum_j M_j D_j \quad (4)$$

em que 10,835 kJ/mol é a constante empírica do método, C_k e D_j representam as contribuições dos grupos de primeira e segunda ordem, respectivamente, e N_k e M_j correspondem às suas multiplicidades. Na aproximação de primeira ordem, o termo correspondente aos grupos de segunda ordem é omitido.

Essa maior granularidade pode melhorar a representação de efeitos estruturais que não são descritos exclusivamente por contribuições de primeira ordem. A aplicabilidade estrutural do método a moléculas associadas ao biodiesel também é observada em estudos de outras propriedades termodinâmicas: Glišić e Skala (2009) o empregaram para estimar parâmetros críticos de trioleína, dioleoína, monooleína e oleato de metila, ressaltando que a decomposição térmica dessas moléculas dificulta ou inviabiliza a determinação experimental direta das propriedades críticas consideradas; Joshipura e Patel (2015) utilizaram o método CG para obter propriedades críticas de ácidos graxos e seus ésteres metílicos (FAME) como base para modelagem de equilíbrio de fases. Em estudos comparativos, o desempenho relativo dos métodos depende da classe molecular e da propriedade avaliada (Constantinou e Gani, 1994).

A extensão do método para estimativa do fator acêntrico e volume molar líquido foi proposta pelos próprios autores em publicação complementar (Constantinou, Gani e O'Connell, 1995), consolidando o conjunto de propriedades acessíveis por essa abordagem.

2.4.4 Método de Marrero e Gani

O método de Marrero e Gani (2001) constitui uma abordagem hierárquica de maior detalhamento, incorporando grupos de primeira, segunda e terceira ordem. Essa estrutura foi desenvolvida para ampliar a representação de compostos de maior complexidade estrutural e diferenciar estruturas que não são distinguidas apenas por grupos de primeira ordem, incluindo, em sua base de desenvolvimento, compostos contendo de 3 a 60 átomos de carbono — faixa que abrange os ácidos graxos C18 e seus ésteres metílicos (Marrero e Gani, 2001).

O método permite estimativas em três níveis hierárquicos. Neste trabalho, entretanto, foram empregadas contribuições até a segunda ordem, não tendo sido utilizados grupos de terceira ordem. Para a entalpia padrão de formação, a expressão correspondente à aproximação aplicada neste estudo é dada por:

$$\Delta H_f^\circ = 5,549 + \sum k N_k Ck^{(1)} + \sum j M_j Dj^{(2)} \quad (5)$$

em que 5,549 kJ/mol é a constante empírica do método, $Ck^{(1)}$ e $Dj^{(2)}$ representam as contribuições dos grupos de primeira e segunda ordem, respectivamente, e N_k e M_j correspondem às suas multiplicidades. Na aproximação de primeira ordem, o termo de segunda ordem é omitido.

Os grupos de segunda e terceira ordem previstos na formulação original podem incorporar parte dos efeitos de conectividade e vizinhança associados à estrutura molecular, ampliando a capacidade do método de distinguir estruturas que as contribuições de primeira ordem não diferenciam. Neste trabalho, entretanto, foram utilizadas somente contribuições até a segunda ordem.

2.4.5 Limitações dos métodos de contribuição de grupos

Apesar de suas vantagens, métodos de contribuição de grupos apresentam limitações intrínsecas, principalmente decorrentes de sua natureza empírica. Como os parâmetros são ajustados a partir de um conjunto de dados experimentais, a qualidade das estimativas depende fortemente da cobertura e da representatividade dessa base de calibração. Compostos com funcionalidades pouco representadas ou combinações estruturais fora do domínio de calibração podem apresentar desvios elevados (Joback e Reid, 1987; Marrero e Gani, 2001).

Nos métodos baseados predominantemente em contribuições de primeira ordem, grupos estruturalmente equivalentes recebem contribuições constantes, o que

limita a representação de efeitos associados à vizinhança, à posição e ao ambiente molecular. Métodos hierárquicos, como o de Marrero e Gani, incluem correções de ordem superior capazes de representar parte desses efeitos, embora sua precisão permaneça condicionada aos grupos disponíveis e à base utilizada na parametrização. Como consequência, as estimativas de entalpias de formação para ácidos graxos C18 com diferentes graus de insaturação podem não capturar plenamente os efeitos associados à posição, à estereoquímica, ao ambiente molecular e às interações não aditivas envolvendo as ligações duplas. Esses efeitos podem se refletir nas entalpias de reação de esterificação, como discutido no Capítulo 4.

Nas esterificações de ácidos graxos, parte dessa limitação também decorre do cancelamento das contribuições estruturais. As ligações duplas presentes no ácido graxo permanecem no éster correspondente e, como recebem contribuições semelhantes nos dois lados da reação, seus efeitos podem ser amplamente cancelados no cálculo de ΔH_r° . Assim, os métodos podem distinguir as entalpias de formação dos compostos insaturados, mas ainda apresentar baixa sensibilidade ao efeito da insaturação sobre a diferença energética da reação.

Além disso, como a entalpia de reação é calculada pela diferença entre as entalpias de formação de produtos e reagentes, os erros das entalpias de formação podem se acumular ou se cancelar no cálculo de ΔH_r° . Isso pode resultar em desvios relativos elevados mesmo quando as entalpias de formação individuais são razoavelmente bem estimadas.

Buscando reduzir parte dessas limitações, Hukkerikar et al. (2012) propuseram o método GC+, que combina contribuição de grupos com índice de conectividade atômica e incorpora análise de incerteza sistemática. Essa abordagem permite associar intervalos de confiança às estimativas, informação que não está disponível nos métodos clássicos de JR, CG e MG, embora não elimine as limitações intrínsecas à natureza empírica dos MCG.

2.5 MÉTODOS QUÂNTICOS SEMIEMPÍRICOS

2.5.1 Características gerais

Métodos quânticos semiempíricos são modelos de química quântica que buscam aproximar o comportamento eletrônico de moléculas utilizando simplificações formais das equações de mecânica quântica, combinadas com parâmetros ajustados a dados experimentais ou cálculos de nível teórico superior. Em comparação com

métodos *ab initio*, como Hartree–Fock, e com métodos baseados na teoria do funcional da densidade (DFT), as abordagens semiempíricas apresentam custo computacional consideravelmente menor, permitindo o tratamento de sistemas maiores e de um número mais elevado de estruturas (Jensen, 2016).

Na estimativa de entalpias de formação, métodos semiempíricos calculam energias eletrônicas e incluem correções empíricas para reproduzir dados de termodinâmica experimental. Assim, a qualidade das estimativas depende tanto da forma das aproximações teóricas quanto do conjunto de dados utilizado na parametrização de cada método, variando conforme a classe molecular e a propriedade considerada (Jensen, 2016; Rozanska et al., 2014).

2.5.2 Métodos AM1 e PM3

O método AM1 (Austin Model 1) foi desenvolvido como uma evolução de modelos semiempíricos anteriores, com foco na melhoria da descrição de energias de formação e geometria molecular de compostos orgânicos. O PM3 pertence à mesma família de métodos fundamentados na aproximação NDDO, diferenciando-se do AM1 principalmente pela estratégia de parametrização e pelo conjunto de parâmetros empregados. Ambos os métodos utilizam formalismos baseados em integrais simplificadas e parametrização extensiva, de forma a tornar os cálculos viáveis para moléculas de porte moderado (Jensen, 2016).

O desempenho de AM1 e PM3 depende da classe molecular e da propriedade avaliada, podendo apresentar limitações na estimativa de entalpias de formação de determinados compostos orgânicos.

2.5.3 Método PM7 e relação entre custo e precisão

O método PM7 representa uma geração mais recente de modelos semiempíricos, incorporando um conjunto de parâmetros revisto e expandido, com o objetivo de melhorar a descrição de energias e propriedades estruturais em ampla faixa de compostos (Stewart, 2013). O PM7 foi desenvolvido com uma parametrização mais recente e ampliada em relação a métodos como AM1 e PM3, embora seu desempenho dependa da classe molecular e da propriedade considerada.

Rozanska et al. (2014) avaliaram a precisão do PM7 para frequências vibracionais, entropias e capacidades caloríficas de compostos orgânicos no estado de gás ideal, empregando estratégia automatizada de otimização de geometria via MOPAC para grande número de espécies, e encontraram desvios compatíveis com aplicações

de engenharia. Em escala ampliada, Rozanska et al. (2015) aplicaram procedimentos automatizados de simulação a um conjunto extenso de compostos orgânicos, indicando a viabilidade de abordagens computacionais sistemáticas para a geração de propriedades moleculares.

Chan (2017) analisou a aplicação de métodos semiempíricos a sistemas de maior porte, indicando que o PM7 pode produzir geometrias compatíveis com determinadas aplicações e que, para energias de reação relativas em sistemas com muitos átomos, o cancelamento de erros sistemáticos pode contribuir para a redução dos desvios, ainda que os erros absolutos nas entalpias de formação individuais possam permanecer relevantes.

Embora métodos de estrutura eletrônica de maior custo computacional, como Hartree–Fock e DFT, possam, em princípio, oferecer maior precisão, seu custo computacional é consideravelmente mais elevado, especialmente quando se pretende analisar conjuntos extensos de moléculas e realizar buscas conformacionais (Jensen, 2016). Nesse cenário, PM7 surge como um compromisso interessante entre custo e precisão, particularmente atraente para aplicações em engenharia química nas quais se necessita de estimativas rápidas para múltiplas espécies e cenários (Rozanska et al., 2015).

Diferentemente dos métodos puramente aditivos de contribuição de grupos, o PM7 determina a energia a partir de uma descrição aproximada da estrutura eletrônica e da geometria molecular completa. Dessa forma, alterações no número e na posição das ligações duplas podem modificar a distribuição eletrônica, a geometria e a energia calculada de maneira não necessariamente equivalente no ácido e no éster correspondente. Essa característica pode conferir ao método maior sensibilidade relativa às diferenças de insaturação, embora não assegure, por si só, maior exatidão nos valores absolutos de entalpia de formação ou de reação.

2.5.4 Busca conformacional com CREST e otimização no MOPAC

A qualidade das estimativas obtidas por métodos semiempíricos depende da conformação molecular considerada e da capacidade do procedimento de otimização de localizar mínimos energeticamente relevantes na superfície de energia potencial. Moléculas flexíveis podem apresentar diversos mínimos locais nessa superfície, correspondentes a diferentes confôrmeros, cujas geometrias e energias podem variar de maneira relevante. Esse aspecto é particularmente importante para moléculas com

cadeias carbônicas longas, como ácidos graxos, ésteres e triglicerídeos, que apresentam elevado número de graus de liberdade rotacionais.

O CREST (Conformer-Rotamer Ensemble Sampling Tool) emprega metadinâmica molecular associada ao Hamiltoniano GFN2-xTB para gerar um conjunto de confôrmeros energeticamente favoráveis para cada molécula (Pracht, Bohle e Grimme, 2020). Essa abordagem permite uma amostragem mais ampla das conformações acessíveis, reduzindo a dependência de uma única geometria inicial e a possibilidade de que o cálculo posterior permaneça restrito a um mínimo local pouco representativo. As estruturas geradas pelo CREST podem ser posteriormente reotimizadas no MOPAC com Hamiltonianos semiempíricos, como AM1, PM3 e PM7 (Stewart, 2023). É importante considerar que a ordenação energética obtida pelo GFN2-xTB pode ser alterada após a otimização semiempírica, uma vez que cada Hamiltoniano descreve a superfície de energia potencial por meio de aproximações e parametrizações próprias. A avaliação de múltiplas geometrias iniciais aumenta a probabilidade de identificação de um mínimo energeticamente favorável no nível semiempírico empregado, justificando conceitualmente o uso de conjuntos de confôrmeros em vez de uma única estrutura de partida.

Os critérios de seleção das estruturas, o número de confôrmeros avaliados e os Hamiltonianos empregados neste trabalho são apresentados no Capítulo 3.

2.6 REFERÊNCIA DFT COM CORREÇÃO EMPÍRICA

Para compostos cujos dados experimentais de ΔH_f° em fase gasosa não foram identificados nas bases consultadas — como é o caso dos ácidos graxos C18 insaturados e seus ésteres metílicos —, a avaliação dos métodos de estimativa requer referências obtidas por métodos de nível teórico superior. Neste trabalho, adota-se como referência para o Nível 2 a metodologia de Silva, Custódio e Reis (2015), que combina cálculos DFT no nível B3LYP/6-311+G(d,p) com uma correção empírica paramétrica ajustada especificamente para ácidos graxos e seus derivados.

