

## Modelos de redes neurais artificiais para produção de hidrogênio por consórcios de microrganismos

### Artificial neural networks to predict hydrogen production from microorganisms consortia

#### RESUMO

**Byanka Damian Mizerkowski**  
[Byanka\\_dm@yahoo.com.br](mailto:Byanka_dm@yahoo.com.br)  
Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil

**Elis Regina Duarte**  
[erduarte@utfpr.edu.br](mailto:erduarte@utfpr.edu.br)  
Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil

**Eduardo Bittencourt Sydney**  
[eduardosydney@utfpr.edu.br](mailto:eduardosydney@utfpr.edu.br)  
Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil

**Walter Jose Martinez Burgos**  
[lng.wjmartinez@gmail.com](mailto:lng.wjmartinez@gmail.com)  
Universidade Federal do Paraná, Curitiba, Paraná, Brasil

Microrganismos decompositores são fundamentais para fermentação da matéria orgânica, tratando efluentes e produzindo compostos de interesse como alternativa energética, tal como o hidrogênio. Redes neurais artificiais (RNAs) constituem ferramentas robustas para a modelagem do processo, portanto o objetivo desse trabalho foi a construção de RNAs para predição de produção de hidrogênio por consórcios de organismos. Avaliou-se três consórcios: Esg (efluente industrial), Lod (lodo de efluentes domésticos) e Vac (resíduos bovinos); com quatro variáveis de entrada (temperatura, pH, % de inóculo e razão C:N do meio), e uma de saída (volume de hidrogênio produzido). As RNAs foram desenvolvidas em ambiente MATLAB, do tipo *feedforward*, 3 camadas e algoritmo de treinamento *backpropagation* Levenberg-Marquardt, testando-se número de neurônios e as partições de dados. Para os consórcios individuais, a rede com 20 neurônios na camada oculta e uso de 80% dos dados para treinamento apresentou melhor adequação, com coeficientes de correlação entre dados alvo e saída entre 0,87 e 0,97, enquanto para o conjunto total de dados, o uso de 10 neurônios foi de  $R=0,88$ . As RNAs mostraram-se eficientes para descrever o processo, mas fatores como o aumento do banco de dados e inserção de identificação de espécies pode aumentar a confiabilidade do modelo proposto.

**PALAVRAS-CHAVE:** Tratamento de efluentes. Alternativas energéticas. Modelagem.

**Recebido:** 19 ago. 2019.

**Aprovado:** 01 out. 2019.

**Direito autorial:** Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



#### ABSTRACT

Fermentative microorganisms decompose organic matter to treat effluents and produce energy alternative products, such as hydrogen. Artificial neural networks (ANNs) are robust tools to model such process. This work aims at predicting hydrogen production by microorganism consortia through artificial neural network. Three consortia were analyzed: Esg (industrial effluent), Lod (domestic effluent sludge), Vac (cattle manure); four input variables (temperature, pH, % of inoculum, C:N medium ratio) and one output (volume of produced hydrogen). Three layers feedforward neural networks were developed using MATLAB and trained by Levenberg-Marquardt backpropagation algorithm, to test number of neurons and partition of data. For each individual consortia, correlation coefficients were highest ( $R$  from 0.87 to 0.97) for 20 neuron network and partitioning data into 80% for training, 10% for validation and test. Considering the whole set of data resulted in  $R=0.88$  with a 10 neuron network and same data partition. Results show that ANNs are efficient to

predict hydrogen production by microorganisms, but factor such as increasing the data bank e inserting other input variables might increase reability.

