# COMPARAÇÃO DO DESEMPENHO ENTRE MODELOS LINEARES E REDES NEURAIS PARA PREVISÃO DO PREÇO DO CAFÉ

**Carolina Deina**
Universidade Tecnológica Federal do Paraná – Campus Pato Branco
Via do Conhecimento, s/n - KM 01 - Fraron, Pato Branco - PR, 85503-390
caroldeina@gmail.com

**Matheus Henrique do Amaral Prates**
Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa
Av. Monteiro Lobato, s/n - Km 04 – Ponta Grossa – PR, 84016-210
matheushaprates22@gmail.com

**Carlos Henrique Rodrigues Alves**
Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa
Av. Monteiro Lobato, s/n - Km 04 – Ponta Grossa – PR, 84016-210
carloshenriquera@gmail.com

**Marcella Scoczynski Ribeiro Martins**
Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa
Av. Monteiro Lobato, s/n - Km 04 – Ponta Grossa – PR, 84016-210
marcella_engcomp@yahoo.com.br

**Hugo Valadares Siqueira**
Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa
Av. Monteiro Lobato, s/n - Km 04 – Ponta Grossa – PR, 84016-210
hugosiqueira@utfpr.edu.br

**Flavio Trojan**
Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa
Av. Monteiro Lobato, s/n - Km 04 – Ponta Grossa – PR, 84016-210
trojan@utfpr.edu.br

## RESUMO

O café é uma iguaria que por séculos agrada ao paladar, sendo o Brasil o principal produtor e exportador dessa mercadoria no mundo. A comercialização de grãos afeta diretamente a economia do país, gerando empregos e incremento no produto interno bruto nacional. Neste sentido, o presente trabalho buscou analisar o desempenho de ferramentas para a previsão do preço desse item. A metodologia baseou-se em uma comparação entre métodos lineares e redes neurais artificiais: Suavização Exponencial, modelo Auto Regressivo (AR), Auto Regressivo Integradas e de Médias Moveis (ARIMA), Perceptron de Múltiplas Camadas (MLP) e Máquinas de Aprendizado Extremo (ELM). Os dados foram coletados do Centro de Estudos Avançados em Economia Aplicada no Brasil (CEPEA) sobre dois tipos de café. Após a normalização dos dados, foram identificados os atrasos mais relevantes para então realizar as previsões. Os resultados computacionais mostram que as redes neurais são capazes de alcançar desempenhos superiores aos demais.

**PALAVRAS CHAVE. Previsão de preço de café, modelos lineares, redes neurais artificiais.**

**Tópicos (Inteligência Computacional, Simulação, Metaheurísticas)**

**ABSTRACT**

Coffee is a delicacy that for centuries delight the human palate, Brazil being the main producer and exporter of this delicacy in the world. The commercialization of grains affects the country's economy, generating jobs and increasing national gross domestic product. Thus, this work sought to analyze the performance of forecasting tools to estimate the price of this item. The methodology was based on a comparison between the linear methods and artificial neural networks: Exponential Smoothing, Auto Regressive model (AR) and Auto Regressive Integrated Moving Average model (ARIMA), multilayer perceptron (MLP) and extreme learning machines (ELM). The data were collected from the Center for Advanced Studies in Applied Economics in Brazil (CEPEA) on two types of coffee. After the normalization of the data, the delays were identified and the predictions were made. The computational results showed that the neural networks are capable of reach higher performances than the other models.

**KEYWORDS. Coffee price forecasting, linear models, artificial neural networks.**

**Paper topics (Computational Intelligence, Simulation, Metaheuristics)**

## 1. Introdução

O café não é uma iguaria apreciada somente em países que têm tradição no comércio dessa mercadoria. Na Alemanha, por exemplo, o café algumas vezes é mais consumido do que a mais tradicional bebida do país, a cerveja. Devido a importância da comercialização desse grão, a representação do café é expressiva para a economia mundial, não somente na produção e exportação do grão, mas também em relação ao consumo. No Brasil, no ano de 2018 foram comercializadas internamente cerca de 20,5 milhões de sacas de 60 kg, o que fez do país o segundo maior consumidor dessa bebida no mundo, ficando atrás apenas dos Estados Unidos [ICO 2019].

