

CONTROLE PREDITIVO POR MODELO DE REDE NEURAL APLICADO A PROCESSOS DE FERMENTAÇÃO E DESTILAÇÃO

Mohamad Al Bannoud (UNIFESP) E-mail: mohamed_bannoud@hotmail.com

Brunno Ferreira dos Santos (PUC-Rio) E-mail: bsantos@puc-rio.br

Tiago Dias Martins (UNIFESP) E-mail: tdmartins@unifesp.br

Resumo: Neste trabalho foram utilizadas Redes Neurais Artificiais (RNA) em combinação com controle preditivo baseado em modelo (Model Predictive Control, MPC) para os processos de fermentação e destilação. Foi avaliada a relação entre a saturação e o valor da função objetivo da RNA utilizada durante a simulação do controle. As RNAs mostraram ser capazes de representar bem o processo, com o controle sendo robusto e sem a presença de offset. A saturação das RNAs não influenciou no resultado do controlador, sendo seu valor tipicamente menor que 0,4 para as estruturas com funções de ativação lineares e maior que 0,6 para as sigmóides. O valor da saturação está mais relacionado com a estrutura e o ajuste de pesos durante o treinamento da RNA, sem haver uma relação direta com o desempenho do controlador.

Palavras-chave: Controle Preditivo, Inteligência Artificial, Redes Neurais Artificiais, Fermentação, Destilação.

NEURAL NETWORK MODEL PREDICTIVE CONTROL APPLIED TO FERMENTATION AND DISTILLATION PROCESSES

Abstract: In this work, Artificial Neural Networks (ANN) were used in combination with model-based predictive control (Model Predictive Control, MPC) for the fermentation and distillation processes. The relationship between saturation and the value of the ANN objective function used during the control simulation was evaluated. The ANNs showed to be able to represent the process well, with the control being robust and without the presence of offset. The RNA saturation did not influence the controller result, its value being typically less than 0.4 for structures with linear activation functions and greater than 0.6 for the sigmoid. The saturation value is more related to the structure and weight adjustment during ANN training, without a direct relationship with controller performance.

Keywords: Predictive Control, Artificial Intelligence, Artificial Neural Networks, Fermentation, Distillation.

1. Introdução

Com o avanço tecnológico e globalização, as empresas se encontram em um cenário no qual necessitam otimizar os processos industriais, com o intuito de gerar produtos com preços competitivos que atendam as especificações dos compradores. Para garantir a qualidade dos produtos gerados, é necessário utilizar tecnologias capazes de manter o valor das variáveis que influenciam o processo industrial dentro de um limite operacional desejado, geralmente próximo à especificação realizada no projeto da planta (regime permanente, *setpoint*). Uma dessas tecnologias é o controle de processos, importante para evitar fermentos aos funcionários, manter a qualidade do produto de forma contínua e a produção desejada com custo mínimo (SMITH; CORRIPIO, 2008).

Um método de controle que vem sendo estudado há várias décadas é o controle preditivo baseado em modelo (Model Predictive Control, MPC). Esse tipo de controle

se destaca em comparação aos métodos clássicos por ser capaz de se antecipar aos erros, gerando uma resposta mais rápida e mais robusta frente aos distúrbios (CHEN et al., 2007), além de ser capaz de superar os problemas de restrições e condições operacionais da planta.

Uma alternativa aos modelos fenomenológicos e sistemas de equações não-lineares é o uso de Redes Neurais Artificiais (RNA). Modelos baseados em RNA possuem a vantagem de não necessitarem de um detalhamento completo do processo (PSICHOGIOS; UNGAR, 1992), muitas vezes precisando apenas das informações sobre as entradas e saídas. Diversos estudos mostraram que as RNAs podem ser utilizados em problemas da engenharia (CHOJI et al., 2021; IRIZAWA et al., 2021; ANTONIOLLI et al., 2021). O uso de RNAs como modelo para controle MPC apresentou bons resultados em processo de fermentação (AGUDELO, 2017) e destilação (BAHAR et al., 2004).

Uma característica importante para avaliar a RNA é o valor de saturação (RAKITIANSKAIA; ENGELBRECHT, 2015). A saturação de uma RNA ocorre quando a saída dos neurônios da camada oculta são muito próximas das extremidades limítrofes da função de ativação. Redes muito saturadas possuem duas desvantagens principais: a primeira é que os valores dos pesos dos neurônios mudam muito lentamente durante a etapa de aprendizagem, gerando uma dificuldade para que a RNA atinja a função objetivo esperada. Em segundo, a saturação prejudica a capacidade de generalização da RNA, pois as novas entradas de teste contêm exemplos diferentes daqueles utilizados no conjunto de treinamento, e redes saturadas tendem a ter uma maior confiança nos dados de treinamento em detrimento da generalização para os dados de teste. Apesar disso, esse aspecto das RNAs é raramente abordado na literatura.

O objetivo desse trabalho foi utilizar modelos baseados em RNAs para serem aplicados em um controle MPC com restrições. Foram utilizados dois processos para os estudos de caso: o primeiro foi um processo de fermentação, e o segundo foi um de destilação. Pretendeu-se, também, avaliar o desempenho das RNAs em cada caso e o desempenho dos controladores com o MPC com base nas integrais dos erros, sendo eles os valores de erro absoluto integral (IAE), erro absoluto de tempo integral (ITAE), erro quadrado integral (ISE) e o erro quadrado de tempo integral (ITSE). Além disso, analisou-se a relação entre a saturação das redes e os resultados do controle, uma vez que se espera que redes pouco saturadas possuem melhor capacidade de generalização, e isso potencialmente contribui para o resultado do controle de processos utilizando MPC com modelos baseados em RNA.

