

OTIMIZAÇÃO VIA MACHINE LEARNING DA EFICIÊNCIA FOTOCATALÍTICA DE ZNO VERDE A PARTIR DE CASCA DE NOZ-PECÃ: CORRELAÇÃO ENTRE PARÂMETROS DE SÍNTESE E DEGRADAÇÃO DA RODAMINA B

NATAN DILL¹; JOÃO LOPES²; TIAGO VOLKMER³; MATEUS FERRER⁴;

¹Universidade Federal de Pelotas- UFPel – natanh28@gmail.com

²Universidade Federal de Pelotas- UFPel – lopes.a.joaopedro@gmail.com

³Universidade Federal de Pelotas- UFPel – tiagovolkmer@gmail.com

⁴Universidade Federal de Pelotas- UFPel – mateusmferrer@gmail.com

1. INTRODUÇÃO

A busca por métodos sustentáveis na síntese de materiais tem ganhado destaque frente aos desafios ambientais das rotas convencionais, que frequentemente envolvem reagentes tóxicos, alto consumo energético e geração de resíduos. Neste contexto, a síntese verde surge como uma alternativa promissora, ao usar extratos vegetais como a casca de noz-pecã (*Carya illinoensis*), aliando baixo custo, baixa toxicidade e valorização de resíduos agrícolas [1].

O óxido de zinco (ZnO) obtido por rotas verdes apresenta propriedades fotocatalíticas relevantes para aplicações ambientais, como a degradação de poluentes orgânicos resistentes. Um exemplo é a Rodamina B, corante amplamente encontrado em efluentes industriais e que apresenta alta resistência à degradação natural [2].

No entanto, a otimização desses processos é complexa, envolvendo múltiplas variáveis inter-relacionadas - como pH, temperatura, concentração de precursor e tempo - o que torna os métodos tradicionais de estatística limitados para capturar relações não lineares. Diante dessa complexidade, algoritmos de Machine Learning (ML) têm se mostrado ferramentas promissoras na modelagem e predição de processos químicos não lineares. Técnicas como Random Forest (RF), Artificial Neural Network (ANN) e Extreme Gradient Boosting (XGBoost) são capazes de identificar padrões complexos e prever resultados com base em múltiplas variáveis experimentais, otimizando recursos e acelerando o desenvolvimento de novos materiais e processos. A aplicação de algoritmos de ML sobre dados de fotocatalise possibilita a predição da eficiência de degradação com base em parâmetros de síntese do material, ranqueando variáveis críticas e reduzindo a quantidade de experimentos necessários [3].

Neste contexto, este trabalho tem como objetivo empregar e comparar os algoritmos RF, ANN e XGBoost para modelar a eficiência fotocatalítica de ZnO verde, correlacionando parâmetros de síntese com a degradação de Rodamina B, a fim de identificar o modelo de melhor desempenho preditivo e as melhores condições experimentais.

2. METODOLOGIA

2.1 Construção do banco de dados

Para a construção do banco de dados, foi realizado um delineamento fatorial 2³ com ponto central, sendo as variáveis de entrada a temperatura de calcinação do ZnO (400, 450 e 500 °C), o pH (4, 7 e 10) e a concentração do catalisador (25, 50 e

75 mg), já a variável de resposta foi a eficiência da degradação (0 - 100%). Os experimentos planejados forneceram a base para a etapa seguinte, na qual foi avaliado o desempenho fotocatalítico do ZnO.

2.2 Fotocatálise

A fotocatalise foi realizada com o objetivo de acompanhar a degradação da Rodamina B ($C\ 1 \times 10^{-5}\ M$, $V\ 50\ ml$) em diferentes condições. As quantidades de catalisador nas condições pré-determinadas foram colocadas em uma caixa de luz UV equipada com 5 lâmpadas. As amostras foram agitadas constantemente a temperatura ambiente durante 3 horas. Nos primeiros 15 minutos, as amostras foram agitadas na ausência de luz UV e uma coleta foi realizada para análise de cada amostra. Após esse período, as lâmpadas UV foram ligadas, e coletas foram realizadas com tempos de 5, 15, 30, 45, 60, 75, 90, 120 e 180 minutos para monitorar a degradação do corante.