O consumo interno e a exportação desse grão movimentam a balança de comercio internacional e os agricultores são os principais responsáveis por contribuírem por isso. Porém, há um problema relacionado às previsões do preço desse bem, que segundo [Cuaresma et al. 2018], está relacionado à volatilidade que algumas vezes pode ser explicada pelas variações de tendências nos preços no mercado mundial. Com isso, torna-se relevante estudar estratégias e ferramentais que possam auxiliar na previsão desses preços. Ainda segundo os mesmos autores, as metodologias de previsão de preços de *commodities* são cada vez mais comuns. Quando os resultados gerados pelas previsões são consistentes, não auxiliam apenas no processo de tomada de decisão, mas também em toda a cadeia produtiva.

Alguns estudos importantes têm sido desenvolvidos sobre esse tópico em diferentes lugares no mundo. [Milas et al. 2004] realizaram previsões de preços utilizando correlação de erro linear e não linear para quatro tipos de café: arábica (brasileiro) arábica (colombiano) arábicas suaves (outros países da América Latina) e café robusta (África e sudeste da Ásia). [Ayyanathan e Kannammal 2014] utilizaram máquinas de vetores suporte e redes neurais para prever o preço de ações de uma empresa líder em exportação de café no mercado indiano. [Fousekis 2017] investigou o preço do café arábica e robusta através de previsões para valor de risco (VaR), comercializados em Nova York e Londres, respectivamente. [Yashavanth et al. 2017] aplicaram o vetor auto regressivo (VAR) e o método ARIMA para prever o preço mensal do café em diferentes centros consumidores da índia. [Pradkhan 2017] utilizou modelos de regressão para identificar causalidades de diferentes atividades que impactam no preço em diferentes *commodities* nos Estados Unidos. [David et al. 2017] utilizaram os modelos estatísticos ARIMA e um modelo ARIMA fracionado para fazer previsões de preços de *commodities* baseadas em quatro séries temporais. [Cuaresma et al. 2018] realizam a previsão de preços do café utilizando métodos

econométricos com entradas de variáveis macroeconômicas e financeiras. [Silva 2018] utilizou a metodologia ARIMA para prever o preço do café arábica brasileiro.

Assim, percebe-se que os modelos de Redes Neurais Artificiais (RNA) se apresentam como uma potencial ferramenta eficaz para o problema, mas ainda com pouca utilização nesse contexto. Frequentemente esses modelos são utilizados para realizar previsão de preços em outros cenários, como em [Oshodi et al. 2017], [Fauziah e Gunaryati 2017] e [Ramyar e Kianfar 2019], que buscaram prever os custos de construções, preços do açúcar e preço do petróleo. Para esses casos foram comparados os desempenhos dos modelos lineares de previsão com RNA's, deixando evidências de que as últimas podem melhorar o resultado final.

Procurando preencher uma lacuna no tema relacionada à utilização de modelos de RNA`s para a "previsão do preço do café" esse trabalho tem como objetivo comparar o desempenho de modelos de redes neurais perceptron de múltiplas camadas (MLP) e Máquina de Aprendizagem Extremo (ELM) com os modelos lineares de Suavização Exponencial (SE), Auto Regressivo (AR) e Auto Regressivo Integrado de Médias Móveis (ARIMA).

O trabalho está organizado da seguinte forma: na seção "Referencial Teórico", são apresentados os conceitos sobre os métodos de SE, AR, ARIMA, MLP e ELM; a Seção "Metodologia" são estruturados os métodos para coleta, tratamento e previsão de dados; a Seção "Resultados e Discussão" apresenta os resultados e as comparações entre as metodologias apresentadas; enquanto que a Seção "Conclusões" procura-se tecer as considerações sobre a contribuição do estudo.

## 2. Referencial teórico

Nessa seção serão descritos os conceitos principais e processos de modelagem dos métodos lineares de previsão, sendo (SE, AR e ARIMA) e as arquiteturas de Redes Neurais Artificiais abordadas (MLP e ELM).