2. Métodos

Para alcançar o objetivo proposto nesse trabalho, foi necessário completar quatro etapas. A primeira foi definir o modelo matemático representativo do processo para cada estudo de caso. A segunda etapa foi gerar os dados de entrada e saída do processo e realizar o treinamento da RNA para que seja capaz de representar o processo com baixos erros. A terceira etapa foi definir o controle MPC, selecionando os parâmetros ajustáveis no controlador, bem como os distúrbios para simular a aplicação do controle. Por último, foi realizado as simulações com as RNAs treinadas como modelo para o controle MPC. Os resultados foram avaliados baseando-se nos critérios de IAE, saturação e o erro quadrático médio (MSE) obtido pela RNA. Todas as etapas são detalhadas a seguir.

O primeiro estudo de caso foi baseado em um processo contínuo de fermentação de primeira geração. A função de transferência utilizada para representar o processo foi a obtida por Violaro et al. (2018), dada pela Equação 1, que relaciona a concentração de

substrato na saída dos fermentadores em série S_4 com a vazão de alimentação do mosto F_a .

$$S_4(s) = \frac{0,161781}{4,360s+1} F_a(s) \quad (1)$$

Em que: S_4 é a variável de saída e F_a a variável de entrada. A unidade de tempo do primeiro caso é de hora (h). Os valores em estado estacionário para S_4 e F_a são, respectivamente, 0,979 kg/m³ e 104,48 m³/h.

O segundo estudo de caso foi baseado em um modelo de separação contínua de metanol-água por coluna de destilação. A Equação 2, retirada do trabalho de Wood e Berry (1973), mostra a relação entre a composição do produto de topo x_D e a vazão de refluxo R .

$$x_D(s) = \frac{12,8e^{-1s}}{16,7s+1} R(s) \quad (2)$$

Em que: x_D é a variável de saída e R a variável de entrada. A unidade de tempo para esse segundo estudo de caso é de minutos (min). Os valores em estado estacionário para x_D e R são, respectivamente, 96 wt.% metanol e 1,95 lb/min.

As funções de transferência das Equações 1 e 2 foram utilizadas para gerar os dados a serem utilizados na etapa de treinamento da RNA. Os dados foram gerados realizando distúrbios aleatórios nas variáveis manipuladas (F_a e R), gerando a dinâmica em malha aberta para as variáveis a serem controladas (S_4 e x_D). Em cada observação, a entrada para a RNA era constituída com os valores da variável manipulada e controlada nos instantes (t) e ($t - \Delta t$), e tendo como saída o valor da variável controlada no instante ($t + \Delta t$). Em ambos os estudos de caso, o intervalo de tempo para a discretização Δt foi de 0,2 unidades de tempo.

Com os dados obtidos, foi realizada a etapa de treinamento. Foram treinadas diversas estruturas de RNAs, utilizando três tipos de algoritmos de aprendizagem, número de neurônios na camada oculta variando entre 1 e 20 e quatro tipos diferentes de funções de ativação, resultando em um total de 960 RNAs treinadas para cada estudo de caso. Para os algoritmos de aprendizagem foram utilizados a retropropagação de gradiente conjugado com reiniciação de Powell-Beale (*traincgb*) (POWELL, 1977), gradiente descendente com impulso e retropropagação de taxa de aprendizagem adaptativa (*traingdx*) (HAGAN et al., 1996) e a retropropagação de Levenberg-Marquardt (*trainlm*) (MARQUARDT, 1963). A função objetivo a ser minimizada foi o erro quadrático médio (*Mean Squared Error*, MSE), representada pela Equação 3.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3)$$

Em que: N é o número de observações, y_i são os valores a serem atingidos e \hat{y}_i são os valores obtidos pela RNA. A RNA com uma camada oculta é representada matematicamente pela Equação 4 (HAYKIN, 2001).

$$\hat{y} = \phi^{(2)} \left[W^{(2)} \phi^{(1)} \left(W^{(1)} x + b^{(1)} \right) + b^{(2)} \right] \quad (4)$$

Sendo \hat{y} o *output* da RNA, $\phi^{(1)}$ e $\phi^{(2)}$ as funções de ativação, $W^{(1)}$ e $W^{(2)}$ as matrizes de pesos, $b^{(1)}$ e $b^{(2)}$ os vetores de bias e x o vetor com os *inputs* da RNA.

Foram utilizadas 4 funções de ativação diferentes, sendo representadas matematicamente pelas Equações 5 a 8, sendo n o sinal de entrada para as funções de ativação.

$$\tanh(n) = \frac{2}{1 + e^{-2n}} - 1 \quad (5)$$

$$\log \text{sig}(n) = \frac{1}{1 + e^{-n}} \quad (6)$$

$$\text{purelin}(n) = n \quad (7)$$

$$\text{satlins}(n) = \begin{cases} -1, n < -1 \\ n, -1 \leq n \leq 1 \\ 1, 1 < n \end{cases} \quad (8)$$

Com as redes treinadas, foi calculado o valor de saturação dos neurônios da camada oculta (RAKITIANSKAIA; ENGELBRECHT, 2015), visando analisar se esse parâmetro influencia no sucesso do uso da RNA como modelo para o processo. As Equações 9 a 11 representam o cálculo para a saturação (RAKITIANSKAIA; ENGELBRECHT, 2015).

$$\bar{g}_b = \begin{cases} \sum_{k=1}^{f_b} \frac{g(\text{net})_k}{f_b}, f_b > 0 \\ 0, f_b \leq 0 \end{cases} \quad (9)$$

$$\bar{g}'_b = \frac{2(\bar{g}_b - g_L)}{g_U - g_L} - 1 \quad (10)$$

$$\varphi_B = \frac{\sum_{b=1}^B |\bar{g}'_b| f_b}{\sum_{b=1}^B f_b} \quad (11)$$

Sendo \bar{g}_b a média dos valores dos sinais de saída, f_b o número dos sinais de saída, $g(\text{net})$ as saídas da função de ativação para todos os valores de net , \bar{g}'_b é a saída na escala [-1,1] para a função de ativação g no domínio $[g_L, g_U]$, B o número total de classes para o histograma e φ_B o valor de saturação. Se o valor de \bar{g}'_b for distribuído de maneira uniforme, a saturação φ_B se aproxima de 0,5. Quanto mais próximo de 0,5 for o valor de φ_B , menos a RNA está saturada.