A metodologia de Silva, Custódio e Reis (2015) foi desenvolvida para ácidos graxos de diferentes comprimentos de cadeia e seus respectivos ésteres metílicos, incluindo compostos de cadeia longa. A aplicação da correção reduziu os desvios absolutos médios para aproximadamente 5,3 kJ/mol nos ácidos graxos e 3,4 kJ/mol nos ésteres avaliados. O desempenho reportado pelos autores fornece suporte para o emprego da metodologia como referência computacional no Nível 2. Entretanto, os

desvios médios obtidos no conjunto avaliado não representam diretamente uma incerteza validada para os compostos C18 insaturados considerados neste trabalho. Os cálculos realizados em nível B3LYP/6-311+G(d,p) fornecem geometrias e energias em um nível de teoria mais elevado que o dos métodos semiempíricos avaliados, enquanto a correção paramétrica é empregada no aprimoramento das estimativas de entalpia de formação. A metodologia abrange ácidos graxos e ésteres tanto saturados quanto insaturados. No conjunto inicial de descritores, os autores consideraram os números de átomos de carbono, de hidrogênio e de ligações duplas; entretanto, como o número de ligações duplas é dependente da composição em carbono e hidrogênio para as moléculas avaliadas, o coeficiente correspondente foi fixado em zero, reduzindo a correção a uma função dessas duas variáveis (Silva, Custódio e Reis, 2015).

É importante notar que a referência DFT não equivale a um dado experimental: trata-se de uma estimativa computacional em fase gasosa, sujeita às limitações do método DFT e da correção empírica paramétrica. Portanto, as comparações no Nível 2 são interpretadas como avaliações da consistência entre métodos de diferentes custos computacionais, e não como validações contra valores absolutos definitivos.

2.7 COMPOSTOS E REAÇÕES ASSOCIADOS À PRODUÇÃO DE BIODIESEL

A produção de biodiesel pode envolver a esterificação de ácidos graxos com álcoois de cadeia curta, formando ésteres e água, e a transesterificação de triglicerídeos com esses álcoois, formando ésteres alquílicos de ácidos graxos e glicerol. Quando o álcool empregado é o metanol, os produtos são denominados ésteres metílicos de ácidos graxos (FAME). O ácido oleico e o ácido linoleico estão entre os principais ácidos graxos presentes em óleos vegetais empregados como matérias-primas para a produção de biodiesel.

Apesar de sua relevância industrial, dados experimentais de ΔH_f° em fase gasosa para o ácido oleico, o ácido linoleico, o metil oleato e o metil linoleato não foram identificados nas principais bases de dados e compilações termoquímicas consultadas (Linstrom e Mallard, 2023; Haynes, 2014; Green e Perry, 2008). Para triglicerídeos insaturados, a determinação experimental de propriedades que exigem informações em fase vapor é particularmente difícil, pois essas moléculas podem sofrer decomposição térmica antes de alcançar as condições necessárias para a medição direta (Glišić e Skala, 2009). Esse cenário reforça a motivação para o emprego de mé-

todos de estimativa e para a adoção de referências computacionais quando apropriado.

3 MATERIAIS E MÉTODOS

3.1 ORGANIZAÇÃO DOS COMPOSTOS

O conjunto de compostos neste trabalho está organizado em dois níveis metodológicos. Os compostos do Nível 1 são apresentados na Tabela 1.

Tabela 1 – Compostos do Nível 1

Composto	Fórmula
ácido acético	$C_2H_4O_2$
ácido propanoico	$C_3H_6O_2$
ácido butírico	$C_4H_8O_2$
ácido pentanoico	$C_5H_{10}O_2$
ácido hexanoico	$C_6H_{12}O_2$
ácido heptanoico	$C_7H_{14}O_2$
ácido octanoico	$C_8H_{16}O_2$
ácido dodecanoico	$C_{12}H_{24}O_2$
ácido tetradecanoico	$C_{14}H_{28}O_2$
ácido hexadecanoico	$C_{16}H_{32}O_2$
metanol	CH_4O
etanol	C_2H_6O
glicerol	$C_3H_8O_3$
propanol	C_3H_8O
butanol	$C_4H_{10}O$
pentanol	$C_5H_{12}O$
hexanol	$C_6H_{14}O$
heptanol	$C_7H_{16}O$
octanol	$C_8H_{18}O$
acetato de metila	$C_3H_6O_2$
propanoato de metila	$C_4H_8O_2$
butanoato de metila	$C_5H_{10}O_2$
pentanoato de metila	$C_6H_{12}O_2$
hexanoato de metila	$C_7H_{14}O_2$
heptanoato de metila	$C_8H_{16}O_2$
octanoato de metila	$C_9H_{18}O_2$
dodecanoato de metila	$C_{13}H_{26}O_2$
tetradecanoato de metila	$C_{15}H_{30}O_2$
hexadecanoato de metila	$C_{17}H_{34}O_2$
triacetina	$C_9H_{14}O_6$
tributirina	$C_{15}H_{26}O_6$

Fonte: elaborado pelo autor.

Nível 2 — Compostos insaturados de cadeia longa (4 moléculas). Reúne os ácidos graxos C18 insaturados — ácido oleico (C18:1, $\Delta 9$ -cis) e ácido linoleico (C18:2, $\Delta 9,12$ -cis,cis) — e seus ésteres metílicos correspondentes, o metil oleato e o metil linoleato. Para esses compostos, dados experimentais de ΔH_f° em fase gasosa são escassos ou ausentes nas principais bases de dados e compilações termoquímicas. Os valores de referência adotados neste nível são entalpias de formação calculadas em nível B3LYP/6-311+G(d,p) com correção empírica paramétrica, conforme a metodologia de Silva, Custódio e Reis (2015), que apresenta desvio absoluto médio residual de 5,3 kJ/mol frente a dados experimentais para ácidos graxos. Foram avaliados JR, CG, MG e PM7; AM1 e PM3 não foram aplicados aos quatro compostos C18 insaturados. O PM7 foi selecionado para este nível por apresentar o menor desvio relativo médio (MRD) global entre os métodos semiempíricos no conjunto de 31 moléculas do Nível 1 (Seção 4.2).

3.2 CURADORIA E ORGANIZAÇÃO DA BASE DE DADOS

Para o Nível 1, os critérios de inclusão dos dados de referência foram: (i) valor em fase gasosa; (ii) rastreável a bases de dados ou compilações termoquímicas (NIST Chemistry WebBook, CRC Handbook of Chemistry and Physics ou Perry's Chemical Engineers' Handbook); (iii) temperatura de 298,15 K; (iv) unidade kJ/mol. Os valores adotados foram harmonizados para fase gasosa, temperatura de 298,15 K, unidade de kJ/mol e estado padrão especificado pelas fontes. Quando mais de um valor compatível foi identificado, a seleção considerou a rastreabilidade, a qualidade do método experimental ou da derivação termoquímica, a incerteza reportada e a recomendação da fonte.

Para o Nível 2, os valores de referência são entalpias calculadas por Silva, Custódio e Reis (2015) em nível B3LYP/6-311+G(d,p) com correção empírica. Esses valores são tratados como referência de alto nível de teoria, não como dados experimentais: as comparações realizadas no Nível 2 avaliam a consistência entre métodos preditivos de diferentes custos computacionais, e não a proximidade a um valor experimental definitivo.

Os dados foram organizados em planilha eletrônica contendo, para cada espécie: identificador, fórmula molecular, grupo, nível metodológico, número de átomos, origem do valor de referência, ΔH_f° de referência e ΔH_f° calculados pelos seis métodos, além de métricas de desvio por método. A planilha contém resultados de JR, CG,

MG, AM1, PM3 e PM7 para as 31 moléculas do Nível 1, e resultados de JR, CG, MG e PM7 para as quatro moléculas do Nível 2.

Foram calculados desvios absolutos e relativos entre os valores calculados e os de referência, por método e por molécula. A partir desses dados foram obtidas as métricas de desempenho por nível e por grupo molecular, detalhadas na Seção 3.6.

3.3 ΔH_f° POR CONTRIBUIÇÃO DE GRUPOS

O cálculo das entalpias de formação pelos métodos de contribuição de grupos baseia-se na decomposição da estrutura de cada molécula em grupos funcionais elementares, seguindo as convenções estabelecidas por cada método.

Para o método de Joback e Reid (1987), a entalpia padrão de formação é obtida conforme a Equação 3, em que 68,29 kJ/mol é uma constante empírica, N_k é o número de ocorrências do grupo k na molécula e $H_{f,k}$ é a contribuição de entalpia padrão de formação do grupo k , tabelada em unidades de kJ/mol.

Para o método de Constantinou e Gani (1994), a entalpia de formação é obtida conforme a Equação 4, sendo 10,835 kJ/mol a constante correspondente e C_k e D_j as contribuições dos grupos de primeira e segunda ordem, respectivamente, com multiplicidades N_k e M_j , tabeladas para o método.

Para o método de Marrero e Gani (2001), as estimativas foram calculadas com contribuições de primeira e segunda ordem. Não foram empregados grupos de terceira ordem no conjunto de dados deste trabalho. A entalpia padrão de formação é obtida conforme a Equação 5, em que 5,549 kJ/mol é a constante empírica do método; $C_k^{(1)}$ e $D_j^{(2)}$ são as contribuições dos grupos de primeira e segunda ordem, respectivamente, com multiplicidades N_k e M_j . Quando, para determinado composto, não se identificam grupos de segunda ordem aplicáveis, o respectivo somatório não contribui para o resultado.

Para as moléculas insaturadas do Nível 2, cada ligação dupla interna do tipo $-\text{CH}=\text{CH}-$ foi decomposta em dois centros $=\text{CH}-$ (carbono monossustituído com ligação dupla), de modo que os compostos monoinsaturados (C18:1) apresentam dois grupos $=\text{CH}-$ e os di-insaturados (C18:2), quatro. A identificação e a contagem dos grupos seguiram a definição de cada método (JR, CG e MG), que pode diferir quanto aos grupos previstos e às respectivas contribuições tabeladas. As contribuições associadas a esses grupos foram contabilizadas conforme as tabelas de cada método.

3.4 ΔH_f° POR MÉTODOS SEMIEMPÍRICOS

As entalpias padrão de formação pelos métodos semiempíricos foram obtidas por meio de um protocolo composto pela geração de estruturas tridimensionais iniciais, busca conformacional com o CREST e posterior otimização das estruturas selecionadas no MOPAC.

As estruturas tridimensionais iniciais foram geradas com o RDKit por meio do procedimento de incorporação molecular, e submetidas a uma pré-otimização com o campo de força UFF. As geometrias resultantes foram utilizadas como entrada para a busca conformacional.

A busca conformacional foi realizada com o CREST versão 3.0.2, associado ao xTB versão 6.7.1 e ao Hamiltoniano GFN2-xTB, em fase gasosa. Foram utilizados os parâmetros `--gfn2 --v3 --ewin 6.0 --rthr 0.125 --mstruct 20000`. A amostragem conformacional foi conduzida a 300 K. As espécies avaliadas são neutras e de camada fechada, correspondendo, por padrão, a carga total igual a zero e multiplicidade singlete.

Para cada molécula, foram selecionados os dez confôrmeros de menor energia identificados pelo CREST. Quando o conjunto conformacional apresentava menos de dez estruturas distintas, todos os confôrmeros disponíveis foram utilizados. Cada estrutura selecionada foi submetida separadamente à otimização de geometria no MOPAC versão 23.2.2.

No Nível 1, o procedimento foi aplicado às 31 moléculas com os Hamiltonianos AM1, PM3 e PM7. As otimizações foram realizadas em fase gasosa, utilizando o algoritmo Eigenvector Following, por meio da palavra-chave EF, juntamente com PRECISE. Para cada combinação entre molécula e Hamiltoniano, foi adotada como estrutura representativa aquela que apresentou o menor valor de FINAL HEAT OF FORMATION após a otimização. O respectivo calor de formação final foi utilizado nas análises estatísticas; esse calor de formação refere-se à temperatura termodinâmica de referência de aproximadamente 298 K adotada pelo MOPAC, distinta da temperatura de 300 K empregada na amostragem conformacional com o CREST/xTB.

Com base no desempenho comparativo de AM1, PM3 e PM7 para o conjunto completo de 31 moléculas do Nível 1, o PM7 foi selecionado para a extensão ao Nível 2 por apresentar o menor desvio relativo médio (MRD) global entre os métodos semiempíricos (Seção 4.2). Para o ácido oleico, o ácido linoleico, o metil oleato e o me-

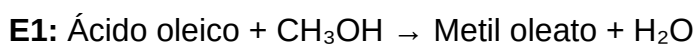
til linoleato, foi aplicado o mesmo procedimento de busca conformacional, seleção de até dez estruturas, otimização no MOPAC e escolha da estrutura com o menor valor final de calor de formação. AM1 e PM3 não foram aplicados no Nível 2.

