

Previsão do Preço do Açúcar: Uma Análise comparativa entre o Modelo Linear Autorregressivo e Redes Neurais

Tathiana M. Barchi¹, Priscila Bassetto², Lucas H. Biuk³, Lucas F. P. Costa¹, Erickson D. P. Puchta², Fernanda Cristina Correa^{3,4}, Mauren L. S. C. Andrade¹, Yslene Rocha Kachba², Hugo Valadares Siqueira^{2,4,5}

¹Graduate Program in Computer Sciences (PPGCC)

²Graduate Program in Production Engineering (PPGEP)

³Graduate Program in Electric Engineering (PPGEE)

⁴Department of Electronic Engineering

Federal University of Technology - Paraná (UTFPR)

Ponta Grossa - PR, Brazil

tathiana@alunos.utfpr.edu.br, pri_bass@hotmail.com,

lucasbiuk1902@gmail.com, luccos@alunos.utfpr.edu.br, ericksonpuchta@gmail.com,

{fernandacorrea, mlsguario, yslener, hugosiqueira}@utfpr.edu.br

Resumo—O Brasil desempenha um papel importante na produção mundial de subprodutos gerados a partir da cana-de-açúcar, principalmente açúcar e etanol. O açúcar é uma das mais antigas *commodities* comercializadas pelo país. Realizar uma previsão adequada dos preços de tal produto tem impacto direto no sistema econômico, pois auxilia os gestores nos planejamentos estratégicos e minimiza os riscos futuros com uma avaliação mais precisa de suas tendências de mercado. Assim, o presente estudo tem como objetivo realizar a previsão do preço do açúcar, usando os modelos Autorregressivo (AR), Perceptron de Múltiplas Camadas (MLP), Máquinas de Aprendizado Extremo (ELM) e Redes Neurais com Estados de Eco (ESN). Os resultados experimentais evidenciaram que os modelos que apresentaram melhores desempenhos para um passo a frente nas bases CEPEA e NY No. 11, foram ELM e MLP, respectivamente.

Palavras-chaves—séries temporais, preço do açúcar, modelo linear, redes neurais artificiais

I. INTRODUÇÃO

O açúcar é uma das mais antigas *commodities* comercializadas pelo Brasil. A partir do século XV, tal produto exerceu grande influência em diversas economias do mundo [1], [2]. Hoje representa um importante ativo para as economias de diversos países e um dos produtos mais regulamentados nos mercados internacionais [3].

A cana-de-açúcar desempenha um importante papel na economia agrícola e industrial, sendo o Brasil um dos maiores produtores mundiais [4]. A cadeia produtiva gera milhares de empregos diretos e indiretos e também um alto giro de ativos econômicos mundo afora [5], [6]. Além disso, trata-se de uma das matérias-primas mais eficientes para a produção de combustíveis no país.

O Brasil é destaque na produção mundial de subprodutos gerados a partir da cana-de-açúcar, sobretudo açúcar e etanol.

Os autores agradecem aos órgãos brasileiros pelo apoio financeiro, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Código Financiador 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processos n° 40558/2018-5, 315298/2020-0, e Fundação Araucária, processo n° 51497. Os autores também agradecem à Universidade Tecnológica Federal do Paraná (UTFPR).

De acordo com o Observatório de Complexidade Econômica – OEC (2020), em 2019, o açúcar bruto foi o 170º produto mais negociado do mundo, totalizando suas negociações em US\$ 21,8 bilhões. Além disso, representa 0,12% do comércio mundial total. No mesmo ano, os maiores exportadores foram Brasil (US\$ 5,33 bilhões), Tailândia (US\$ 3,04 bilhões), Índia (US\$ 1,72 bilhões), França (US\$ 1 bilhão) e Austrália (US\$ 998 milhões) [7]. Tal produto foi o 10º mais exportado pelo Brasil, tendo como principais destinos a Argélia (US\$ 634 milhões), Bangladesh (US\$ 474 milhões), Nigéria (US\$ 418 milhões), Arábia Saudita (US\$ 402 milhões) e China (US\$ 391 milhões).