**KEYWORDS:** Effluent treatment. Alternative energy. Modelling.

## INTRODUÇÃO

A digestão anaeróbica de matéria orgânica é um processo amplamente utilizado para tratamento biológico de efluentes e, atualmente, a produção de hidrogênio associada tem sido estudada como alternativa para fornecimento de combustível limpo (Chen *et al.*, 2005). As principais vias de produção fermentativa de hidrogênio a partir de efluentes ocorrem no metabolismo de bactérias acidófilas anaeróbicas (tais como *Clostridium sp.*) ou anaeróbicas facultativas (como *Enterobacter sp.*) (Chen *et al.*, 2005; Hamilton *et al.*, 2018; Seppala *et al.*, 2011), atuando em consórcios de microrganismos com composição variável e comportamento diferenciado de acordo com suas origens (Rajhi *et al.*, 2013; Mohan *et al.*, 2007; Maintinguer *et al.*, 2008). Esforços de pesquisa significativos têm sido feitos para descrever a complexidade dos processos biológicos de produção de hidrogênio por meios fermentativos, fornecendo ferramentas robustas, tais como modelos computacionais, para a previsão do comportamento e da bioprodução de hidrogênio. O uso de Redes Neurais Artificiais (RNAs) pode ser uma proposta valiosa para prever a bioprodução de hidrogênio por consórcio de microrganismos fermentadores. RNAs mostraram eficiência para descrever uma série de comportamentos e dinâmicas de produção de hidrogênio por organismos fermentadores (Sridevi *et al.*, 2014; Prakasham *et al.*, 2011). O presente manuscrito apresenta a construção de redes neurais artificiais para prever a produção de hidrogênio por consórcios de microrganismos provenientes de diferentes efluentes.

## MATERIAL E MÉTODOS

Foram analisados os resultados de volume de hidrogênio produzidos (variável-alvo) por três tipos de consórcios de microrganismos: Lod (efluente industrial), Esg (efluente de tratamento de esgotos domésticos), e Vac (efluente de gado). As variáveis de entrada foram temperatura ( $37-39^{\circ}\text{C}\pm 5^{\circ}\text{C}$ ), pH ( $5,5-6\pm 1-2$ ), % de inóculo ( $15\%\pm 5\%$ ) e razão carbono:nitrogênio do meio ( $3,5-5\pm 1-3$ ), totalizando 27 corridas para cada consórcio).

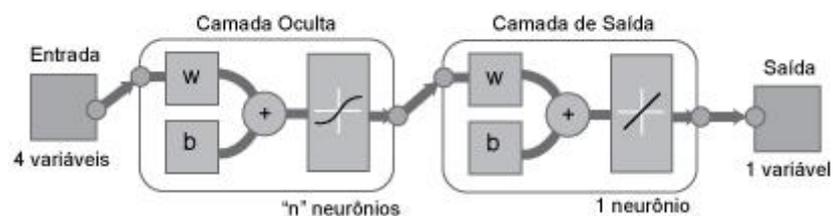
As redes neurais foram desenvolvidas em ambiente MATLAB, utilizando o pacote *Neural Network Toolbox* para modelos de ajuste de curva entrada-saída (*Input-Output curve fitting*). Foram analisadas redes do tipo *feed-forward* com três camadas (entrada, oculta e saída), utilizando função sigmoideal para transferência de dados para camada oculta e função linear para transferência para camada de saída. O treinamento foi feito por meio do algoritmo Levenberg-Marquardt de retropropagação (*backpropagation*), com máximo de 1000 iterações e avaliação de performance através do erro quadrático médio (EQM). Foram testados o

número de neurônios na camada oculta (5, 10, 15 ou 20) e as parcelas de dados utilizadas para treinamento, validação e teste (60:20:20%, 70:15:15% ou 80:10:10%).

## RESULTADOS E DISCUSSÃO

Na Figura 1 é apresentado o tipo de rede neural artificial desenvolvida no presente trabalho, onde 'w' significa peso (*weight*) e 'b' significa o erro/tendência (*bias*). A adequação de uma rede para a predição do volume de hidrogênio produzido seguiu a análise das respostas de cada modelo analisado.

Figura 1 – Representação gráfica da configuração das redes neurais artificiais estudadas



Fonte: Autoria própria (2019).