### 2.1 Suavização Exponencial

As técnicas de Suavização Exponencial (SE) são amplamente utilizadas em decorrência da sua facilidade de implementação, aplicação, ajuste e alcance de resultados eficientes de previsão [HYNDMAN e ATHANASOPOULOS 2018]. Dentre os modelos de SE, existem três principais os quais são considerados como base para todas as outras variações: simples, de Holt e de Holt-Winters [HYNDMAN e ATHANASOPOULOS 2018]. Esses modelos podem ser aplicados em séries de dados históricos que apresentem ciclicidade, sazonalidade e tendência [TAYLOR 2003].

O método de SE simples é utilizado quando há ausência de tendência e sazonalidade na série temporal. Como regra o modelo exige apenas que existam dados passados da variável a ser prevista e um parâmetro de aproximação com valor entre 0 e 1, chamado de constante de suavização do ciclo ($\alpha$) [MENEGHINI et al. 2018].

Já em relação ao método de suavização exponencial dupla de Holt, sua utilização se dá quando a série temporal apresenta ciclicidade e tendência, porém sem sazonalidade. Esse método, além do parâmetro ($\alpha$), utiliza também o parâmetro de suavização de tendência ($\beta$) com valores que variam entre 0 e 1 [MENEGHINI et al. 2018], conforme a Equação (1). As Expressões (2) e (3) apresentam a estimativa de nível e a tendência, respectivamente [HYNDMAN e ATHANASOPOULOS 2018].

$$\hat{Z}_{t+n} = L_t + n\,T_t \tag{1}$$

$$L_t = \alpha Z_t + (1-\alpha)\,.\,(L_{t-1} + T_{t-1}) \qquad 0 \le \alpha \le 1 \tag{2}$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)\,.\,T_{t-1} \qquad 0 \le \beta \le 1 \tag{3}$$

em que: $\hat{Z}_{t+n}$ é valor previsto para o período $t+n$, $(1 - \alpha)$ é taxa exponencial dos pesos de ponderação dos dados, $Z_t$ é o valor real para o período de tempo *t*, *n* representa períodos de tempos e $(1 - \beta)$ é taxa exponencial dos pesos de ponderação dos dados.

Quando a série temporal, além de ter comportamento cíclico e tendencioso, também apresenta um aspecto sazonal, usa-se o método de suavização exponencial de Holt Winters, que é suavizado pelo parâmetro (γ) [HYNDMAN e ATHANASOPOULOS 2018]. Para o cálculo, também consideraram dois tipos de efeitos sazonais (multiplicativo e aditivo). Detalhadamente, informações a respeito desses métodos podem ser encontradas em [Hyndman e Athanasopoulos 2018] e [Winters 1960].

### 2.2 Modelo Auto Regressivo

O modelo Autorregressivo (AR) de primeira ordem (*p*) é uma alternativa para previsões de séries temporais que necessitam que a série de dados apresente sua variância, auto-covariância e média invariáveis em relação ao tempo [HYNDMAN e ATHANASOPOULOS 2018]. A representação do modelo AR, em que a série de dados seja $y_t$, descrita pelos seus dados passado e pelo ruído aleatório $\varepsilon_t$, é representada na Equação (4).

$$y_t = \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} + \varepsilon_t \tag{4}$$

sendo $y_{t-1} + \cdots + y_{t-p}$ os termos passados da série, $\varphi_p$ os coeficientes livres do modelo e $\varepsilon_t$ o erro intrínseco ao processo de previsão.

Veja que $\varepsilon_t$ pode ser visto uma sequência de choques aleatórios e independentes uns dos outros gerados a partir do processo.

A definição dos valores dos coeficientes livres do modelo é feita de forma determinística por meio da aplicação direta das equações de Yule-Walker [Box et al. 1994], as quais garantem o mínimo erro quadrático médio.

