DOI: 10.48011/asba.v2i1.1713

Identificação de Sistemas Dinâmicos Não Lineares Multivariáveis com Redes de Estado de Eco

Bernardo B. Schwedersky * **Rodolfo C. C. Flesch** **
Hiago A. S. Dangui *

* *Programa de Pós-Graduação em Automação e Sistemas, Universidade Federal de Santa Catarina, Florianópolis, Brazil (e-mail: [bernardo.schwedersky, hiagodangui]@gmail.com).*
** *Departamento de Automação e Sistemas, Universidade Federal de Santa Catarina, Florianópolis, Brazil (e-mail: rodolfo.flesch@ufsc.br)*

Abstract: This paper presents an echo state network based identification method for MIMO nonlinear dynamic systems. The ESN is a recurrent neural network capable to model nonlinear dynamics with computationally efficient training algorithms. A echo state network architecture for process identification tasks and a procedure to create and tune the main hyperparameters are proposed. The proposed method is evaluated using both a simulated pH neutralization process and a real refrigerating compressor test rig. In both cases the proposed model architecture presented better results than the linear, Hammerstein and extreme learning machines models, considered as baseline. The evaluation results show that the proposed ESN-based system identification method can be used directly to obtain MIMO models for nonlinear processes using a computationally efficient training algorithm.

Resumo: Este artigo apresenta um método de identificação baseado em redes de estado de eco para sistemas dinâmicos não lineres multivariáveis. A rede de estado de eco é uma rede neural recorrente, capaz de modelar dinâmicas não lineares, utilizando algoritmos de treinamento computacionalmente eficientes. Uma arquitetura de rede de estado de eco para identificação de sistemas e o processo para criação e ajuste dos hiperparâmetros do modelo são propostos. O método proposto e avaliado por meio de um processo de neutralização de pH simulado e uma bancada real de avaliação de compressores de refrigeração. Em ambos os casos, a estratégia para identificação de sistemas proposta apresentou melhores resultados que os métodos de comparação, que foram modelos lineares, modelos Hammerstein e máquinas de aprendizado extremo. Os resultados demonstraram que a estratégia proposta pode ser usada diretamente para obtenção de modelos multivariáveis não lineares utilizando um algoritmo de treinamento computacionalmente eficiente.

Keywords: System identification; echo state networks; nonlinear systems; recurrent neural networks; ridge regression.

Palavras-chaves: Identificação de sistemas; redes de estado de eco; sistemas não lineares; redes neurais recorrentes; regressão com regularização.

1. INTRODUÇÃO

A modelagem de sistemas reais possui importância fundamental na engenharia. Tarefas como a análise de processos existentes, projeto de novos processos e controladores, otimização, simulação, detecção de falhas, entre outras, são baseadas em modelos de processos e diretamente impactadas pela qualidade dos modelos obtidos. Em situações nas quais a obtenção de um modelo fenomenológico é difícil, técnicas de identificação de sistemas são tipicamente empregadas (Nelles, 2001).

Muitos processos apresentam comportamentos não lineares estáticos e dinâmicos, especialmente quando amplas áreas de operação são consideradas. Ferramentas de aprendizado profundo (DL, do inglês *deep learning*), em especial as redes neurais recorrentes (RNN, do inglês *recurrent neural networks*), configuram-se como uma opção interessante para a modelagem experimental de sistemas, por permitirem a identificação de processos nos quais existe pouco ou nenhum conhecimento da física que rege o processo (Isermann and Münchof, 2011; Goodfellow et al., 2016).

A aplicação de RNNs na identificação de sistemas dinâmicos é um tema vasto e muito investigado na literatura. Logo após a proposição das redes *perceptron* de múltiplas camadas (MLP, do inglês *MultiLayer Perceptron*), surgiram os primeiros estudos utilizando RNN para identificação, explorando arquiteturas MLP (Fernandez et al., 1990; Ayoubi, 1994). Apesar de as redes MLP recorrentes ainda serem empregadas para problemas de identificação de sistemas, como é o caso de Boussaada et al. (2018),

que as empregam para identificar o comportamento da radiação solar diária, novas arquiteturas de RNN vêm sendo exploradas, buscando uma maior qualidade do modelo resultante e uma maior facilidade na sua obtenção. Uma arquitetura que se destaca, devido ao seu treinamento computacionalmente eficiente, é a rede de estado de eco (ESN, do inglês *Echo State Network*), proposta em (Jaeger, 2001).

Algumas abordagens para identificação de sistemas não lineares baseadas em ESN foram apresentadas na literatura. A referência Jaeger (2003) explorou a utilização da rede ESN em uma abordagem *online*. Por sua vez, em Rodan and Tino (2010) foram exploradas arquiteturas simplificadas de ESN para identificação de sistemas dinâmicos. O estudo recente apresentado em Yang et al. (2019) explora métodos para obtenção *online* de modelos baseados ESN com algoritmos recursivos esparsos. No domínio das aplicações, alguns trabalhos recentes vêm explorando a utilização desse modelo para identificação de diversos tipos de sistemas dinâmicos, como em sistemas de geração eólica Chen et al. (2019), poços de petróleo Antonelo et al. (2017) e sistemas de refrigeração Schwedersky et al. (2018).

De forma geral, a utilização da ESN em problemas de identificação de sistemas não lineares ainda é incipiente, e os trabalhos que a usam apresentam arquiteturas e métodos de ajuste do modelo específicos para cada aplicação. Este artigo apresenta uma arquitetura geral e um método de ajuste para identificação de sistemas dinâmicos multivariáveis e não lineares por meio de redes de estado de eco. Na seção 2 é proposta uma arquitetura simplificada geral, sendo apresentada a forma de construção de modelos ESN, e proposto um método de ajuste dos principais hiperparâmetros. A aplicação do método proposto é demonstrada por meio de dois estudos de caso, o primeiro consistindo na identificação de um processo de neutralização de pH e o segundo em uma bancada de ensaios de compressores herméticos de refrigeração, apresentados na seção 3. Os resultados obtidos são comparados com os apresentados por técnicas tradicionais de identificação de sistemas, como modelos lineares e não lineares, na seção 4, seguidos pelas conclusões, na seção 5.