A seleção das estruturas foi realizada exclusivamente com base nos resultados obtidos após a otimização, sem utilização dos valores experimentais ou computacionais de referência como critério de escolha. Esses valores foram empregados apenas na etapa posterior de avaliação dos desvios.

Não foram realizados cálculos de frequências vibracionais. Consequentemente, as estruturas selecionadas correspondem às geometrias otimizadas de menor calor de formação entre os confôrmeros avaliados, mas não foram confirmadas por análise vibracional como mínimos verdadeiros da superfície de energia potencial.

3.5 ΔH_f° DAS ESTERIFICAÇÕES

As entalpias de reação de esterificação foram calculadas para duas reações relevantes à síntese de biodiesel com matérias-primas insaturadas:



A entalpia de reação foi obtida pela expressão geral de Hess, conforme a Equação 2.

Para cada método (JR, CG, MG e PM7), as entalpias de formação dos ácidos graxos insaturados e dos ésteres metílicos foram tomadas dos cálculos do Nível 2. O metanol foi estimado por cada método com o respectivo valor do Nível 1 (JR: $-216,20$; CG: $-216,53$; MG: $-215,29$; PM7: $-204,76$ kJ/mol). Para a água, adotou-se um procedimento distinto conforme o tipo de método: nos métodos de contribuição de grupos (JR, CG e MG), que não dispõem de grupos para representar a molécula de água, utilizou-se o valor experimental $\Delta H_f^\circ(\text{H}_2\text{O}, \text{gás}) = -241,826$ kJ/mol; no PM7, a água foi calculada pelo próprio método, resultando em $\Delta H_f^\circ(\text{H}_2\text{O}, \text{gás}) = -238,83$ kJ/mol. Dessa forma, as entalpias de reação obtidas pelo PM7 são internamente consistentes, ao passo que, nos métodos de contribuição de grupos, a água entra como valor experimental fixo enquanto os demais compostos são estimados — caráter híbrido justificado pela ausência de uma decomposição em grupos aplicável à água.

A entalpia de reação de **referência** foi obtida em **fase gasosa**, pela mesma expressão de Hess, combinando as entalpias de formação gasosas de referência dos

compostos C18 (referência B3LYP corrigida de Silva, Custódio e Reis, 2015) com as entalpias de formação gasosas experimentais de metanol e água. Esse procedimento produz ΔH_r° de referência de +26,41 kJ/mol (E1) e +2,21 kJ/mol (E2), grandezas fase-consistentes com os métodos avaliados (que também operam em fase gasosa) e que constituem a base de desvio adotada no Capítulo 4. Complementarmente, Silva, Custódio e Reis (2015) reportam, em **fase líquida**, ΔH_r° de +46,63 kJ/mol (E1) e +21,24 kJ/mol (E2), valores que incorporam correções de mudança de fase e são apresentados apenas como referência complementar para a fase condensada, não como base de cálculo dos desvios nem como representação direta das condições operacionais do processo.

Para as métricas de avaliação das entalpias de reação, o desvio absoluto (DA) é adotado como métrica primária. Para reações cujo valor de referência apresenta pequeno módulo, como E2, o desvio relativo (DR) torna-se fortemente influenciado pelo denominador e não deve ser interpretado isoladamente como medida de desempenho; é reportado apenas como informação complementar, fenômeno análogo ao discutido para esterificações saturadas de cadeia longa em trabalhos anteriores (Fernandes e Monteiro Filho, 2024).

3.6 MÉTRICAS ESTATÍSTICAS DE AVALIAÇÃO

Para quantificar o desempenho de cada método, foram calculadas as seguintes métricas de desvio entre valores calculados ($\Delta H_{calc,i}^\circ$) e de referência ($\Delta H_{ref,i}^\circ$) para um conjunto de n compostos:

Desvio absoluto (DA) e desvio relativo (DR) por composto:

$$DA_i = |\Delta H_{calc,i}^\circ - \Delta H_{ref,i}^\circ| \quad (6)$$

$$DR_i = 100 \times |\Delta H_{calc,i}^\circ - \Delta H_{ref,i}^\circ| / |\Delta H_{ref,i}^\circ| \quad (7)$$

Erro absoluto médio (MAE):

$$MAE = (1/n) \sum |\Delta H_{calc,i}^\circ - \Delta H_{ref,i}^\circ| \quad (8)$$

Neste trabalho adota-se preferencialmente a denominação erro absoluto médio (MAE). A denominação desvio absoluto médio (MAD) é mantida apenas ao repro-

duzir a terminologia de Silva, Custódio e Reis (2015), na qual corresponde à mesma média dos desvios absolutos.

Raiz do erro quadrático médio (RMSE):

$$RMSE = \sqrt{[(1/n) \sum (\Delta H_{calc,i}^{\circ} - \Delta H_{ref,i}^{\circ})^2]} \quad (9)$$

Erro médio de viés (MBE):

$$MBE = (1/n) \sum (\Delta H_{calc,i}^{\circ} - \Delta H_{ref,i}^{\circ}) \quad (10)$$

Um MBE positivo indica tendência do método a fornecer valores mais positivos (ou menos negativos) que a referência; um MBE negativo indica tendência a valores mais negativos. Como a maioria das entalpias de formação avaliadas é negativa, evita-se a interpretação isolada em termos de superestimação ou subestimação, que pode ser ambígua nesse contexto.

Desvio relativo médio (MRD):

$$MRD = (100/n) \sum |(\Delta H_{calc,i}^{\circ} - \Delta H_{ref,i}^{\circ}) / \Delta H_{ref,i}^{\circ}| \quad (11)$$

Foram calculadas métricas globais do Nível 1, métricas por classe molecular no Nível 1, métricas globais do Nível 2, o desvio absoluto (DA) de cada reação e as métricas agregadas MAE, RMSE e MBE para as esterificações E1 e E2. Essas métricas foram calculadas por nível metodológico e por grupo molecular. Nota importante para entalpias de reação: quando o módulo do valor de ΔH_r° de referência é pequeno, o DR e as médias dele derivadas (MRD) tornam-se altamente sensíveis ao denominador e devem ser interpretados em conjunto com métricas absolutas. Nesses casos, o DA é adotado como métrica de avaliação primária.

3.7 FERRAMENTAS E TRATAMENTO DOS DADOS

O fluxo de cálculo das entalpias de formação pelos métodos semiempíricos compreendeu a geração de estruturas iniciais com o RDKit, pré-otimização com o campo de força UFF, busca conformacional com o CREST versão 3.0.2 associado ao

xTB versão 6.7.1 e ao Hamiltoniano GFN2-xTB (Pracht, Bohle e Grimme, 2020), e otimização de geometria no MOPAC versão 23.2.2 (Stewart, 2023) com o algoritmo Eigenvector Following (palavras-chave EF e PRECISE) e os Hamiltonianos AM1, PM3 e PM7, em fase gasosa. Não foram realizados cálculos de frequências vibracionais. Os cálculos pelos métodos de contribuição de grupos (JR, CG e MG) foram conduzidos a partir da decomposição de cada molécula segundo as definições de grupos do respectivo método; as multiplicidades e as contribuições foram registradas em planilha eletrônica, na qual os somatórios de JR, CG e MG foram implementados com os parâmetros tabelados nas respectivas fontes metodológicas. Nessa mesma planilha foram também implementadas as fórmulas das entalpias de reação e das métricas estatísticas de desempenho.

Os valores de referência do Nível 2 foram obtidos da literatura (Silva, Custódio e Reis, 2015) e adotados diretamente, sem recálculo, em fase gasosa e em kJ/mol: $\Delta H_f^\circ(\text{ácido oleico}) = -694,4$; $\Delta H_f^\circ(\text{ácido linoleico}) = -561,9$; $\Delta H_f^\circ(\text{metil oleato}) = -627,1$; e $\Delta H_f^\circ(\text{metil linoleato}) = -518,8$ kJ/mol.

4 RESULTADOS E DISCUSSÃO

4.1 VISÃO GERAL E ESCOLHA DAS MÉTRICAS

Os resultados deste capítulo estão organizados segundo os dois níveis metodológicos definidos no Capítulo 3. O **Nível 1** compreende a avaliação comparativa quantitativa frente a dados experimentais para as 31 moléculas representativas (ácidos carboxílicos saturados C2–C16, álcoois C1–C8 e glicerol, ésteres metílicos saturados C3–C16, e triglicerídeos de cadeia curta). O **Nível 2** apresenta os resultados para os quatro compostos insaturados de cadeia longa (ácido oleico, ácido linoleico, metil oleato e metil linoleato) frente à referência DFT de Silva, Custódio e Reis (2015), incluindo o cálculo comparativo de entalpias de reação de esterificação.

Para entalpias de formação (ΔH_f°), utiliza-se como métrica primária o desvio relativo médio (MRD), por ser adimensional e permitir comparação direta entre métodos em moléculas de grandezas muito distintas. Como métricas complementares, reportam-se MAE (erro absoluto médio, em kJ/mol), RMSE e MBE.

Para entalpias de reação (ΔH_r°), o desvio absoluto (DA, em kJ/mol) é adotado como métrica primária, pois o DR é matematicamente inadequado quando a entalpia de reação de referência é pequena em módulo — a anomalia de denominador discutida na Seção 3.5 (Fernandes e Monteiro Filho, 2024). A referência de avaliação adotada é a entalpia de reação em **fase gasosa**, recalculada por lei de Hess a partir das entalpias de formação gasosas do Nível 2 ($E1 = +26,41$; $E2 = +2,21$ kJ/mol), por ser fase-consistente com os métodos avaliados (JR, CG, MG e PM7), que produzem ΔH_r° em fase gasosa. Justamente porque a referência de E2 é pequena em módulo (+2,21 kJ/mol), o DA é preferível ao DR — situação que reforça a opção metodológica da Seção 3.5. Os valores em fase líquida reportados por Silva, Custódio e Reis (2015) incorporam correções de mudança de fase e constituem uma referência complementar para a fase condensada. Entretanto, não representam diretamente as condições operacionais de um processo industrial ou de um reator específico.

4.2 ΔH_r° DOS COMPOSTOS — NÍVEL 1

4.2.1 Desempenho global

A Tabela 2 apresenta o MRD global dos seis métodos para o conjunto completo de 31 moléculas do Nível 1, com os semiempíricos avaliados sobre o melhor conformero de cada molécula (protocolo CREST+MOPAC multiconformacional). Os três métodos de contribuição de grupos (JR, CG e MG) e o método PM7 apresentam de-

semprenhos globais próximos no conjunto avaliado, na faixa de MRD de 1,5–2,2%. PM3 (MRD de 3,01% e RMSE de 40,29 kJ/mol) e AM1 (MRD global de 8,86%) apresentaram desempenho global inferior ao dos métodos de contribuição de grupos e do PM7, com AM1 exibindo os maiores desvios. O PM7 apresentou o menor módulo de MBE entre os métodos semiempíricos, indicando menor tendência sistemática global de deslocamento em relação às referências no conjunto avaliado ($|MBE| = 3,01$ kJ/mol, frente a 30,02 kJ/mol do AM1 e 8,94 kJ/mol do PM3). Entre os semiempíricos, o PM7 também apresentou o menor MRD, MAE e RMSE globais no Nível 1; isso não demonstra superioridade universal do PM7, mas indica melhor desempenho agregado no conjunto de 31 moléculas avaliadas. As métricas globais complementares são reunidas na Tabela 3, e a Figura 1 permite visualizar a comparação dos MRDs globais.

