



# Previsão do preço de açúcar utilizando Redes Neurais Artificiais.

## *Sugar price forecasting using Artificial Neural Networks.*

Henrique Nazário Rocha, Hugo Valadares Siqueira.

### RESUMO

Este trabalho aborda uma breve descrição sobre previsão de dados para um conjunto de séries temporais para o preço de açúcar. Os países que foram usados para base de amostra de dados foram: Austrália, China, Brasil, Índia, México, Paquistão, Rússia, Tailândia, EUA e União Europeia. A União Europeia tem uma política agrícola comum que controla os preços dos países produtores que a constituem. Os modelos de previsão utilizados são arquiteturas de redes neurais artificiais e metodologias lineares da família Box & Jenkins. Os resultados computacionais indicam que a rede neural com estados de eco se destaca na solução do problema.

**Palavras-chave:** Redes Neurais Artificiais, Previsão de Séries Temporais, Preço de Açúcar.

### ABSTRACT

This work approaches a brief description of forecasting data for a set of time series for sugar price. The countries used for the sample database were: Australia, China, Brazil, India, Mexico, Pakistan, Russia, Thailand, USA and European Union. The European Union has a regular agricultural policy that controls the prices of the producing countries that compose it. The forecasting models used are artificial neural network architectures and linear methodologies from the Box & Jenkin family. The computational results indicate that the echo states network stands out in solving the problem.

**Keywords:** Neural Networks, Time Series Forecast, Sugar Price.

## 1 INTRODUÇÃO

O Brasil tem como grande parte da sua riqueza hereditária a produção de açúcar, a qual começou há mais de duzentos anos e dura até hoje como um dos principais produtos da sua economia (DANTAS, LUIZ FL LEGEY & ANTONELLA, 2013). O país é um dos grandes produtores e fornecedores de açúcar do mundo (MORAES & MIRIAN, 2015).

O consumo de açúcar tem aumentado ao longo dos anos. Vários fatores contribuem para isso, como a elaboração de produtos que necessitam do seu uso, por exemplo doces, refrigerantes, balas, etc. Estes são alimentos altamente dependentes de açúcar bruto. Trata-se, assim, de um produto de consumo básico e uma commodity essencial produzida em várias partes do mundo (SILVA ET AL., 2018).

Modelos da Econometria, conjunto de ferramentas estatísticas com o objetivo de entender a relação entre variáveis econômicas, que também aborda análise de séries temporais, são frequentemente usados para relacionar o preço do açúcar com as épocas que foram cotados. Neste processo é feita uma comparação entre métodos para conseguir verificar o modelo que mais se adequa à proposta. Além disso, uma combinação dos modelos normalmente pode ser feita para obter uma gama maior de resultados, o que enriquece a análise geral (OBE & D. K. SHANGODOYIN, 2010).

\* Bacharelado em Engenharia Elétrica, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; nazario.utfpr@gmail.com

† Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; hugosiqueira@utfpr.edu.br



Este trabalho aborda a previsão de séries temporais relativas ao preço do açúcar. Os países que foram usados para base de amostra de dados foram: Austrália, China, Brasil, Índia, México, Paquistão, Rússia, Tailândia, EUA, União Europeia. A União Europeia tem uma política agrícola comum que controla os preços dos países produtores que a constituem. Todos os dados utilizados podem ser acessados no site [indexmundi](https://www.indexmundi.com/) (<https://www.indexmundi.com/>). Para tal, lançar-se-á mão de redes neurais artificiais, os quais já mostraram sua capacidade de superar modelos tradicionais como os da família Box & Jenkins (SILVA ET AL, 2018, DELATORRE ET AL., 2015). Tais metodologias não-lineares com alto poder de processamento e mapeamento (HAYKIN, 1999), os quais são brevemente descritos a seguir.