A degradação foi monitorada por espectrofotometria UV-Vis a $\lambda = 554\ nm$.

2.3 Modelos de Machine Learning

Os dados do planejamento experimental foram coletados em planilha eletrônica e processados utilizando a biblioteca Pandas do Python 3.10 (Google Colab). Posteriormente, foram normalizados com a função MinMaxScaler da biblioteca Scikit-learn. Os dados foram divididos em 80% para treino e 20% para teste, aplicando validação cruzada ($cv = 5$) para todos os algoritmos.

Para desenvolver um modelo preditivo generalizado, testaram-se 3 algoritmos de aprendizado selecionados devido à sua capacidade de capturar relações não-lineares e de operar eficientemente com conjuntos de dados de tamanho moderado: Random Forest (RF), Artificial Neural Network (ANN) e Extreme Gradient Boosting (XGBoost). Cada modelo foi ajustado por meio de diferentes hiperparâmetros, mostrados na Tabela 1, buscando maior precisão e minimizando sobre ajustes (overfitting) [OVIEDO, 2025].

As bibliotecas NumPy e Scikit-learn foram usadas para implementação dos modelos e avaliação de métricas, as quais foram baseadas nos valores de Coeficiente de Determinação (R^2) e na Raiz do Erro Quadrático Médio (RMSE), garantindo uma análise sólida para comparação entre os modelos.

Tabela 1: Configurações de hiperparâmetros e métricas dos modelos de ML.

Modelo	Hiperparâmetros testados	Intervalos de variação
Random Forest (RF)	Número de árvores (n_estimators), profundidade máxima (max_depth)	n_estimators = [10, 50, 100, 200] max_depth = [3, 5, 7, 10]
XGBoost (XGB)	Número de árvores (n_estimators), profundidade máxima (max_depth)	n_estimators = [10, 50, 100, 200] max_depth = [3, 5, 7, 10]
Rede Neural Artificial (ANN)	Número de camadas ocultas, número de neurônios por camada, função de ativação	camadas = [1, 2, 3], neurônios = [16, 32, 64], ativação = [Relu, Tanh]

3. RESULTADOS E DISCUSSÃO

A Tabela 2 apresenta os resultados do delineamento fatorial 2^3 com ponto central, revelando a forte influência dos parâmetros de síntese na eficiência fotocatalítica. Constata-se também que todas as condições testadas resultaram uma degradação superior a 50% da Rodamina B em 90 minutos, confirmando o potencial do material sintetizado. Destaca-se o experimento 8 (500 °C, pH 10, 75 mg L⁻¹), que atingiu a maior eficiência de degradação (98,97%), sugerindo uma combinação ideal entre temperatura de calcinação, alcalinidade do meio e concentração do catalisador.

Tabela 2: Resultados de %R para o tempo de 90 minutos de fotocatalise.

Experimento	Temperatura (°C)	pH	[ZnO] (g L ⁻¹)	%R
1	400	4	25	54.642
2	400	4	75	97.852
3	400	10	25	94.338
4	400	10	75	91.5
5	500	4	25	61.414
6	500	4	75	96.342
7	500	10	25	93.288
8	500	10	75	98.966
9	450	7	50	96.428
10	450	7	50	94.116

Para prever essa complexidade não linear, três algoritmos de Machine Learning (ML) foram avaliados. Como mostra a Tabela 3, os modelos Random Forest (RF) e XGBoost exibiram desempenho preditivo superior e muito parecido, ambos com coeficiente de determinação (R^2) acima de 0,93 e erro médio (RMSE) próximos de 9,8 no conjunto de teste. O XGBoost mostrou-se ligeiramente superior, devido à sua eficiência algorítmica em identificar relações complexas em conjuntos de dados de tamanho moderado. A Artificial Neural Network (ANN), por outro lado, apresentou o menor desempenho com um $R^2 = 0,87$ na fase de teste, possivelmente por requerer volumes maiores de dados para generalização adequada neste estudo específico.

