

## ANÁLISE DE REGRESSÃO LINEAR DAS PROPRIEDADES TERMODINÂMICAS DOS ISÔMEROS DE PROPANOL E BUTANOL COMO COMBUSTÍVEL ALTERNATIVO

Anna Beatriz Vidal Pavão<sup>1</sup>; Pedro Lucas Silva e Silva<sup>2</sup>; Rafaela Gonçalves Faustino<sup>3</sup>; José de Brito Monteiro Neto<sup>4</sup>; Luana Vitoria Araújo dos Santos<sup>5</sup>; Guilherme Ribeiro Araújo<sup>6</sup>; Raquel Maria Trindade Fernandes<sup>7</sup>; Alamgir Khan<sup>8</sup>

<sup>1</sup>[bia.pavao133@gmail.com](mailto:bia.pavao133@gmail.com);

<sup>2</sup>[pedrolucasmgl63@gmail.com](mailto:pedrolucasmgl63@gmail.com);

<sup>3</sup>[rafaela.faustino27@gmail.com](mailto:rafaela.faustino27@gmail.com);

<sup>4</sup>[britomonteiro712@gmail.com](mailto:britomonteiro712@gmail.com);

<sup>5</sup>[luana.v.santos02@gmail.com](mailto:luana.v.santos02@gmail.com);

<sup>6</sup>[guilherme.araujo1201@gmail.com](mailto:guilherme.araujo1201@gmail.com);

<sup>7</sup>[raquelfernandes@professor.uema.br](mailto:raquelfernandes@professor.uema.br);

<sup>8</sup>[alamgir@cecen.uema.br](mailto:alamgir@cecen.uema.br)

**Palavras-Chave:** RNA, Termodinâmica, Combustível.

### Introdução

O desenvolvimento da humanidade, ao longo de muitos anos, garantiu melhores índices de conforto e longevidade devido a avanços na agricultura, na medicina, dentre outros. A partir destes, a densidade populacional no planeta vem aumentando e, com isso, também aumenta a procura por mais recursos energéticos, causando impactos ambientais que vêm sendo discutido mundialmente, mediante a conscientização da gravidade da questão.

Nesse sentido, a crescente preocupação com as questões ambientais e a conscientização mundial sobre a promoção do desenvolvimento em bases sustentáveis vêm estimulando a realização de pesquisas de desenvolvimento tecnológico que visam à incorporação dos efeitos da aprendizagem e a consequente redução dos custos de geração dessas tecnologias (FREITAS & DATHEIN, 2013).

Atualmente, os recursos naturais e renováveis tem sido o foco de inúmeras pesquisas, impulsionadas pelo aumento das preocupações com o meio ambiente, devido aos problemas ecológicos e do aquecimento global, gerados pela utilização de combustíveis fósseis. O aproveitamento correto das fontes renováveis é um excelente modo de substituir as “energias sujas” e evitar danos ao planeta (AZEVEDO, 2013).

Deste modo, o principal objetivo deste trabalho é analisar a relação entre as propriedades termodinâmicas e efeitos das estruturas isoméricas das moléculas do propanol e do butanol e a produção de combustíveis alternativos, utilizando métodos semiempíricos, juntamente com técnicas estatísticas multivariadas e Redes Neurais Artificiais (RNA).

As redes neurais artificiais (RNAs) são sistemas computacionais inspirados na estrutura e no funcionamento do cérebro humano. Elas consistem em camadas de neurônios artificiais que processam informações, permitindo que as RNAs aprendam e façam previsões com base em grandes volumes de dados.

No contexto das aplicações em química e física, as RNAs têm sido utilizadas para modelar propriedades termodinâmicas de combustíveis alternativos, como o propanol e o butanol, como foi o caso neste estudo. Por exemplo, Goodfellow, Bengio e Courville (2016) destacam que as RNAs podem ser treinadas para prever com alta precisão parâmetros

termodinâmicos, como entalpia, entropia e capacidade calorífica, mesmo em condições variáveis e complexas.

### Material e Métodos

Neste trabalho foram utilizados os bancos de dados das propriedades termodinâmicas de a Entropia (S), a Entalpia (H), Capacidade calorífica (C) e Energia livre de Gibbs (G) computados pelo nosso grupo (SERRA, K. F. da C. Et.al., 2023).