Tabela 2 – Desvio relativo médio global por método no Nível 1

Método	MRD (%)
PM7	1,47%
CG	1,96%
JR	2,06%
MG	2,19%
PM3	3,01%
AM1	8,86%

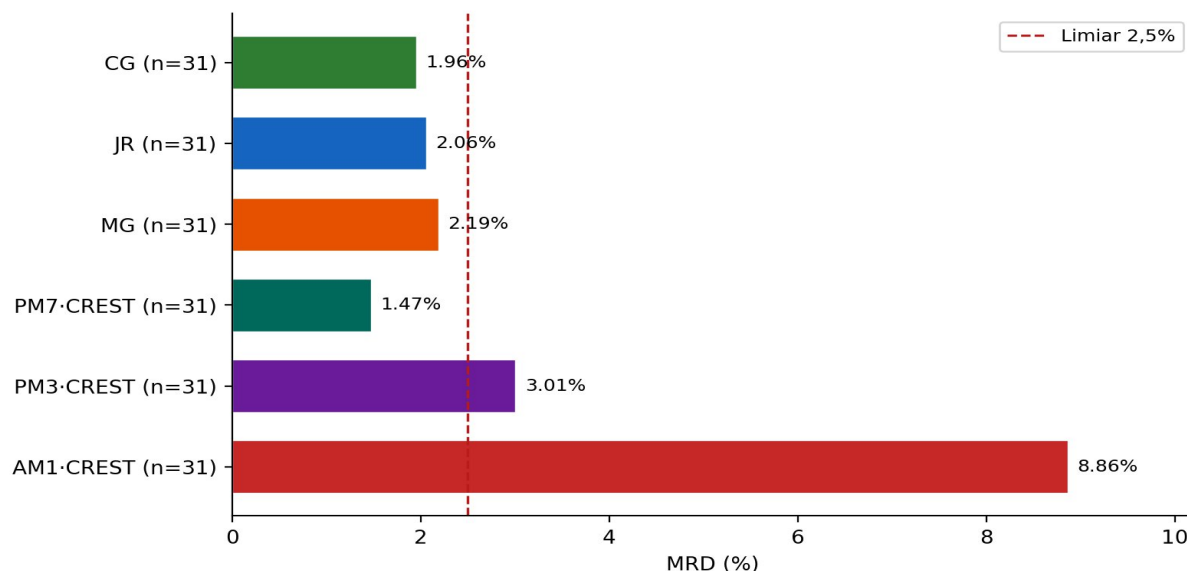
Fonte: elaborado pelo autor.

Tabela 3 – Métricas globais de desempenho dos métodos no Nível 1

Método	MAE (kJ/mol)	RMSE (kJ/mol)	MBE (kJ/mol)	MRD (%)
PM7	9,41	17,84	3,01	1,47%
CG	11,64	20,01	-0,79	1,96%
JR	12,48	21,25	-1,28	2,06%
MG	12,71	18,92	-0,98	2,19%
PM3	18,79	40,29	8,94	3,01%
AM1	42,63	53,80	-30,02	8,86%

Fonte: elaborado pelo autor.

Figura 1 – Desvio relativo médio global por método no Nível 1



Fonte: elaborado pelo autor.

Embora PM7 apresente o menor MRD global entre os métodos avaliados, a ordenação do desempenho varia consideravelmente entre as classes moleculares. PM3 apresentou o menor MRD para os ácidos, MG apresentou os menores valores para os álcoois e triglicerídeos, e PM7 apresentou o menor MRD para os ésteres metílicos. Portanto, não foi observada uma hierarquia uniforme de desempenho entre as diferentes classes. No conjunto global do Nível 1, JR, CG, MG e PM7 apresentaram métricas agregadas próximas, com MRDs situados entre 1,47% e 2,19%. O PM3 apresenta MRD moderado no conjunto global, porém RMSE relativamente elevado (40,29 kJ/mol), o que evidencia a influência de erros pontuais de grande magnitude. Os triglicerídeos (triacetina e tributirina) podem contribuir de forma relevante para esse comportamento, conforme indicam os desvios por classe no Apêndice A (Tabela 12).

4.2.2 Desempenho por grupo molecular

Tabela 4 – Desvio relativo médio por método e classe molecular no Nível 1

Grupo molecular	JR	CG	MG	AM1	PM3	PM7
Ácidos (n=10)	0,73%	0,72%	1,30%	4,78%	0,56%	0,74%
Álcoois (n=9)	1,83%	1,94%	1,61%	17,57%	4,21%	1,68%

Grupo molecular	JR	CG	MG	AM1	PM3	PM7
Ésteres (n=10)	2,96%	2,63%	3,24%	5,45%	2,68%	1,41%
Triglicerídeos (n=2)	5,20%	4,84%	3,98%	7,13%	11,44%	4,51%
MRD Global (n=31)	2,06%	1,96%	2,19%	8,86%	3,01%	1,47%

Fonte: elaborado pelo autor

A Tabela 4 e a Figura 2 mostram o desempenho por classe molecular. A análise revela que nenhum método é universalmente superior: cada abordagem apresenta vantagens específicas por família de compostos. A Tabela 5 apresenta os valores individuais de entalpia, e a Tabela 6 reúne os respectivos desvios relativos.

Ácidos carboxílicos (n=10, C2–C16). JR e CG atingem MRDs próximos de 0,72–0,73% e MG alcança 1,30%. De forma notável, PM3 apresenta o menor MRD entre todos os métodos para ácidos (~0,56%), comportamento que contrasta com seu desempenho inferior em outros grupos. PM7 situa-se em 0,74%, e AM1 em 4,78%. O grupo funcional ácido carboxílico está bem parametrizado tanto nos MCG quanto no PM3 e no PM7, possivelmente em razão da grande disponibilidade de dados experimentais de ácidos de cadeia curta utilizados no desenvolvimento desses métodos.

AM1 apresentou desempenho particularmente inferior para a classe dos álcoois, com MRD de 17,57%. Os resultados deste trabalho não permitem atribuir esse comportamento a uma interação molecular, ligação química ou parâmetro específico. O glicerol constitui o caso mais desafiador, para o qual o AM1 apresentou desvio absoluto próximo de 99 kJ/mol.

Para os ésteres metílicos do conjunto avaliado, o PM7 apresentou o menor MRD entre os métodos analisados, de 1,41%. JR e CG registraram MRDs entre 2,6% e 3,0%, e MG apresentou MRD de 3,2%.

Triglicerídeos de cadeia curta (n=2: triacetina e tributirina). Os MCG apresentam desvios maiores do que para compostos funcionais simples: JR e CG situam-se em 5,2% e 4,8%, com MG apresentando o menor desvio nesse grupo (MRD ≈ 3,98%). Entre os semiempíricos, PM7 alcança 4,51%, comparável aos MCG, enquanto AM1 (7,13%) e sobretudo PM3 (11,44%) degradam o desempenho — o que reforça a maior dificuldade de parametrização semiempírica para moléculas multifuncionais de maior número de átomos. A classe dos triglicerídeos possui apenas duas moléculas

(n = 2) e deve ser interpretada com cautela: métricas agregadas nessa classe são sensíveis a valores individuais.

Tabela 5 – Entalpias padrão de formação (kJ/mol) para os compostos do Nível 1

Molécula	Ref.	JR	CG	MG	AM1	PM3	PM7
ácido acético	-432,20	-434,88	-431,35	-426,86	-431,10	-426,95	-427,59
ácido propanoico	-455,70	-455,52	-452,12	-447,69	-456,47	-445,32	-446,54
ácido butírico	-475,80	-476,16	-472,88	-468,52	-482,53	-468,02	-468,51
ácido pentanoico	-491,30	-496,80	-493,64	-489,35	-506,38	-490,54	-489,36
ácido hexanoico	-511,90	-517,44	-514,41	-510,18	-533,64	-511,17	-510,95
ácido heptanoico	-536,20	-538,08	-535,17	-531,01	-561,98	-535,97	-532,29
ácido octanoico	-556,00	-558,72	-555,93	-551,84	-587,85	-556,07	-553,70
ácido dodecanoico	-642,00	-641,28	-638,98	-635,15	-699,24	-642,60	-642,11
ácido tetradecanoico	-693,70	-682,56	-680,51	-676,81	-756,10	-693,31	-689,36
ácido hexadecanoico	-737,10	-723,84	-722,04	-718,47	-812,23	-737,09	-733,68
metanol	-200,94	-216,20	-216,53	-215,29	-238,71	-217,14	-204,76
etanol	-234,95	-236,84	-237,30	-236,12	-262,35	-238,03	-233,76
glicerol	-577,90	-567,22	-582,61	-582,73	-676,93	-599,79	-599,08
propanol	-254,60	-257,48	-258,06	-256,95	-288,59	-257,81	-255,67
butanol	-275,10	-278,12	-278,82	-277,78	-319,60	-282,99	-276,09
pentanol	-295,70	-298,76	-299,59	-298,61	-349,23	-306,99	-300,65
hexanol	-316,20	-319,40	-320,35	-319,44	-378,88	-332,58	-322,83
heptanol	-336,80	-340,04	-341,11	-340,26	-407,52	-355,33	-344,15
octanol	-357,30	-360,68	-361,88	-361,09	-436,22	-378,98	-365,54
acetato de metila	-411,90	-422,53	-424,85	-424,39	-403,84	-393,96	-405,06
propanoato de metila	-427,50	-443,17	-440,32	-443,61	-428,91	-412,35	-423,66
butanoato de metila	-450,70	-463,81	-461,08	-464,44	-455,48	-434,99	-445,41
pentanoato de metila	-471,10	-484,45	-481,84	-485,27	-480,10	-457,24	-466,29
hexanoato de metila	-492,20	-505,09	-502,61	-506,10	-505,92	-479,98	-487,80
heptanoato de metila	-515,50	-525,73	-523,37	-526,93	-534,28	-500,76	-509,07
octanoato de metila	-533,90	-546,37	-544,13	-547,76	-560,90	-523,47	-531,35
dodecanoato de metila	-617,90	-628,93	-627,18	-631,07	-672,84	-615,26	-618,32
tetradecanoato de metila	-656,90	-670,21	-668,71	-672,73	-728,54	-658,05	-661,36
hexadecanoato de metila	-665,60	-711,49	-710,24	-714,39	-785,13	-696,31	-705,61
triacetina	-1245,00	-1186,21	-1193,08	-1202,40	-1145,47	-1103,97	-1192,06
tributirina	-1388,90	-1310,05	-1312,36	-1325,77	-1301,98	-1228,35	-1322,60

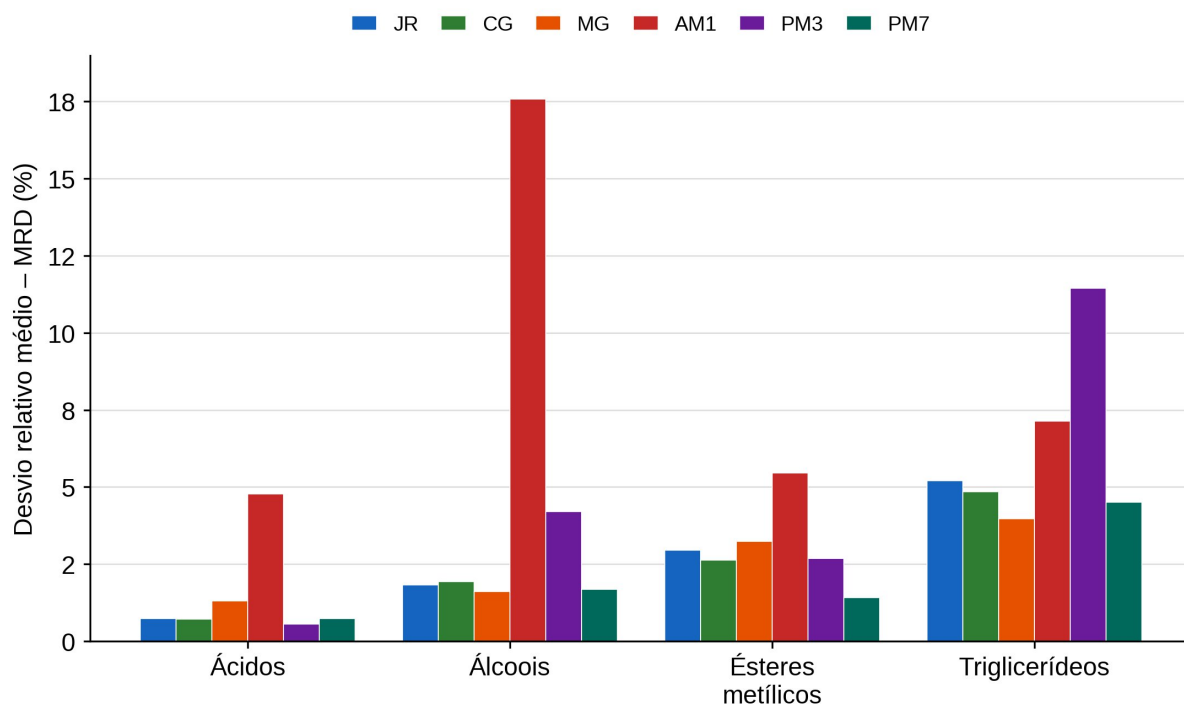
Fonte: elaborado pelo autor.