## 2 MODELOS DE PREVISÃO

Nesta seção apresentamos os modelos lineares auto-regressivo (AR) e auto-regressivo de médias móveis (ARMA), assim como as arquiteturas de redes neurais que serão utilizadas para prever o preço do açúcar. Tais propostas foram selecionadas por se enquadrarem ou como modelos clássicos e bem sedimentados na literatura ou por apresentarem interessantes resultados em trabalhos correlatos.

### 2.1 PERCEPTRON MÚLTIPLAS CAMADAS (MLP)

Um neurônio artificial consiste de uma unidade de processamento de informação. Sua inspiração deriva do estudo de sistema nervoso, no caso, sobre a propagação e geração de pulsos elétricos pela membrana celular dos neurônios biológicos. A representação do mesmo é dada por um modelo matemático simplificado do neurônio biológico, que possui pesos sinápticos ajustáveis, com o intuito de replicar a plasticidade sináptica (HAYKIN, 1999). O neurônio artificial contém funções de receber, processar e transmitir sinais, sendo capazes de reconhecer padrões de dados através do processo de treinamento, e realizar generalizações baseada no aprendizado adquirido (DA SILVA, SPATTI D & FLAUZINO, 2010).

O perceptron de múltiplas camadas (MLP - *multilayer perceptron*) é uma rede neural que contém no mínimo três camadas de neurônios. É classificada como uma rede feedforward, sendo constituída de uma camada de entrada, camadas ocultas e uma camada de saída, sendo que a quantidade de camadas ocultas pode variar. Tal rede consegue classificar ou realizar a previsão de diversos tipos de dados (DA SILVA, SPATTI D & FLAUZINO, 2010).

### 2.2 MÁQUINAS DE APRENDIZADO EXTREMO (ELM)

A topologia de uma máquina de aprendizado extremo (ELM - Extreme Learning Machine) é bastante semelhante às MLP's, mas com a diferença de que na camada intermediária não ocorre treinamento, ou seja, os pesos sinápticos são obtidos de forma aleatória e apenas os pesos da camada de saída tem seus valores ajustados (ELMAN, 1990). Assim o esforço de processamento para o treinamento é menor. Por esse fator de aleatoriedade são classificadas como máquinas desorganizadas (SIQUEIRA, BOCCATO, ATTUX & LYRA, 2012).

O treinamento da rede consiste em uma regressão linear. Neste trabalho, na topologia configurada foi usado o método da pseudo-inversa de Moore-Penrose para obtenção dos valores ótimos para os pesos sinápticos (HUANG, ZHOU, DING & ZHANG, 2012).

\* Bacharelado em Engenharia Elétrica, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; [nazario.utfpr@gmail.com](mailto:nazario.utfpr@gmail.com)

† Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; [hugosiqueira@utfpr.edu.br](mailto:hugosiqueira@utfpr.edu.br)



### 2.3 REDE NEURAL DE ELMAN E JORDAN

Assim como as redes neurais já citadas (MLP e ELM), a rede de Elman também surge a partir da inspiração do neurônio biológico, mas difere no seu processo de combinação de neurônios. Esta é classificada como uma rede neural recorrente (RNN), que é dividida em duas partes. A primeira recebe as entradas de dados como os outros modelos, e a segunda, por sua vez, chamada de unidade de contexto, consiste nas saídas da camada oculta. Logo, existe um processo cíclico dos dados (ELMAN, 1990).

A rede de Jordan foi criada a partir de uma MLP em 1986, sendo esta a primeira rede neural recorrente. A arquitetura foi criada com intuito de usar seu processo recorrente em aplicações de reconhecimento de séries temporais, mas atualmente é aplicada a todos os tipos de problemas comumente resolvidos com redes neurais artificiais (JORDAN, 1990). Diferentemente do modelo de rede de Elman descrito anteriormente, a inicialização e recepção dos dados para as unidades de contexto são alimentadas pelas saídas dos neurônios localizados na camada de saída e não da camada oculta. Este tipo de característica traz resultados distintos (HAYKIN, 1999).

