

MODELAGEM DA CINÉTICA DE EXTRAÇÃO SUPERCRÍTICA DE ARTEMISININA VIA REDES NEURAIS ARTIFICIAIS

Ian de Alencar Irizawa (Unifesp) E-mail: ian.dealencar@yahoo.com.br
Prof. Dr. Tiago Dias Martins (Unifesp) E-mail: tdmartins@unifesp.br
Profa. Dra. Priscilla Carvalho Veggi (Unifesp) E-mail: pveggi@unifesp.br

Resumo: A artemisinina é o composto majoritário sintetizado a partir da *Artemisia annua L.* É de grande interesse para as indústrias farmacêuticas, cosméticas e/ou alimentícia. Diversos grupos de pesquisa vêm estudando modelos matemáticos para descrever o comportamento cinético do processo. Neste sentido, este trabalho teve como objetivo o desenvolvimento de uma rede neural artificial para modelar a cinética de extração supercrítica de artemisinina. Foram utilizados oito experimentos em diferentes condições operacionais como base de dados. Foram empregadas duas estratégias para o treinamento das redes. A primeira utilizou como variáveis de entrada a pressão, temperatura, vazão de solvente e massa de extrato nos tempos t e $t-1$. A variável de saída foi a massa de extrato no $t+1$. Já a segunda estratégia utilizou a variável tempo no lugar das variáveis massa de extrato nos tempos t e $t-1$, utilizados na primeira estratégia, e a variável de saída foi a massa de extrato no tempo t . A melhor rede foi proveniente da segunda estratégia, com 7 neurônios na 1ª camada intermediária e 1 neurônio na 2ª camada intermediária (estrutura 4-7-1-1) conseguindo esta, prever e descrever o perfil cinético de maneira precisa em relação aos dados experimentais.

Palavras-chave: artemisinina, fluido supercrítico, extração, inteligência artificial.

Kinetic modeling of artemisinin supercritical extraction using artificial neural networks

Abstract: Artemisinin is the major compound synthesized from *Artemisia annua L.* It is of great interest for the pharmaceutical, cosmetic and/or food industries. Several research groups have been studying mathematical models to describe the process's kinetic behavior. Thus, this work aimed to develop an artificial neural network to model the kinetics of supercritical extraction of artemisinin. Eight experiments at different operational conditions were used as dataset. Two strategies were used to train the network. The first used as input variables: pressure, temperature, solvent flow, and extracted mass at times t and $t-1$. The output variable was the extracted mass at the time $t+1$. The second strategy used 'time' as input - to substitute the variables mass of extract at times t and $t-1$ used in the first strategy, and the output variable was the mass of extract at time t . The best network was one structure from the second strategy, with 7 neurons in the 1st intermediate layer and 1 neuron in the 2nd intermediate layer (structure 4-7-1-1), which was able to predict and describe the kinetic profile accurately when compared to the experimental data.

Keywords: artemisinin, supercritical fluid, extraction, artificial intelligence.

1. Introdução

As plantas são utilizadas pelos seres humanos desde a pré-história para os mais diferentes fins, como, por exemplo, medicinais. Uma matriz vegetal que tem atraído grande atenção da comunidade científica é a *Artemisia* (*Artemisia annua L.*). Esta possui um composto bioativo de grande interesse, a artemisinina, que se destaca por seu caráter efetivo antimalárico contra cepas de *Plasmodium spp* (HELLER e ROEPE, 2019).

A obtenção de compostos bioativos utilizando a extração com fluidos supercríticos (SFE – *Supercritical Fluid Extraction*), tendo o dióxido de carbono como principal solvente, é uma tecnologia com significativas vantagens em relação aos métodos de extração

convencionais. Na SFE o solvente é facilmente separado do soluto, geralmente possui rápida extração, utiliza pouco ou nenhum solvente orgânico (ANUKIRUTHIKA *et al.*, 2019), além de operar em baixas temperaturas, evitando assim a degradação térmica de compostos bioativos do extrato (HEWAVITHARANA *et al.*, 2020).

A partir dos dados da cinética de extração da SFE, os parâmetros dos modelos fenomenológicos, como, Crank (1975), Tan e Liou (1989) e Sovová *et al.* (1994), podem ser ajustados. Porém, algumas dificuldades são encontradas, especialmente no que diz respeito à generalização do modelo para diferentes condições de processo. Geralmente, diferentes condições levam à diferentes valores dos parâmetros, o que torna difícil sua aplicação no contexto industrial, pois muitas vezes a dependência dos parâmetros com relação às condições operacionais (como temperatura e pressão) não é simples.

Como alternativa, podem ser utilizadas as Redes Neurais Artificiais (RNAs), que possuem a flexibilidade para se adaptar a situações diferentes, e consistem em uma única equação, podendo ser utilizadas para diferentes condições de processo, como mostrado no trabalho de Soares *et al.* (2016).

As RNAs são compostas por camadas de neurônios que se interligam, assim como o cérebro humano. As redes são compostas por uma camada de entrada, intermediárias e uma de saída, sendo a informação transmitida de forma unidirecional. Cada neurônio possui n terminais de entrada (x_1, \dots, x_n) com um peso atrelado a ele (w_1, \dots, w_n). Tais pesos são combinados com um parâmetro bias (b_i) por meio de um somatório, gerando o coeficiente de ativação (α_{ai}) (BRAGA *et al.*, 2000).

$$\alpha_{ai} = \sum_{i=1}^n w_{ij} b_i \quad (1)$$

Com o valor do coeficiente de ativação, calcula-se o sinal de saída do neurônio k (y_k), conforme a equação 2:

$$y_k = \varphi(\alpha_{ai} + b_k) \quad (2)$$

Em que: b_k é o bias externo ao neurônio k , φ é a função de ativação e y_k é o sinal de saída do neurônio k .