2. IDENTIFICAÇÃO COM REDE NEURAL DE ESTADO DE ECO

A ESN é uma arquitetura de RNN baseada na computação de reservatório Lukoševičius and Jaeger (2009), um paradigma alternativo ao dos modelos tradicionais de RNN. Esse paradigma se baseia na utilização de um sistema dinâmico não linear com pesos fixos gerados aleatoriamente, por exemplo uma rede neural, para mapear as entradas em um espaço de alta dimensionalidade, no qual a tarefa de classificação ou de regressão é realizada eficientemente Lukoševičius (2012).

O conjunto de estados desse sistema dinâmico compõe a estrutura chamada de reservatório (do inglês, *reservoir*). Tal estrutura pode ser entendida como um *kernel* temporal, projetando a entrada em um espaço dinâmico não linear. Durante a operação desse reservatório, os seus estados formam uma trajetória dependente dos estímulos externos e também da memória de estímulos passados. A saída da rede é computada pela camada de saída, que lê os estados instantâneos do reservatório. Uma representação geral da rede ESN é apresentada na Figura 1.

2.1 Estrutura do modelo

O funcionamento geral de uma rede ESN é centrado no reservatório, o qual é constituído de neurônios artificiais conectados de forma recorrente, característica responsável por seu comportamento dinâmico. As conexões entre os neurônios do reservatório são fixas e geradas de forma aleatória, logo os únicos elementos treináveis da rede são os pesos e bias que compõem o mecanismo de saída da mesma.

O funcionamento da rede ESN pode ser descrito por um par de equações. A equação (1) descreve o comportamento dinâmico dos estados do reservatório, sendo a equação de atualização dos estados, enquanto a equação (2) define o mecanismo de leitura do reservatório, o qual realiza o mapeamento entre os estados internos do reservatório e a saída da rede, sendo denominada equação de saída.

$$\mathbf{x}_e(k) = \mathrm{f}\big((\mathbf{1}-\alpha)\mathbf{x}_\mathrm{e}(\mathrm{k}-1) + \alpha(\mathbf{W}^{\mathrm{res}}_{\mathrm{in}}\mathbf{u}_\mathrm{e}(\mathrm{k}) + \mathbf{W}^{\mathrm{res}}_{\mathrm{res}}\mathbf{x}_e(k-1) + \mathbf{W}^{\mathrm{res}}_{\mathrm{bias}}\mathbf{1})\big), \quad (1)$$

$$\mathbf{y}_e(k) = \mathrm{g}\big(\mathbf{W}^{\mathrm{out}}_{\mathrm{res}}\mathbf{x}_\mathrm{e}(\mathrm{k}) + \mathbf{W}^{\mathrm{out}}_{\mathrm{bias}}\mathbf{1}\big), \quad (2)$$

onde: $\mathbf{u}(k)$ é o vetor de entrada no tempo k, com dimensão N_u; $\mathbf{x}(k)$ é o vetor de estados do reservatório, com dimensão N_x; $\mathbf{y}_e(k)$ é o vetor de saídas da rede, com dimensão N_y; α é a taxa de vazamento (do inglês, *leak rate*); $\mathrm{f}(\cdot)$ é a função de ativação dos neurônios, usualmente tangente hiperbólica; $\mathbf{W}^{\mathrm{res}}_{\mathrm{in}}$ e $\mathbf{W}^{\mathrm{res}}_{\mathrm{bias}}$ são matrizes com pesos que ligam as entradas e bias com o reservatório, respectivamente; $\mathbf{W}^{\mathrm{res}}_{\mathrm{res}}$ é a matriz que representa as conexões recorrentes entre os neurônios do reservatório; $\mathbf{W}^{\mathrm{out}}_{\mathrm{res}}$ é a matriz de pesos entre o reservatório e a saída; e $\mathrm{g}(\cdot)$ é uma função de pós-processamento da saída.

A criação do reservatório de neurônios é uma tarefa fundamental para o sucesso do modelo, pois os pesos das conexões entre esses neurônios são fixos. As conexões não treináveis, $\mathbf{W}^{\mathrm{res}}_{\mathrm{res}}$, $\mathbf{W}^{\mathrm{res}}_{\mathrm{in}}$, $\mathbf{W}^{\mathrm{res}}_{\mathrm{out}}$ e $\mathbf{W}^{\mathrm{res}}_{\mathrm{bias}}$, são geradas usando uma distribuição específica, geralmente uniforme. Os principais parâmetros definidos pelo projetista são a taxa de conexão ($c^{\mathrm{para}}_{\mathrm{de}}$) e o escalamento ($v^{\mathrm{para}}_{\mathrm{de}}$), os quais definem a esparsidade e a força das conexões, respectivamente. De forma geral, essas matrizes são criadas com alta esparsidade Lukoševičius (2012).

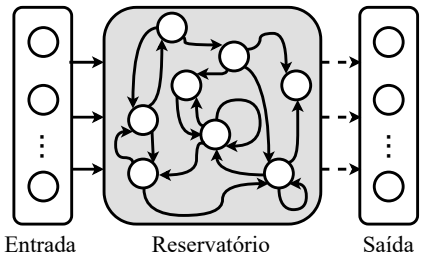


Figura 1. Arquitetura do modelo ESN. Linhas sólidas representam conexões fixas, enquanto tracejados representam conexões treináveis do modelo.

A matriz de pesos recorrentes, $\mathbf{W}_{\text{res}}^{\text{res}}$, que define as conexões entre os neurônios dentro do reservatório, recebe atenção especial. O principal parâmetro projetado é denominado raio espectral, $\rho(\mathbf{W}_{\text{res}}^{\text{res}})$, o qual corresponde ao maior autovalor de $\mathbf{W}_{\text{res}}^{\text{res}}$. De forma geral, a matriz $\mathbf{W}_{\text{res}}^{\text{res}}$ é criada usando uma esparsidade pré-definida e então é escalada, fazendo com que o maior autovalor seja igual ao raio espectral desejado.