O conceito do controle MPC é baseado no cálculo otimizado de ações de controle para as variáveis manipuladas com base no comportamento futuro de uma planta. A otimização é obtida resolvendo um problema de minimização que busca diminuir os erros entre os valores das variáveis controladas e o *setpoint* desejado, além de obedecer a critérios de restrição que representam as limitações do processo simulado. O resultado obtido é uma sequência de ações ótima nas variáveis manipuladas em um horizonte de

controle H_c , resultando nos melhores valores para as variáveis controladas no horizonte de predição H_p (KLAUCO, 2019). A formulação matemática para o MPC utilizado nesse trabalho, para ambos os estudos de caso, é dada pelas Equações 12 a 17.

$$\min J = \sum_{k=0}^{H_p-1} (y_k - SP_k)^2 Q_p + \sum_{k=0}^{H_c-1} \Delta u_k^2 Q_c \quad (12)$$

Submetido às seguintes restrições:

$$X_{k+1} = f(X_k, u_k), k = 0, 1, \dots, H_p - 1 \quad (13)$$

$$y_k = g(X_k, u_k), k = 0, 1, \dots, H_p - 1 \quad (14)$$

$$u_{\min} \leq u_k \leq u_{\max}, k = 0, 1, \dots, H_p - 1 \quad (15)$$

$$\Delta u_{\min} \leq \Delta u_k \leq \Delta u_{\max}, k = 0, 1, \dots, H_p - 1 \quad (16)$$

$$\Delta u_k = 0, k \geq H_c \quad (17)$$

Sendo J a função custo, H_p o horizonte de predição, H_c o horizonte de controle, y_k os valores da variável controlada, SP_k os valores de *setpoint*, Q_p o valor de penalização para o erro entre os valores da variável controlada e *setpoint*, Q_c a penalização para a variação da variável manipulada, u_k a variável manipulada no horizonte de predição, Δu_k a variação dos valores da variável manipulada ($\Delta u_k = u_k - u_{k-1}$), u_{\min} e u_{\max} os valores máximo e mínimo, respectivamente, para a variável manipulada, Δu_{\min} e Δu_{\max} os valores de variação máximo e mínimo, respectivamente, para a variável manipulada e X_k o vetor com as variáveis de estado do processo.

Para cada caso, foi necessário definir os valores de u_{\min} , u_{\max} , Δu_{\min} , Δu_{\max} , H_p , H_c , Q_p e Q_c . Os valores utilizados para os modelos dos processos de fermentação e destilação são mostrados na Tabela 1, sendo obtidos via heurística.

A cada intervalo de tempo Δt , o sistema composto pelas Equações 12 a 17 foi resolvido, com a minimização da função objetivo sendo realizada pelo método de ponto interior (POWELL, 1978; WALTZ et al., 2006). O valor de $\Delta u(k=0)$ obtido pela minimização é enviado e utilizado na simulação do processo.

Tabela 1 - Valores dos parâmetros para o controle MPC

Parâmetro	Fermentação	Destilação	Parâmetro	Fermentação	Destilação
u_{\min}	50 kg/m ³	1,30 lb/min	H_c	1	1
u_{\max}	150 kg/m ³	2,50 lb/min	H_p	16	16
Δu_{\min}	- 5 kg/m ³	- 0,05 lb/min	Q_p	1000	100
Δu_{\max}	+ 5 kg/m ³	+ 0,05 lb/min	Q_c	100	10

Para avaliar o comportamento do controlador MPC com as RNAs, foram definidos distúrbios nos *setpoints* durante as simulações. Para o processo de fermentação, três degraus foram realizados ao longo de 150h de simulação. O primeiro ocorreu em 10h, com magnitude de $-0,5 \text{ kg/m}^3$. O segundo ocorreu em 50h, com magnitude de $+0,8 \text{ kg/m}^3$. O último distúrbio ocorreu em 100h, com magnitude de $+1 \text{ kg/m}^3$. Para o processo de destilação, a simulação realizada foi de 300min, com três degraus no *setpoint*: O primeiro degrau foi feito em 10min com magnitude de $-2 \text{ wt. \% metanol}$, o segundo foi feito em 110min com magnitude de $+1 \text{ wt. \% metanol}$ e o último foi feito em 180min, com magnitude de $+3 \text{ wt. \% metanol}$. Para avaliar o desempenho do controle, foram analisados os valores de quatro critérios de erro: IAE, ITAE, ISE e ITSE, sendo representados, respectivamente, pelas Equações 18 a 21.

$$IAE = \int_0^{\infty} |e(t)| dt \quad (18)$$

$$ITAE = \int_0^{\infty} t |e(t)| dt \quad (19)$$

$$ISE = \int_0^{\infty} e(t)^2 dt \quad (20)$$

$$ITSE = \int_0^{\infty} t \cdot e(t)^2 dt \quad (21)$$

3. Resultados e discussões

Nessa seção são apresentados os resultados das melhores redes de cada caso. Também é apresentado o mapa de calor de todas as simulações, bem como a simulação do controle MPC com RNA para a estrutura que apresentou o menor valor de IAE.