Um dos grandes motivos da rápida difusão dessa nova forma de modelagem é a facilidade na aprendizagem e adaptação a fenômenos não lineares, com resultados altamente satisfatórios. Isso é proveniente a partir do ajuste dos pesos de cada neurônio, criando assim relações entre as variáveis de entrada e saída. Além disso, as RNAs se mostraram flexíveis, com capacidade de, mesmo com dados incorretos ou banco de dados incompleto, apresentarem resultados acertados. Sua fácil implementação a partir de processos já existentes, bem como execução em tempo real são outras vantagens dessa tecnologia (SZYMCZAK *et al.*, 2009).

As RNAs têm sido bastante utilizadas em diversas áreas da engenharia química. E no que se refere à extração supercrítica, uma das estratégias utilizadas foi a predição da derivada inicial da curva de extração de linalol de sementes de coentro (ZEKOVIĆ *et al.*, 2017), e na extração supercrítica de ácidos graxos de sementes da framboesa (PAVLIĆ *et al.*, 2020). A modelagem de toda a curva cinética de extração foi realizada por BAHMANI *et al.* (2018) na extração do óleo essencial do estragão Soares *et al.* (2016).

Sendo assim, o presente trabalho teve como objetivo modelar, empregando RNAs, a cinética de extração de artemisinina a partir da *Artemisia annua L.* em diferentes condições operacionais. Além disso, comparar a influência de diferentes algoritmos de otimização, funções de ativação e verificar sua capacidade de interpolação.

2. Métodos

Os dados experimentais utilizados neste trabalho foram coletados por Condori (2005), que realizou um estudo da extração supercrítica da artemisinina a partir da matriz vegetal *Artemisia Annu L.* utilizando CO₂ como solvente. Foram utilizados dados da massa de extrato obtida em função do tempo de processo, em oito experimentos realizados em diferentes condições de pressão (200 e 400 bar), temperatura (30 e 50 °C) e vazão de solvente ($5,5 \cdot 10^{-5}$ e $11,1 \cdot 10^{-5}$ kg CO₂/s). Uma vez que o número de dados experimentais era limitado, e que os dados coletados não se apresentavam em um intervalo constante de tempo, utilizou-se o software *TableCurve*® para gerar 100 pontos, distribuídos uniformemente, para cada um dos experimentos entre 0 e 130 minutos (1 ponto a cada 1,3 minutos).

Duas estratégias foram utilizadas para a modelagem das curvas. As propostas estão apresentadas na Tabela 1. A proposta I consistiu na obtenção da curva cinética de extração a partir de informações da condição do processo e da inclinação da curva de extração. Essa estratégia é bastante empregada em trabalhos que envolvem controle de processos e pode ser facilmente expandida para estudos que envolvem o desenvolvimento de modelos semelhantes. A proposta II consistiu na modelagem da curva de extração utilizando-se as condições do processo e o tempo, como variáveis de entrada. Essa estratégia também é comumente utilizada na modelagem de processos e diversos estudos podem ser encontrados na literatura.

Tabela 1 – Resumo de estratégias propostas.

PROPOSTA	VARIÁVEIS DE ENTRADA	VARIÁVEL DE SAÍDA
I	Massa de extrato para um tempo t (m_t); Massa de extrato com um atraso do tempo t (m_{t-1}); Pressão (P); Temperatura (T); Vazão de solvente (CO ₂).	Massa de extrato no tempo $t+1$ (m_{t+1}).
II	Tempo (t), Pressão (P); Temperatura (T); Vazão de solvente (CO ₂);	Massa de extrato no tempo t (m_t).

A Figura 1 ilustra as duas estratégias propostas. A Figura 1a apresenta a estrutura geral da RNA com duas camadas intermediárias para a proposta I. Já a Figura 1b ilustra uma RNA com duas camadas intermediárias para a proposta II.

2.1 Definição do modelo neural

Para a definição do modelo neural duas etapas foram realizadas: obtenção da rede neural e simulação. A obtenção da RNA consistiu na divisão aleatória dos dados em três categorias: treinamento, validação e teste, nos quais 70% dos dados foram alocados para a etapa de treinamento, 15% para a etapa de validação e 15% para a etapa de testes. Dos oito experimentos disponíveis, sete foram utilizados na etapa de obtenção da RNA, totalizando 686 pontos. Um experimento foi separado para a etapa exclusiva de simulação.

(a)

(b)

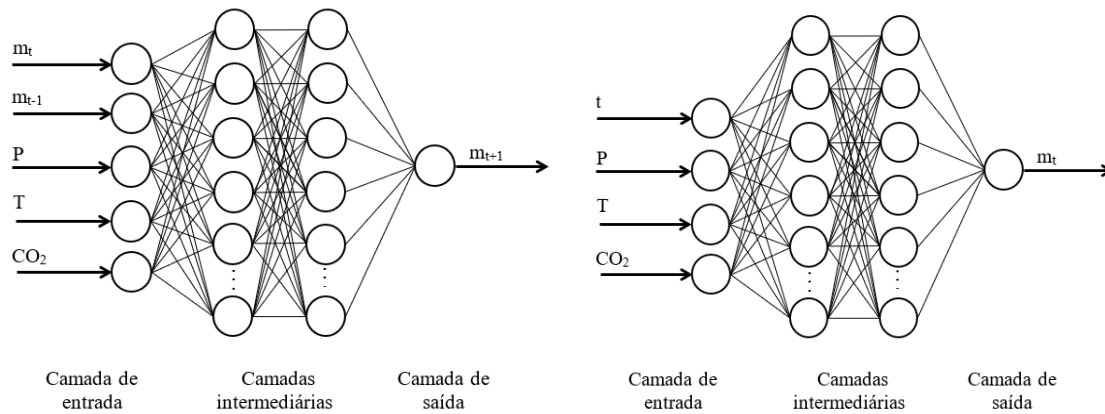


Figura 1 – (a) Estruturas da RNA da Proposta I com duas camadas intermediárias; (b) Estrutura da RNA da Proposta II com duas camadas intermediárias.