Um adequado conhecimento sobre o mercado e as safras é muito importante em um sistema econômico, pois existe uma estreita associação entre as produções agrícolas e os preços. Uma queda inesperada na produção reduz o excedente comercializável e a renda dos agricultores e leva ao aumento dos preços. Em contrapartida, um excesso de produção pode levar a uma queda nos preços e tem um efeito adverso na renda dos agricultores. Desse modo, o impacto no preço de uma mercadoria essencial tem um papel significativo na determinação da taxa de inflação, salários, vencimentos e várias políticas em uma economia. No caso de culturas comerciais como a cana-de-açúcar, o nível de produção afeta o custo da matéria-prima das indústrias usuárias e suas vantagens competitivas no mercado [8].

Diante de tal premissa, devido à incerteza nos preços internacionais do açúcar e às relações apresentadas, torna-se importante realizar um estudo de previsão do seu preço para auxiliar os gestores nos planejamentos estratégicos, minimizando os riscos futuros e apresentando avaliações de maior acurácia de suas tendências.

Recentemente, Silva et al. [9] propuseram investigar a aplicação de redes neurais para prever os preços dos derivados da cana-de-açúcar: açúcar, etanol hidratado e anidro etanol. Os resultados computacionais mostraram que a aplicação de máquinas de aprendizado extremo apresentou melhores resul-

tados gerais, mostrando-se como um candidato viável para esse tipo de tarefa. No trabalho de Suresh e Priya [10], os autores buscaram prever a área canavieira, a produção e a produtividade de Tamil Nadu por meio do ajuste de modelos univariados autorregressivo e médias móveis. Ainda, Amrouk e Eckelei [11] combinaram um modelo Bayesiano (BMA) para abordar a incerteza na previsibilidade de preços do açúcar. Os resultados mostraram que a previsibilidade do BMA é geralmente alta, em comparação a outras de abordagens de séries temporais. Já a investigação de Obe e Shangodoyin [12], empregou uma técnica heurística para desenvolver uma Rede Neural Artificial (RNA) para prever a produção de cana-de-açúcar na Nigéria. Os resultados indicaram que o modelo desenvolvido se ajusta bem aos dados reais e pode ser usado para fins de previsão com alta precisão.

Assim, o presente estudo tem como objetivo realizar a previsão do preço do açúcar no Brasil utilizando um modelo linear e arquiteturas de redes neurais artificiais: modelo Autorregressivo (AR), Perceptron de Múltiplas Camadas (MLP), Máquinas de Aprendizado Extremo (ELM) e Redes Neurais com estados de Eco (ESN).

O restante do trabalho está organizado da seguinte forma: a Seção II apresenta os modelos de previsão abordados neste trabalho; a Seção III discute as particularidades das séries utilizadas e etapas envolvidas no processo de previsão; a Seção IV mostra os resultados computacionais e uma análise crítica dos mesmos; a Seção V apresenta as principais conclusões e perspectivas futuras.

II. MODELOS DE PREVISÃO

O processo de observação e quantificação numérica de uma categoria de fenômenos que gera uma sequência de dados distribuídos no tempo é denominada série temporal [13]. Esta também pode ser definida como uma ordem de números reais registrados em intervalos de tempo, como anual, mensal, semanal, diário ou até mesmo em milissegundos. Logo, a única variável independente na série é o tempo [14].

Realizar a previsão destas séries requer o uso de métodos estatísticos, quantitativos, ou mesmo de inteligência computacional, os quais se baseiam na extrapolação de características de observações passadas e no seu inter-relacionamento, fornecendo previsões precisas se o futuro apresentar comportamento similar ao passado [15]. Na literatura existem inúmeros modelos de previsão. No entanto, para realizar o emprego do método correto é necessário analisar o comportamento da série que se deseja prever [16].

A literatura especializada da área é rica em abordagens para modelagem e previsão de séries temporais, sobretudo modelos lineares da família Box & Jenkins [17]. Entretanto, uma classe de métodos capaz de realizar tais previsões são os modelos inspirados na natureza do sistema nervoso, composto por neurônios altamente conectados e capazes de realizar adaptações a estímulos desconhecidos, o que é caracterizado como aprendizado, abordagem essa conhecida como Redes Neurais Artificiais (RNAs) [18], [19]. A seguir, são descritos os métodos de previsão aplicados na execução deste trabalho.

A. Modelo Autorregressivo

O modelo Auto-Regressivo, do inglês *Autoregressive* (AR), estabelece que a observação da série temporal atual é uma

combinação linear das observações anteriores da série temporal adicionadas de um ruído branco [20].