Para o caso da fermentação, a Tabela 2 mostra os resultados para as dez melhores estruturas que apresentaram os valores mais baixos da função objetivo. Já a Tabela 3 mostra os resultados para as estruturas que apresentaram o melhor desempenho no controle, refletida pelos menos valores no critério de erro IAE. Em ativações é indicado as funções de ativação utilizadas na camada oculta e de saída, respectivamente. A coluna neurônios indica a quantidade de neurônios utilizados na camada oculta.

As RNAs que alcançaram os valores mais baixos de MSE (Tabela 2) foram obtidas utilizando como algoritmo de aprendizagem o método *trainlm* e funções de ativação linear na camada intermediária e de saída (Purelin-Purelin). Em termos de número de neurônios, não foi possível observar uma tendência de diminuição do MSE com o aumento do número de neurônios. O maior valor de saturação foi de 0,535, com as demais sendo menores que esse valor. A média de saturação obtida foi de 0,406. Percebe-se que a maior parte das RNAs possuíam uma saturação menor que 0,5. Nota-se que os valores de ISE, IAE, ITSE e ITAE não foram os melhores em comparação com os valores da Tabela 3, indicando que uma RNA com valor de MSE muito baixo não necessariamente tem valor de IAE baixo durante a simulação do controle. Isso ocorre pois uma RNA com valor muito baixo de MSE pode indicar um sobreajuste nos dados de treinamento diminuindo sua capacidade de generalização, o que prejudica a eficiência do modelo e, conseqüentemente, sua aplicação no MPC, uma vez que na simulação existem dinâmicas diferentes daquelas disponíveis nos dados de treinamento.

Analisando-se as redes que possuíam os melhores desempenhos no critério de erro IAE, observa-se que *traincgb* e *traindx* foram os algoritmos que apresentaram os melhores resultados. As funções de ativação foram todas do tipo Purelin-Purelin, e nota-se que também não há uma tendência em melhorar o desempenho com o aumento do

número de neurónios. A melhor RNA (traincgb - Purelin-Purelin 4-4-1) possuiu uma saturação de 0,286, apesar de haver boas redes nos quais a saturação foi maior que 0,5. A média do valor de saturação obtido foi de 0,417, sendo praticamente igual a média obtida das melhores redes que apresentaram os valores mais baixos de MSE.

Tabela 2 - Estruturas de RNAs que apresentaram os menores valores de MSE para a fermentação

Algoritmo	Ativações	Neurónios	MSE	φ_B	IAE	ITAE	ISE	ITSE
trainlm	Purelin-Purelin	19	$1,02 \cdot 10^{-10}$	0,306	7,99	$5,30 \cdot 10^2$	3,91	$2,85 \cdot 10^2$
trainlm	Purelin-Purelin	4	$8,83 \cdot 10^{-11}$	0,485	7,67	$5,05 \cdot 10^2$	3,95	$2,89 \cdot 10^2$
trainlm	Purelin-Purelin	5	$2,97 \cdot 10^{-9}$	0,428	7,67	$5,05 \cdot 10^2$	3,95	$2,89 \cdot 10^2$
trainlm	Purelin-Purelin	5	$3,21 \cdot 10^{-10}$	0,378	7,67	$5,05 \cdot 10^2$	3,95	$2,89 \cdot 10^2$
trainlm	Purelin-Purelin	3	$2,19 \cdot 10^{-10}$	0,529	7,67	$5,05 \cdot 10^2$	3,95	$2,89 \cdot 10^2$
trainlm	Purelin-Purelin	6	$2,48 \cdot 10^{-8}$	0,386	7,67	$5,05 \cdot 10^2$	3,95	$2,89 \cdot 10^2$
trainlm	Purelin-Purelin	1	$6,40 \cdot 10^{-11}$	0,535	7,67	$5,05 \cdot 10^2$	3,95	$2,89 \cdot 10^2$
trainlm	Purelin-Purelin	19	$8,56 \cdot 10^{-11}$	0,334	10,9	$7,81 \cdot 10^2$	5,05	$3,75 \cdot 10^2$
trainlm	Purelin-Purelin	20	$1,71 \cdot 10^{-7}$	0,336	1490	$1,68 \cdot 10^5$	$2,73 \cdot 10^4$	$3,47 \cdot 10^6$
trainlm	Purelin-Purelin	15	$2,54 \cdot 10^{-10}$	0,329	7,67	$5,05 \cdot 10^2$	3,95	$2,89 \cdot 10^2$

Tabela 3 - Estruturas de RNAs que apresentaram os menores valores de IAE na simulação para a fermentação

Algoritmo	Ativações	Neurónios	MSE	φ_B	IAE	ITAE	ISE	ITSE
traincgb	Purelin-Purelin	4	$1,16 \cdot 10^{-5}$	0,286	6,36	$4,17 \cdot 10^2$	3,57	$2,60 \cdot 10^2$
traingdx	Purelin-Purelin	11	$6,34 \cdot 10^{-5}$	0,323	6,47	$4,23 \cdot 10^2$	3,62	$2,64 \cdot 10^2$
traingdx	Purelin-Purelin	9	$2,28 \cdot 10^{-4}$	0,285	6,53	$4,28 \cdot 10^2$	3,52	$2,56 \cdot 10^2$
traincgb	Purelin-Purelin	8	$2,21 \cdot 10^{-4}$	0,360	7,08	$4,72 \cdot 10^2$	3,42	$2,50 \cdot 10^2$
traingdx	Purelin-Purelin	7	$3,96 \cdot 10^{-5}$	0,459	7,14	$4,67 \cdot 10^2$	3,94	$2,86 \cdot 10^2$
traincgb	Purelin-Purelin	1	$2,14 \cdot 10^{-6}$	0,534	7,23	$4,76 \cdot 10^2$	3,75	$2,73 \cdot 10^2$
traincgb	Purelin-Purelin	17	$2,68 \cdot 10^{-5}$	0,312	7,26	$5,00 \cdot 10^2$	3,53	$2,57 \cdot 10^2$
traincgb	Purelin-Purelin	2	$3,21 \cdot 10^{-6}$	0,727	7,37	$4,85 \cdot 10^2$	3,86	$2,82 \cdot 10^2$
trainlm	Purelin-Purelin	14	$2,67 \cdot 10^{-10}$	0,352	7,51	$4,95 \cdot 10^2$	3,94	$2,88 \cdot 10^2$
traingdx	Purelin-Purelin	1	$2,97 \cdot 10^{-5}$	0,530	7,51	$4,91 \cdot 10^2$	4,00	$2,91 \cdot 10^2$