A partir desses dados, utilizou-se Redes Neurais Artificiais (RNA), com foco na arquitetura Multi-Layer Perceptron (MLP), para prever e analisar a relação entre essas variáveis nos álcoois propanol (1-propanol e 2-propanol) e butanol (1-butanol, 2-butanol e terc-butanol). Aonde, a propriedade termodinâmica de interesse, a energia livre de Gibbs (G), foi considerada como a variável dependente (Y) como *alvo*. As variáveis independentes (X) incluíram a entropia (S), a entalpia (H) e a capacidade calorífica (C) como *input*. Então, assim a arquitetura MLP configurada para o estudo foi composta de:

- Camada de Entrada: Incluiu os neurônios correspondentes às variáveis independentes (S, H e C).
- Camadas Ocultas: Foram ajustadas para otimizar o desempenho, utilizando um número de camadas e neurônios baseado em experimentos de validação cruzada.
- Camada de Saída: Composta por um único neurônio que representava a variável dependente, a energia livre de Gibbs (G).

A tabela 1 é uma representação das variáveis utilizadas, seus valores de entradas e suas funções de ativação.

**Tabela 1:** Variáveis utilizadas

| TIPO DE REDE    | PESO DE DECAIMENTO |                 |
|-----------------|--------------------|-----------------|
| MLP             | Camada oculta      | CAMADA DE SAÍDA |
| MÍN. OCULTO: 3  | Mín.: 0,0001       | MÍN.: 0,0001    |
| MÁX. OCULTO: 10 | MÁX.: 0,001        | MÁX.: 0,001     |

**Fonte:** Autoral, 2024.

Neste estudo, foram geradas 2000 redes neurais diferentes durante o processo de treinamento. Dentre as 2000 redes testadas, foram selecionadas as 5 melhores para a Energia Livre de Gibbs em cada base utilizada. O refinamento das redes é de fundamental importância para uma análise mais precisa dos dados.

Além disso, o tamanho da amostra de treinamento (train sample size) foi 90%, o que significa que 90 % amostras dos dados foram usadas para treinar a rede neural. Já o tamanho da amostra de teste (test sample size) foi 10%, indicando que 10 % amostras dos dados foram reservadas para testar a performance do modelo treinado, permitindo avaliar sua capacidade de generalizar para novos dados. As funções de ativação foram *Identity*, *Logistic*, *Tanh*, *Exponential* e *Sine* foram aplicadas às saídas dos neurônios das redes, tanto nas camadas ocultas (hidden neurons) quanto na camada de saída (output neurons).

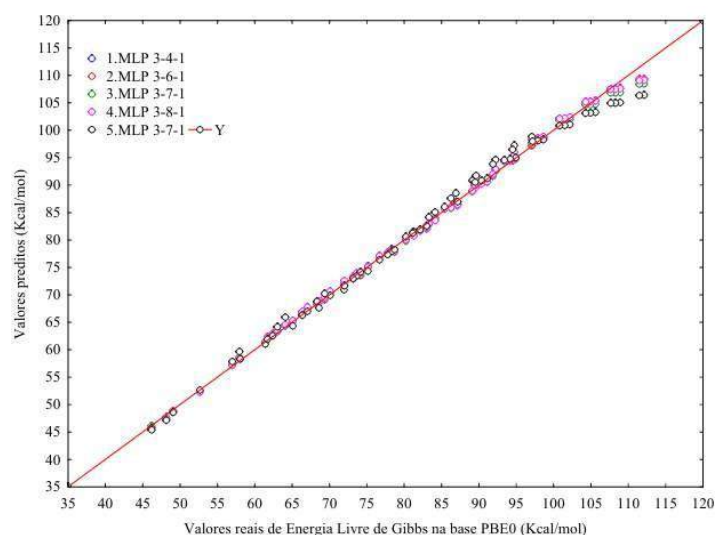
Para a realização do estudo, os softwares MS Excel e Statistica7 foram utilizados. O Excel foi usado principalmente para a organização e visualização dos dados, visto que ele

permite a criação de tabelas, gráficos e realizar cálculos básicos, o que ajuda a estruturar e entender os dados coletados. O Statistica foi empregado para análises estatísticas mais avançadas já que ele oferece uma gama de ferramentas para análise de dados, incluindo testes estatísticos, análises multivariadas e modelagem preditiva.