### 2.3 Modelo Autorregressivo Integrado de Médias Móveis

O método ARIMA surgiu da uma combinação da autoregressão (AR) com um modelo de média móvel (MA), acrescentados da diferenciação (I) [OSHODI et al. 2017]. Proposto por [Box et al. 1994], ele teve como objetivo captar o comportamento da autocorrelação ou da correlação seriada entre os dados de uma série temporal, para realização da atividade de previsão. Diferente de AR e MA, o ARIMA é uma alternativa para ser aplicado em séries não estacionárias que possuam tendência, desde que satisfaça a condição de diferenciação dos dados, tornando-os estacionários [OSHODI et al. 2017].

Para realização das previsões com este método, alguns passos devem ser seguidos [Box et al. 1994]: primeiro, é necessário deixar a série estacionária, sendo o número de diferenciações definida por (*d*); posteriormente, deverá ser encontrado os valores *p* e *q* (ordem do AR e do MA) em paralelo com algum método de estimação dos valores dos coeficientes livres; e tendo os parâmetros do ARIMA definidos, realizam-se então as previsões. O modelo ARIMA (*p, d, q*) pode ser descrito na Equação (5).

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{5}$$

em que $\varphi_0$ é uma constante no modelo estimado, $\varphi_1, \ldots, \varphi_p$ são as ponderações para os valores passados de $y_t$, $\theta_1, \ldots, \theta_q$ são os parâmetros que possibilitam escrever a série em função de choques passados e $\varepsilon_t$ são os termos de erros, similares a Equação (4).

Nota-se que os termos $\varphi_p$ ponderam as entradas provenientes do modelo AR e $\theta_q$ as do modelo MA. Ainda, é muito usual considerar $\varphi_0 = 0$.

Esse método de previsão tem gerado bons resultados e tem sido usado em problemas de previsão em diferentes contextos [ZHANG 2003]. Também pode ser visualizado em [Hyndman e Athanasopoulos 2018] mais detalhadamente.

Diferentemente dos modelos AR, entretanto, a definição dos parâmetros livres do ARIMA depende da solução de um conjunto de equações não lineares. Dessa forma, não há solução determinística direta para seu cálculo que garanta o mínimo MSE. Na literatura, o mais usual é lançar mão de estimadores de máxima verossimilhança.

### 2.4 Perceptron de Múltiplas camadas

De acordo com [Silva et al. 2018a] um *perceptron* de múltiplas camadas (MLP) tem a capacidade de transformar por aproximação qualquer função contínua, não-linear, diferençável e limitada. Por isso, a MLP é considerado um aproximador universal.

Numa MLP, as camadas são formadas por neurônios artificias. As camadas intermediárias recebem os sinais de entrada da primeira camada e, então, mapeiam estes de forma não linear para outro espaço, que é definido de acordo com o problema. O sinal transformado adentra a camada de saída e, por meio de combinações lineares ou não, esta fornece a resposta da rede [HAYKIN 1999] [CASTRO 2006].

A rede não apresenta recorrência, ou seja, tem uma condição de *feedforward*. As camadas disjuntas são conectadas, embora os neurônios das mesmas sejam incomunicáveis.

O processamento de sinal da rede é dado pela Equação (6) [Silva et al. 2018a]:

$$y = f_s\left(\sum_{k=0}^{K} W_{1k}^{0}\left(f\sum_{n=0}^{N} W_{kn}^{i}\,u_n + b_n\right)\right) \tag{6}$$

na qual **u** representa o vetor de entradas da rede, **b** é um vetor contendo o bias, *f* é a função de ativação, **w** é o vetor de pesos sinápticos que junto com as entradas formam um vetor de sinal de saída **y**.

Na Equação (5) tem-se $W_{kn}^{i}$ são os peso da camada intermediária, *k*= 0, 1, ..., *K* indica o neurônio em 0, 1, ..., *N* índice de cada entrada e $W_{1k}^{0}$ são conexões de saída do neurônio.

O treinamento de um MLP é o processo de ajuste dos pesos sinápticos. O algoritmo de retropropagação de erro (*backpropagation*), o mais comumente utilizado, opera através de um conjunto de padrões de dados inseridos na rede. A cada entrada de dados a rede analisa suas saídas e então compara as respostas obtidas com as desejadas para calcular o valor do erro e assim ajustar os pesos [AYYANATHAN e KANNAMMAL 2014].