Utilizando-se o software MATLAB R2016b, foram testadas diversas combinações de estrutura de RNAs, funções de ativação e algoritmos de otimização para obtenção do melhor resultado nesta etapa. O treinamento se concentrou na otimização dos pesos e *bias* das RNAs, utilizando os algoritmos de otimização: Levenberg-Marquardt (*trainlm*), Resilient-Backpropagation (*trainrp*), Gradient Descent Backpropagation (*traingd*), Scale Conjugate Gradient (*trainscg*) e Bayesian Regularization Backpropagation (*trainbr*). Foi utilizada como função objetivo, a média do erro quadrático, MSE (Eq. 3):

$$F_{obj} = \frac{1}{N} \sum_{i=1}^N (m_i^{EXP} - m_i^{RNA})^2 \tag{3}$$

em que *m* é a massa extraída, *N* é o número total de observações, e os sobrescritos *EXP* e *RNA* correspondem a experimental e calculado pela RNA, respectivamente.

Ao final de cada etapa, o coeficiente de correlação de Pearson (ρ) também foi calculado (Eq. 4).

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{cov(x,y)}{var(x) var(y)} \tag{4}$$

Sendo:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \tag{5}$$

$$\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^n y_i \tag{6}$$

Dentre as funções de transferência disponíveis, foram utilizadas a linear (*purelin*), logarítmica-sigmóide (*logsig*) e tangente hiperbólica (*tansig*). Tais funções são as mais utilizadas e são empregadas para conferir não-linearidade ao modelo.

A etapa de simulação consistiu em usar as RNAs treinadas visando observar se elas eram capazes de generalizar o aprendizado. O experimento realizado à $T = 400 \text{ }^\circ\text{C}$, $P = 30 \text{ bar}$ e $Q = 5,5 \cdot 10^{-5} \text{ kg CO}_2/\text{s}$ foi utilizado exclusivamente para essa etapa e não foi apresentado na etapa de obtenção das RNAs. A melhor RNA foi determinada a partir da análise dos valores de função objetivo da etapa de obtenção, em conjunto com o ρ da etapa de simulação.

Para a proposta I, além das condições operacionais, somente a condição inicial do processo ($m(t=0)$ e $m(t=1,3 \text{ min})$) foi fornecida e, a partir dela, o restante da curva foi calculado utilizando-se somente os valores calculados pela rede, retroalimentados como entrada. Já na proposta II, a rede foi simulada em um intervalo de 0,25 min para prever a massa extraída. Por fim, gráficos de massa extraída vs tempo foram construídos.

Vale ressaltar que uma vantagem de se utilizar a proposta II é que a curva cinética pode ser obtida para qualquer valor de tempo. Por outro lado, a RNA da proposta I prevê a massa extraída a cada 1,3 minutos, já que o intervalo de tempo considerado era constante.

Por fim, a melhor RNA foi utilizada para simular todos os experimentos considerados, visando-se determinar a capacidade de interpolação dos modelos obtidos. Isso pode ser feito porque, nessa etapa a curva é construída apenas a partir da condição inicial do processo (Proposta I), ou interpolando-se valores de tempo não apresentados na etapa de obtenção (Proposta II).

3. Resultados e Discussão

Para cada uma das duas propostas, foram treinadas e simuladas mais de 500 estruturas diferentes com uma e duas camadas intermediárias, das quais as 15 melhores foram selecionadas.

3.1 Proposta I

Para a Proposta I foram utilizadas como entradas da RNA: pressão, vazão de solvente, temperatura e massas de extrato no tempo t e no tempo $t-1$. No treinamento das estruturas foram testados diferentes algoritmos de treinamento e funções de ativação, a fim de se atingir o melhor resultado. Na Tabela 2, estão apresentados os quinze melhores resultados obtidos para a Proposta I, contendo os valores de F_{obj} das etapas de treinamento, validação e teste, ρ da etapa de simulação e o número de parâmetros de cada estrutura.

Pode-se notar que a combinação logsig-purelin foi a que apresentou o melhor desempenho, visto que ela é predominante dentre os melhores resultados obtidos. Outros dois pontos relevantes são que, de forma geral, a função objetivo do treinamento diminuiu com o aumento do número de parâmetros e que os valores de ρ para a simulação não tem relação direta com a F_{obj} na etapa de obtenção das RNAs. Isso mostra que a avaliação dos modelos obtidos não é uma tarefa simples e diversos parâmetros estatísticos, e configurações de RNAs, precisam ser levados em consideração.

Além disso, verificou-se que redes com menor número de parâmetros (menores que 50) apresentaram, em média, valores de ρ na etapa de simulação superiores em relação às redes com mais de 50 parâmetros (média de 0,504 contra 0,433). Porém, não apresenta uma relação diretamente proporcional ao número de parâmetros. Isso mostra que RNAs que possuem muitos parâmetros são mais suscetíveis ao problema de sobreajuste (BRAGA *et al.*, 2000).