## Resultados e Discussão

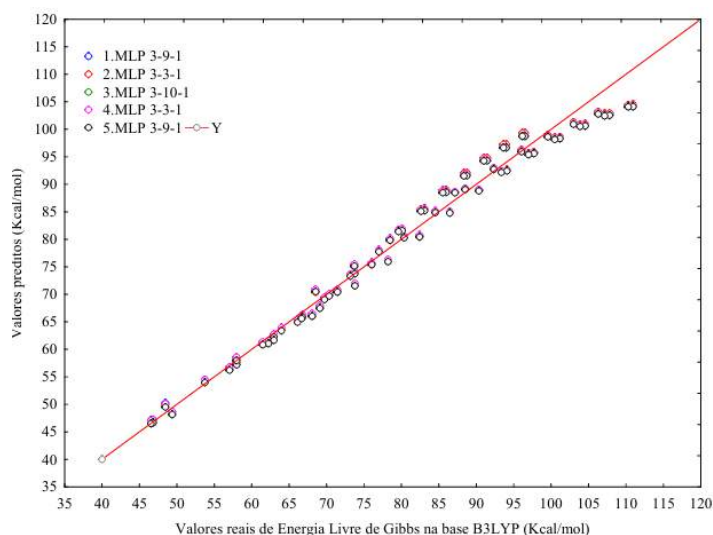
Foram criados gráficos que ilustram a relação entre os valores reais (dados) e os valores preditos de Energia Livre de Gibbs, utilizando diferentes bases computacionais (PBE0, B3LYP, PM6) e redes neurais artificiais do tipo Multilayer Perceptron (MLP) e gráficos que ilustram a relação entre os valores reais e os erros das redes utilizadas.

**Figura 1:** Relação entre os valores reais e preditos de Energia Livre de Gibbs para a base PBE0



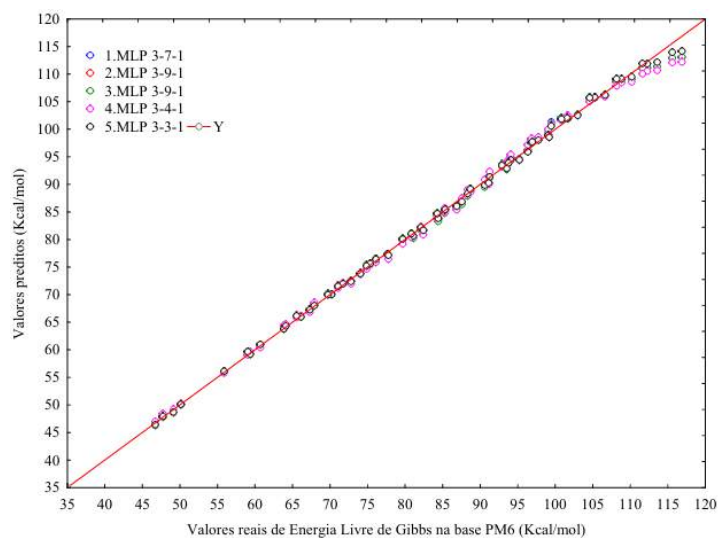
Fonte: Autoral, 2024.

**Figura 2:** Relação entre os valores reais e preditos de Energia Livre de Gibbs para a base B3LYP



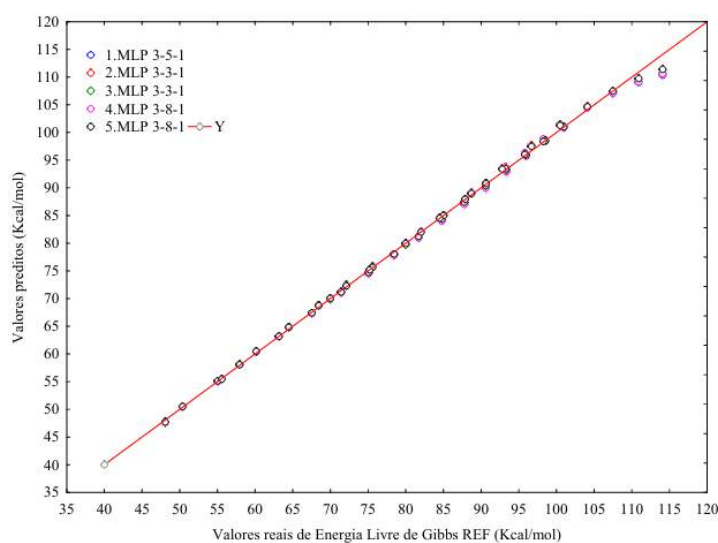
Fonte: Autoral, 2024.