Uma condição importante para o sucesso da ESN é a propriedade de estado de eco (do inglês, *echo state property*). Um reservatório possui a propriedade de estado de eco se os estados do mesmo convergirem assintoticamente, sob a influência de uma entrada. Para redes ESN que empregam neurônios sem vazamento, a condição suficiente para obtenção da propriedade de estado de eco é $\rho(|\mathbf{W}_{\text{res}}^{\text{res}}|) < 1$, em que $|\mathbf{W}_{\text{res}}^{\text{res}}|$ representa a matriz formada pelo valor absoluto de cada elemento da matriz original. Para o caso em que são usadas unidades com vazamento, a condição suficiente é $\rho(\mathbf{M}) < 1$, sendo $\mathbf{M} = |\mathbf{W}_{\text{res}}^{\text{res}}| + (1-\alpha)\mathbf{I}$, com $\mathbf{I}$ representando a matriz identidade de mesma dimensão de $\mathbf{W}_{\text{res}}^{\text{res}}$. Mais detalhes acerca da verificação dessa condição e da inicialização dos modelos são encontrados em Yildiz et al. (2012) e Lukoševičius (2012).

A escala de tempo do reservatório pode ser ajustada usando o parâmetro α, denominado taxa de vazamento. Esse parâmetro define a quantidade do estado passado que será preservada pelo reservatório com unidades integradoras com vazamento (do inglês, *leaky integrator units*). Para $\alpha \in (0, 1]$ a memória do reservatório pode ser ajustada, sendo que valores pequenos da taxa de vazamento levam a um reservatório com maior memória, enquanto valores próximos de 1 resultam em pouca memória.

2.2 Treinamento do modelo

A saída da rede ESN é formada por uma combinação linear dos estados internos do reservatório, somados a um bias. O mecanismo de leitura do reservatório, $\mathbf{W}_{\text{res}}^{\text{out}}$, e a matriz de bias associada à saída, $\mathbf{W}_{\text{bias}}^{\text{out}}$, são as únicas matrizes treináveis do modelo, sendo o método de regressão com regularização (do inglês, *ridge regression*) uma alternativa comumente utilizada para o ajuste dessas matrizes Lukoševičius (2012). Como alternativa para o treinamento de $\mathbf{W}_{\text{res}}^{\text{out}}$ pelo método de regressão com regularização, o método de regressão LASSO (do inglês *Least Absolute Shrinkage and Selection Operator*) pode ser aplicado, de forma a regularizar o modelo pela seleção das variáveis utilizadas pelo mecanismo de leitura, resultando em um modelo com uma quantidade de parâmetros menor, o que pode ser interessante em algumas aplicações Qiao et al. (2018).

Pelo fato de serem ajustados apenas os pesos do mecanismo de leitura do reservatório, a tarefa de aprendizado consiste em um problema de mínimos quadrados ordinário, cuja solução é computacionalmente mais eficiente que a dos algoritmos usados no treinamento de redes recorrentes convencionais. É possível, também, estimar os pesos do mecanismo de leitura do reservatório de forma recursiva, pelo método dos mínimos quadrados recursivo Jaeger (2003). Além da eficiência computacional, o treinamento da rede ESN evita diversas desvantagens apresentadas pelos algoritmos baseados em métodos de gradiente, em especial o fenômeno do desaparecimento do gradiente, o qual dificulta o treinamento de redes com múltiplas camadas escondidas Hochreiter (1998); Lukoševičius and Jaeger (2009).

Considerando $\mathbf{W}^{\text{out}} = \left[\mathbf{W}_{\text{bias}}^{\text{out}}\ \mathbf{W}_{\text{res}}^{\text{out}}\right]$, a saída da ESN pode ser escrita como

$$\mathbf{Y}_e = \mathbf{W}^{\text{out}}\mathbf{X}_e, \tag{3}$$

sendo que a matriz $\mathbf{Y}_e \in \Re^{N_y \times T}$ é definida como

$$\mathbf{Y}_e = \left[\mathbf{y}_e(k)\ \mathbf{y}_e(k+1)\ \cdots\ \mathbf{y}_e(k+T-1)\right], \tag{4}$$

$\mathbf{X}_e \in \Re^{(1+N_u+N_x)\times T}$ é uma matriz constituída pelos estados do reservatório x_e e uma entrada constantemente ativada, sendo formada como

$$\mathbf{X}_e = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mathbf{x}_e(k) & \mathbf{x}_e(k+1) & \cdots & \mathbf{x}_e(k+T-1) \end{bmatrix}, \tag{5}$$

ambas geradas por um reservatório excitado por $\mathbf{u}_e(k)$ durante um tempo de treinamento de T amostras. $\mathbf{Y}_e$ é processada por uma função $\mathrm{g}(\cdot)$, que geralmente é uma função linear em tarefas de regressão Shi and Han (2007).

O processo de aprendizado consiste em encontrar um valor ótimo para $\mathbf{W}^{\text{out}} \in \Re^{N_y \times (1+N_x)}$, o qual minimize o erro médio quadrático entre $\mathbf{y}_e(k)$ e $\mathbf{y}(k)$, para os instantes k, $k+1, \ldots, k+T-1$. A matriz $\mathbf{Y}$ contém os valores desejados para as saídas do treinamento, tendo a mesma forma de $\mathbf{Y}_e$.

A matriz $\mathbf{W}^{\text{out}}$ é obtida pelo método de regressão com regularização, sendo definida como Lukoševičius (2012)

$$\mathbf{W}^{\text{out}} = \underset{\mathbf{W}^{\text{out}}}{\operatorname{argmin}} \frac{1}{N_{\text{y}}} \sum_{i=1}^{N_{\text{y}}} \left(\sum_{n=1}^{T} \left(y_{e_i}(n) - y_i(n)\right)^2 + \beta \left\| w_i^{\text{out}} \right\|^2 \right). \tag{6}$$

A minimização da função objetivo obtém a matriz $\mathbf{W}^{\text{out}}$ que minimiza o erro e, ao mesmo tempo, considera um termo de regularização β, o qual penaliza pesos altos. A solução desse problema é descrita como:

$$\mathbf{W}^{\text{out}} = (\mathbf{X}_e^T\mathbf{X}_e + \beta\mathbf{I})^{-1}\mathbf{X}_e^T\mathbf{Y}. \tag{7}$$

2.3 Procedimento para identificação

Para a identificação de sistemas dinâmicos não lineares, a ESN é aplicada considerando um esquema apresentado no diagrama de blocos da Figura 2. O modelo baseado na ESN é alimentado por exemplos que consistem em pares de entrada e saída do processo, $\mathbf{u}(k)$ e $\mathbf{y}(k)$, respectivamente. A camada de entrada pode ser usada para escalamento dos dados de entrada, caso não haja normalização prévia. Os exemplos são processados pela camada de entrada e excitam o reservatório. As saídas desse reservatório, em conjunto com as saídas do processo, são usadas para ajustar os pesos da camada de saída, conforme apresentado na seção 2.2.