No modelo AR, pressupõe-se que a previsão da série é dada pela combinação linear dos termos passados da série ($x_{t-1} + \dots + x_{t-p}$), multiplicados pelos coeficientes livres do modelo (ϕ_p), adicionado o erro de especificação (ε_t), que é associado ao erro inerente ao processo de modelagem ou previsão, conforme Equação (1):

$$\hat{x}_t = \sum_{i=1}^p \phi_i x_{t-i} + \varepsilon_t, \quad (1)$$

em que \hat{x}_t é valor previsto pelo modelo no tempo t e ϕ_i são os p coeficientes livres.

Nesse modelo, os valores dos coeficientes livres foram calculados por equações lineares de forma fechada, aplicando as equações de Yule-Walker [17].

B. Perceptron de múltiplas camadas

Um Perceptron de Múltiplas Camadas (*Multilayer Perceptron* - MLP) é um modelo de rede neural artificial considerado aproximador universal por sua capacidade de aproximar qualquer função não-linear, contínua, limitada e diferenciável [21]. Tem por característica principal a existência de ao menos uma camada intermediária (escondida) de neurônios entre as camadas de entrada e saída da rede, sendo convencional o uso de apenas uma ou até duas camadas escondidas. É um modelo versátil para aplicação em áreas como previsão de séries temporais, classificação de dados e reconhecimento de padrões [9], [22].

Nesta rede, camadas disjuntas são conectadas, enquanto neurônios das mesmas camadas não se comunicam, o que a permite definir sua arquitetura como *feedforward*, ou seja, a rede não apresenta recorrência [23].

O treinamento de uma MLP é o processo de ajuste dos pesos das ligações entre neurônios [24]. O objetivo é encontrar o melhor conjunto de pesos, com base em alguma métrica de erro, usando um processo iterativo, sendo um procedimento caracterizado como otimização linear irrestrita [19].

Independentemente do modelo de otimização escolhido para realizar o treinamento, é muito usual aplicar o algoritmo de retropropagação de erro (*backpropagation*), para o cálculo do vetor gradiente da função custo formado pela saída da rede e o sinal desejado [25]. Neste trabalho, utilizou-se o *backpropagation* tradicional com otimização via gradiente descendente em uma MLP de uma camada intermediária [21].

C. Máquinas de Aprendizado Extremo

Máquinas de Aprendizado Extremo (ELM, *Extreme Learning Machines*), propostas em [26], possuem apenas uma camada intermediária. É uma abordagem de rede *feedforward*, que combina simplicidade no processo de treinamento e desempenho. Os pesos são gerados aleatoriamente, mas apenas aqueles da camada de saída são ajustados [27].

Semelhante à MLP, a ELM é capaz de aproximar qualquer mapeamento não-linear, contínuo, limitado e diferenciável com precisão arbitrária [27]. A precisão necessária é alcançada com a inserção de novos neurônios na camada oculta, aumentando o potencial de aproximação da rede.

Para a ELM, considerando $\mathbf{u}_n = [u_n, u_{n-1}, \dots, u_{n-K+1}]^T$ como o vetor que contém o sinal de entrada, $\mathbf{W}^h \in \mathfrak{R}^{N_k \times K}$

como os coeficientes (pesos) da camada intermediária, \mathbf{b} para o viés (*bias*) de cada neurônio artificial e $\mathbf{f}^h(\cdot)$ sendo as funções de ativação dos neurônios, os neurônios da camada oculta têm suas ativações dadas pela Equação 2 [28]:

$$\mathbf{x}_n^h = \mathbf{f}^h(\mathbf{W}^h \mathbf{u}_n + \mathbf{b}). \quad (2)$$

Dessa forma, a Equação 3 gera as saídas:

$$\mathbf{y} = \mathbf{W}^s \mathbf{x}_n^h, \quad (3)$$

em que \mathbf{W}^s são os pesos da camada de saída.

O treinamento se resume a encontrar a solução de um problema de mínimos quadrados, uma tarefa com solução determinística e baixo custo computacional. A forma mais tradicional é lançar mão do algoritmo da pseudo-inversa de Moore-Penrose [29].