**Figura 3:** Relação entre os valores reais e preditos de Energia Livre de Gibbs para a base PM6



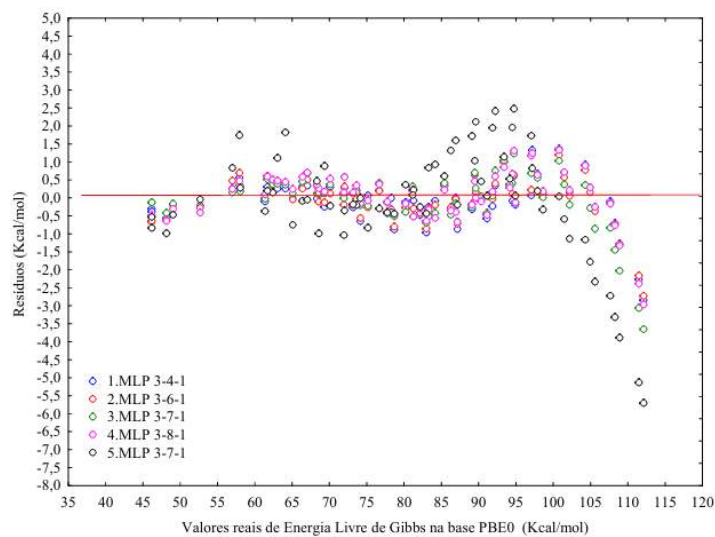
Fonte: Autoral, 2024.

**Figura 4:** Relação entre os valores reais e preditos de Energia Livre de Gibbs para os valores de referência



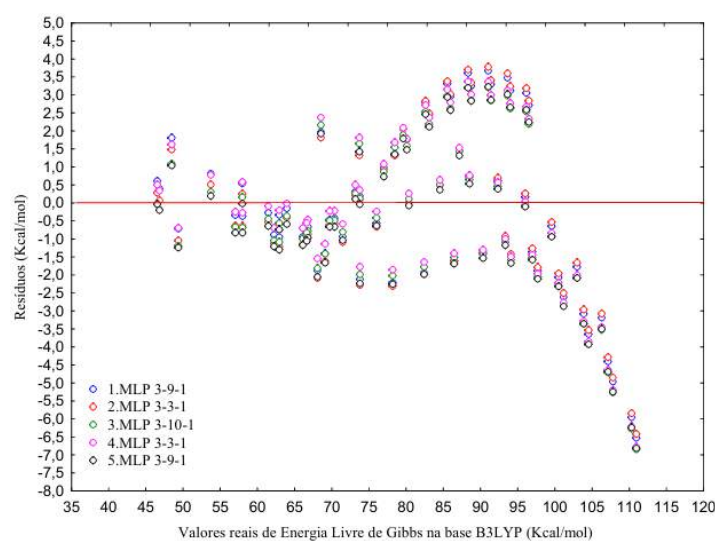
Fonte: Autoral, 2024.

**Figura 5:** Relação entre os valores reais de Energia Livre de Gibbs e os erros das redes testadas para a base PBE0



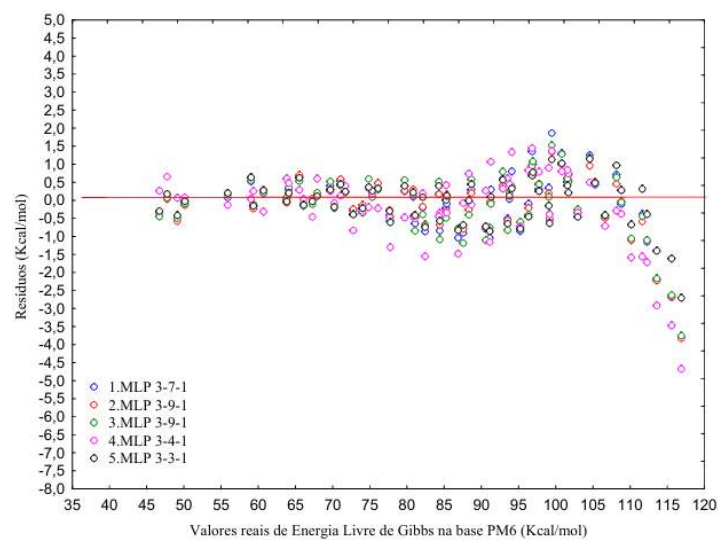
Fonte: Autoral, 2024.