O procedimento proposto para a construção do modelo identificado a partir de dados, é sumarizado pelas etapas abaixo:

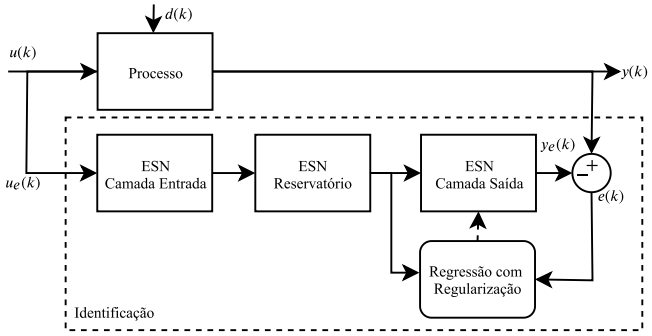


Figura 2. Identificação de sistemas com modelo ESN. O modelo ESN tem seus parâmetros estimados pelo método de regressão com regularização a partir das séries de dados obtidas em um ensaio de identificação.

(1) Obtenção da série de dados: os pares de entrada e saída são obtidos por meio da aplicação de um sinal de excitação na entrada do processo. Para sistemas lineares, o sinal de excitação geralmente aplicado é o sinal binário pseudo aleatório (PRBS, do inglês, *Pseudo-Random Binary Signal*), cujo espectro pode ser facilmente escolhido de acordo com a dinâmica do processo. Para sistemas não lineares, além do espectro frequencial, a amplitude deve ser escolhida cuidadosamente, de modo a cobrir as condições de operação de interesse. Para isso, uma alternativa é utilizar um sinal PRBS em que cada degrau apresenta uma amplitude diferente, o que leva a um PRBS modulado em amplitude (APRBS) Nelles (2001).
(2) Divisão da série de dados: o conjunto de dados obtido durante o ensaio de identificação deve ser dividido em três porções distintas, usadas para treinar, desenvolver e testar modelos. O primeiro conjunto é usado diretamente para obtenção dos parâmetros ajustáveis dos modelos. O conjunto de desenvolvimento é usado para verificar a capacidade dos modelos, sendo útil para a seleção de arquiteturas e hiperparâmetros no processo de obtenção do modelo final. O conjunto de teste é usado para a avaliação entre os diversos modelos, sendo utilizado apenas após a seleção dos modelos finais. Cada conjunto deve ser composto por porções contíguas do ensaio de identificação.
(3) Criação do modelo e ajuste dos hiperparâmetros: o modelo é criado conforme detalhado na seção 2.1. Os principais hiperparâmetros (quantidade de neurônios, taxa de conexão, taxa de vazamento e escalamento da matriz do reservatório) são inicializados em valores dentro da região admissível. Para o ajuste desses parâmetros, o qual é muitas vezes complexo, é realizado um procedimento de busca em grade (*grid search*), partindo da arquitetura inicial, e avaliando a influência de cada parâmetro no desempenho do modelo para um conjunto de validação, levando em conta uma métrica de desempenho adequada. Detalhes de aplicação desta etapa são apresentados nos estudos de caso da seção 3.
(4) Avaliação do modelo: a qualidade do modelo final desenvolvido é verificada por meio da aplicação do conjunto de testes no modelo final.

3. SIMULAÇÃO E AVALIAÇÃO EXPERIMENTAL

Nesta seção o método para identificação de sistemas baseado em ESN, proposto na seção 2, é avaliado. São apresentados dois problemas de identificação de sistemas não lineares, para os quais é explorado o procedimento de obtenção da série de dados, criação e ajuste dos hiperparâmetros do modelo ESN. O primeiro problema, apresentado na seção 3.1, consiste na identificação de um processo de neutralização de pH. O segundo problema compreende a identificação de uma bancada multivariável real, usada para ensaios de compressores herméticos de refrigeração, apresentado na seção 3.2. O desenvolvimento dos modelos segue as etapas apresentadas em 2.3, sendo que os três primeiros passos são apresentados para cada estudo de caso, e o quarto passo é apresentado na seção 4.

3.1 Processo de neutralização de pH

O processo de neutralização de pH é vastamente explorado na literatura para fins de identificação e controle, com a formulação usada neste artigo consistindo na apresentada em Henson and Seborg (1994). Este processo desperta interesse e é bastante estudado devido à sua relevância na indústria química e ao seu comportamento não linear desafiador, que o torna um bom estudo de caso para a avaliação de técnicas de identificação de sistemas. O maior desafio para algoritmos de identificação consiste em assumir que apenas ácidos fortes e bases fortes podem ser empregados e que o sistema deve operar na faixa altamente não linear do processo, próxima do pH neutro, com valores entre 6 e 8 Gomez et al. (2004).

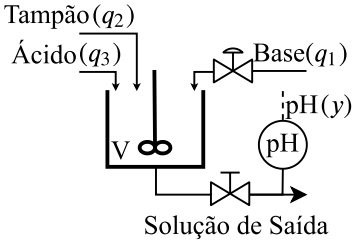


Figura 3. Diagrama de processo e instrumentos de um processo de neutralização de pH.