Tabela 6 – Desvios relativos das entalpias padrão de formação no Nível 1

Molécula	JR (%)	CG (%)	MG (%)	AM1 (%)	PM3 (%)	PM7 (%)
ácido acético	0,62	0,20	1,24	0,26	1,22	1,07
ácido propanoico	0,04	0,79	1,76	0,17	2,28	2,01
ácido butírico	0,08	0,61	1,53	1,41	1,64	1,53
ácido pentanoico	1,12	0,48	0,40	3,07	0,15	0,40
ácido hexanoico	1,08	0,49	0,34	4,25	0,14	0,19
ácido heptanoico	0,35	0,19	0,97	4,81	0,04	0,73
ácido octanoico	0,49	0,01	0,75	5,73	0,01	0,41
ácido dodecanoico	0,11	0,47	1,07	8,92	0,09	0,02
ácido tetradecanoico	1,61	1,90	2,43	9,00	0,06	0,63
ácido hexadecanoico	1,80	2,04	2,53	10,19	0,00	0,46
metanol	7,59	7,76	7,14	18,80	8,06	1,90
etanol	0,80	1,00	0,50	11,66	1,31	0,50
glicerol	1,85	0,82	0,84	17,14	3,79	3,66
propanol	1,13	1,36	0,92	13,35	1,26	0,42
butanol	1,10	1,35	0,97	16,18	2,87	0,36
pentanol	1,03	1,31	0,98	18,10	3,82	1,67
hexanol	1,01	1,31	1,02	19,82	5,18	2,10
heptanol	0,96	1,28	1,03	21,00	5,50	2,18
octanol	0,95	1,28	1,06	22,09	6,07	2,31
acetato de metila	2,58	3,14	3,03	1,96	4,36	1,66
propanoato de metila	3,67	3,00	3,77	0,33	3,54	0,90
butanoato de metila	2,91	2,30	3,05	1,06	3,48	1,17
pentanoato de metila	2,83	2,28	3,01	1,91	2,94	1,02
hexanoato de metila	2,62	2,11	2,82	2,79	2,48	0,89
heptanoato de metila	1,98	1,53	2,22	3,64	2,86	1,25
octanoato de metila	2,34	1,92	2,60	5,06	1,95	0,48
dodecanoato de metila	1,79	1,50	2,13	8,89	0,43	0,07
tetradecanoato de metila	2,03	1,80	2,41	10,91	0,18	0,68
hexadecanoato de metila	6,89	6,71	7,33	17,96	4,61	6,01
triacetina	4,72	4,17	3,42	7,99	11,33	4,25
tributirina	5,68	5,51	4,55	6,26	11,56	4,77

Fonte: elaborado pelo autor.

Figura 2 – Desvio relativo médio por classe molecular e método no Nível 1



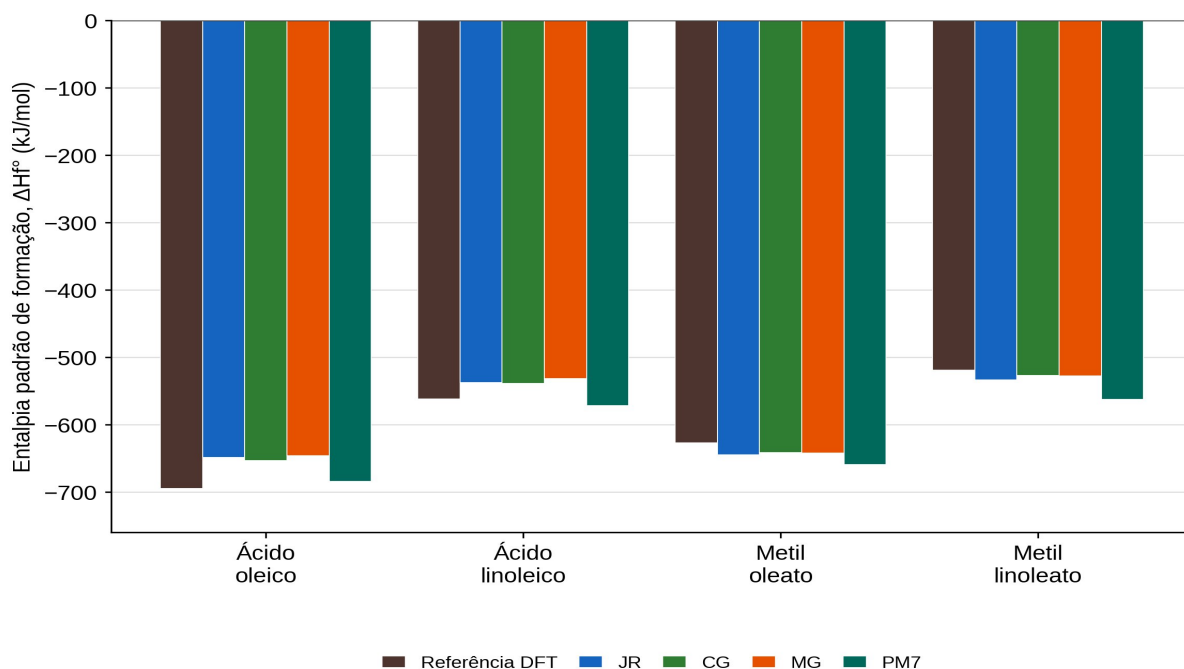
Fonte: elaborado pelo autor.

4.3 ΔH_f° DOS COMPOSTOS C18 — NÍVEL 2

4.3.1 Resultados e desvios frente à referência DFT

As métricas globais do Nível 2 são sintetizadas na Tabela 7. A Tabela 8 apresenta as entalpias de formação calculadas pelos métodos de contribuição de grupos e pelo PM7, e os desvios em relação à referência B3LYP/6-311+G(d,p) com correção empírica de Silva, Custódio e Reis (2015) para as quatro moléculas insaturadas do Nível 2. Os desvios individuais são apresentados na Tabela 9, e a Figura 3 permite visualizar as diferenças entre as referências e os métodos.

Figura 3 – Entalpias padrão de formação dos compostos C18 insaturados do Nível 2



Fonte: elaborado pelo autor com base nos resultados deste trabalho e em Silva, Custódio e Reis (2015).

Tabela 7 – Métricas globais de desempenho dos métodos no Nível 2

Método	MAE (kJ/mol)	RMSE (kJ/mol)	MBE (kJ/mol)	MRD (%)
CG	21,70	25,03	10,49	3,48%
MG	25,56	29,83	13,79	4,10%
JR	25,56	28,28	9,45	4,13%
PM7	23,89	27,95	-18,82	4,18%

Fonte: elaborado pelo autor.

Tabela 8 – Entalpias padrão de formação de referência e calculadas para os compostos do Nível 2

Composto	Ref.	JR	CG	MG	PM7
Ácido oleico (C18:1)	-694,40	-648,80	-653,16	-645,91	-684,26
Ácido linoleico (C18:2)	-561,90	-537,48	-538,75	-531,69	-571,76
Metil oleato (C19:1)	-627,10	-644,73	-641,36	-641,83	-658,88
Metil linoleato (C19:2)	-518,80	-533,40	-526,95	-527,61	-562,56

Fonte: elaborado pelo autor.

Tabela 9 – Desvios relativos das entalpias padrão de formação no Nível 2

Composto	JR (%)	CG (%)	MG (%)	PM7 (%)
Ácido oleico (C18:1)	6,57	5,94	6,98	1,46
Ácido linoleico (C18:2)	4,35	4,12	5,38	1,75
Metil oleato (C19:1)	2,81	2,27	2,35	5,07
Metil linoleato (C19:2)	2,81	1,57	1,70	8,44
MRD global (n = 4)	4,13%	3,48%	4,10%	4,18%

Fonte: elaborado pelo autor.

4.3.2 Análise por subgrupo e padrões de desvio

No conjunto restrito de quatro compostos do Nível 2, o método CG apresentou os menores valores agregados de MAE, RMSE e MRD. O PM7, embora selecionado com base em seu desempenho entre os métodos semiempíricos no Nível 1, apresentou viés negativo nesse conjunto, fornecendo valores de entalpia de formação mais negativos que a referência. Esses resultados devem ser interpretados considerando o pequeno tamanho amostral e a natureza computacional da referência utilizada. Os métodos apresentam MRDs globais no Nível 2 superiores aos observados no Nível 1 (2,0–2,2%): JR (4,13%), MG (4,10%) e PM7 (4,18%) ficam próximos, e CG apresenta o menor desvio (3,48%). Esse aumento reflete, em parte, a distância estrutural entre os compostos insaturados C18 e o domínio de parametrização dos métodos, e, em parte, as limitações da própria referência DFT, cujo desvio absoluto médio residual reportado por Silva, Custódio e Reis (2015) para o conjunto de ácidos graxos avaliado pelos autores é de 5,3 kJ/mol — valor que descreve o desempenho médio no conjunto de validação do estudo original e não constitui um intervalo de incerteza validado para os quatro compostos C18 analisados neste trabalho.

O padrão de desvio varia com o subgrupo:

Ácidos graxos C18 insaturados (ácido oleico e ácido linoleico). Os três MCG apresentam desvios numericamente maiores neste subgrupo: JR (5,46%), CG (5,03%) e MG (6,18%). Todos subestimam o módulo de ΔH_f° (preveem valores menos negativos do que a referência DFT), padrão compatível com limitações de parametrização ou de representação estrutural para os ácidos graxos insaturados; entretanto, os dados disponíveis não permitem identificar isoladamente a origem dos desvios. O PM7, ao contrário, é o método mais preciso para os ácidos insaturados (MRD de 1,61%), superestimando levemente o módulo de ΔH_f° — comportamento oposto ao dos MCG.

Ésteres metílicos insaturados (metil oleato e metil linoleato). Os desvios dos MCG são numericamente menores: JR (2,81%), CG (1,92%) e MG (2,02%), comparáveis aos observados para ésteres saturados no Nível 1; CG apresenta o menor valor nesse subgrupo (1,92%). O PM7, porém, inverte o padrão observado nos ácidos: apresenta o maior desvio para os ésteres insaturados (MRD de 6,75%), superestimando o módulo de ΔH_f° . O comportamento diferenciado observado entre ácidos graxos e ésteres metílicos evidencia que o desempenho dos métodos depende da classe molecular. Entretanto, os dados disponíveis não permitem identificar isoladamente se essa diferença decorre da parametrização, da representação eletrônica, do tratamento conformacional ou de outros fatores estruturais.

Hierarquia global no Nível 2: CG (3,48%) < MG (4,10%) $\approx$ JR (4,13%) $\approx$ PM7 (4,18%). Diferentemente do Nível 1 — onde PM7 apresentava o menor desvio global — no Nível 2 o CG mostra vantagem modesta sobre os demais para os compostos insaturados testados, com o PM7 limitado por seu desempenho fraco nos ésteres insaturados, ainda que seja o melhor método para os ácidos insaturados.

4.4 ΔH_f° DAS ESTERIFICAÇÕES — NÍVEL 2

4.4.1 Resultados das entalpias de reação

As entalpias de reação foram calculadas para as duas esterificações do Nível 2 pela lei de Hess, combinando, para cada método, as ΔH_f° dos quatro compostos C18 com as ΔH_f° de metanol e água obtidas pelo mesmo método (para o PM7, $\Delta H_f^\circ(\text{H}_2\text{O}) = -238,83 \text{ kJ/mol}$ e $\Delta H_f^\circ(\text{CH}_3\text{OH}) = -204,76 \text{ kJ/mol}$; para os MCG, os respectivos valores de contribuição de grupos, com $\Delta H_f^\circ(\text{H}_2\text{O}) = -241,83 \text{ kJ/mol}$). A referência de reação adotada é a entalpia em fase gasosa, recalculada por lei de Hess

a partir das entalpias de formação gasosas do Nível 2 (referência B3LYP corrigida de Silva, Custódio e Reis, 2015, para os compostos C18, combinada com os valores gasosos de metanol e água): $\Delta H_r^\circ(\text{E1}) = +26,41 \text{ kJ/mol}$ e $\Delta H_r^\circ(\text{E2}) = +2,21 \text{ kJ/mol}$. Essa escolha torna a referência fase-consistente com os métodos avaliados, que produzem ΔH_r° gasoso. Os valores em fase líquida reportados por Silva, Custódio e Reis (2015) incorporam correções de mudança de fase e constituem uma referência complementar para a fase condensada. Entretanto, não representam diretamente as condições operacionais de um processo industrial ou de um reator específico. O desvio absoluto (DA) é a métrica primária, conforme a Seção 3.5. A Tabela 10 apresenta as entalpias de reação, a Tabela 11 reúne os desvios e as métricas agregadas, e a Figura 4 permite comparar os métodos nas duas esterificações.