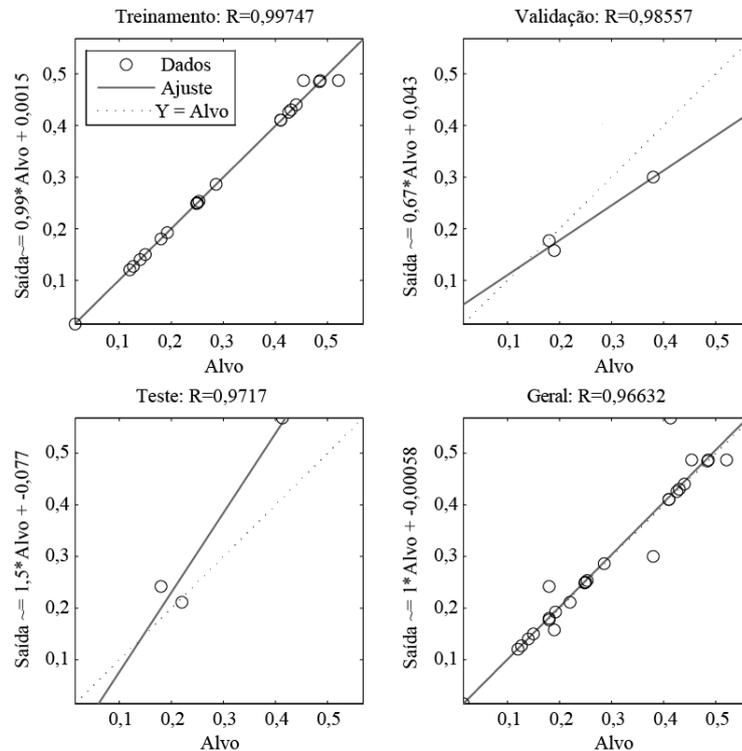
Primeiramente, foram avaliados os valores de coeficiente de correlação (R) entre os dados alvo e os dados de saída. Os dados alvo correspondem aos dados originais, enquanto os dados de saída são aqueles calculados por meio da rede proposta. Para as redes com valores de R mais próximos de 1, analisou-se a distribuição dos dados nos gráficos de dispersão entre alvo e saída.

Quanto mais próxima a disposição dos dados ao redor da linha de ajuste linear do gráfico, melhor a correlação entre os dados reais e os dados calculados pela rede proposta. O próximo fator de análise foi a distribuição do erro (erro=alvo-saída), que indica a adaptabilidade da rede neural para o banco de dados fornecido. Por fim, a performance de treinamento foi avaliada comparando o comportamento do erro quadrático médio ao longo das iterações, para as etapas de treinamento, validação e teste.

Tais resultados, referentes à análise do consórcio Esg, são apresentados nas Figuras 2 e 3. Nesse caso, a rede neural com 20 neurônios na camada oculta e partição dos dados com 80% para treinamento, e 10% para cada uma das respectivas etapas foi avaliada como mais adequada para descrever o processo. Os coeficientes de correlação variaram de um máximo de 0,99747 para a etapa de treinamento até do mínimo de 0,96632 para os dados gerais, como indicado na Figura 2. Além disso, observou-se distribuição adequada dos dados em torno da reta de ajuste linear.

A distribuição dos erros (Figura 3a) indicou que a maioria dos erros ocorreram em torno do valor 0,003834, com a ocorrência de extremos variando entre -0,1488 e 0,07427. Esses valores indicam os limites da rede para extrapolação, além de sugerirem a necessidade de incremento da base de dados. Quanto mais próxima a distribuição dos erros da distribuição normal, melhor é a constituição da base de dados.