A Figura 1 mostra o comportamento da variável controlada e manipulada. É possível observar que o controle MPC com modelo de RNA conseguiu manter a variável controlada no *setpoint* e levar a variável controlada para o novo *setpoint* após os distúrbios realizados durante todo o tempo de simulação. A dinâmica foi robusta, com S_4 atingindo o *setpoint* sem oscilação. Além disso, é possível observar que o modelo representou bem o processo, uma vez que não houve *offset*.

A Figura 2 ilustra os mapas de calor para o caso de fermentação. São dispostas 3 três variáveis: o resultado de IAE para a simulação com o controle MPC, o MSE obtido no treinamento da RNA e o valor de saturação da rede. As funções de ativação Tanh, Logsig, Purelin e Satlins foram abreviadas para, respectivamente, Tan, Log, Pur e Sat. Utilizando-se a função de ativação Tanh na camada intermediária (Fig. 2a) nota-se que

uma diminuição do valor de MSE não é requisito obrigatório para obter o valor mais baixo de ISE na simulação do processo. Além disso, percebe-se que essas redes possuíam um valor de saturação alto, acima de 0,7 para a maioria das redes. As redes com melhor desempenho para o controle foram aquelas com funções de ativação do tipo Purelin e Satlins na camada de saída. Para as redes do tipo Satlins (Figura 2b), é possível observar que aquelas com valores mais baixos de MSE durante a etapa de treinamento implicaram em valores menores de IAE.

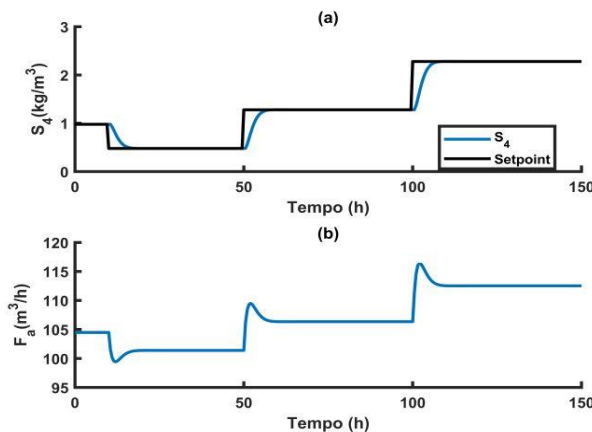


Figura 1 - Resultado do controle para o processo de fermentação (a) variável controlada e (b) variável manipulada

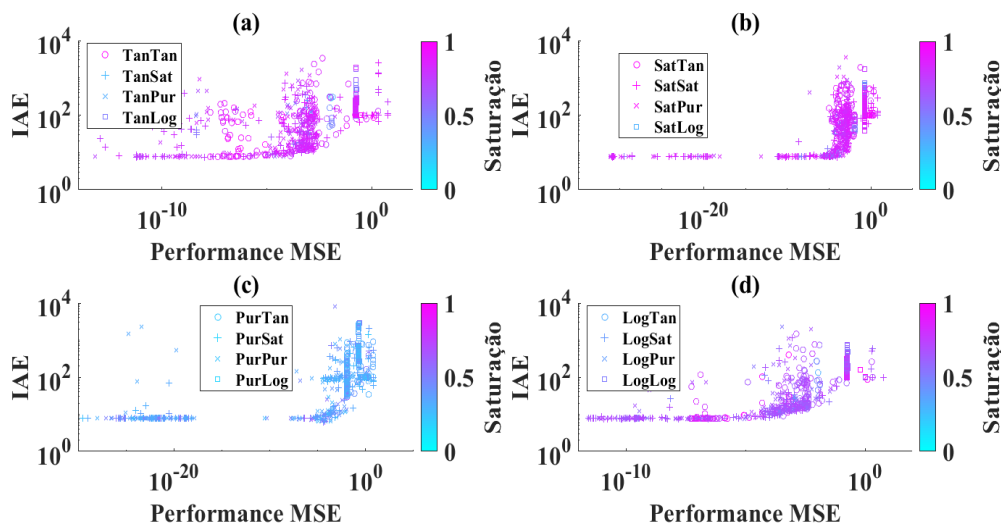


Figura 2 - IAE, MSE e saturação para todas as simulações do processo de fermentação para funções de ativação da camada intermediária de (a) Tan (b) Sat (c) Pur e (d) Log

Destaca-se que os melhores valores ocorreram com uma combinação de funções Satlins e Purelin na camada de saída. A maioria das estruturas com função de ativação do tipo Satlins e Logsig na saída não foram capazes de realizar o controle do processo de maneira adequada. A maior parte das estruturas possuíam uma saturação acima de 0,6.

A Figura 2c mostra que a saturação para as redes com purelin na camada intermediária foi baixa, com valores abaixo de 0,4. O comportamento é semelhante com o que ocorreu na Figura 2b, entretanto é possível observar alguns pontos nos quais tiveram valores baixos de MSE, porém altos valores de IAE. As estruturas com Logsig na camada intermediária como função de ativação (Figura 2d) tiveram os maiores valores de MSE, com várias estruturas não sendo capazes de representar o modelo e garantir um controle satisfatório. De maneira geral, as melhores estruturas foram aquelas com funções de ativação do tipo lineares (Purelin e Satlins), que garantiram valores baixos de MSE e IAE.