Para [Silva et al. 2018a] um ponto extremamente importante nesse processo é a definição do ponto de parada do processo. Comumente o critério do erro mínimo de treinamento é utilizado para interromper o processo quando atingido o valor de erro esperado.

A escolha do conjunto de pesos ideais é feita em muitos casos pela técnica da validação cruzada. Além de ajudar na escolha da topologia ideal para um problema específico, essa técnica usa um conjunto de amostras quaisquer da rede em cada iteração. Aquele que apresentar menor erro de validação tem seus pesos adotados como os capazes de maximizar a capacidade de generalização da rede [HAYKIN 1999].

### 2.5 Máquinas de Aprendizado Extremo

As máquinas de aprendizado extremo (*Extreme Learning Machines* – ELM) são RNAs *feedforward* que assim como as MLPs são capazes de aproximar qualquer função continua, não-linear, diferençável e limitada. A arquitetura dessa rede é muito semelhante com a MLP, apresentando apenas uma camada intermediaria [HUANG et al. 2004].

A grande diferença entre elas ocorre na forma em que a rede é treinada, que apesar da sua forma simples, seu desempenho computacional é rápido e seu resultado é eficaz [MICHE et al. 2010]. Enquanto na MLP os pesos da camada oculta e da camada de saída são ajustados por metodologias de otimização não linear, na ELM os pesos da oculta são gerados e de forma randômica e mantidos sem ajuste, e os pesos da de saída são determinadas de forma analítica [HUANG et al. 2006].

Quando são inseridos novos neurônios na camada oculta, aumenta-se o potencial da rede alcançando a precisão necessária. Conforme [Silva et al. 2018a] esses neurônios têm suas ativações dadas pela equação (7) e a Equação (8) gera a saída.

$$x_n^h = f^h(\mathbf{W^h u}_n + \mathbf{b}) \tag{7}$$

$$\mathbf{y} = \mathbf{W}^{out}\, x_n^h \tag{8}$$

de forma que $\mathbf{u}_n = [u_n, u_{n-1}, \dots, u_{n-K-1}]^T$ é o vetor que contém o sinal de entrada, $\mathbf{W}^h$ são os coeficientes (pesos) da camada intermediária, $\mathbf{b}$ é polarização de cada neurônio $f^h(.)$ especifica as funções de ativação dos neurônios e $\mathbf{W}^{out}$ é a matriz contendo os pesos da camada de saída.

Para determinar os pesos da última camada de forma analítica, precisa encontrar a matriz inversa de um sistema de equações lineares em que não há uma solução exata. Uma forma muito comum para resolver este problema, é usando a pseudo inversa Moore-Penrose [Silva et al. 2018a], a qual pelo cálculo da norma mínima (Equação 9) encontra uma aproximação para este tipo de sistema de equações não-lineares.

$$\mathbf{W}^{out} = (\mathbf{X}_h^T \mathbf{X}_h)^{-1}\, \mathbf{X}_h^T \mathbf{d} \tag{9}$$

em que $\mathbf{d}$ é o vetor com as saídas desejadas e $\mathbf{X}_h$ a matriz contendo as saídas da camada oculta.

## 3. Metodologia

Para atingir os objetivos deste trabalho, uma aplicação numérica foi realizada através da coleta de dados brutos sobre os preços praticados para dois tipos específicos de café: Arábica e Robusta. Os dados foram coletados na base disponível no site do Centro de Estudos Avançados em Economia Aplicada da Escola Superior de Agronomia Luiz de Queiroz (ESALQ) - Universidade de São Paulo (USP) [CEPEA 2019].

Os dados referem-se aos preços médios mensais do Indicador do café do ESALQ / BVMF (commodities e câmbio futuro). A média mensal (em reais) corresponde a um saco de 60 kg, negociado no mercado spot do estado de São Paulo. Em relação ao café Robusta, foram coletadas amostras entre o período de novembro de 2001 até dezembro de 2018, totalizando 206 dados mensais de preços praticados. Para o Arábica, os dados correspondem ao período entre setembro de 1996 e dezembro de 2018, totalizando 271 amostras.