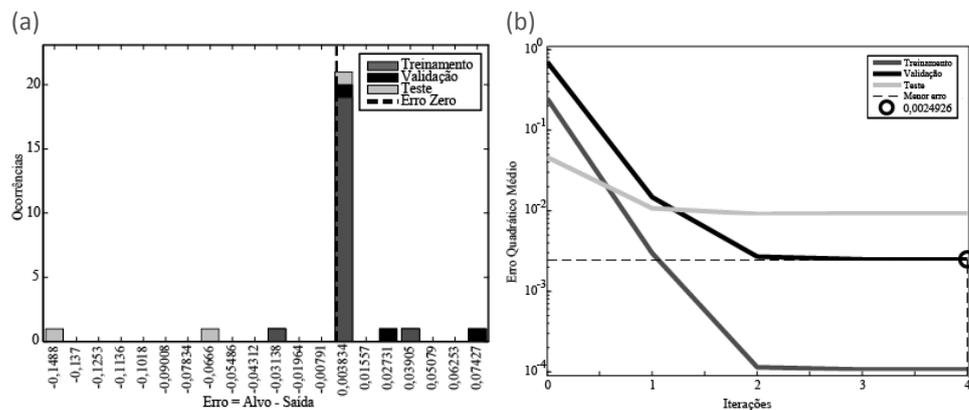
Figura 2 – Ajuste linear para a configuração 20:80 para o consórcio Esg



Fonte: Autoria própria (2019).

A performance de treinamento (Figura 3b) mostrou que a convergência da rede, coincidente com o menor valor sustentado do erro quadrático médio, ocorreu com 4 iterações. Ainda, o comportamento dos valores de EQM entre as etapas de treinamento, validação e teste foi similar após a segunda iteração. Isso indica que a rede neural proposta não difere das demais, ou seja, não houve sobreajuste dos dados no treinamento. Por fim, a rede apresentou EQM de cerca de 0,0025.

Figura 3 – Histograma de erros (a) e performance de treinamento para a configuração 20:80 para o consórcio Esg



Fonte: Autoria própria (2019).

Para os demais consórcios e para o conjunto total de dados, os resultados das redes com melhor adequação são apresentados na Tabela 1. Observou-se que a configuração com 20 neurônios na camada oculta e partição com 80% dos dados para treinamento foram mais eficientes para descrever os processos. Já para o conjunto total de dados, o melhor desempenho foi observado para a rede com 10 neurônios na camada oculta.

Tabela 1 – Valores dos coeficientes de correlação e EQM para as redes neurais para os consórcios Lod e Vac, e para o conjunto geral dos dados

Consórcio	Configuração	Coeficientes de correlação				EQM
		Treinem.	Valid.	Teste	Geral	
Lod	20:80	0,92916	0,97386	0,99694	0,8685	0,0088
Vac	20:80	0,99958	0,95067	0,99998	0,90503	0,0854
Conjunto	10:80	0,90298	0,88015	0,90144	0,88653	0,0082

Fonte: Autoria própria (2019).

## CONCLUSÃO

As configurações de redes neurais artificiais abordadas nesse trabalho mostraram-se eficientes para descrever a produção de hidrogênio por consórcios de microrganismos fermentativos, a partir de dados de entrada de temperatura, pH, percentagem de inoculo e razão C:N do meio. Considerando cada consórcio separadamente, a configuração com 20 neurônios na camada oculta e partição dos dados com 80% para treinamento apresentou os melhores ajustes. Quando o conjunto total de dados foi considerado, agrupando todos os dados dos três consórcios sem distinção, a configuração com 10 neurônios na camada oculta e mesma partição de dados ofereceu o melhor ajuste. Assim, conclui-se que a partição de dados, com 80% para treinamento, 10% para validação e 10% para teste, é a mais eficaz para esse tipo de dados. Os valores do erro quadrático médio encontrados para as configurações de redes neurais foram de 0,0024926 para o consórcio Esg, de 0,0088124 para o consórcio Lod, 0,0854011 para o consórcio Vac e 0,015003 para o conjunto total de dados.

A quantidade de neurônios na camada oculta pode variar, o que não representa um fator problemático já que afeta, principalmente, o tempo de computação do modelo. A distribuição de erros dos modelos mostrou, em alguns casos, a ocorrência de extremos. O aumento da quantidade de dados ao modelo normalmente diminui esse problema, ainda mais quando supra variações não abordadas pelos bancos de dados disponíveis. Esse é um fator determinante para o uso de redes neurais artificiais: quanto maior e melhor o banco de dados, melhor o ajuste encontrado.