Para o caso de destilação, a Tabela 4 mostra os resultados das RNAs que apresentaram os valores de MSE mais baixos. A Tabela 5 mostra as RNAs que, combinadas com o MPC, resultaram em um controle que apresentaram os valores mais baixos de IAE nas simulações.

Tabela 4 - Estruturas de RNAs que apresentaram os menores valores de MSE para a destilação

Algoritmo	Ativações	Neurônio	MSE	ϕ_B	IAE	ITAE	ISE	ITSE
trainlm	Logsig-Purelin	5	1,40.10⁻⁶	0,665	4,70.10¹	5,71.10³	8,44.10¹	1,14.10⁴
trainlm	Tansig-Satlins	1	1,43.10 ⁻⁶	0,594	4,69.10 ¹	5,69.10 ³	8,42.10 ¹	1,14.10 ⁴
trainlm	Purelin-Purelin	18	1,53.10 ⁻⁶	0,345	4,68.10 ¹	5,67.10 ³	8,42.10 ¹	1,14.10 ⁴
trainlm	Logsig-Satlins	12	1,55.10 ⁻⁶	0,677	4,68.10 ¹	5,68.10 ³	8,42.10 ¹	1,14.10 ⁴
trainlm	Satlins-Purelin	13	1,60.10 ⁻⁶	0,870	4,88.10 ¹	6,06.10 ³	8,69.10 ¹	1,19.10 ⁴
trainlm	Satlins-Satlins	9	1,60.10 ⁻⁶	0,815	4,68.10 ¹	5,67.10 ³	8,40.10 ¹	1,14.10 ⁴
trainlm	Satlins-Satlins	11	1,62.10 ⁻⁶	0,873	4,66.10 ¹	5,63.10 ³	8,41.10 ¹	1,14.10 ⁴
trainlm	Satlins-Purelin	7	1,64.10 ⁻⁶	0,857	4,67.10 ¹	5,66.10 ³	8,42.10 ¹	1,14.10 ⁴
trainlm	Satlins-Purelin	12	1,65.10 ⁻⁶	0,764	1,73.10 ³	3,24.10 ⁵	1,21.10 ⁴	2,51.10 ⁶
trainlm	Logsig-Purelin	7	1,65.10 ⁻⁶	0,678	4,69.10 ¹	5,67.10 ³	8,43.10 ¹	1,14.10 ⁴

Entre as dez melhores redes treinadas (Tabela 4), todas foram obtidas com o algoritmo de aprendizagem de trainlm, e funções de ativação diversas, sendo quatro delas do tipo Satlins-Purelin. Da mesma forma que ocorreu no caso da fermentação, não houve forte influência do número de neurônios no valor do MSE. A maioria das redes apresentaram um valor de saturação maior que 0,6. A média de saturação para as dez redes foi de 0,714. Nos valores de ISE, IAE, ITAE e ITSE, os valores foram levemente maiores em comparação aos da Tabela 5. Entre as redes treinadas, aquelas que apresentaram os menores valores para IAE no controle (Tabela 5) foram aquelas treinadas pelos algoritmos traincgb e traingdx, com as funções de ativação sendo combinações de Purelin e Satlins.

Da mesma forma que ocorreu na fermentação, um valor alto de MSE não necessariamente garante que a RNA está inadequada para ser utilizada no controle. Para as dez redes com os valores mais baixos de MSE, a média obtida para o MSE foi de 1,57.10⁻⁶, enquanto que para as dez redes com menor valor de IAE a média para o MSE foi de 7,52.10⁻⁵. Isso indica que o valor de MSE para as redes com melhor desempenho no controle foi cerca de 50 vezes maior aos valores de MSE das redes que alcançaram o menor valor de MSE durante o treinamento. A média de saturação da Tabela 5 foi de 0,471, sendo bem diferente da média da Tabela 4. Entretanto, houveram redes com

saturação 0,274 e 0,806, que apesar de divergir bastante da média, possuíram valores baixos de IAE. Isso é um indicativo que a saturação da rede não é um parâmetro confiável para garantir que a RNA irá se comportar adequadamente como um modelo para o controle MPC.

Tabela 5 - Estruturas de RNAs que apresentaram os menores valores de IAE na simulação para a destilação

Algoritmo	Ativações	Neurônio	MSE	φ_B	IAE	ITAE	ISE	ITSE
traincgb	Purelin-Purelin	4	$8,76 \cdot 10^{-5}$	0,448	$4,24 \cdot 10^1$	$5,17 \cdot 10^3$	$8,20 \cdot 10^1$	$1,11 \cdot 10^4$
traingdx	Purelin-Satlins	6	$8,29 \cdot 10^{-5}$	0,274	$4,30 \cdot 10^1$	$5,22 \cdot 10^3$	$8,20 \cdot 10^1$	$1,11 \cdot 10^4$
traingdx	Purelin-Satlins	2	$7,85 \cdot 10^{-5}$	0,473	$4,31 \cdot 10^1$	$5,22 \cdot 10^3$	$8,22 \cdot 10^1$	$1,11 \cdot 10^4$
traincgb	Purelin-Purelin	2	$1,64 \cdot 10^{-4}$	0,338	$4,31 \cdot 10^1$	$5,22 \cdot 10^3$	$8,26 \cdot 10^1$	$1,12 \cdot 10^4$
traingdx	Satlins-Purelin	2	$6,54 \cdot 10^{-5}$	0,806	$4,33 \cdot 10^1$	$5,26 \cdot 10^3$	$8,23 \cdot 10^1$	$1,11 \cdot 10^4$
traingdx	Purelin-Purelin	1	$1,17 \cdot 10^{-4}$	0,610	$4,33 \cdot 10^1$	$5,26 \cdot 10^3$	$8,20 \cdot 10^1$	$1,11 \cdot 10^4$
traincgb	Purelin-Purelin	2	$5,06 \cdot 10^{-5}$	0,478	$4,33 \cdot 10^1$	$5,25 \cdot 10^3$	$8,25 \cdot 10^1$	$1,12 \cdot 10^4$
traingdx	Purelin-Purelin	3	$3,27 \cdot 10^{-5}$	0,511	$4,35 \cdot 10^1$	$5,26 \cdot 10^3$	$8,24 \cdot 10^1$	$1,11 \cdot 10^4$
traingdx	Purelin-Purelin	5	$5,17 \cdot 10^{-5}$	0,366	$4,37 \cdot 10^1$	$5,27 \cdot 10^3$	$8,22 \cdot 10^1$	$1,11 \cdot 10^4$
traingdx	Purelin-Purelin	3	$2,17 \cdot 10^{-4}$	0,407	$4,37 \cdot 10^1$	$5,31 \cdot 10^3$	$8,19 \cdot 10^1$	$1,11 \cdot 10^4$