**Figura 6:** Relação entre os valores reais de Energia Livre de Gibbs e os erros das redes testadas para a base B3LYP



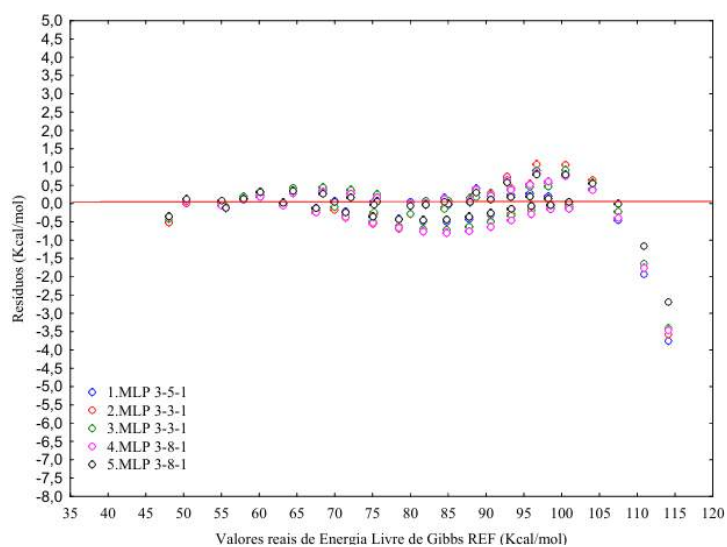
Fonte: Autoral, 2024.

**Figura 7:** Relação entre os valores reais de Energia Livre de Gibbs e os erros das redes testadas para a base PM6



**Fonte:** Autoral, 2024.

**Figura 8:** Relação entre os valores reais de referência da Energia Livre de Gibbs e os erros das redes testadas



**Fonte:** Autoral, 2024.

A partir da análise das figuras de 1 à 4, é possível perceber que as redes neurais artificiais do tipo MLP, mesmo configuradas de forma relativamente simples, conseguem captar de maneira eficaz a relação entre as variáveis independentes (S, H, C) e a Energia Livre de Gibbs. No entanto, a eficácia dessas previsões varia dependendo da base computacional utilizada, o que sugere que a escolha da base é um fator crítico para a precisão das previsões.

Além disso, a variação no número de neurônios nas camadas ocultas das MLPs influencia o desempenho preditivo, com certas configurações fornecendo melhores previsões em determinadas bases.

Já nas figuras 5, 6, 7 e 8, os resíduos observados mostram que, por mais que as redes tenham conseguido capturar algumas das características dos dados e tenham conseguido se aproximar dos valores reais de entrada, ainda tem uma grande variabilidade, logo ainda há espaço para melhorias.

O coeficiente de determinação ( $R^2$ ) também foi calculado (Tabela 1 a Tabela 4), entre os valores de target (valores reais ou esperados) e os valores de output (valores previstos pelo modelo) de cada rede neural treinada e comparado com o erro quadrático médio (RMSE), que foi calculado a partir da média dos quadrados dos resíduos (erros das redes treinadas).