### 2.4 REDES NEURAS COM ESTADO DE ECO (ESN)

A rede neural com estados de eco (*Echo State Network* - ESN) é considerada uma máquina desorganizada, conceito que foi inserido por Alan Turing (em 1948) (TURING, 1948), em que definiu como sendo mecanismos aleatórios em sua construção, mas capazes de executar tarefas singulares mediante ao aprendizado e treinamento do mesmo (JAEGER, 2001).

Em geral, as ESN's são constituídas de 3 (três) camadas: de entrada; a segunda, que é chamada de reservatório de dinâmicas, na qual os neurônios artificiais estão totalmente interconectados, o que gera uma característica não-linear; e a terceira camada, a de saída, responsável por combinar as saídas do reservatório de dinâmicas (HAYKIN, 1999).

O reservatório, por sua vez, pode ser elaborado seguindo propostas diversas. Neste trabalho abordou-se as propostas de Jaeger e Ozturk et al. (SIQUEIRA, 2012). Acerca do treinamento, utilizou-se a mesma ideia descrita na Seção 1.2.

### 2.5 MODELO AUTO-REGRESSIVO (AR)

Modelos lineares e econométricos são há décadas aplicadas na previsão e modelagem de séries temporais. Propostas como suavização exponencial e a metodologia Box & Jenkins são representantes desta classe. Neste trabalho, abordaremos alguns modelos desta última classe (SIQUEIRA, 2012).

Um modelo Auto-Regressivo (AR) envolve uma combinação linear dos atrasos de uma variável em seus próprios valores defasados, ou seja, referentes à observação dos seus valores anteriores. O mesmo é representado na Eq. 1.

$$X_t = c + \sum_{i=1}^P \varphi_i X_{t-i} + \varepsilon_t \quad (1)$$

O modelo AR é formalizado para analisar séries temporais estacionárias e utiliza coeficientes para representar a função que se deseja realizar a previsão. Para determinar os coeficientes ótimos é necessário

\* Bacharelado em Engenharia Elétrica, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; nazario.utfpr@gmail.com

† Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; [hugosiqueira@utfpr.edu.br](mailto:hugosiqueira@utfpr.edu.br)



resolver uma relação de recorrência, conhecida como Equações de Yule-Walker (DINDA & O'HALLARON, 2000).

## 2.6 MODELO AUTO-REGRESSIVO DE MÉDIAS MÓVEIS (ARMA)

Além do AR, outros modelos lineares também são bastante utilizados. Um dos comumente empregados na previsão de séries temporais é o Médias Móveis (MA), que pode ser observado na Eq. 2.

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2)$$

Neste, diferentemente do AR que combina de forma ponderada os valores anteriores da série de dados, são combinados sinais que simbolizam ruídos brancos (HONG, 1999).

Por sua vez, um Modelo Auto-Regressivo e de Médias Móveis (ARMA, do inglês Autoregressive Moving and Average) é uma combinação dos dois métodos (DINDA & O'HALLARON, 2000), AR e MA, representado na Eq. 3.

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}. \quad (3)$$

## 3 METODOLOGIA

Para realizar tarefas de previsão, um ponto crítico é determinar quais os melhores atrasos devem ser utilizados como entrada para o modelo. Note que a utilização de uma grande quantidade de atrasos como entrada não necessariamente leva a menores erros (GUYON & ELISSEEFF, 2003).

Pode-se, então, testar todas as possíveis combinações de entrada, mas isso é muito dispendioso computacionalmente, o que pode tornar o treinamento do modelo uma tarefa extremamente demorada. O método Wrapper com seleção progressiva é uma forma de contornar esse problema (SIQUEIRA, 2012).

Neste trabalho foi usado o método Wrapper para seleção de entradas, para todas as metodologias, tanto de redes neurais artificiais quanto para os modelos lineares AR e ARMA (HAYKIN, 1999). Variou-se também para as RNAs o número de camadas intermediárias (MLP) e a quantidade de neurônios (SIQUEIRA, 2012). Cada topologia com sua configuração determinada pelo Wrapper teve 30 (trinta) treinamentos, e testes executados. A disposição da quantidade de dados para cada amostra foi de 70% para treinamento, 15% para validação, e por fim 15% para realizar os testes. A métrica de erro usada foi o erro quadrado médio (MSE - *Mean Squared Error*) (SIQUEIRA, 2012).