D. Redes Neurais com Estado de Eco

Redes Neurais Recorrentes (*Recurrent Neural Networks* - RNN) são aquelas que apresentam *loops* de realimentação (*feedback*), ou seja, algumas saídas são reinseridas como entradas. Este procedimento origina uma “memória interna”, que somada a presença de realimentação, as quais podem trazer respostas de saída mais precisas, acarretam em mais informações disponíveis para formar o sinal de saída [19], [30].

Entretanto, uma RNN totalmente treinada pode apresentar dificuldades durante o ajuste: o treinamento pode levar à instabilidade, dificuldades na manipulação da função de custo, convergência local e um custo computacional elevado [31].

Para superar esses problemas, [30] introduziu as Redes Neurais com Estado de Eco (ESNs, *Echo State Networks*). Nesta proposta, a camada oculta (reservatório dinâmico) constituída por neurônios não-lineares com *loops* recorrentes, permanece sem ajuste. Portanto, como no caso da ELM, o processo de treinamento se resume ao cálculo dos pesos de saída por meio de uma solução de mínimos quadrados, uma abordagem simples e eficiente, usualmente tratada com o mesmo operador de Moore-Penrose mencionado [32].

A capacidade de memória está presente em uma ESN se a propriedade do estado de eco, conforme definido por [30], for respeitada. O autor afirma que se algumas condições forem satisfeitas na matriz do reservatório \mathbf{W} , o histórico de entrada recente rege os estados internos dos neurônios presentes nesta camada. Portanto, esses estados se tornam uma transformação não-linear, criando respostas de rede com eco. A consequência imediata é que todos os pesos podem ser definidos com antecedência [33].

O vetor de estado de eco \mathbf{x}_n é formado pela saída dos neurônios do reservatório, sendo calculado pela Equação (4):

$$\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{W}^e \mathbf{u}_{n+1} + \mathbf{W} \mathbf{x}_n), \quad (4)$$

em que o vetor $\mathbf{u}_n = [u_n, u_{n-1}, \dots, u_{n-K-1}]^T$ contém as amostras de entrada que são transmitidas aos neurônios ocultos por meio de uma combinação linear com os elementos da matriz de pesos de entrada \mathbf{W}^e e $\mathbf{f}(\cdot) = [f_1(\cdot), f_2(\cdot), \dots, f_N(\cdot)]$ são as funções de ativação dos neurônios no reservatório.

Dessa forma, a saída da rede é definida pela Equação (5):

$$\mathbf{y}_{n+1} = \mathbf{W}^s \mathbf{x}_{n+1}, \quad (5)$$

sendo $\mathbf{W}^s \in \mathbb{R}^{L \times N}$ a matriz ponderada de saída.

A ESN apresenta semelhanças com a ELM, uma vez que o processo de treinamento é eficiente devido a rápida convergência e solução analítica para os pesos da camada de saída, bem como pelo fato de ser um aproximador universal [30].

III. ESTUDO DE CASO

O estudo de casos deste trabalho envolve a modelagem e previsão do preço do açúcar. Para isso, foram utilizados dois bancos de dados representados em dólar americano, sendo o primeiro do Centro de Estudos Avançados em Economia Aplicada (CEPEA), com um corte temporal mensal de Janeiro de 2005 a Fevereiro de 2021, totalizando 194 amostras, e o segundo da Bolsa de Nova York (NY), compreendendo um período mensal de Julho de 2000 a Maio de 2021, totalizando 251 amostras.

Em relação à análise de dados, foi feito o teste aumentado de Dickey-Fuller [34] para verificar se estes eram estacionários. Através dos resultados, notou-se que ambas as bases de dados não eram estacionárias. Neste caso, constatou-se a presença de tendência, que foi corrigida através de diferenciação, a qual é revertida após o processo de previsão. Em seguida, foi feita a separação da base de dados em três grupos, treinamento, validação e teste, conforme Tabela I.

Tabela I
DESCRIÇÃO DOS DADOS E PORCENTAGEM DOS CONJUNTOS DE TREINAMENTO, VALIDAÇÃO E TESTE

Base	Amostras (total)	% Trein.	% Val.	% Teste
CEPEA	194	68,39%	18,14%	13,47%
NY No.11	251	64,80%	19,20%	16%

Para seleção das melhores entradas para as redes neurais foi utilizada função de autocorrelação parcial (FACP). Este filtro selecionou os 4 primeiros atrasos para a base CEPEA e os 3 primeiros para NY.