E1: ácido oleico + metanol $\rightarrow$ metil oleato + água.

E2: ácido linoleico + metanol $\rightarrow$ metil linoleato + água.

Tabela 10 – Entalpias padrão de reação das esterificações E1 e E2

Reação	Referência gasosa	JR	CG	MG	PM7
E1	26,41	-21,56	-13,50	-22,46	-8,69
E2	2,21	-21,55	-13,50	-22,46	-24,87

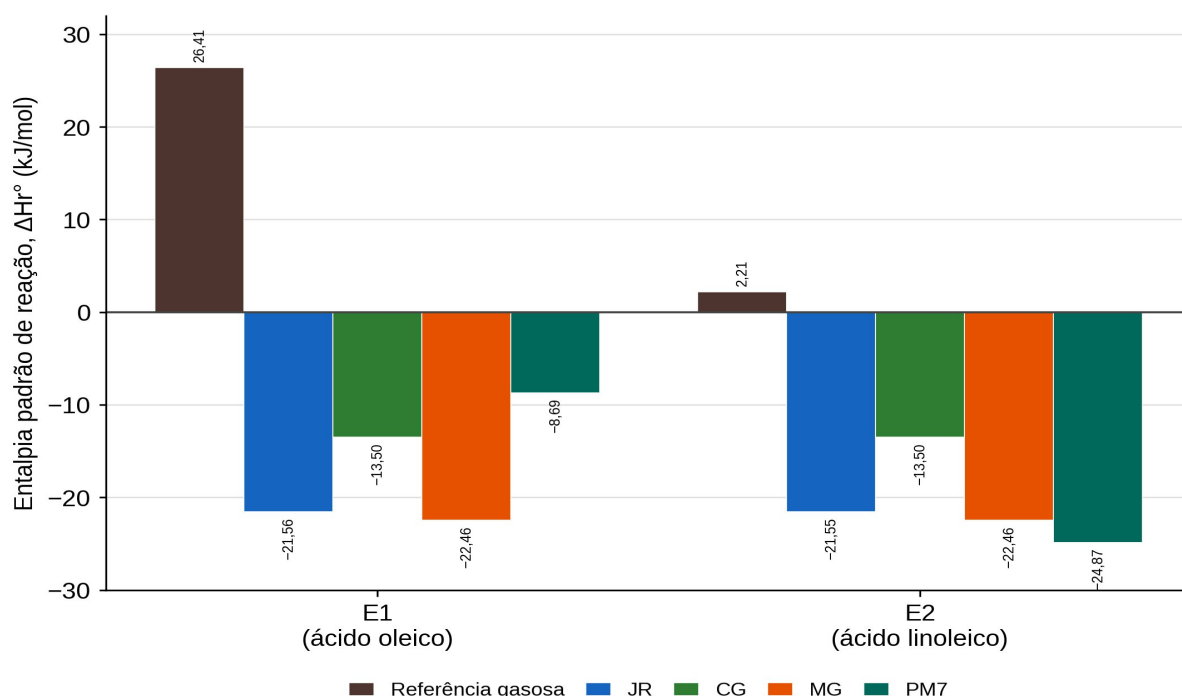
Fonte: elaborado pelo autor.

Tabela 11 – Desvios absolutos e métricas agregadas das esterificações E1 e E2

Métrica	JR	CG	MG	PM7
DA — E1	47,97	39,91	48,87	35,10
DA — E2	23,76	15,71	24,67	27,08
MAE	35,87	27,81	36,77	31,09
RMSE	37,85	30,32	38,71	31,35
MBE	-35,87	-27,81	-36,77	-31,09

Fonte: elaborado pelo autor.

Figura 4 – Entalpias padrão de reação por método para as esterificações E1 e E2



Fonte: elaborado pelo autor com base nos resultados deste trabalho e em Silva, Custódio e Reis (2015).

4.4.2 Análise dos desvios e fenômeno de sinal

O resultado mais notável das esterificações do Nível 2 é que todos os métodos preveem sinal negativo para ambas as reações ($\Delta H_r^\circ \approx -9$ a -25 kJ/mol), enquanto a referência gasosa indica sinal positivo (endotérmico) para ambas — +26,41 kJ/mol para a esterificação do ácido oleico (E1) e +2,21 kJ/mol para a do ácido linoleico (E2). Esse **erro de sinal** afeta tanto os métodos de contribuição de grupos quanto o PM7, configurando uma limitação dos métodos preditivos de baixo custo para a energética dessas reações em fase gasosa. A divergência de sinal é robusta para E1, cuja referência gasosa é claramente positiva (+26,41 kJ/mol) e distante de zero. Como o valor de referência de E2 é próximo de zero, o sinal nominal é sensível a pequenas variações nas entalpias de formação utilizadas em seu cálculo. Por esse motivo, a interpretação deve considerar prioritariamente os desvios absolutos e a proximidade da reação em relação à termoneutralidade.

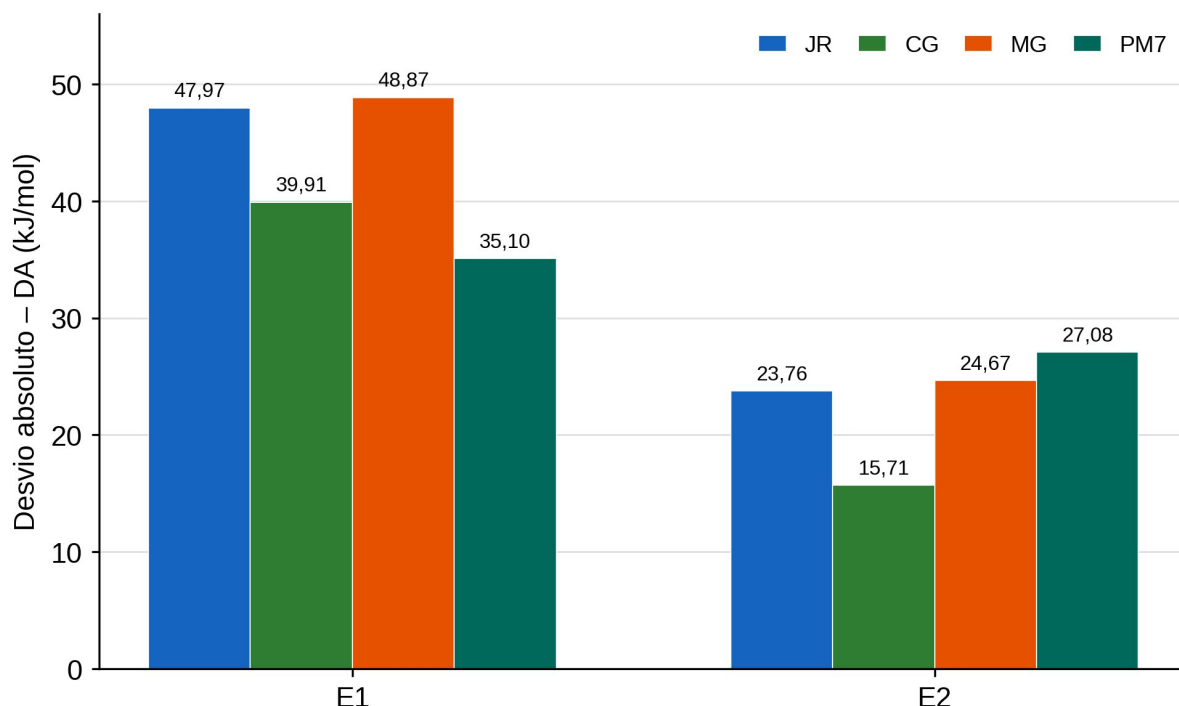
Comportamento semelhante foi relatado anteriormente para a esterificação gasosa do ácido hexadecanoico, conforme Fernandes e Monteiro Filho (2024).

Invariância dos MCG com o grau de insaturação. Uma observação central é que os três MCG preveem praticamente a mesma ΔH_r° para E1 e E2, independentemente do número de duplas ligações: JR prevê $-21,56$ e $-21,55$ kJ/mol, CG prevê $-13,50$ kJ/mol para ambas, e MG prevê $-22,46$ kJ/mol para ambas. Essa invariância decorre da estrutura aditiva dos métodos: as contribuições dos grupos $=CH-$ aparecem igualmente no ácido graxo (reagente) e no éster metílico (produto), cancelando-se exatamente no cálculo de ΔH_r° . Em contraste, a referência gasosa mostra uma diferença de $24,20$ kJ/mol entre as duas esterificações ($+26,41$ vs $+2,21$ kJ/mol), diferença que pode estar associada a limitações de parametrização, representação eletrônica ou tratamento conformacional, sem que o protocolo empregado permita separar quantitativamente essas possíveis contribuições — efeitos que os três métodos de contribuição de grupos avaliados (JR, CG e MG) não capturam, nas decomposições empregadas neste trabalho.

O PM7 apresentou sensibilidade à diferença estrutural entre E1 e E2 e reproduziu parcialmente a separação relativa entre as duas referências, recuperando $16,18$ dos $24,20$ kJ/mol esperados. Entretanto, os resultados não permitem identificar a origem física dessa concordância nem afirmar que o efeito eletrônico associado à insaturação tenha sido corretamente descrito, especialmente porque os valores absolutos e os sinais das entalpias de reação permaneceram incorretos.

Considerando exclusivamente as duas esterificações avaliadas ($n=2$), o CG apresentou os menores valores agregados de MAE e RMSE, com MAE de $27,8$ kJ/mol, seguido por PM7 ($31,1$ kJ/mol), JR ($35,9$ kJ/mol) e MG ($36,8$ kJ/mol). Trata-se, porém, de uma ordenação restrita a apenas duas reações, e não de uma classificação geral de desempenho dos métodos: a vantagem do CG decorre sobretudo de seu menor desvio em E2 ($DA_{CG} = 15,7$ kJ/mol), reação cuja referência é pequena em módulo. Para E1, o PM7 apresenta o menor desvio absoluto ($35,1$ kJ/mol, contra $39,9$ do CG), por prever ΔH_r° menos negativo. Nenhum método, contudo, aproxima-se da referência endotérmica em termos absolutos. A Figura 5 apresenta os desvios absolutos por reação e método.

Figura 5 – Desvios absolutos por reação e método para as esterificações E1 e E2



Fonte: elaborado pelo autor com base nos resultados deste trabalho e em Silva, Custódio e Reis (2015).

4.4.3 Diferenciação por grau de insaturação

Com a referência gasosa adotada, ambas as esterificações são endotérmicas em fase gasosa, sendo a do ácido oleico (E1, +26,41 kJ/mol) mais endotérmica do que a do ácido linoleico (E2, +2,21 kJ/mol), esta última praticamente termoneutra. Nos três métodos de contribuição de grupos avaliados (JR, CG e MG), nas decomposições e ordens empregadas neste trabalho, as contribuições associadas às insaturações se cancelam entre reagentes e produtos, resultando em valores iguais ou praticamente idênticos dentro do arredondamento empregado. O PM7 apresentou sensibilidade à diferença estrutural entre E1 e E2 e reproduziu parcialmente a separação relativa entre as duas referências, recuperando 16,18 dos 24,20 kJ/mol esperados. Entretanto, os resultados não permitem identificar a origem física dessa concordância nem afirmar que o efeito eletrônico associado à insaturação tenha sido corretamente descrito, especialmente porque os valores absolutos e os sinais das entalpias de reação permaneceram incorretos.

Os valores em fase líquida reportados por Silva, Custódio e Reis (2015) incorporam correções de mudança de fase e constituem uma referência complementar para a fase condensada. Entretanto, não representam diretamente as condições operacionais de um processo industrial ou de um reator específico.

4.5 SÍNTESE CRUZADA E IMPLICAÇÕES

Os resultados dos dois níveis revelam três padrões estruturantes:

1. Proximidade de desempenho entre JR, CG, MG e PM7 para ΔH_f° de compostos com dados experimentais (Nível 1). Para as 31 moléculas com dados rastreáveis, os quatro métodos apresentam MRDs na faixa de 1,5–2,2%. Nenhum método é universalmente superior: PM7 apresenta o menor MRD global e se destaca em ésteres, PM3 em ácidos, e MG em álcoois e triglicerídeos.