A função de ativação usada foi a tangente hiperbólica nas RNAs. A variação do número de neurônios artificiais na camada intermediária foi dada a partir de 5 (cinco) com incremento a cada 5 (cinco) neurônios, terminando com o limite de 20 (vinte). O coeficiente de regularização (CR) foi usado em ELMs e ESN com o fim de melhorar sua capacidade de generalização (SIQUEIRA, 2012).

O pré-processamento usado para tratar os dados de preço de açúcar é chamado de dessazonalização (z-score) para retirada da componente sazonal das séries. Tal método deixa a série com média zero e desvio padrão aproximadamente unitário, podendo ser classificada como estacionária no sentido amplo (SIQUEIRA, 2012).

\* Bacharelado em Engenharia Elétrica, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; nazario.utfpr@gmail.com

† Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; hugosiqueira@utfpr.edu.br





## 4 RESULTADOS

Na Tab. 1, cada valor apresentado é o melhor obtido nas análises de sua respectiva configuração usando o valor MSE de teste, considerando como horizonte a previsão um passo à frente. Ademais, os números destacados em vermelho representam o menor MSE para cada país analisado, ou seja, qual método apresenta o menor erro.

Com os resultados obtidos foi possível constatar que o MSE foi baixo em praticamente todos os algoritmos. Com a variação da configuração dos modelos, a performance dos modelos teve um decréscimo tênue. Além disso, foi possível verificar que a rede ESN com o reservatório de dinâmicas de Ozturk et al. foi a que apresentou o melhor desempenho geral, tendo a melhor resposta em 4 (quatro) das 8 (oito) análises realizadas neste trabalho.

Durante a elaboração deste trabalho, um total de 8 (oito) modelos foram compilados com suas permutações de configuração 30 (trinta) vezes, cada um teve um tempo de processamento que, a depender do modelo, poderia terminar sua execução em minutos ou até em dias.

Tabela 1 - Resultado do conjunto teste, valores em MSE.

Modelo	Brasil	U.Europeia	Índia	México	Paquistão	Rússia	Tailândia	U.S.A.
AR	0,045259	0,002286	0,033419	0,053161	0,028244	0,049265	0,017333	0,015241
ARMA	0,219269	0,002369	0,032867	0,170936	0,026320	0,052821	0,040447	0,047841
ELM	0,050441	0,002150	0,025853	0,044971	0,023259	0,051062	0,018800	0,015192
ELMAN	0,056957	0,009178	0,044000	0,080399	0,035000	0,133999	0,014159	0,029335
ESN (Jaeger)	0,048973	0,002220	0,026543	0,049104	0,025862	0,053276	0,014631	0,014345
ESN (Ozturk)	0,043765	0,002349	0,024477	0,049344	0,021228	0,054008	0,014631	0,014246
Jordan	0,064202	0,002900	0,033414	0,062956	0,033608	0,026965	0,013601	0,027726
MLP	0,052504	0,003187	0,030551	0,359174	0,27267	0,054867	0,014364	0,014948

Fonte: Autoria própria

## 5 CONCLUSÃO

O presente relatório apresentou uma aplicação de redes neurais artificiais (RNAs) e modelos lineares (AR e ARMA) para previsão de séries temporais com foco na cotação do preço de açúcar. Os países selecionados para a amostra de dados foram: Austrália, China, Brasil, Índia, México, Paquistão, Rússia, Tailândia, EUA e União Europeia.

Diferentes tipos de configurações foram definidas, com o intuito de se obter uma gama de resultados, sendo a maioria deles resultados adequados ao problema em questão. A análise dos modelos propostos levou a uma visão ampla sobre seus devidos potenciais.