O processo de neutralização de pH consiste na reação entre uma base q_1 (NaOH), uma solução tampão q_2 ($NaHCO_3$) e um ácido q_3 (HNO_3). Essa reação é realizada em um tanque de volume constante com misturamento contínuo. O modelo fenomenológico no domínio do tempo contínuo desse reator é descrito por duas equações diferenciais ordinárias, definidas como:

$$\frac{dW_a(t)}{\mathrm{dt}} = \frac{q_1(t)(W_{a_1} - W_a(t)}{V} + \frac{q_2(t)(W_{a_2} - W_a(t))}{V} + \frac{q_3(t)(W_{a_3} - W_a(t))}{V} \quad (8)$$

$$\frac{dW_b(t)}{\mathrm{dt}} = \frac{q_1(t)(W_{b_1} - W_b(t)}{V} + \frac{q_2(t)(W_{b_2} - W_b(t))}{V} + \frac{q_3(t)(W_{b_3} - W_b(t))}{V}, \quad (9)$$

sendo que W representa concentrações. A saída do processo, pH, é representada por y, sendo definida pela seguinte equação algébrica:

$$W_a(t) + 10^{y(t)-14} - 10^{-y(t)} + W_b(t)\frac{1 + 2 \times 10^{y(t)-K_2}}{1 + 10^{K_1-y(t)} + 10^{y(t)-K_2}} = 0. \quad (10)$$

Os parâmetros do modelo fenomenológico são definidos na Tabela 1. A saída do processo, y, é manipulada pela vazão q_1 (ml/s), assumindo que as vazões q_2 e q_3 são mantidas constantes. Os valores iniciais para a simulação da operação do processo são apresentados na Tabela 2.

Obtenção e divisão da série de dados A simulação do modelo fenomenológico foi utilizada como o modelo real. Para a resolução das equações diferenciais ordinárias, definidas em (9), foi utilizado o método de integração numérica Runge-Kutta 45, e foi adicionado um ruído branco com variância $0,05$ na equação de saída (10), de forma a representar um ruído de medição.

Uma simulação em malha aberta foi realizada, com o objetivo de obter um conjunto de dados para a identificação do sistema. Foi considerado um período de amostragem de 10 segundos para o modelo. Para excitação do processo foi usado um sinal APRBS, projetado de forma a levar o sistema para próximo das condições de operação desejadas para operação do processo, representando regiões de pH entre 5 e 10.

Criação do modelo e ajuste dos hiperparâmetros Neste trabalho, a influência dos hiperparâmetros na qualidade final do modelo é explorada experimentalmente. Para tal, foi considerada uma estrutura padrão, a partir da qual foram realizadas variações individuais nos hiperparâmetros, sendo monitorado o erro absoluto percentual médio (MAPE, do inglês *Mean Absolute Percentage Error*) como métrica de desempenho. Os valores iniciais para essa estrutura devem ser escolhidos dentro da faixa admissível para

Tabela 1. Parâmetros usados para simulação do modelo fenomenológico de um reator de neutralização de ph

Parâmetro	Valor
K_1	$6,35$
K_2	$10,25$
V	2900 ml
W_{a_1}	$-3,05 \times 10^{-3}$ mol
W_{a_2}	-3×10^{-2} mol
W_{a_3}	3×10^{-3} mol
W_{b_1}	5×10^{-5} mol
W_{b_2}	3×10^{-2} mol
W_{b_3}	0

Tabela 2. Parâmetros usados para simulação do modelo fenomenológico de um reator de neutralização de ph

Variável	Valor Inicial
q_1	$15,55$ ml/s
q_2	$0,55$ ml/s
q_3	$16,60$ ml/s
W_a	$-4,32 \times 10^{-4}$ mol
W_b	$5,28 \times 10^{-4}$ mol

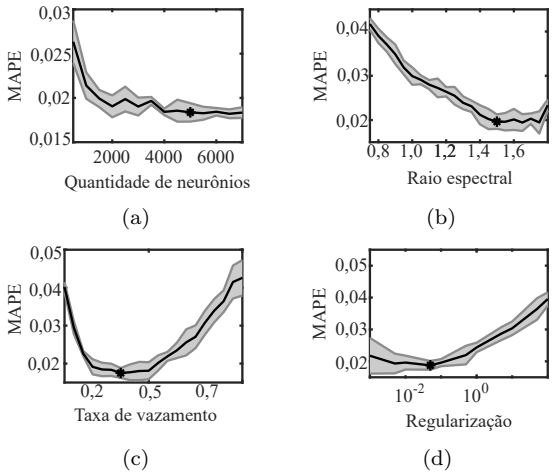


Figura 4. Impacto de hiperparâmetros na qualidade do modelo ESN para o processo de neutralização de pH. A linha preta representa a média do MAPE, a região cinza representa dois desvios-padrões e asterisco indica os valores finais selecionados.

cada parâmetro. É indicada a utilização de reservatórios com grande quantidade de unidades, na ordem de milhares, e valores pequenos para $c_{\text{res}}^{\text{res}}$, gerando um reservatório com conexões esparsas. Para o raio espectral, devem ser usados valores menores que 1, por ser uma condição suficiente para garantia da propriedade de estado de eco. Por sua vez, a taxa de vazamento pode ser inicializada em valores intermediários, na faixa entre $(0, 1]$. Por fim, para o parâmetro de regularização, é indicado a utilização de um valor inicial pequeno, em geral menor que 1.

A estrutura padrão consistiu em um reservatório com 3000 unidades e taxa de vazamento $0,2$. A matriz $\mathbf{W}_{\text{res}}^{\text{res}}$ foi inicializada considerando uma taxa de conexão $c_{\text{res}}^{\text{res}} = 0,01$ e um raio espectral $v_{\text{res}}^{\text{res}} = 0,9$. O treinamento foi realizado considerando um parâmetro de regularização $\beta = 0,1$. Cada hiperparâmetro foi avaliado dentro de uma região adequada, tendo sido treinados 20 modelos em cada configuração. Uma representação gráfica desse experimento é apresentada na Figura 4, na qual são apresentados o valor médio e o intervalo compreendido por dois desvios-padrões da métrica EQM obtida testando os modelos com dados do conjunto de validação, em função dos quatro principais hiperparâmetros.