O próximo passo experimental foi determinar a quantidade de neurônios para a camada intermediária de cada rede neural. Este número foi determinado através de uma varredura no qual para cada configuração foram feitos 30 rodadas de simulação, considerando um range de 3 até 100 neurônios [9].

Para avaliar os resultados obtidos, foram adotadas as métricas erro quadrático médio (*Mean Squared Error* - MSE) e erro percentual absoluto médio (*Mean Absolute Percentage Error* - MAPE). O MSE calcula a média dos quadrados dos erros, ou seja, a diferença quadrática média entre valor real e valor estimado. Este sempre apresenta valores não negativos, os quais são melhores quanto mais próximos a zero. O MSE é expresso pela Equação (6):

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2, \quad (6)$$

sendo e_t a diferença entre valor real e valor predito no instante t .

O MSE é uma das métricas padrão para avaliação de performance em modelos de previsão. Entretanto esta penaliza diferenças muito grandes entre real e predito, ou seja, é sensível a escala, ao contrário do MAPE, que calcula a porcentagem de erro, expresso pela Equação (7):

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{e_t^2}{x_t} \right|, \quad (7)$$

no qual x_t é o valor real.

IV. RESULTADOS

A seguir, serão descritos os resultados computacionais obtidos nesta investigação. Foram realizadas as previsões utilizando os modelos AR, ELM, ESN, e MLP, os quais foram desenvolvidos na linguagem Matlab. Por conta dos aspectos estocásticos envolvidos na inicialização dos parâmetros, foram realizadas 30 rodadas independentes para cada modelos.

Assim, definidas as técnicas de previsão e a aplicação dos dados passados para obtenção dos parâmetros necessários, podem-se obter as projeções futuras da série temporal. O horizonte de previsão é o comprimento do tempo, contado a partir de uma origem predeterminada. Dessa forma, quanto maior for o horizonte pretendido, menor a confiabilidade na previsão prevista [17]. Neste estudo considerou-se o horizonte de 1, 2 e 3 passos à frente [35].

Logo, ao analisar os resultados da execução dos modelos sobre a base de dados CEPEA, pode-se observar que a ELM apresentou melhor desempenho frente aos outros nos três horizontes de previsão testados, conforme Tabela II.

Tabela II
PERFORMANCE DOS MODELOS SOBRE A BASE CEPEA

Modelos		Horizontes de previsão		
		1	2	3
AR	MSE	0,3753	1,3788	2,7244
	MAPE	4,2589	8,0012	11,1335
ELM	N. Neurônios	20	15	10
	MSE	0,1775	0,8654	1,6523
	MAPE	2,3989	5,6702	8,2228
ESN	N. Neurônios	20	7	5
	MSE	1,7918	4,6818	6,0345
	MAPE	7,0319	11,8156	15,0568
MLP	N. Neurônios	60	5	5
	MSE	0,1939	0,9817	2,3587
	MAPE	2,5766	5,6802	9,5909

Ainda sobre esta base, nota-se que o modelo MLP apresentou um MSE 9% maior em relação ao modelo ELM para previsão um passo à frente, sendo assim, o segundo melhor preditor. O pior resultado, neste caso, foi dado pela ESN, que apresentou um MSE quase dez vezes maior.

A Figura 1 apresenta os valores reais e previstos pelo ELM para previsão um passo à frente.

Utilizando os dados da base da Bolsa de Nova York, percebe-se que a MLP apresentou melhor resultado na previsão para quase todos os cenários de erro, exceto pelo MSE de 2 passos adiante, como mostra a Tabela III.

Os valores reais e previstos para um passo a frente obtidos pela MLP são apresentados na Figura 2.

De modo geral, assim como na base CEPEA, as redes neurais ELM e MLP apresentaram os melhores resultados com valores aproximados para os diferentes horizontes de previsão testados, no entanto, a MLP utilizou mais neurônios, resultando em maior esforço computacional.