**Tabela 2** - Comparação entre o  $R^2$  e o RMSE da Energia Livre de Gibbs utilizando métodos DFT

| MÉTODOS                      | REDES      | R2     | RMSE        |
|------------------------------|------------|--------|-------------|
| <b>PBE0</b>                  | MLP 3-4-1  | 99,85% | <b>0,65</b> |
|                              | MLP 3-6-1  | 99,86% | <b>0,64</b> |
|                              | MLP 3-7-1  | 99,79% | <b>0,8</b>  |
|                              | MLP 3-8-1  | 99,83% | <b>0,71</b> |
|                              | MLP 3-7-1  | 99,24% | <b>1,53</b> |
| <b>B3LYP</b>                 | MLP 3-9-1  | 98,21% | <b>2,27</b> |
|                              | MLP 3-3-1  | 98,18% | <b>2,28</b> |
|                              | MLP 3-10-1 | 98,29% | <b>2,25</b> |
|                              | MLP 3-3-1  | 98,28% | <b>2,25</b> |
|                              | MLP 3-9-1  | 98,28% | <b>2,27</b> |
| <b>PM6</b>                   | MLP 3-7-1  | 99,77% | <b>0,86</b> |
|                              | MLP 3-9-1  | 99,81% | <b>0,81</b> |
|                              | MLP 3-9-1  | 99,78% | <b>0,86</b> |
|                              | MLP 3-4-1  | 99,67% | <b>1,06</b> |
|                              | MLP 3-3-1  | 99,87% | <b>0,64</b> |
| <b>VALORES DE REFERÊNCIA</b> | MLP 3-5-1  | 99,86% | <b>0,59</b> |
|                              | MLP 3-3-1  | 99,83% | <b>0,65</b> |
|                              | MLP 3-3-1  | 99,85% | <b>0,6</b>  |
|                              | MLP 3-8-1  | 99,84% | <b>0,62</b> |
|                              | MLP 3-8-1  | 99,92% | <b>0,44</b> |

Fonte: Autoral, 2024.

No estudo realizado, foi possível observar que o erro quadrático médio dos resíduos da propriedade G2 (Energia Livre de Gibbs na base B3LYP) foi o mais elevado, quando comparado ao encontrado em outras bases, o que nos diz que as predições realizadas pelas redes neurais apresentaram uma maior variação em comparação com os valores reais para essa base específica. Entretanto o coeficiente de determinação ( $R^2$ ) teve uma média de 98,2%, o que indica que, mesmo com um erro quadrático médio mais elevado, a maioria da variação nos dados ainda foi explicada de forma satisfatória pelas redes neurais.

### Conclusões

O presente estudo investigou o uso das Redes Neurais Artificiais (RNAs) para analisar parâmetros termodinâmicos de álcoois, mais especificamente o butanol e propanol, visando sua aplicação como combustíveis alternativos benignos. A análise dos dados obtidos mostrou que as redes neurais podem ser ferramentas eficazes para prever propriedades como a Entalpia, Entropia, Capacidade Calorífica e Energia Livre de Gibbs, proporcionando uma compreensão mais detalhada da eficiência desses combustíveis.

A partir dos resultados, conclui-se que tanto o butanol quanto o propanol apresentam características promissoras como combustíveis alternativos, com o butanol demonstrando ligeira superioridade em termos de eficiência termodinâmica. Desta forma, estudos adicionais são indispensáveis para validar esses achados em condições práticas e ampliar o conhecimento sobre a viabilidade desses combustíveis em larga escala.

### Agradecimentos

Gratidão ao programa PIBIC/UEMA pela concessão de bolsa, laboratório de Físico-Química (LFQ) do curso de Química e ao grupo de pesquisa NEQTA, obrigada. Agradeço à minha família e a Deus por tudo.

### Referências

FREITAS, G.C.; DATHEIN, R. As energias renováveis no Brasil: uma avaliação acerca das implicações para o desenvolvimento socioeconômico e ambiental. **Revista Nexos Econômicos**. v. 7, n. 1, p. 71-94, 2013.

AZEVEDO, P.J.S. Uma análise dos efeitos da crise econômico-financeira sobre as políticas de incentivo às energias renováveis. [Dissertação] Universidade do Porto, 2013.

GOODFELLOW, I., BENGIO, Y., & COURVILLE, A. (2016). *Deep Learning*. Cambridge: MIT Press.

SERRA, K. F. da C.; KHAN, A.; FERNANDES, R. M. T.; VAZQUEZ, P. A. M.; KHAN, A. Abordagem de análise estatística multivariada para investigar as quantidades termodinâmicas do combustível alternativo benigno. *Journal of the Serbian Chemical Society*, 2023. Disponível em: <https://doi.org/10.2298/JSC230530090S>. Acesso em: 22 de setembro de 2024.