2. Limitação dos métodos para ΔH_f° de compostos insaturados (Nível 2). Os desvios aumentam para 3,5–4,2% nos compostos insaturados C18. No caso das entalpias de formação, essa perda de desempenho dos MCG está associada sobretudo ao domínio de parametrização: cadeias longas insaturadas são estruturalmente distantes e sub-representadas nas bases experimentais utilizadas na calibração desses métodos. Os MCG tendem a subestimar o módulo de ΔH_f° dos ácidos graxos insaturados, ao passo que o PM7 o superestima nos ésteres insaturados. CG apresenta o melhor desempenho global no Nível 2 (MRD = 3,48%), enquanto o PM7 é o melhor método para os ácidos insaturados (1,61%). A adequação desses desvios para aplicações de triagem depende da sensibilidade do processo, da finalidade do cálculo e do nível de precisão exigido em cada aplicação; o desvio absoluto médio residual de 5,3 kJ/mol reportado por Silva, Custódio e Reis (2015) para o conjunto de ácidos graxos avaliado pelos autores deve ser considerado na interpretação.

3. Diferenciação de ΔH_r° por grau de insaturação: o limite dos MCG e a diferenciação apresentada pelo PM7. Para as entalpias de reação, a invariância dos MCG decorre do cancelamento das contribuições dos grupos $=CH-$ do ácido e do éster na aplicação da lei de Hess. O PM7 apresentou sensibilidade à diferença estrutural entre E1 e E2 e reproduziu parcialmente a separação relativa entre as duas referências, recuperando 16,18 dos 24,20 kJ/mol esperados. Entretanto, os resultados não permitem identificar a origem física dessa concordância nem afirmar que o efeito eletrônico associado à insaturação tenha sido corretamente descrito, especialmente porque os valores absolutos e os sinais das entalpias de reação permaneceram incorretos.

Isso constitui uma limitação dos três métodos de contribuição de grupos avaliados (JR, CG e MG), nas decomposições e ordens empregadas neste trabalho, para avaliações energéticas de rotas de biodiesel com matérias-primas de composição variável em insaturação.

Implicação prática: Para estimativas rápidas de ΔH_f° de compostos insaturados C18 sem dados experimentais, CG oferece o menor desvio global entre os MCG e o PM7 é o mais preciso para os ácidos insaturados. Para avaliar entalpias de reação de esterificações envolvendo graus de insaturação distintos, os três métodos de contribuição de grupos avaliados (JR, CG e MG) não diferenciam as esterificações por grau de insaturação, por construção; o PM7 oferece diferenciação qualitativa, mas aplicações que exijam maior precisão quantitativa devem recorrer à validação por métodos de maior nível teórico, como os de Silva, Custódio e Reis (2015), ou por dados experimentais, quando disponíveis.

5 CONCLUSÕES

5.1 SÍNTESE GERAL

Este trabalho realizou uma análise comparativa sistemática entre três métodos de contribuição de grupos (Joback e Reid, Constantinou e Gani, e Marrero e Gani) e três métodos semiempíricos (AM1, PM3 e PM7, via protocolo CREST+MOPAC) para a estimativa de entalpias padrão de formação e de reação em sistemas relevantes à produção de biodiesel. A avaliação foi estruturada em dois níveis complementares: uma avaliação comparativa quantitativa frente a dados experimentais para 31 moléculas representativas (Nível 1) e uma extensão preditiva controlada a quatro compostos insaturados de cadeia longa (Nível 2), utilizando referências obtidas por DFT segundo a metodologia de Silva, Custódio e Reis (2015).

Os principais resultados são apresentados a seguir, organizados pelos dois níveis metodológicos.

5.2 SÍNTESE DO NÍVEL 1

Para as 31 moléculas do Nível 1 com dados experimentais rastreáveis — ácidos carboxílicos saturados (C2–C16), álcoois (C1–C8 e glicerol), ésteres metílicos (C3–C16) e triglicerídeos de cadeia curta (triacetina e tributirina) — os três métodos de contribuição de grupos e o método PM7 via protocolo CREST → MOPAC apresentaram MRDs globais próximos, na faixa de 1,5–2,2%. PM3 e AM1 apresentaram desempenho global inferior ao dos métodos de contribuição de grupos e do PM7, com AM1 exibindo os maiores desvios entre os métodos analisados.

A análise por grupo molecular confirmou que nenhum método é universalmente superior:

- Para **ácidos carboxílicos**, JR, CG e PM7 atingem MRDs próximos de 0,72–0,74%, com PM3 apresentando o menor desvio nesse grupo (~0,56%) — contrastando com seu desempenho inferior nas demais famílias.
- Para álcoois, MG apresenta o menor MRD (1,61%), seguido de perto pelo PM7 (1,68%); entre os semiempíricos, o PM3 apresenta desvio moderado (4,21%) e o AM1 o maior desvio da classe (17,57%). Os resultados deste trabalho não permitem atribuir esse comportamento do AM1 a um parâmetro ou interação molecular específica.
- Para **ésteres metílicos**, PM7 se destaca com MRD de 1,41%, seguido de CG (2,63%) e JR (2,96%).

- Para **triglicerídeos de cadeia curta** (triacetina e tributirina), MG apresenta os menores desvios ($\sim 3,98\%$), resultado que pode estar relacionado ao maior detalhamento estrutural proporcionado pelas contribuições de primeira e segunda ordem empregadas; entretanto, como o subgrupo contém apenas duas moléculas, esse resultado não permite estabelecer uma vantagem geral do método para triglicerídeos. O PM7 (4,51%) é comparável aos MCG.

O método AM1 apresentou os maiores desvios entre os métodos analisados, com MRD global de 8,86% e desempenho particularmente inferior para os álcoois, classe na qual o MRD atingiu 17,57%; esse desempenho desaconselha seu uso para estimativas termoquímicas em sistemas de biodiesel no conjunto avaliado.

5.3 SÍNTESE DO NÍVEL 2 — ΔH_f°

Para as quatro moléculas insaturadas do Nível 2 (ácido oleico, ácido linoleico, metil oleato e metil linoleato), avaliadas frente à referência B3LYP/6-311+G(d,p) de Silva, Custódio e Reis (2015), os métodos apresentaram MRDs globais de 3,48% (CG), 4,10% (MG), 4,13% (JR) e 4,18% (PM7) — maiores do que os observados no Nível 1; a adequação desses desvios para aplicações de triagem depende da sensibilidade do processo, da finalidade do cálculo e do nível de precisão exigido em cada aplicação, e deve considerar o desvio absoluto médio residual de 5,3 kJ/mol reportado por Silva, Custódio e Reis (2015) para o conjunto de ácidos graxos avaliado pelos autores, que não constitui um intervalo de incerteza validado para os quatro compostos C18 deste trabalho.

Os métodos de contribuição de grupos apresentaram maiores desvios para os ácidos graxos insaturados do que para os ésteres correspondentes, enquanto o PM7 apresentou comportamento oposto.

5.4 SÍNTESE DO NÍVEL 2 — ΔH_r°

Para as duas reações de esterificação do Nível 2 — ácido oleico + metanol (E1) e ácido linoleico + metanol (E2) — adotou-se como referência de avaliação a entalpia de reação em **fase gasosa**, recalculada por lei de Hess a partir das entalpias de formação gasosas do Nível 2 (referência B3LYP corrigida de Silva, Custódio e Reis, 2015, combinada com os valores gasosos de metanol e água). Essa referência fornece ΔH_r° de +26,41 kJ/mol (E1) e +2,21 kJ/mol (E2), indicando que a esterificação gasosa de E1 é endotérmica, enquanto E2 é praticamente termoneutra. A escolha da

fase gasosa é deliberada: os métodos avaliados produzem ΔH_r° gasoso, e a comparação só é legítima entre grandezas da mesma fase.

Todos os métodos — MCG e PM7 — preveem valores negativos (−9 a −25 kJ/mol) para ambas as reações. Para E1, isso configura um claro **erro de sinal** frente à referência endotérmica (+26,41 kJ/mol), robusto por estar a referência distante de zero; comportamento semelhante foi relatado anteriormente para a esterificação gasosa do ácido hexadecanoico, conforme Fernandes e Monteiro Filho (2024). Como o valor de referência de E2 é próximo de zero, o sinal nominal é sensível a pequenas variações nas entalpias de formação utilizadas em seu cálculo. Por esse motivo, a interpretação deve considerar prioritariamente os desvios absolutos e a proximidade da reação em relação à termoneutralidade.

O contraste central desta etapa está entre a **invariância dos MCG** e a **diferenciação apresentada pelo PM7** quanto ao grau de insaturação. Os três MCG fornecem valores iguais ou praticamente idênticos para E1 (C18:1) e E2 (C18:2), devido ao cancelamento das contribuições dos grupos =CH– na lei de Hess. O PM7 apresentou sensibilidade à diferença estrutural entre E1 e E2 e reproduziu parcialmente a separação relativa entre as duas referências, recuperando 16,18 dos 24,20 kJ/mol esperados. Entretanto, os resultados não permitem identificar a origem física dessa concordância nem afirmar que o efeito eletrônico associado à insaturação tenha sido corretamente descrito, especialmente porque os valores absolutos e os sinais das entalpias de reação permaneceram incorretos.

Considerando as duas esterificações avaliadas ($n=2$), o CG apresenta o menor desvio absoluto médio (27,8 kJ/mol), seguido por PM7 (31,1 kJ/mol), JR (35,9 kJ/mol) e MG (36,8 kJ/mol). Essa ordenação restringe-se às duas reações analisadas e não deve ser lida como classificação geral de desempenho dos métodos. Embora nenhum método se aproxime da referência em termos absolutos, apenas o PM7 reproduz a diferença de ΔH_r° entre E1 e E2 — característica que os três métodos de contribuição de grupos avaliados não reproduzem, por construção.

Restrição de fase das conclusões sobre ΔH_r° . As conclusões acima restringem-se à **fase gasosa**, fase em que operam os métodos avaliados e na qual foi construída a referência de desvio. Os valores em fase líquida reportados por Silva, Custódio e Reis (2015) incorporam correções de mudança de fase e constituem uma referência complementar para a fase condensada. Entretanto, não representam dire-

tamente as condições operacionais de um processo industrial ou de um reator específico. A extensão das conclusões à fase líquida ou a condições reais de processo exigiria incorporar correções de vaporização, efeitos de fase ou modelos de solvatação — fora do escopo da avaliação aqui realizada.

5.5 TRABALHOS FUTUROS

Ampliação da avaliação termoquímica. As conclusões deste trabalho sobre o comportamento dos métodos frente à insaturação e aos triglicerídeos apoiam-se em um número reduzido de moléculas e de reações, o que limita seu alcance. A continuidade natural consiste em ampliar o conjunto molecular avaliado, incorporando outros ácidos graxos insaturados — como o ácido linolênico (C18:3) e o ácido erúico (C22:1) —, ésteres etílicos e triglicerídeos de maior cadeia e complexidade, a exemplo da trioleína e da trilinoleína, componentes majoritários dos óleos vegetais de interesse industrial. Para esses compostos, ainda são escassas referências experimentais confiáveis ou de nível teórico superior; sua obtenção, pela metodologia de Silva, Custódio e Reis (2015) ou por protocolos DFT equivalentes, é condição para uma avaliação de Nível 2 expandida e estatisticamente mais robusta.

Aperfeiçoamento da etapa conformacional. No plano metodológico, o protocolo CREST–MOPAC pode ser aprimorado pela incorporação de cálculos vibracionais que confirmem que as geometrias otimizadas correspondem a mínimos verdadeiros, pela análise sistemática da dispersão conformacional e pela comparação entre a seleção do confórmero mais estável — adotada neste trabalho — e médias ponderadas por distribuição de Boltzmann. A essas etapas deve somar-se uma quantificação explícita da incerteza associada às estimativas, ainda ausente na avaliação atual.

Continuidade no projeto Grimperium. O protocolo avaliado neste trabalho tem continuidade direta no Grimperium (Fernandes, 2026), projeto computacional já em desenvolvimento que organiza, em um fluxo computacional automatizado, as etapas aqui exercitadas de forma manual. Nesse fluxo, uma molécula é fornecida por representação estrutural, como SMILES; a busca conformacional é conduzida com CREST/GFN2-xTB; os confórmeros são otimizados e têm o calor de formação calculado com PM7 no MOPAC; seleciona-se a estrutura de menor valor termoquímico entre as avaliadas; calculam-se descritores moleculares; e, por fim, aplica-se uma correção por aprendizado de máquina (Ramakrishnan et al., 2015). A infraestrutura

de geração de dados PM7, de treinamento e de predição em lote já se encontra implementada, mas sua validação dirigida a sistemas lipídicos permanece como trabalho futuro.