Em relação aos algoritmos de treinamento, notou-se que o trainlm foi o que apresentou melhores resultados ao compararmos com os demais. Este resultado é esperado, pois cada algoritmo possui um equacionamento diferente para o ajuste dos pesos, o que afeta o resultado final. Por isso, existe a importância de se testar diversos algoritmos de otimização a fim de se encontrar o melhor modelo. Esse comportamento é comumente observado na literatura, que mostra essa importância na procura pelo melhor modelo de RNA (MARTINS *et al.*, 2020 e OTTAIANO *et al.*, 2020).

Tabela 2 - Desempenho RNA da Proposta I.

Estrutura	Função de treinamento	Funções de transferência	F _{obj} Treinamento	F _{obj} Validação	F _{obj} Teste	ρ Simulação	# parâmetros
5-1-1	trainlm	logsig-purelin	9,57E-07	1,70E-05	7,80E-07	0,531	8
5-2-1	traingd	tansig-tansig	3,79E-03	3,94E-03	4,01E-03	0,658	15
5-2-1	trainlm	tansig-purelin	1,32E-06	2,78E-07	1,39E-05	0,586	15
5-2-1	trainlm	logsig-purelin	3,64E-06	2,39E-06	1,49E-06	0,419	15
5-3-1	trainlm	logsig-purelin	1,03E-06	1,52E-06	1,04E-05	0,420	22
5-5-1	trainlm	logsig-purelin	4,31E-06	6,66E-07	2,00E-06	0,477	36
5-6-1	trainlm	logsig-purelin	1,76E-06	2,71E-04	2,90E-06	0,438	43
5-8-1	trainlm	logsig-purelin	1,29E-06	5,46E-07	6,48E-07	0,385	57
5-14-1	trainlm	logsig-purelin	6,94E-07	8,75E-07	1,23E-06	0,374	99
5-20-1	trainlm	logsig-purelin	1,02E-06	1,05E-06	8,73E-06	0,632	141
5-22-1	trainscg	tansig-tansig	1,25E-04	1,09E-04	6,06E-05	0,360	155
5-23-1	trainlm	logsig-purelin	8,35E-07	7,40E-07	9,90E-07	0,427	162
5-26-1	trainlm	logsig-purelin	7,84E-07	1,04E-06	7,69E-07	0,423	183
5-31-1	trainlm	logsig-purelin	1,02E-06	1,11E-06	5,22E-07	0,435	218
5-35-1	trainlm	logsig-purelin	7,30E-07	3,51E-07	6,40E-07	0,428	246

A respeito do número de camadas intermediárias, ao se utilizar duas camadas intermediárias, os resultados foram inferiores quando comparados aos obtidos com RNAs de uma camada intermediária, e sequer estão entre as 15 melhores estruturas. Além disso, é relevante citar que a estrutura 5-2-1 apareceu três vezes dentro dos quinze melhores resultados, devendo esta ser a rede com maior potencial de generalização.

Pode-se verificar que, mesmo com as diferentes funções de transferência e estruturas, a Proposta I não apresentou resultados promissores, pois o melhor resultado apresentou ρ igual a 0,6582 (estrutura 5-2-1 – traingd - tansig-tansig), destacada na Tabela 2.

A Figura 2 apresenta os resultados obtidos pela rede neural 5-2-1 nas etapas de treinamento, teste e validação. O gráfico consiste nos valores experimentais como função dos valores obtidos pela RNA. Pode-se observar que nas três etapas, os valores preditos pela RNA foram semelhantes aos dados experimentais. Os erros obtidos foram baixos (inferiores a 1%). No entanto, mesmo apresentando bom desempenho na etapa de obtenção, nota-se que não apresentou resultados que possam ser considerados satisfatórios, quando confrontados com uma série de dados desconhecida na etapa de simulação. Esse resultado mostra a importância de se executar uma etapa de simulação com as características empregadas neste trabalho, com dados independentes daqueles utilizados no treinamento. Usualmente, estudos envolvendo RNAs concentram a escolha do melhor modelo analisando-se os resultados obtidos nas etapas de teste e validação. Porém, este estudo mostrou que essa prática pode levar a resultados errôneos quanto à escolha do melhor modelo. Ottaiano *et al.* (2020) apresentaram este mesmo comportamento em seu estudo e mostraram a importância de se realizar a etapa de simulação.

Além disso, apesar dessa RNA ter apresentado valores maiores de F_{obj}, nas etapas de obtenção, em relação a outras estruturas de RNA, foi ela que apresentou a melhor predição e generalização dos dados apresentados. Isso demonstra que as redes neurais, muitas vezes, podem sofrer com o sobreajuste (*overfitting*). Fenômeno no qual as redes ficam viciadas nos dados apresentados em seu treinamento e acabam minando sua capacidade de generalização e predição.