A Figura 3 mostra o comportamento das variáveis controladas e manipuladas para a simulação da rede que apresentou o menor valor de IAE (traincgb - Purelin-Purelin-4-4-1). Pode-se observar que a dinâmica da variável controlada foi robusta, com oscilação e overshoot praticamente nulos. A variável controlada conseguiu atingir o *setpoint* sem a presença de offset, da mesma forma que ocorreu no processo de fermentação. A RNA foi capaz de representar bem o processo e a aplicação em conjunto com o MPC permitiu realizar o controle com baixos erros. A Figura 4 ilustra o mapa de calor relacionando os valores de MSE da RNA, saturação e IAE obtido nas simulações. A Figura 4a mostra os valores de cada simulação considerando as estruturas com a função de ativação Tanh na camada intermediária. As Figuras 4b, 4c e 4d ilustram os resultados considerando como função de ativação na camada intermediária as estruturas Satlins, Purelin e Logsig, respectivamente. As funções de ativação Tanh, Logsig, Purelin e Satlins foram abreviadas para, respectivamente, Tan, Log, Pur e Sat.

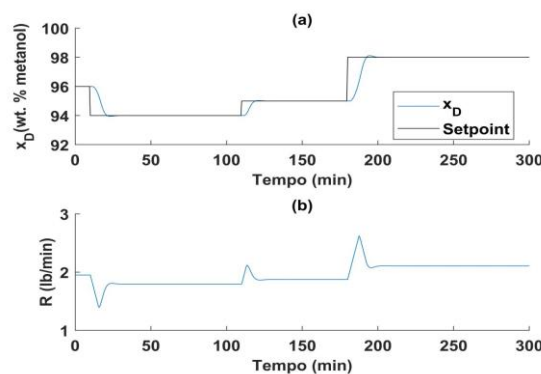


Figura 3 - Resultado do controle NNMPC para o caso de destilação (a) variável controlada e (b) variável manipulada

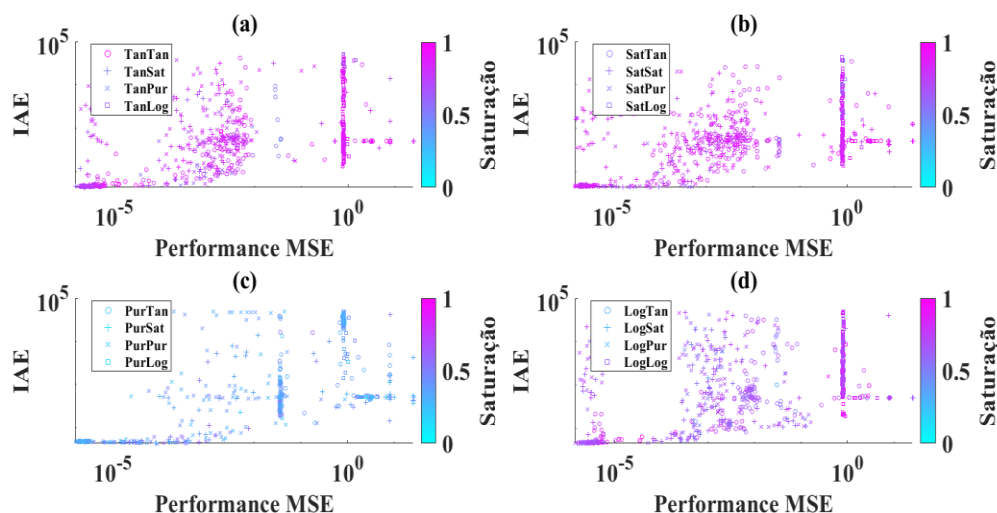


Figura 4 - IAE, MSE e saturação para todas as simulações do processo de destilação para funções de ativação da camada intermediária de (a) Tan (b) Sat (c) Pur e (d) Log

Diferente do que ocorreu na fermentação, o resultado do mapa de calor para o processo de destilação apresentou muitos pontos espalhados, indicando que houve estruturas que resultaram em um controlador ineficiente mesmo com valores baixos de MSE. Alguns comportamentos foram semelhantes, como resultados melhores para as combinações de funções de ativação Purelin e Satlin, bem como os valores reduzidos de saturação para as estruturas que possuíam Purelin na primeira camada. Há uma melhora nos valores de IAE conforme o valor de MSE diminui, apesar de isso não garantir que a RNA possua uma capacidade de generalização adequada para ser aplicada junto com o MPC.