O destaque negativo foi o desempenho ruim e inesperado obtido pela ESN, o qual foi o pior em ambas as bases. Pelo fato

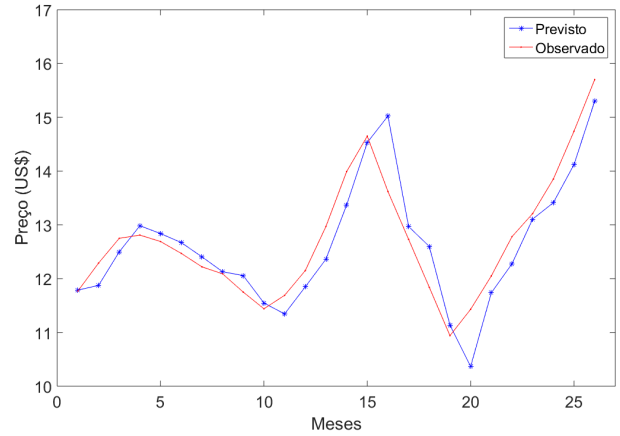


Figura 1. CEPEA - Previsão um passo à frente com a ELM

Tabela III
PERFORMANCE DOS MODELOS SOBRE A BASE NY No. 11

Modelos		Horizontes de previsão		
		1	2	3
AR	MSE	1,9814	2,6907	2,9197
	MAPE	7,8490	10,2440	11,4419
ELM	N. Neurônios	10	7	5
	MSE	1,6278	1,6278	3,0901
	MAPE	6,9625	10,3343	11,7938
ESN	N. Neurônios	10	3	7
	MSE	12,2056	41,6003	91,4865
	MAPE	18,5696	33,3506	49,0946
MLP	N. Neurônios	70	100	50
	MSE	1,4023	2,2246	2,8104
	MAPE	6,4845	9,3301	11,0463

de ser uma rede recorrente com capacidade de aproximação universal e por possuir memória, era esperado que ela fosse superior, pelo menos, ao modelo AR [36]. Isso se torna ainda mais surpreendente pelos bons resultados que tal modelo já alcançou em problemas de previsão [22], [24].

Conforme esperado, o modelo AR não esteve entre aqueles com melhor resultado. Técnicas lineares, como a aplicada neste estudo, geralmente se mostram menos flexíveis para problemas que envolvam padrões não-lineares. Entretanto, destaca-se que muitas áreas de pesquisa ainda lançam mão destas abordagens para tratar problemas de mundo real [9], [37].

Em suma, as redes *feedforward* mostraram desempenho superior em um problema que apresenta difícil solução, sobrepondo a proposta linear e a recorrente. Isto explica sua extensa utilização em previsão de séries temporais devido sua notável capacidade de reconhecimento de padrões e generalização resultante da sua aptidão em capturar não-linearidade entre as componentes de uma série [9], [19], [25].

Por fim, destacamos que para os resultados constantes nas Tabelas II e III aplicou-se o teste de Friedman [38] para averiguarmos se os resultados são significativamente diferentes. Os p-valores encontrados foram da ordem de 10^{-6} , o que indica que modificar o preditor leva a resultados distintos [29].

V. CONCLUSÃO

O açúcar é uma *commodity* que influencia em grande medida o produto interno bruto dos países produtores. O Brasil é o líder mundial em produção e exportação deste bem. Neste

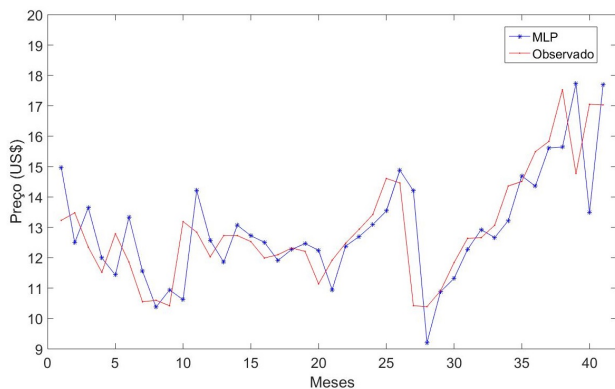


Figura 2. NY No.11 - Previsão um passo à frente com a MLP

sentido, o presente trabalho propôs a utilização de modelo linear e redes neurais para previsão do preço do açúcar: AR, ELM, ESN e MLP.

Os experimentos computacionais consideraram diferentes horizontes de previsão - 1, 2 e 3 passos à frente. Em relação à seleção de entradas para os modelos neurais utilizou-se a função de autocorrelação parcial.