De forma semelhante ao apontado na literatura Lukoševičius (2012), os hiperparêmetros cujo ajuste apresenta impacto relevante na qualidade do modelo são o raio espectral e a taxa de vazamento, ambos associados à construção do reservatório. Para tais parâmetros, o ajuste em uma faixa específica pode resultar em melhorias expressivas na qualidade do modelo. O tamanho do reservatório apresenta influência menos significativa. O parâmetro de regularização, associado ao método de treinamento do mecanismo de leitura do reservatório, também apresenta influência no desempenho, com valores pequenos resultando em um modelo com alta variância, e valores altos resultando em bias elevado.

Levando em conta essa análise, o modelo final foi obtido modificando a estrutura padrão. O reservatório teve sua quantidade de unidades aumentada até que não fosse mais

observada melhoria significativa na qualidade do modelo, utilizando um valor final de 5000 unidades, e teve sua taxa de vazamento definida como $0,25$. A matriz $\mathbf{W}_{\text{res}}^{\text{res}}$ foi inicializada considerando uma taxa de conexão $c_{\text{res}}^{\text{res}} = 0,01$ e um raio espectral $v_{\text{res}}^{\text{res}} = 1,5$. O parâmetro de regularização foi definido como $\beta = 0,005$. Os valores utilizados para a estrutura final estão destacados com um asterisco na Figura 4

3.2 Bancada de ensaios de compressores herméticos

A bancada de ensaios de compressores herméticos tem como finalidade a realização de ensaios em amostras de compressores de refrigeração, submetendo-os a diferentes condições de operação características de sistemas de refrigeração. A bancada contém válvulas tanto na linha de entrada de fluido refrigerante quanto na linha de saída, com o objetivo de manipular as pressões de sucção e de descarga do compressor. Essas duas válvulas são conectadas por meio de um reservatório pulmão, o qual tem como função desacoplar parcialmente as duas pressões. O diagrama de processo e instrumentos é mostrado na Figura 5.

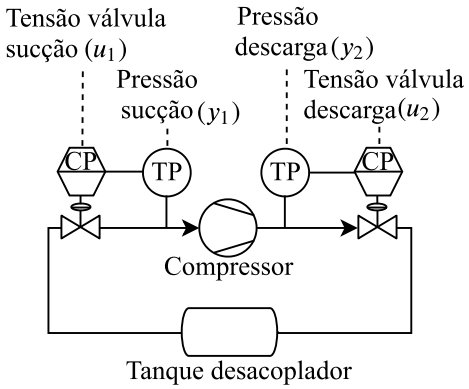


Figura 5. Diagrama simplificado de processo e instrumentos para a bancada de ensaios de compressores herméticos de refrigeração.

A válvula localizada na linha de sucção é normalmente fechada, fazendo com que a pressão na sucção do compressor seja diretamente proporcional à tensão aplicada à válvula. Por sua vez, a associada à descarga do compressor é normalmente aberta, fazendo com que a pressão de descarga seja também diretamente proporcional à tensão aplicada à respectiva válvula. É importante ressaltar que, pela natureza do processo, existe um acoplamento entre as pressões de sucção e descarga, fazendo com que mudanças na posição da válvula associada à sucção afetem também a pressão de descarga. Já o acoplamento no sentido descarga-sucção é desprezível, sendo quase totalmente compensado pelo reservatório pulmão.

Obtenção e divisão da série de dados Para este estudo de caso foi utilizada uma bancada real de ensaios, da qual foram obtidos os dados para identificação. Essa bancada adquire as pressões e atua nas válvulas com um período de amostragem de $0,1$ s. Utilizando conhecimento prévio acerca das regiões de operação do processo e da duração das dinâmicas envolvidas, foram gerados três sinais APRBS distintos, capazes de excitar o sistema de modo

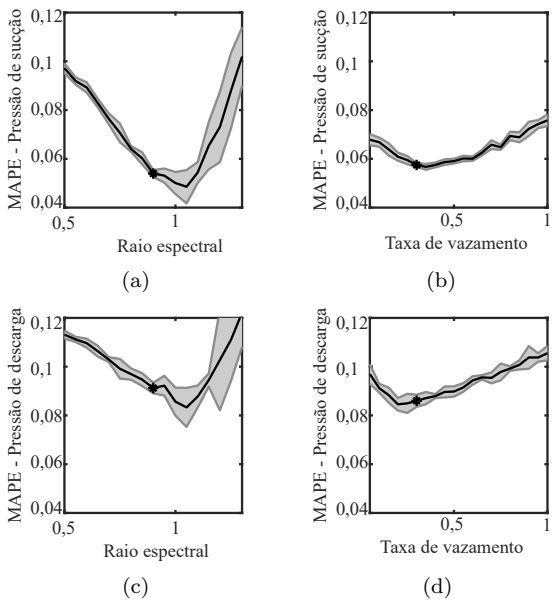


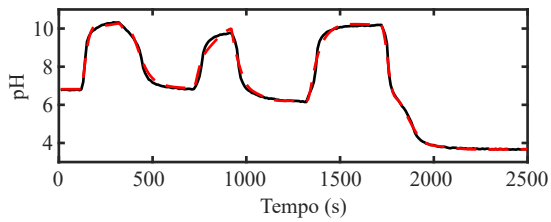
Figura 6. Impacto de hiperparâmetros na qualidade do modelo ESN para a bancada de ensaios de compressores. A linha preta representa a média do MAPE, a região cinza representa dois desvios-padrões e asterisco indica os valores finais selecionados.

que as várias regiões de operação sejam varridas em cada um deles. Cada variável manipulada foi estimulada em momentos distintos, mantendo a outra variável em diferentes patamares constantes. Os ensaios realizados foram divididos, de forma semelhante ao estudo de caso apresentado anteriormente, em três conjuntos distintos, usados para treinamento, desenvolvimento e teste dos modelos.