Da mesma forma que ocorreu na fermentação, o valor de saturação não mostrou influência nos valores de MSE e IAE. Os valores foram menores que 0,4 para as estruturas com o Purelin na camada intermediária, e maior que 0,6 para as outras funções de ativação. Em especial as estruturas com Tanh na camada intermediária apresentaram valores mais altos de saturação. As estruturas do tipo sigmóides são mais facilmente saturadas do que as outras (LECUN et al., 2012), já as estruturas com função de ativação Purelin possuem uma saturação menor, já que essa função não possui restrição de valor mínimo e máximo. Nesse estudo, esse comportamento foi ressaltado nos resultados de ambos os processos de destilação e fermentação.

4. Conclusão

Neste trabalho foram utilizadas RNAs para representar o modelo de um processo de fermentação e de destilação. As RNAs foram usadas em conjunto com o controle MPC com restrições, objetivando manter o valor da variável controlada no *setpoint* mediante certos distúrbios na simulação. Em ambos os estudos de caso, o controle foi capaz de atingir o *setpoint* sem a presença de offset e de maneira robusta, com overshoot e oscilação praticamente inexistente. Isso mostrou a eficiente capacidade das RNAs atuarem em conjunto com o controle MPC para representar os processos. As simulações mostraram que existe uma tendência na diminuição do erro IAE para redes com baixo MSE, apesar de isso não garantir que a determinada estrutura consiga controlar de maneira adequada com baixo valor de IAE, necessitando que mais de uma estrutura seja testada. A saturação das RNAs não influenciou no resultado, sendo seu valor

tipicamente menor que 0,4 para as estruturas com funções de ativação lineares e maior que 0,6 para as sigmóides, indicando que seu valor está mais relacionado com a estrutura da RNA e não com os valores de MSE e IAE.

Referências

- ANTONIOLLI, E.A.; SANTOS, J.A.A.; GUEDES, A.L.; PASA, L.A. & SCHMIDT, C.A.P. *Revisão sistemática: aplicação de redes neurais para previsão do consumo de energia elétrica*. Engenharia e Tecnologia, Vol. 13, n. 2, 2021.
- AGUDELO, W. E. H. *Modelagem e controle de processos para a produção de etanol de primeira e segunda geração aplicando redes neurais*. 2017. Tese (Doutorado em Engenharia Química) – Universidade Estadual de Campinas, Campinas, 2017.
- BAHAR, A.; OZGEN, C.; LEBLEBICIOGLU, K. & HALICI, U. *Artificial Neural Network Estimator Design for the inferential Model Predictive Control of an Industrial Distillation Column*. Industrial and Engineering Chemistry Research, Vol. 43, n. 19, p. 6102-6111, 2004.
- CHEN, X. S.; ZHAI, J.Y.; LI, S. H. & MINER, Q. L. *Application of model predictive control in ball mill grinding circuit*. Minerals Engineering, Vol. 20, n. 11, p.1099-1108, 2007.
- CHOJI, T.T.; OTTAIANO, G.Y.; SEOLATTO, A.A.; SILVA, E.A & MARTINS, T.D. *Cinética da biossorção da mistura Ni²⁺-Cr³⁺ via redes neurais artificiais*. Engenharia e Tecnologia, Vol. 13, n. 1, 2021.
- HAGAN, M.T.; DEMUTH, H. B. & BEALE, M.H. *Neural Network Design*, Boston, MA: PWS Publishing, 1996.
- HAYKIN, S. *Redes neurais, princípios e prática*. 3. ed. – Porto Alegre: Bookman, 2001.
- IRIZAWA, I.A.; MARTINS, T.D. & VEGGI, P.C. *MODELAGEM DA CINÉTICA DE EXTRAÇÃO SUPERCRÍTICA DE ARTEMISININA VIA REDES NEURAIAS ARTIFICIAIS*. Engenharia e Tecnologia, Vol. 13, n. 2, 2021.
- KLAUCO, M. & KVASNICA, M. *MPC-Based Reference Governors Theory and Case Studies*. Suíça: Editora Springer Nature Switzerland AG, 2019.
- LECUN, Y. A.; BOTTOU, L.; ORR, G. B. & MULLER, K. R. *Efficient backprop*. In: Neural Networks: Tricks of the trade. Springer, 2012, p. 9-48.
- MARQUARDT, D. *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*. SIAM Journal on Applied Mathematics, Vol. 11, n. 2, p. 431–441, 1963.
- POWELL, M.J.D. *Restart procedures for the conjugate gradient method*. Mathematical Programming, Vol. 12, p. 241–254, 1977.
- POWELL, M. J. D. *The Convergence of Variable Metric Methods For Nonlinearly Constrained Optimization Calculations*. Nonlinear Programming 3 Academic Press, p. 27-63, 1978.
- PSICHOGIOS, D. C. & UNGAR, L. H. *A hybrid neural network-first principles approach to process modeling*. AIChE Journal, Vol. 38, n. 10, p. 1499-1511, 1992.
- RAKITIANSKAIA, A. & ENGELBRECHT, A. *Measuring Saturation in Neural Networks*. IEEE Symposium Series on Computational Intelligence, p. 1423-1430, 2015.
- SMITH, C. A. & CORRIPIO, A. B. *Princípios E Prática do Controle Automático de Processo*. Rio de Janeiro: Editora LTC, 2008.
- VIOLARO, F. A. R.; RIVERA, E. C. & ALVAREZ, L. A. *Multivariable Model Predictive Control of a Continuous Fermentation Unit for First-Generation Ethanol Production*. Chemical Engineering Transactions, Vol. 65, p. 67-72, 2018.
- WALTZ, R. A.; MORALES, J. L.; NOCEDAL, J. & ORBAN, D. *An interior algorithm for nonlinear optimization that combines line search and trust region steps*. Mathematical Programming, Vol. 107, n. 3, p. 391–408, 2006.
- WOOD, R. K. & BERRY, M. W. *Terminal composition control of a binary distillation column*. Chemical Engineering Science, Vol. 28, p. 1707-1717, 1973.