Após a coleta, estes dados foram analisados graficamente com o objetivo de identificar seu comportamento, que de acordo com [Bouzada 2012] poderá ser composto de tendência, ciclicidade, sazonalidade, e termo aleatório (flutuação). Com base no comportamento da série, define-se o modelo de suavização que será utilizado e, caso os dados não sejam estacionários, faz-se a normalização dos mesmos [Hyndman e Athanasopoulos 2018]. Em seguida, definem-se os parâmetros dos modelos.

As previsões pelos modelos SE, AR, e ARIMA foram realizadas com o auxílio do *software* Minitab®, enquanto que as previsões pelos modelos MLP e ELM foram feitas utilizando Python®. Os computadores utilizados são munidos de processador Core I5 7ª geração e 8 GB de RAM. Ao final, os resultados das previsões pelos modelos propostos foram comparados pelo Erro Quadrático Médio (MSE) [Silva et al. 2018a] e estão descritos na próxima seção.

## 4. Resultados e Discussão

Nesta seção serão apresentados os resultados obtidos com a análise e tratamento dos dados históricos dos preços praticados do café das qualidades Arábica e Robusta. Posteriormente serão apresentadas as previsões utilizando os métodos propostos.

### 4.1 Análise e tratamento dos dados

Inicialmente, foi identificada a regularidade nos dados e realizou-se a plotagem dos mesmos em um gráfico, com o objetivo de identificar as componentes dos preços mensalmente praticados para os cafés Arábica e Robusta (Figura 1).

Com base na definição de componentes temporais dadas por [Bouzada 2012] verificou-se que as duas séries possuem tendência de crescimento ao longo do tempo, ciclos de variação nos preços praticados, e ausência de sazonalidade. Com isso, aplicou-se o método da diferenciação as séries para torná-las estacionárias. Tendo os dados diferenciados, foi construído o gráfico (Figura 2) para análise da função de autocorrelação parcial (FACP) [SIQUEIRA et al. 2018].

É possível perceber que a série de preços do café Arábica possui coeficiente autocorrelação significativo até o segundo atraso, enquanto que a série do café Robusta apresenta até o quarto.

Na realização das previsões pelos modelos SE, AR, ARIMA e ELM foram separados os dados referentes aos anos de 2017 e 2018 para teste, o restante utilizou-se para treinamento (ajuste dos métodos). Para a MLP, as séries até o ano de 2015 foram utilizadas para treinamento, o período de 2015 a 2016 usou-se para validação e os dados de 2017 e 2018 foram utilizados para testar o modelo.

A escolha dos modelos de SE, dependem apenas da identificação do padrão da Figura 1. Pela análise realizada, as séries temporais apresentaram apenas comportamento cíclico e com tendência, sem sazonalidade, requisitos para utilização SE dupla de Holt [MENEGHINI et al. 2018]. Esse modelo tem como parâmetros as constantes de suavização de ciclo e tendência, sendo respectivamente, $\alpha$ e $\beta$. Não há um método absoluto para a definição desses valores [MONTGOMERY et al. 1990]. Portanto, realizou-se testes paramétricos empíricos até encontrar os valores que apresentaram os melhores resultados para as previsões. Os melhores valores encontrados foram para a série temporal café arábica $\alpha$=0,9 e $\beta$=0,3 e para o café Robusta $\alpha$=0,9 $\beta$=0,6.

Segundo [Montgomery et al. 1990] valores grandes para a constante $\alpha$, fazem com que a reação do modelo seja rápida diante as mudanças. Como as duas séries são bastante variáveis ao longo dos meses (Figura 1), isso pode explicar por que o valor de $\alpha$ que apresentou os melhores resultados se aproximou de 1.