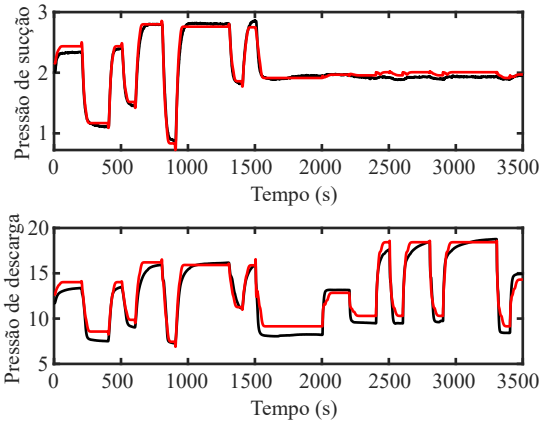
Criação do modelo e ajuste dos hiperparâmetros De forma semelhante à apresentada para o processo de neutralização de pH, para criação do modelo baseado em ESN para a bancada de ensaios de compressores herméticos foi selecionada uma arquitetura inicial, a partir da qual os principais hiperparâmetros foram ajustados. O modelo ESN foi construído com uma estrutura multivariável, devido à natureza do processo, considerando as tensões das duas válvulas como entrada e retornando as pressões na sucção e descarga do compressor como saída do modelo. Para realização do ajuste de hiperparâmetros, é necessária uma métrica adequada, com a qual seja possível avaliar de forma igual as saídas do modelo. Para tal, foi utilizado o MAPE, calculado para cada saída do modelo.

A estrutura padrão consistiu em um reservatório com 3000 unidades e taxa de vazamento $0,7$. A matriz $\mathbf{W}_{\text{res}}^{\text{res}}$ foi inicializada considerando uma taxa de conexão $c_{\text{res}}^{\text{res}} = 0,01$ e um raio espectral $v_{\text{res}}^{\text{res}} = 0,8$. O treinamento foi realizado considerando um parâmetro de regularização $\beta = 0,01$. O ajuste dos hiperparâmetros foi realizado por meio da mesma avaliação experimental apresentada para o processo de neutralização de pH. Os resultados desse experimento são sumarizados na Figura 6, na qual são apresentados, por simplicidade, apenas os dois principais hiperparâmetros, raio espectral e taxa de vazamento.

O raio espectral apresenta boa sintonia com valores inferiores a 1, a partir do qual o desvio-padrão aumenta



(a) Processo de neutralização de pH



(b) Bancada de ensaios de compressores

Figura 7. Comparação entre a saída real dos processos (linha preta) e a saída obtida pelo modelo baseado em ESN (linha vermelho).

significativamente. Para a taxa de vazamento existe um comportamento do erro diferente para cada variável, com y_2 apresentando melhor sintonia com $\alpha = 0,2$, enquanto y_1 apresenta menor MAPE com $\alpha = 0,35$. A estrutura padrão foi modificada considerando uma sintonia em que o MAPE médio das duas variáveis de interesse seja mínimo. Para o reservatório foram utilizadas 5000 unidades e taxa de vazamento $0,3$. A matriz $\mathbf{W}_{\text{res}}^{\text{res}}$ foi inicializada considerando uma taxa de conexão $c_{\text{res}}^{\text{res}} = 0,01$ e um raio espectral $v_{\text{res}}^{\text{res}} = 0,9$. Para o parâmetro de regularização foi utilizado $\beta = 1$.

4. RESULTADOS E DISCUSSÕES

Os modelos treinados foram avaliados utilizando a porção de testes das séries de dados. Essa porção é contígua e compreende uma série completa de degraus nas variáveis de entrada. Esses resultados são apresentados graficamente na Figura 7, sendo que na Figura 7.a é apresentado o resultado do modelo ESN comparado com a saída do processo de neutralização de pH, e na Figura 7.b são apresentadas as saídas do processo e do modelo ESN para a bancada de ensaios de compressores.

Os resultados foram compilados em tabelas para ambos os estudos de caso. Como métricas de avaliação foram utilizadas o erro quadrático médio (EQM), o critério de correlação R^2 e o erro percentual absoluto médio (MAPE). Na Tabela 3 são apresentados os resultados para o processo de neutralização de pH, enquanto na Tabela 4 são detalhados os resultados para ambas as saídas da bancada de ensaios de compressores. Para fins de comparação, resultados obtidos com modelos lineares de primeira e segunda ordem, além de com um modelo

Tabela 3. Comparação entre as métricas de desempenho obtidas pelos modelos avaliados para o processo de neutralização de pH.

Modelo	EQM	R^2	MAPE
Linear (1ª ordem)	0,37	0,93	0,066
Linear (2ª ordem)	0,33	0,94	0,060
ELM-Hammerstein	0,22	0,96	0,059
ESN	0,03	0,99	0,015

Tabela 4. Comparação entre as métricas de desempenho obtidas pelo modelos avaliados para a bancada de teste de compressores.

	Pressão de Sucção			Pressão de Descarga		
Modelo	EQM (bar^2)	R^2	MAPE	EQM (bar^2)	R^2	MAPE
Linear (1ª ordem)	0,040	0,79	0,094	6,50	0,47	0,213
Linear (2ª ordem)	0,041	0,78	0,097	6,31	0,49	0,210
ELM-Hammerstein	0,015	0,92	0,051	0,76	0,94	0,059
ESN	0,010	0,95	0,036	0,78	0,94	0,060

não linear baseado em máquinas de aprendizado extremo (ELM, do inglês *Extreme Learning Machine*), do tipo ELM-Hammerstein Tang et al. (2014), foram utilizados. Os modelos do tipo ELM-Hammerstein apresentam como não linearidade estática uma rede neural ELM, gerada a partir de uma distribuição aleatória, com quantidade de neurônios idêntica à utilizada pelos modelos baseados em ESN. Em ambos os estudos de caso, foi escolhida como porção dinâmica linear uma estrutura de primeira ordem.