Nesta seção, H_{298} designa a propriedade-alvo registrada na base de referência utilizada pelo Grimperium. Antes de interpretá-la como entalpia padrão de formação em fase gasosa, deverá ser confirmada sua equivalência com o calor de formação fornecido pelo MOPAC e com a definição de ΔH_f° adotada neste trabalho. Essa verificação deverá considerar unidades, temperatura de referência, pressão padrão, estado físico, convenções de energia e rastreabilidade dos dados, evitando tratar essas propriedades como automaticamente equivalentes.

$$\Delta_{ML} = H_{298}^{ref} - H_{298}^{PM7} \quad (12)$$

O valor corrigido é obtido pela soma do resultado semiempírico com o resíduo aprendido:

$$H_{298}^{corr} = H_{298}^{PM7} + \Delta_{ML} \quad (13)$$

Referência e validação da propriedade-alvo. No estágio atual, a referência de maior nível empregada no treinamento do Grimperium é baseada em dados CBS-QB3 (Fernandes, 2026) e é distinta da referência adotada no Nível 2 deste trabalho, obtida em B3LYP/6-311+G(d,p) com correção empírica segundo Silva, Custódio e Reis (2015); o cálculo de baixo custo é o PM7 obtido após a busca conformacional. Trata-se de uma camada de correção residual aplicada ao resultado semiempírico, e não de um substituto para métodos de maior nível teórico: o modelo busca aproximar a referência e tem potencial para reduzir e corrigir parte dos desvios sistemáticos do PM7, mas deve ser submetido a validação externa antes de qualquer uso quantitativo. A base de referência CBS-QB3 atualmente utilizada no desenvolvimento do Grimperium deverá ter sua proveniência primária, sua definição termoquímica e seus critérios de curadoria integralmente verificados antes da etapa de treinamento definitivo. Essa verificação deverá incluir autoria, método de geração, unidades, temperatura, estado físico, convenção energética, versão, licença e rastreabilidade dos dados.

Domínio de aplicabilidade do modelo. A base atual do Grimperium é composta predominantemente por moléculas CHON de tamanho pequeno e médio. Embora ácidos graxos e ésteres de biodiesel pertençam ao domínio elementar CHO, moléculas de cadeia longa e, sobretudo, triglicerídeos podem situar-se distantes do domínio estrutural predominante no treinamento, de modo que o desempenho global

do modelo não se transfere automaticamente para compostos C18 ou triglicerídeos. Constituem, portanto, trabalhos futuros: a avaliação formal do domínio de aplicabilidade e a identificação de previsões fora dele; a separação de treinamento, validação e teste por famílias ou arcabouços moleculares e séries homólogas, evitando vazamento entre moléculas muito semelhantes; a validação externa específica para ácidos graxos, ésteres e triglicerídeos; a expansão dirigida da base para sistemas lipídicos; a estimativa de incerteza; a comparação com os métodos de contribuição de grupos JR, CG e MG; e a avaliação dirigida a entalpias de reação, e não apenas a propriedades moleculares isoladas.

Integração com simulação de processos. Como trabalho futuro, propõe-se integrar as estimativas do Grimperium ao DWSIM (Medeiros, 2026) para avaliar como diferentes fontes de dados termoquímicos afetam entalpias de reação, balanços de energia e demandas térmicas em simulações de processos. Essa integração deverá preservar a rastreabilidade dos valores, suas unidades, fase, temperatura de referência e incerteza, sem pressupor que a entalpia de formação isolada seja suficiente para caracterizar um componente.

Limitações e perspectiva de integração. A representação realista de correntes de biodiesel, predominantemente líquidas, requer propriedades adicionais, como capacidades caloríficas, entalpias de vaporização, densidade, viscosidade e parâmetros de equilíbrio de fases. Assim, o protocolo CREST–MOPAC e a correção por Δ -learning podem fornecer parte das informações termoquímicas, enquanto o DWSIM serviria para avaliar seus impactos em processos; a consolidação dessa plataforma depende de validação externa e da obtenção consistente das demais propriedades.

REFERÊNCIAS

- CHAN, B.** Use of Low-Cost Quantum Chemistry Procedures for Geometry Optimization and Vibrational Frequency Calculations: Determination of Frequency Scale Factors and Application to Reactions of Large Systems. **Journal of Chemical Theory and Computation**, v. 13, n. 12, p. 6052–6060, 2017. <https://doi.org/10.1021/acs.jctc.7b00721>
- CONSTANTINOU, L.; GANI, R.** New Group Contribution Method for Estimating Properties of Pure Compounds. **AIChE Journal**, v. 40, n. 10, p. 1697–1710, 1994. <https://doi.org/10.1002/AIC.690401011>
- CONSTANTINOU, L.; GANI, R.; O'CONNELL, J.** Estimation of the Acentric Factor and the Liquid Molar Volume at 298 K Using a New Group Contribution Method. **Fluid Phase Equilibria**, v. 103, p. 11–22, 1995. [https://doi.org/10.1016/0378-3812\(94\)02593-P](https://doi.org/10.1016/0378-3812(94)02593-P)
- FERNANDES, Igor Leno de Souza. Grimperium - Corrigindo Química com Machine Learning. [S. l.]: GitHub, 2026. Disponível em: https://github.com/IgorLeno/grimperium_V2. Acesso em: 12 jun. 2026.
- FERNANDES, I. L. S.; MONTEIRO FILHO, E. S.** Comparação entre métodos de contribuição de grupos e métodos semiempíricos para o cálculo de entalpias padrão de formação em sistemas relevantes à produção de biodiesel. In: **XII Congresso Brasileiro de Termodinâmica – CBTermo**, 2024, Curitiba. Anais [...]. Curitiba: [s.n.], 2024.
- GLIŠIĆ, Sandra; SKALA, Dejan.** The prediction of critical parameters for triolein, diolein, monoolein and methyl esters. In: INTERNATIONAL SYMPOSIUM ON SUPERCRITICAL FLUIDS, 9., 2009, Arcachon, França. **New Trends in Supercritical Fluids: Energy, Materials, Processing**. Arcachon: [s. n.], 2009. Topic: Properties, P018. 6 p. ISBN 978-2-9511591-7-4.
- GREEN, D. W.; PERRY, R. H.** **Perry's Chemical Engineers' Handbook**. 8. ed. New York: McGraw-Hill, 2008.
- HAYNES, W. M. (Ed.).** **CRC Handbook of Chemistry and Physics**. 95. ed. Boca Raton: CRC Press, 2014.
- HUKKERIKAR, A.; SARUP, B.; TEN KATE, A.; ABILDSKOV, J.; SIN, G.; GANI, R.** Group-Contribution+ (GC+) Based Estimation of Properties of Pure Components: Improved Property Estimation and Uncertainty Analysis. **Fluid Phase Equilibria**, v. 321, p. 25–43, 2012. <https://doi.org/10.1016/J.FLUID.2012.02.010>
- JENSEN, F.** **Introduction to Computational Chemistry**. 3. ed. Chichester: John Wiley & Sons, 2016.
- JOBACK, K. G.; REID, R. C.** Estimation of Pure-Component Properties from Group-Contributions. **Chemical Engineering Communications**, v. 57, p. 233–243, 1987. <https://doi.org/10.1080/00986448708960487>

JOSHIPURA, M.; PATEL, R. Modeling Vapor Pressure of Fatty Acid and Fatty Acid Methyl Esters Using Cubic Equations of State. **Asia-Pacific Journal of Chemical Engineering**, v. 10, p. 170–177, 2015. <https://doi.org/10.1002/APJ.1859>

LINSTROM, P. J.; MALLARD, W. G. (Eds.). NIST Chemistry WebBook, NIST Standard Reference Database Number 69. Gaithersburg: National Institute of Standards and Technology, 2023. Disponível em: <https://doi.org/10.18434/T4D303>. Acesso em: 12 jun. 2026.

MARRERO, J.; GANI, R. Group-Contribution Based Estimation of Pure Component Properties. **Fluid Phase Equilibria**, v. 183–184, p. 183–208, 2001. [https://doi.org/10.1016/S0378-3812\(01\)00431-9](https://doi.org/10.1016/S0378-3812(01)00431-9)

MEDEIROS, Daniel Wagner Oliveira de. DWSIM: open source chemical process simulator. [S. l.]: DWSIM, 2026. Disponível em: <https://dwsim.org>. Acesso em: 12 jun. 2026.

PRACHT, P.; BOHLE, F.; GRIMME, S. Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. **Physical Chemistry Chemical Physics**, v. 22, n. 14, p. 7169–7192, 2020. <https://doi.org/10.1039/C9CP06869D>

RAMAKRISHNAN, Raghunathan; DRAL, Pavlo O.; RUPP, Matthias; VON LILIENFELD, O. Anatole. Big Data Meets Quantum Chemistry Approximations: The Δ -Machine Learning Approach. *Journal of Chemical Theory and Computation*, v. 11, n. 5, p. 2087–2096, 2015. DOI: 10.1021/acs.jctc.5b00099.

ROZANSKA, X.; STEWART, J. J. P.; UNGERER, P.; LEBLANC, B.; FREEMAN, C.; SAXE, P.; WIMMER, E. High-Throughput Calculations of Molecular Properties in the MedeA Environment: Accuracy of PM7 in Predicting Vibrational Frequencies, Ideal Gas Entropies, Heat Capacities, and Gibbs Free Energies of Organic Molecules. **Journal of Chemical & Engineering Data**, v. 59, n. 10, p. 3136–3143, 2014. <https://doi.org/10.1021/JE500201Y>

ROZANSKA, X.; UNGERER, P.; LEBLANC, B.; SAXE, P.; WIMMER, E. Automatic and Systematic Atomistic Simulations in the MedeA® Software Environment: Application to EU-REACH. **Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles**, v. 70, p. 405–417, 2015. <https://doi.org/10.2516/OGST/2014041>

SILVA, M. V. D.; CUSTODIO, R.; REIS, M. H. M. Determination of Enthalpies of Formation of Fatty Acids and Esters by Density Functional Theory Calculations with an Empirical Correction. **Industrial & Engineering Chemistry Research**, v. 54, n. 51, p. 12914–12922, 2015. <https://doi.org/10.1021/acs.iecr.5b02580>

STEWART, J. J. P. Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-optimization of Parameters. **Journal of Molecular Modeling**, v. 19, p. 1–32, 2013. <https://doi.org/10.1007/s00894-012-1667-x>

STEWART, J. J. P. MOPAC, versão 23.2.2. Colorado Springs: Stewart Computational Chemistry, 2023. Disponível em: <https://github.com/MOPAC/mopac>. Acesso em: 12 jun. 2026.

YONEDA, Y. An Estimation of the Thermodynamic Properties of Organic Compounds in the Ideal Gas State. I. Acyclic Compounds and Cyclic Compounds with a Ring of Cyclopentane, Cyclohexane, Benzene, or Naphthalene. **Bulletin of the Chemical Society of Japan**, v. 52, p. 1297–1314, 1979. <https://doi.org/10.1246/BCSJ.52.1297>

APÊNDICE A — MÉTRICAS POR CLASSE E MÉTODO (NÍVEL 1)

Tabela 12 – MAE, RMSE e MBE por classe molecular no Nível 1

Classe	n	Método	MAE (kJ/mol)	RMSE (kJ/mol)	MBE (kJ/mol)
Ácidos	10	JR	4,40	6,16	0,66
		CG	4,46	6,67	3,49
		MG	7,60	9,36	7,60
		AM1	29,78	39,04	-29,56
		PM3	2,62	4,44	2,49
		PM7	3,80	4,63	3,78
Álcoois	9	JR	5,18	6,74	-2,81
		CG	5,20	6,40	-5,20
		MG	4,31	5,66	-4,31
		AM1	56,50	60,68	-56,50
		PM3	13,35	15,03	-13,35
		PM7	6,16	8,54	-5,89
Ésteres	10	JR	15,86	18,81	-15,86
		CG	14,11	17,46	-14,11
		MG	17,35	20,31	-17,35
		AM1	32,89	48,87	-31,28
		PM3	13,46	15,55	7,08
		PM7	7,91	13,42	-1,07
Triglicerídeos	2	JR	68,82	69,55	68,82
		CG	64,23	65,40	64,23
		MG	52,87	53,86	52,87
		AM1	93,22	93,43	93,22
		PM3	150,79	151,11	150,79
		PM7	59,62	59,99	59,62

Fonte: elaborado pelo autor.

Nota: a classe dos triglicerídeos contém apenas duas moléculas ($n = 2$); as métricas agregadas devem ser interpretadas com cautela.