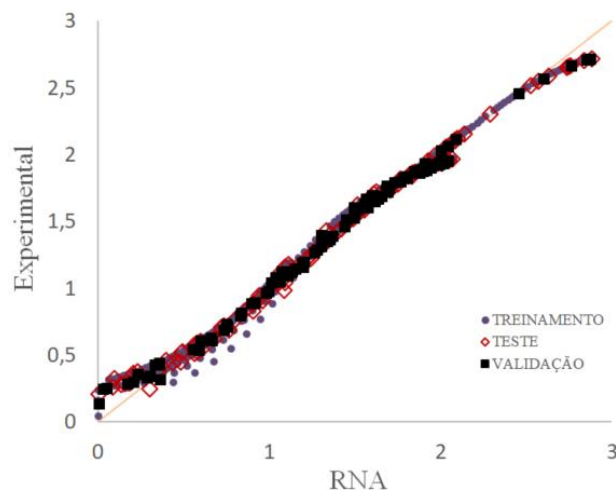


Figura 2 - Gráfico de valores experimentais vs gerados pela RNA 5-2-1 em seu treinamento, teste e validação.

Diante desses resultados, desenvolveu-se uma segunda proposta, considerando a variável tempo como uma das variáveis de entrada.

3.2 Proposta II

A segunda proposta foi desenvolvida com uma alteração nas variáveis de entrada e saída. Foram removidas as variáveis de entrada: ‘massa extraída’ nos tempos t e $t-1$ e adicionada a variável ‘tempo’. Pressão, temperatura e vazão de solvente foram mantidas, visto que elas determinam a condição operacional. Foram utilizados os mesmos algoritmos e funções de ativação da proposta I, alternando entre uma e duas camadas intermediárias.

Na Tabela 3 estão apresentados os 15 melhores resultados obtidos pelas diferentes estruturas testadas na Proposta II, onde se pode notar que RNAs com somente uma camada intermediária não aparecem, pois não apresentaram resultados tão satisfatórios quanto aquelas com duas camadas.

Em um contexto geral, a Proposta II apresentou resultados bastante promissores, com estruturas que atingiram valores baixos de erro médio, além de altos valores de ρ . A estrutura 4-7-1-1, com funções de ativação purelin-tansig-tansig foi a que apresentou os melhores resultados na etapa de simulação, apresentando um erro médio de 5,17% e ρ igual a 0,997. Além disso, é a que possui o menor número de parâmetros, o que facilita no desempenho computacional na simulação da rede.

Pôde-se notar que nessa segunda estratégia foram obtidos valores baixos de funções objetivo nas etapas de treinamento, validação e teste. Porém, vale salientar que tais valores superiores aos obtidos nas mesmas etapas da Proposta I. Assim como na primeira estratégia, também se verificou que a função de treinamento trainlm foi predominante dentre as melhores RNAs. Além disso, pôde-se verificar que o número de parâmetros pouco influenciou nos resultados obtidos por essa estratégia, visto que o menor valor de ρ , mesmo com uma grande variação de número de parâmetros dentre os quinze melhores resultados foi de, aproximadamente 0,885. Os erros médios, em geral, também foram menores quando comparados aos apresentados na Proposta I.

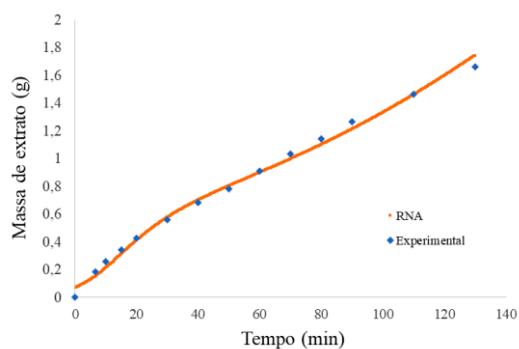
Tabela 3 - Desempenho da RNA na proposta com variável tempo.

Estrutura	Função de treinamento	Função transferência	F _{obj} Treinamento	F _{obj} Validação	F _{obj} Teste	Erro Médio Simulação (%)	ρ Simulação	# parâmetros
4-7-1-1	trainbr	purelin-tansig-tansig	8,74E-04	7,75E-04	8,00E-04	5,17	0,997	45
4-1-15-1	trainlm	logsig-logsig-tansig	2,29E-01	1,95E-01	2,36E-01	23,68	0,885	51
4-4-7-1	trainlm	purelin-tansig-tansig	9,99E-04	1,12E-03	1,49E-03	6,97	0,997	63
4-4-7-1	trainlm	logsig-logsig-tansig	9,99E-05	1,08E-04	1,86E-04	10,39	0,996	63
4-7-3-1	trainlm	logsig-logsig-tansig	9,13E-05	7,37E-05	6,95E-05	30,61	0,996	63
4-4-10-1	trainbr	purelin-tansig-tansig	9,72E-04	1,09E-03	1,05E-03	15,29	0,990	81
4-3-16-1	trainbr	purelin-tansig-tansig	8,88E-04	1,08E-03	1,18E-03	13,01	0,998	96
4-7-8-1	trainbr	purelin-tansig-tansig	8,85E-04	1,49E-03	1,32E-03	12,70	0,993	108
4-2-25-1	trainlm	logsig-logsig-tansig	9,98E-05	6,84E-05	2,32E-04	26,20	0,986	111
4-3-20-1	trainlm	logsig-logsig-tansig	9,97E-05	5,24E-04	1,99E-04	30,32	0,997	116
4-4-21-1	trainlm	logsig-logsig-tansig	9,97E-05	4,97E-05	2,20E-04	34,25	0,990	147
4-4-26-1	trainlm	logsig-logsig-tansig	8,80E-05	1,44E-04	9,70E-05	19,34	0,985	177
4-9-16-1	trainlm	logsig-logsig-tansig	9,70E-05	8,78E-05	3,04E-04	30,15	0,987	222
4-5-31-1	trainlm	logsig-logsig-tansig	9,65E-05	2,43E-04	6,71E-05	35,11	0,995	243
4-35-4-1	trainlm	purelin-tansig-tansig	1,01E-03	8,36E-04	9,58E-04	16,29	0,999	324