As configurações que tiveram melhor desempenho estão listadas na Tabela 3.

Tabela 3: Configurações que tiveram melhor desempenho

Modelo	Melhor configuração encontrada	R2 treino	R2 teste	RMSE treino	RMSE teste
Random Forest (RF)	n_estimators = 50 max_depth = 5	0.9709	0.9317	6.2267	9.9810
XGBoost (XGB)	n_estimators = 100 max_depth = 3	0.9679	0.9343	6.5392	9.7895
Rede Neural Artificial (ANN)	3 camadas (32 neurônios/camada) ativação: ReLU	0.9211	0.8693	10.2446	13.8032

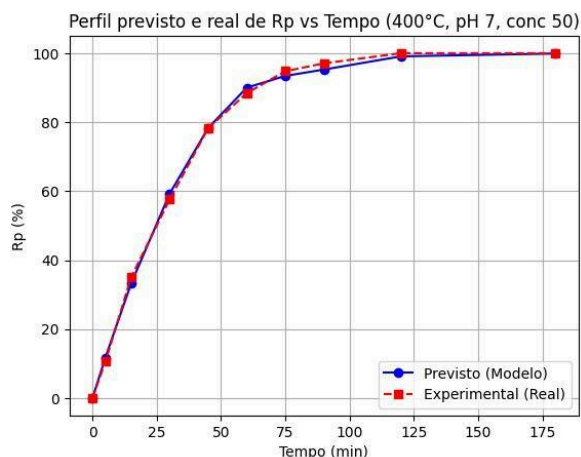


Figura 1: Curva de degradação prevista vs. experimental para XGBoost.

Conforme a Tabela 3, o melhor modelo com os melhores hiperparâmetros ajustados, foi utilizado para tentar prever o desempenho fotocatalítico de um conjunto de variáveis (400, 7, 50) que não estavam no banco de dados da Tabela 2. A Figura 1 mostra as curvas prevista pelo modelo e a real obtida experimentalmente. Como podemos observar, a curva prevista pelo modelo ficou bem próxima a real, validando a capacidade de predição/generalização do modelo escolhido, com Mean Absolute Error (MAE): 1,05.

4. CONCLUSÃO

Este trabalho mostrou que algoritmos de ML, com destaque para o modelo XGBoost, podem modelar e prever com precisão ($R^2 > 0,93$) a eficiência de fotocatalisadores de ZnO sintetizados por rota verde a partir de casca de noz-pecã. O modelo identificou condições ótimas de síntese que atingiram > 98% de degradação da Rodamina B, mostrando que a incorporação entre síntese verde e inteligência artificial é uma estratégia promissora para otimizar nanomateriais sustentáveis em remediação ambiental.

5. REFERÊNCIAS

- [1] NORUZI, M. Biossíntese de nanopartículas de ouro usando extratos vegetais. **Engenharia de Bioprocessos e Biosistemas**, [S. l.], v. 1, n. 1, p. 1-10, ago. 2014. DOI: <https://doi.org/10.1007/s00449-014-1251-0>.
- [2] DE SOUZA, E. F.; PORTO, M. B.; POMPERMAYER, N. B.; BERGAMO, M. H. da S. Comparação dos processos de síntese e do desempenho de fotocatalisadores para a degradação do corante rodamina B. **Engenharia Sanitária e Ambiental**, Rio de Janeiro, v. 23, n. 4, p. 707-718, 2018. DOI: <https://doi.org/10.1590/S1413-41522018149826>.
- [3] OVIEDO, L. R. Nanocatalisador à base de nanozeólita e nanopartículas verdes de óxido de cobre para aplicação em processos avançados de tratamento associado com a inteligência artificial. 2025. Tese (Doutorado em Nanociências) – Programa de Pós-Graduação em Nanociências, Universidade Franciscana, Santa Maria, 2025.