Para ambos os casos, o modelo obtido por meio do método proposto apresentou melhores resultados quando comparado com os demais modelos. Para o problema de neutralização de pH, o modelo ESN apresentou uma alta fidelidade, verificada por inspeção na Figura 7.a e suportada pelas métricas de desempenho apresentadas na Tabela 3. De forma semelhante, o modelo ESN para a bancada de ensaios de compressores apresentou melhores resultados, para ambas as saídas modeladas, quando comparado aos modelos linear e ELM-Hammerstein. Para esse processo, é importante notar que se verificam, principalmente para a pressão de descarga, erros relacionados à característica estática do processo, os quais decorrem de perturbações não medidas, em especial a temperatura do compressor, a qual afeta o regime de operação do mesmo. A análise das métricas de desempenho, sumarizadas na Tabela 4, evidencia que mesmo em um processo real sujeito a perturbações não medidas, o modelo ESN apresentou melhor desempenho, com resultados semelhantes aos obtidos pelo modelo ELM-Hammerstein para a pressão de descarga e melhores para a de sucção. É importante notar que o modelo ESN, que apresenta alta complexidade computacional, apresentou resultados superiores não só aos modelos menos complexos, mas também quando comparado com um modelo com complexidade semelhante, caso do modelo ELM-Hammerstein.

5. CONCLUSÕES

Este artigo apresentou uma estratégia que se baseia na rede de estado de eco para identificação de processos

multivariáveis não lineares. A criação de modelos ESN e o ajuste dos hiperparâmetros do modelo foi discutida, e a aplicação dessa abordagem foi demonstrada em um estudo de caso simulado, o reator de neutralização de pH de Henson and Seborg (1994), e em um processo real, uma bancada de ensaios de compressores de refrigeração.

Foi demonstrado que a criação do modelo ESN, bem como a sintonia dos hiperparâmetros, apresenta grande impacto na qualidade dos modelos obtidos. Apesar de existirem boas práticas para a criação de um reservatório, a escolha dos parâmetros usados para criação do reservatório deve receber atenção, carecendo de um ajuste cuidadoso. Em todos os casos avaliados, os hiperparâmetros que apresentaram maior influência na qualidade do modelo foram o raio espectral e a taxa de vazamento.

Os modelos baseados em ESN alcançaram bons resultados quando comparados com os obtidos por modelos lineares e modelos não lineares do tipo ELM-Hammerstein. Porém o custo computacional para cálculo da saída dos modelos baseado em ESN foi superior ao verificado para os demais modelos. Esses resultados atestam que a aplicação do método baseado em ESN proposto para criação de modelos MIMO para processos não lineares configura-se uma alternativa a técnicas clássicas, podendo ser aplicado sem a necessidade de conhecimento acerca da dinâmica do processo a ser identificado.

AGRADECIMENTOS

Este trabalho foi financiado em parte pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/Brasil) por meio das concessões 140283/2018-8 e 309244/2018-8, e em parte pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/Brasil) por meio do projeto PrInt/UFSC.

REFERÊNCIAS

Antonelo, E.A., Camponogara, E., and Foss, B. (2017). Echo state networks for data-driven downhole pressure estimation in gas-lift oil wells. *Neural Networks*, 85, 106–117.

Ayoubi, M. (1994). Nonlinear dynamic systems identification with dynamic neural networks for fault diagnosis in technical processes. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, volume 3, 2120–2125.

Boussaada, Z., Curea, O., Remaci, A., Camblong, H., and Mrabet Bellaaj, N. (2018). A nonlinear autoregressive exogenous (NARX) neural network model for the prediction of the daily direct solar radiation. *Energies*, 11(3), 620.

Chen, Y., He, Z., Shang, Z., Li, C., Li, L., and Xu, M. (2019). A novel combined model based on echo state network for multi-step ahead wind speed forecasting: A case study of NREL. *Energy conversion and management*, 179, 13–29.

Fernandez, B., Parlos, A., and Tsai, W. (1990). Nonlinear dynamic system identification using artificial neural networks (ANNs). In *1990 IJCNN International Joint Conference on Neural Networks*. (IEEE).

Gomez, J., Jutan, A., and Baeyens, E. (2004). Wiener model identification and predictive control of a pH neutralisation process. *IEE Proceedings-Control Theory and Applications*, 151(3), 329–338.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press, Cambridge.

Henson, M.A. and Seborg, D.E. (1994). Adaptive nonlinear control of a pH neutralization process. *IEEE Transactions on Control Systems Technology*, 2(3), 169–182.

Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(2), 107–116.

Isermann, R. and Münchof, M. (2011). *Identification of Dynamic Systems*. Springer, Berlim.

Jaeger, H. (2001). The "echo state" approach to analysing and training recurrent neural networks. Technical Report 34, German National Research Center for Information Technology, Bonn, Germany.

Jaeger, H. (2003). Adaptive nonlinear system identification with echo state networks. In *Advances in Neural Information Processing Systems (NIPS 2003)*, 609–616.

Lukoševičius, M. (2012). A practical guide to applying echo state networks. In G. Montavon, G.B. Orr, and K.R. Müller (eds.), *Neural Networks: Tricks of the Trade.*, 659–686. Springer, Berlin, 2 edition.

Lukoševičius, M. and Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3), 127–149.

Nelles, O. (2001). *Nonlinear System Identification*. Springer, London.

Qiao, J., Wang, L., and Yang, C. (2018). Adaptive lasso echo state network based on modified bayesian information criterion for nonlinear system modeling. *Neural Computing and Applications*, 1–15.

Rodan, A. and Tino, P. (2010). Minimum complexity echo state network. *IEEE Transactions on Neural Networks*, 22(1), 131–144.

Schwedersky, B.B., Flesch, R.C.C., and Dangui, H.A.C. (2018). Practical nonlinear model predictive control using an echo state network model. In *2018 International Joint Conference on Neural Networks (IJCNN)*, 1–8.

Shi, Z. and Han, M. (2007). Ridge regression learning in ESN for chaotic time series prediction. *Control and Decision*, 22(3), 258.

Tang, Y., Li, Z., and Guan, X. (2014). Identification of nonlinear system using extreme learning machine based hammerstein model. *Communications in Nonlinear Science and Numerical Simulation*, 19(9), 3171–3183.

Yang, C., Qiao, J., Ahmad, Z., Nie, K., and Wang, L. (2019). Online sequential echo state network with sparse RLS algorithm for time series prediction. *Neural Networks*, 118, 32–42.

Yildiz, I.B., Jaeger, H., and Kiebel, S.J. (2012). Re-visiting the echo state property. *Neural Networks*, 35, 1–9.