Considerar um consórcio de microrganismos como um conjunto de espécies, descrito pela sua fonte, facilita o desenvolvimento de experimentos e diminui o esforço laboratorial de identificação de espécies. Entretanto, é importante ressaltar que a identificação das espécies dos consórcios, assim como a abundância de cada espécie, pode facilitar o entendimento da produção de hidrogênio. Essa informação pode ser considerada como variável de entrada para a rede neural, o que pode resultar em melhores ajustes do modelo e previsibilidade de produção por espécie. Em suma, a utilização de redes neurais artificiais mostrou-se como

uma ferramenta robusta e eficaz para descrever a produção de hidrogênio por consórcios de microrganismos fermentadores. Melhorias em relação ao banco de dados, como fornecimento de novos experimento e identificação de espécies, podem melhorar o ajuste das redes descritas nesse trabalho. De forma geral, os resultados aqui apresentados fornecem importantes informações sobre o uso de redes neurais artificiais na área de bioengenharia.

#### AGRADECIMENTOS

Ao Programa de Bolsas de Iniciação Científica – PIBIC 2018/2019, CNPq, Fundação Araucária.

#### REFERÊNCIAS

- CHEN, W.-M.; TSENG, Z.-J.; LEE, K.-S.; CHANG, J.-S. Fermentative hydrogen production with *Clostridium butyricum* CGS5 isolated from anaerobic sewage sludge. *International Journal of Hydrogen Energy*, v. 30, p.1063-1070, 2005.
- HAMILTON, C.; CALUSINSKA, M.; BAPTISTE, S.; MASSET, J.; BECKERS, L.; THONART, P.; HILIGSMANN, S. Effect of the nitrogen source on the hydrogen production metabolism and hydrogenases of *Clostridium butyricum* CWBI1009. *International Journal of Hydrogen Energy*, v.43, p,5451-5462, 2018.
- SEPPALA, J. J.; PUHAKKA, J. J.; YLI-HARJA, O.; KARP, M. T.; SANTALA, V. Fermentative hydrogen production by *Clostridium butyricum* and *Escherichia coli* in pure and cocultures. *International Journal of Hydrogen Energy*, v.36, p.10701-10708, 2011.
- RAJHI, H.; DÍAZ, E. E.; ROJAS, P.; SANZ, J. L. Microbial consortia for hydrogen production enhancement. *Curr. Microbiol.*, DOI 10.1007/s00284-013-0328-3, 2013.
- MOHAN, S. V.; BHASKAR, Y. V.; KRISHNA, P. M.; RAO, N. C.; BABU, V. L.; SARMA, P. N. Biohydrogen production from chemical wastewater as substrate by selectively enriched anaerobic mixed consortia: Influence of fermentation pH and substrate composition. *International Journal of Hydrogen Energy*, v.32, p. 2286-2295, 2007.
- MAINTINGUER, S. I.; FERNANDES, B. S.; DUARTE, I. C. S.; SAAVENDRA, N. K.; ADORNO, M. A. T.; VARESCHE, M. B. Fermentative hydrogen production by microbial consortium. *International Journal of Hydrogen Energy*, v.33, p. 4309-4317, 2008.
- SRIDEVI, K.; SIVARAMAN, E.; MULLAI, P. Back Propagation neural network modelling of biodegradation and fermentative biohydrogen production using distillery wastewater in a hybrid upflow anaerobic sludge blanket reactor. *Bioresource Technology*, v.165, p.233-240, 2014.
- PRAKASHAM, R. S.; SATHISH, T.; BRAHMAIAH, P. Imperative role of neural networks coupled genetic algorithm on optimization of biohydrogen yield. *International Journal of Hydrogen Energy*, v.36, p.4332-4339, 2011.