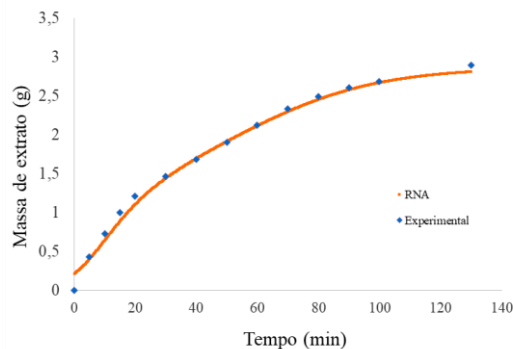
Com o intuito de demonstrar o poder de interpolação da melhor rede neural encontrada e seu alto grau de adaptação a diversas situações, foram simulados os oito experimentos com intervalo de tempo de 0,25 minutos. Portanto, foram fornecidas à rede as condições de pressão, temperatura, vazão do solvente em 519 condições de tempo (intervalo de 0,25 minutos até o tempo 130 min), para que a RNA gerasse o perfil da curva de extração por completo. Esse perfil simulado pela RNA foi comparado com os dados experimentais fornecidos por Condori (2005) em diferentes condições, como mostrado na Figura 3.

Analisando-se a Figura 3, pode-se verificar o grande poder de interpolação da RNA obtida e como a capacidade de predição foi precisa. Os gráficos simulados representam adequadamente o perfil esperado o que confirma o erro médio baixo, obtido em relação aos pontos experimentais. A Figura 3e mostra o perfil obtido na simulação do experimento que não foi apresentado na etapa de treinamento. Nota-se que a RNA conseguiu simular com precisão o comportamento da curva, comprovando sua eficiência.

Por fim, calculou-se o erro médio (%), o MSE e o ρ considerando todos os oito experimentos. Os resultados obtidos foram um erro médio de 4,93%, MSE igual a $2,619 \cdot 10^{-3}$ e ρ igual a 0,997. Esse resultado reforça o resultado apresentado na Figura 3: de que a RNA foi capaz de interpolar adequadamente e prever o comportamento da extração supercrítica da matriz vegetal.



(a)



(b)