Os resultados apresentados mostraram que os modelos ELM e MLP se aproximaram em apresentar os melhores resultados para previsão nas bases CEPEA e NY No. 11, respectivamente. Por outro lado, a rede neural ESN apresentou o pior resultado nas duas bases aplicadas.

Para trabalhos futuros, recomenda-se o desenvolvimento de um modelo híbrido para correção de erro. Ao realizar a hibridização entre modelos lineares e não-lineares considera-se que os problemas do mundo real apresentam uma estrutura complexa com ambos os padrões [39], [40]. Outra possibilidade é a utilização de ensembles.

REFERÊNCIAS

[1] P. D. Griggs, *Global Industry, Local Innovation: the history of cane sugar production in Australia, 1820-1995*. Peter Lang, Bern, Switzerland, 2011.

[2] S. Gudoshnikov, L. Jolly, and D. Spence, *The world sugar market*. Association with the International Sugar Organization, Elsevier, 2004.

[3] I. Benešová, H. Řezbová, L. Smutka, K. Tomšík, A. Laputková *et al.*, "European sugar market—impact of quota system," *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. 63, no. 6, pp. 1825–1838, 2015.

[4] M. L. d. Moraes and M. R. P. Bacchi, "Etanol: do início às fases atuais de produção," *Revista de Política Agrícola*, vol. 23, no. 4, pp. 5–22, 2015.

[5] L. R. A. Alves and M. R. P. Bacchi, "Oferta de exportação de açúcar do brasil," *Revista de Economia e Sociologia Rural*, vol. 42, no. 1, pp. 9–33, 2004.

[6] R. D. Soliani, A. Argoud *et al.*, "Analysis of the logistics infrastructure of the sugar bulk produced to export in the brazilian state of são paulo," *European Journal of Business and Social Sciences*, vol. 4, no. 6, pp. 100–113, 2015.

[7] "Raw sugar (hs: 1701) product trade, exporters and importers." [Online]. Available: <https://oec.world/en/profile/hs92/raw-sugar>

[8] V. Pagani, T. Stella, T. Guameri, G. Finotto, M. V. D. Berg, F. R. Marin, M. Acutis, and R. Confalonieri, "Forecasting sugarcane yields using agro-climatic indicators and canegro model: A case study in the main production region in brazil," *Agricultural Systems*, vol. 154, pp. 45–52, 2017.

[9] N. Silva, I. Siqueira, S. Okida, S. L. Stevan, and H. Siqueira, "Neural networks for predicting prices of sugarcane derivatives," *Sugar Tech*, vol. 21, no. 3, pp. 514–523, 2019.

[10] K. Suresh and S. K. Priya, "Forecasting sugarcane yield of tamil nadu using arima models," *Sugar Tech*, vol. 13, no. 1, pp. 23–26, 2011.

[11] E. M. Amrouk and T. Heckelee, "Forecasting international sugar prices: A bayesian model average analysis," *Sugar Tech*, vol. 22, no. 4, pp. 552–562, 2020.

[12] O. O. Obe and D. K. Shangodoyin, "Artificial neural network based model for forecasting sugar cane production 1," *Journal of Computer Science*, vol. 6, no. 4, pp. 439–445, 2010.

[13] R. Souza, "Modelos Estruturais Para Previsão de Séries Temporais: abordagens clássica e bayesiana," in *Anais do Colóquio Brasileiro de Matemática*, vol. 17, Rio de Janeiro, RJ, Brasil, 1989.

[14] R. Reyhani and A. M. E. Moghadam, "A heuristic method for forecasting chaotic time series based on economic variables," in *2011 Sixth International Conference on Digital Information Management*. IEEE, 2011, pp. 300–304.

[15] S. G. Makridakis and S. C. Wheelwright, *Forecasting Methods for Management*, 5th ed. Nova Jersey, EUA: Wiley, 1989.

[16] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and machine learning forecasting methods: Concerns and ways forward," *PloS one*, vol. 13, no. 3, p. e0194889, 2018.

[17] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[18] L. Castro, "Fundamentals of Natural Computing - Basic Concepts, Algorithms, and Applications," in *Chapman and Hall / CRC computer and information science series*, 2006.

[19] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Hamilton: Prentice Hall, 2008.

[20] O. S. Oshodi, O. A. Ejohwomu, I. O. Famakin, and P. Cortez, "Comparing univariate techniques for tender price index forecasting: Box-jenkins and neural network model," *Construction Economics and Building*, vol. 17, no. 3, pp. 109–123, 2017.