Foi possível verificar que a rede ESN com o reservatório de dinâmicas elaborado por Ozturk et al. teve o melhor desempenho. Esta rede é munida de recorrência no reservatório de dinâmicas, o que lhe confere uma memória intrínseca. Esta propriedade pode ser crucial no desempenho em problemas deste tipo, uma vez que há dependência temporal entre as amostras.

\* Bacharelado em Engenharia Elétrica, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; nazario.utfpr@gmail.com

† Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; hugosiqueira@utfpr.edu.br



SEI-SICITE 2021

Pesquisa e Extensão para um  
mundo em transformação

## AGRADECIMENTOS

Agradecimento à CNPQ pela bolsa concedida para realização deste projeto.

## REFERÊNCIAS

DANTAS, GUILHERME A., LUIZ FL LEGEY, AND ANTONELLA MAZZONE. "Energy from sugarcane bagasse in Brazil: An assessment of the productivity and cost of different technological routes." **Renewable and Sustainable Energy Reviews** **21** (2013): 356-364.

OBE, O. O., and D. K. Shangodoyin. "Artificial Neural Network Based Model for Forecasting Sugar Cane Production 1." (2010).

HAYKIN, S. (1999) NEURAL NETWORKS: A Comprehensive Foundation. **Prentice Hall, Upper Saddle River**.

DA SILVA, I.N., SPATTI D., E FLAUZINO, R. (2010). Redes neurais artificiais para engenharia e ciências aplicadas: curso prático, **Artliber Editora Ltda**, São Paulo, SP, Brasil.

SIQUEIRA, H. V., L. BOCCATO, R. ATTUX, AND C. LYRA FILHO. "Echo State Networks and Extreme Learning Machines: a Comparative Study on Seasonal Streamflow Series. Prediction." **Lecture Notes in Computer Science**, 2012a: 491-500.

HUANG, G.-B., ZHOU, H., DING, X., ZHANG, R. Extreme Learning Machines for Regression and Multiclass Classification. **IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics**, vol. 42, no. 2, pp.513-529, 2012.

ELMAN, Jeffrey L. Finding structure in time. **Cognitive science**, v. 14, n. 2, p. 179-211, 1990.

JORDAN, Michael I. Attractor dynamics and parallelism in a connectionist sequential machine. In: **Artificial neural networks: concept learning**. 1990. p. 112-127.

TURING, Alan M. Intelligent machinery. **National Physical Laboratory**. Essential Turing, v. 2004, p. 410-432, 1948.

JAEGER, Herbert. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. **Bonn, Germany: German National Research Center for Information Technology GMD Technical Report**, v. 148, n. 34, p. 13, 2001.

DINDA, PETER A.; O'HALLARON, DAVID R. Host load prediction using linear models. **Cluster Computing**, v. 3, n. 4, p. 265-280, 2000.

HONG, YONGMIAO. Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. **Journal of the American Statistical Association**, v. 94, n. 448, p. 1201-1220, 1999.

GUYON, Isabelle; ELISSEEFF, André. An introduction to variable and feature selection. **Journal of machine learning research**, v. 3, n. Mar, p. 1157-1182, 2003.

SILVA, NATÁLIA ; SIQUEIRA, IGOR ; OKIDA, SÉRGIO ; STEVAN, SÉRGIO L. ; Siqueira, Hugo . **Neural Networks for Predicting Prices of Sugarcane Derivatives. Sugar Tech**, v. 1, p. 1-10, 2018.

DELATORRE, A. ; OKIDA, S. ; TROJAN, F. ; PONTES, J. C. A. ; STEVAN, S. L. ; Siqueira, H. V. . Avaliação preliminar do desempenho de redes neurais feedforward para a previsão de preços de derivados da cana-de-açúcar. In: **X Congresso Brasileiro de Agroinformática**, 2015, Ponta Grossa. Anais do X SBIAGRO, 2015.

\* Bacharelado em Engenharia Elétrica, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; nazario.utfpr@gmail.com

† Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil; [hugosiqueira@utfpr.edu.br](mailto:hugosiqueira@utfpr.edu.br)