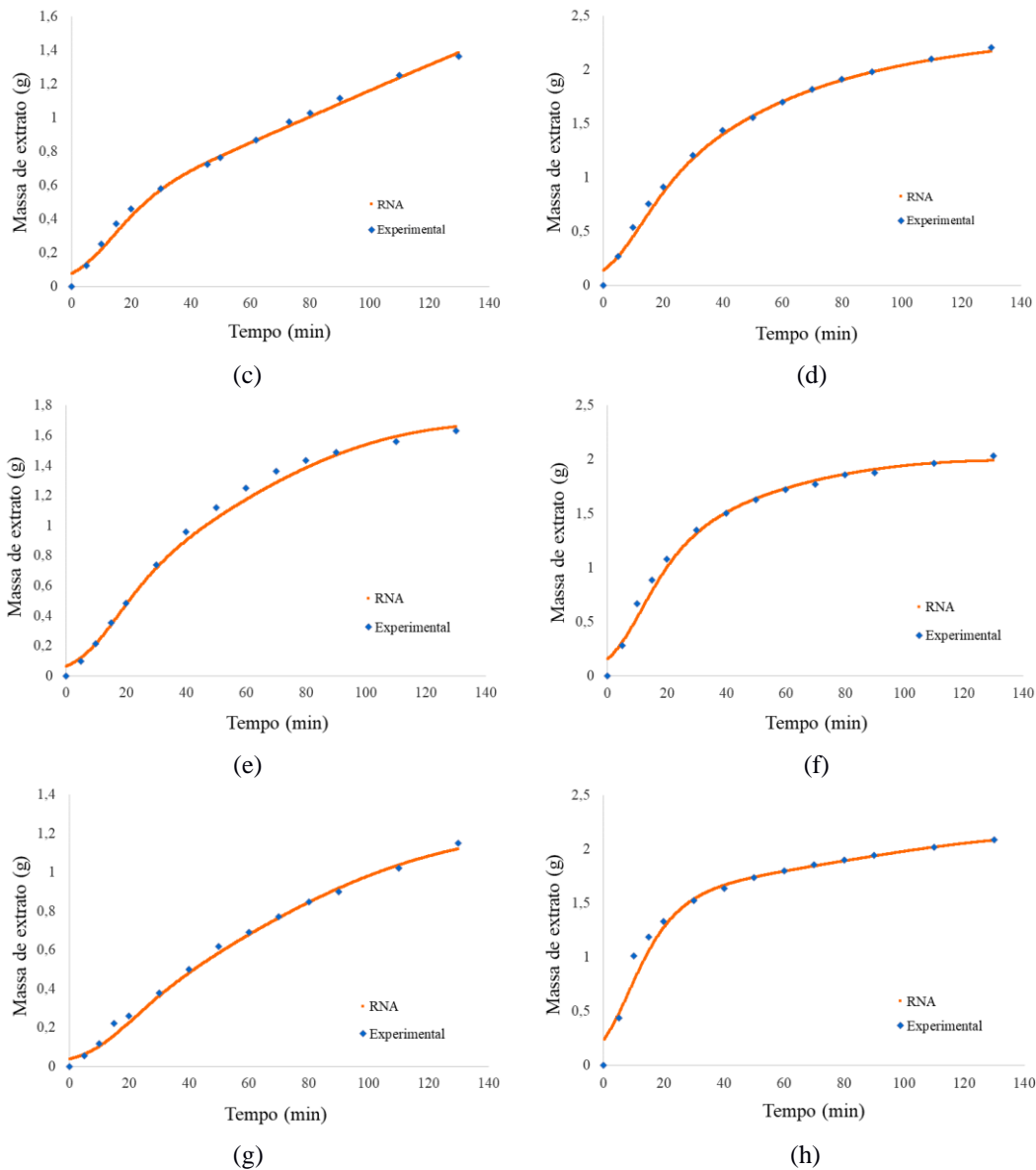


Figura 3 – Gráficos de massa extraída (g) versus tempo (min) da rede neural 4-7-1-1 e dos dados experimentais, para (a) $P = 50$ bar, $T = 400$ °C e $Q = 5,5 \cdot 10^{-5}$ kg/m³; (b) $P = 50$ bar, $T = 400$ °C e $Q = 11,1 \cdot 10^{-5}$ kg/m³; (c) $P = 50$ bar, $T = 200$ °C e $Q = 5,5 \cdot 10^{-5}$ kg/m³; (d) $P = 50$ bar, $T = 200$ °C e $Q = 11,1 \cdot 10^{-5}$ kg/m³; (e) $P = 30$ bar, $T = 400$ °C e $Q = 5,5 \cdot 10^{-5}$ kg/m³ (exclusivo simulação); (f) $P = 30$ bar, $T = 400$ °C e $Q = 11,1 \cdot 10^{-5}$ kg/m³; (g) $P = 30$ bar, $T = 200$ °C e $Q = 5,5 \cdot 10^{-5}$ kg/m³; (h) $P = 30$ bar, $T = 200$ °C e $Q = 11,1 \cdot 10^{-5}$ kg/m³.

Ao se comparar os resultados obtidos nesse estudo, com os estudos de Soares *et al.* (2016) e de Bahmani *et al.* (2018), verificou-se que as melhores redes neurais artificiais obtiveram resultados excelentes. Em relação à função objetivo (MSE), este estudo obteve 0,0026, Soares *et al.* (2016) obteve 0,07 e Bahmani *et al.* (2018), 0,0008, mostrando uma precisão acentuada entre os valores preditos pela RNA e os valores experimentais. Ao compararmos o R^2 , os três estudos apresentam um alto grau de correlação, sendo o valor obtido por Soares *et al.* (2016): 0,999, o de Bahmani *et al.* (2019): 0,996 e neste estudo 0,997. Com isso, podemos afirmar que as redes neurais se mostraram uma ferramenta eficiente para o cálculo e predição de valores em diferentes áreas da engenharia.

Finalmente, ao se comparar os resultados obtidos pela RNA 4-7-1-1 com os modelos fenomenológicos utilizados por Condori (2005), observamos que a RNA foi a que

apresentou o melhor resultado, ou seja, a menor F_{obj} . Esse valor foi obtido calculando-se a F_{obj} para todos os dados disponíveis, assim como fez Condori (2005). O melhor modelo apresentado por Condori (2005) foi o de Martinez (2003) que apresentou um valor de 0,235, conforme apresenta a Tabela 4.

Tabela 4 - Valores de função objetivo obtida via RNA comparada aos modelos fenomenológicos.

Crank (1975)	Tan e Liou (1989)	Goto et al. (1993)	Sovová et al. (1994)	Naik et al. (1989)	Martinez (2003)	RNA 4-7-7-1
1,856	3,744	1,161	$6,809 \cdot 10^{-2}$	$6,831 \cdot 10^{-1}$	$2,358 \cdot 10^{-1}$	$2,619 \cdot 10^{-3}$

Em resumo, este trabalho mostrou que é importante realizar uma extensa procura dentro das possibilidades de construção de um modelo neural, avaliando diferentes variáveis de entrada e saída, algoritmos de otimização, funções de ativação e número de neurônios; e que as RNAs são uma alternativa eficiente para a modelagem de curvas cinéticas de extração supercrítica em diferentes condições operacionais. Trabalhos futuros consistirão na utilização do modelo proposto aqui em estudos de controle de processo.

4. Conclusões

Ao final do trabalho, pode-se concluir que as RNAs têm um grande potencial a ser explorado. Como observado no presente trabalho, o MSE calculado pelos modelos clássicos maior em relação ao obtido pela rede neural, mostrando o poder de interpolação e generalização das RNAs, bem como a capacidade de predição da cinética da extração supercrítica da artemisinina.

Devem-se citar os resultados não satisfatórios com a primeira estratégia que não foi capaz de obter uma correlação maior que 70%. Porém, deve se ressaltar que com uma RNA bem treinada, bem ajustada e com a estratégia correta, foi possível obter uma correlação de mais de 0,997 em relação aos dados experimentais. Além disso, foi possível predizer qualquer ponto dentro da curva cinética com menor erro se comparado aos modelos clássicos.