[21] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Hamilton, ON: Prentice Hall, 2008.

[22] H. Siqueira and I. Luna, "Performance comparison of feedforward neural networks applied to streamflow series forecasting," *Mathematics In Engineering Science And Aerospace*, vol. 10, no. 1, pp. 41–53, 2019.

[23] Y. Kachba, D. M. d. G. Chiroli, J. T. Belotti, T. A. Alves, Y. de Souza Tadano, and H. Siqueira, "Artificial neural networks to estimate the influence of vehicular emission variables on morbidity and mortality in the largest metropolis in south america," *Sustainability*, vol. 12, no. 7, p. 2621, 2020.

[24] H. Siqueira, L. Boccato, R. Attux, and C. Lyra, "Echo state networks and extreme learning machines: A comparative study on seasonal streamflow series prediction," in *Neural Information Processing*, T. Huang, Z. Zeng, C. Li, and C. S. Leung, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 491–500.

[25] I. N. Da Silva, D. H. Spatti, and R. A. Flauzino, "Redes neurais artificiais para engenharia e ciências aplicadas: curso prático," *Revista Controle Automação*, vol. 23, no. 5, pp. 649–652, 2012.

[26] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, vol. 2, 2004, pp. 985–990 vol.2.

[27] —, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.

[28] Y. de Souza Tadano, H. V. Siqueira, and T. A. Alves, "Unorganized machines to predict hospital admissions for respiratory diseases," in *2016 IEEE Latin American Conference on Computational Intelligence (LACCI)*. IEEE, 2016, pp. 1–6.

[29] J. Belotti, H. Siqueira, L. Araujo, S. L. Stevan, P. S. de Mattos Neto, M. H. Marinho, J. F. L. de Oliveira, F. Usberti, M. d. A. Leone Filho, A. Converti *et al.*, "Neural-based ensembles and unorganized machines to predict streamflow series from hydroelectric plants," *Energies*, vol. 13, no. 18, p. 4769, 2020.

[30] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks – with an Erratum note," GMD - German National Research Institute for Computer Science, Tech. Rep. 148, 2010.

[31] I. B. Yildiz, H. Jaeger, and S. J. Kiebel, "Re-visiting the echo state property," *Neural Networks*, vol. 35, pp. 1–9, 2012. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0893608012001852>

[32] H. Siqueira, L. Boccato, R. Attux, and C. Lyra Filho, "Echo state networks for seasonal streamflow series forecasting," in *International Conference on Intelligent Data Engineering and Automated Learning*. Springer, 2012, pp. 226–236.

[33] H. V. Siqueira, L. Boccato, R. Attux, and C. Lyra Filho, "Echo state networks in seasonal streamflow series prediction," *Learning and Nonlinear Models*, vol. 10, no. 3, pp. 181–191, 2012.

[34] D. A. Dickey and W. A. Fuller, "Likelihood ratio statistics for autoregressive time series with a unit root," *Econometrica: journal of the Econometric Society*, pp. 1057–1072, 1981.

- [35] J. Belotti, J. J. Mendes, M. Leme, F. Trojan, S. L. Stevan, and H. Siqueira, "Comparative study of forecasting approaches in monthly streamflow series from brazilian hydroelectric plants using extreme learning machines and box & jenkins models," *Journal of Hydrology and Hydromechanics*, vol. 69, no. 2, pp. 180–195, 2021.
- [36] A. Millea, "Explorations in echo state networks," Master's thesis, University of Groningen, 2014.
- [37] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [38] P. A. Morettin and C. Toloi, *Análise de séries temporais*. Ed. Blucher, 2006.
- [39] V. H. A. Ribeiro, G. Reynoso-Meza, and H. V. Siqueira, "Multi-objective ensembles of echo state networks and extreme learning machines for streamflow series forecasting," *Engineering Applications of Artificial Intelligence*, vol. 95, p. 103910, 2020.
- [40] P. S. de Mattos Neto, J. F. L. de Oliveira, D. S. d. O. S. Júnior, H. V. Siqueira, M. H. D. N. Marinho, and F. Madeiro, "A hybrid nonlinear combination system for monthly wind speed forecasting," *IEEE Access*, vol. 8, pp. 191 365–191 377, 2020.