Referências

- ANUKIRUTHIKA, T.; SAYANTANI, D.; MOSES, J.A. & ANANDHARAMAKRISHNAN, C. *Modern Applications of Supercritical Fluids Extraction in Food Toxicology*. Reference Module in Food Science, 2019. <http://dx.doi.org/10.1016/B978-0-08-100596-5.22939-9>
- BAHMANI L.; ABOONAJMI M.; ARABHOSSEINI A. & MIRSAEEDGHAZI H. *ANN modeling of extraction kinetics of essential oil from tarragon using ultrasound pre-treatment*. Engineering in Agriculture, Environment and Food. Vol. 11, p. 25–29, 2018. <https://doi.org/10.1016/j.eaef.2017.10.003>
- BRAGA A.P.; CARVALHO A.P.L. & LUDERMIR T.B. *Redes Neurais Artificiais: Teoria e Aplicações*. Rio de Janeiro: LTC, 2000.
- CONDORI S.Q. *Determinação de parâmetros de processo nas diferentes etapas da extração supercrítica de produtos naturais: Artemisia annua, Cordia verbenacea, Ocimum selloi Foeniculum Vulgare*. 2005.
- CRANK J. *The mathematics of diffusion (2nd ed.)*. Oxford: Clarendon Press, 1975.
- GOTO, M.; SATO, M. & HIROSE, T. *Extraction of peppermint oil by supercritical carbon dioxide*, Journal of Chemical Engineering of Japan. Vol. 26, p. 401-407, 1993.

<https://doi.org/10.1252/jcej.26.401>

HELLER, L. & ROEPE, P. *Artemisinin-Based Antimalarial Drug Therapy: Molecular Pharmacology and Evolving Resistance*. Tropical medicine and infectious disease. Vol. 4, p. 89, 2019. <https://dx.doi.org/10.3390/tropicalmed4020089>

HEWAVITHARANA, G.G., PERERA, D.N., NAVARATNE, S.B., & WICKRAMASINGHE, I. *Extraction methods of fat from food samples and preparation of fatty acid methyl esters for gas chromatography: A review*. Arabian Journal of Chemistry. Vol. 13, p. 6865-6875, 2020. <https://doi.org/10.1016/j.arabjc.2020.06.039>

MARTINS T.D.; ANNICHINO-BIZZACCHI J.M.; ROMANO A.V.C. & MACIEL FILHO R. *Artificial neural networks for prediction of recurrent venous thromboembolism*. International Journal of Medical Informatics. Vol. 141, p. 104221, 2020. <https://doi.org/10.1016/j.ijmedinf.2020.104221>.

MARTINEZ, J. *Extração de óleos voláteis e outros compostos com CO₂ supercrítico: desenvolvimento de uma metodologia de aumento de escala a partir da modelagem matemática do processo e avaliação dos extratos obtidos*. 2003.

NAIK, S.N.; LENTZ, H. & MAHESHWARI, R.C. *Extraction of perfumes and flavours from plant materials with liquid carbon dioxide under liquid–vapor equilibrium conditions*. Fluid Phase Equilibria. Vol. 49, p. 115-126, 1989. [https://doi.org/10.1016/0378-3812\(89\)80009-3](https://doi.org/10.1016/0378-3812(89)80009-3).

OTTAIANO, G.Y.; CRUZ, I.N.S.; CRUZ, H.S. & MARTINS, T.D. *Estimation of vaporization properties of pure substances using artificial neural networks*. Chemical Engineering Science, Vol. 231, p. 116324, 2021. <https://doi.org/10.1016/j.ces.2020.116324>

PAVLIĆ B.; PEZO L.; MARIĆ B.; TUKULJAC L.; ZEKOVIĆ Z.; SOLAROV B. & TESLIĆ N. *Supercritical fluid extraction of raspberry seed oil: Experiments and modelling*. Journal of Supercritical Fluids, Vol. 157, p. 104687, 2020. <https://doi.org/10.1016/j.supflu.2019.104687>

SOARES L.A.P.; ANDRADE R.R. & MARTINS T.D. *Kinetics of second-generation ethanol fermentation with cell recycle using artificial neural networks*. In: I Congresso Brasileiro em Engenharia de Sistemas em Processos - PSE 2019, 2019, Rio de Janeiro. Anais do I PSE 2019, 2019.

SOARES, L.A.P.; RIBAS, M.C.; SILVA, E.A. & MARTINS, T.D. *Modelagem cinética da extração supercrítica do óleo de candeia por redes neurais artificiais*. In: XXI Congresso Brasileiro de Engenharia Química e XVI Encontro Brasileiro Sobre Ensino em Engenharia Química, 2016, Fortaleza. Anais do XXI Congresso Brasileiro de Engenharia Química e XVI Encontro Brasileiro Sobre Ensino em Engenharia Química, 2016.

SOVOVÁ H.; KOMERS R. & J. KUCERA J. JEZ. *Supercritical carbon dioxide extraction of caraway essential oil*. Chemical Engineering Science. Vol 49, p. 2499-2505, 1994. [https://doi.org/10.1016/0009-2509\(94\)E0058-X](https://doi.org/10.1016/0009-2509(94)E0058-X)

SZYMCZAK S.; BIERNACKA J.M.; CORDELL H. J.; GONZALEZ-RECIO I. R.; KÖNIG I. R.; ZHANG H. & SUN Y. V. *Machine learning in genome-wide association studies*. Genetic Epidemiology. Vol 33, p. S51-S57, 2009.

TAN C.S. & LIOU D.C. *Desorption of ethyl acetate from activated carbon by supercritical carbon dioxide*, Industrial & Engineering Chemistry Research. Vol. 27, p. 988-991, 1988. <https://doi.org/10.1002/gepi.20473>

ZEKOVIĆ Z.; BERA O.; ĐUROVIĆ S. & PAVLIĆ B. *Supercritical fluid extraction of coriander seeds: Kinetics modelling and ANN optimization.* Journal of Supercritical Fluids. Vol. 125, p. 88–95, 2017. <https://doi.org/10.1016/j.supflu.2017.02.006>