O modelo AR teve seus coeficientes estimados a partir da FACP. Para o Arábica, considerou-se de 1 a 2 atrasos e para o Robusta até 4. Em ambos os casos, o modelo AR(1) foi o que chegou aos menores erros. Observa-se que a série foi diferenciada nos termos do usual para modelos ARIMA.

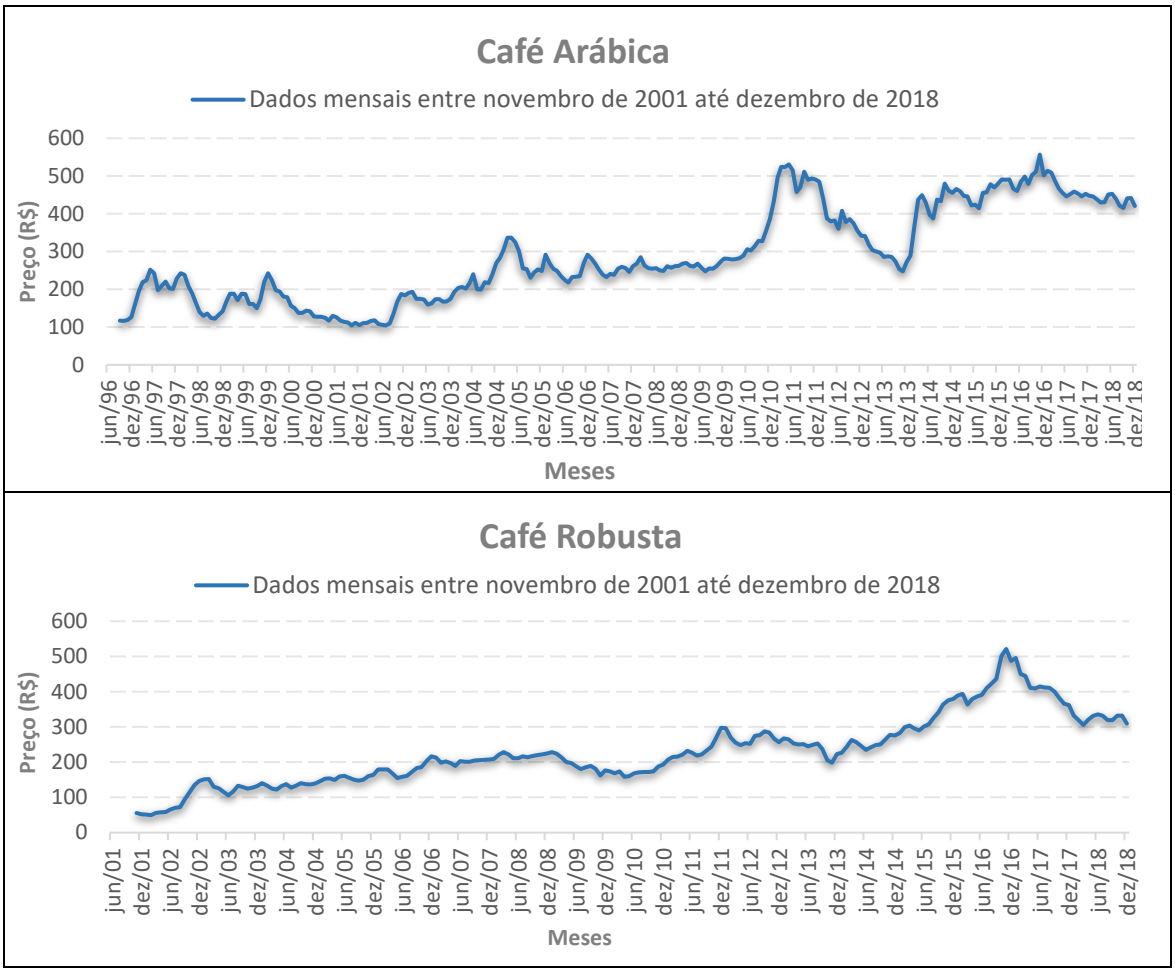



Figura 1 – Dados de preços praticados para o café Arábica e Robusta referente ao período de setembro de 1996 até dezembro de 2018 e novembro de 2001 até 2018, respectivamente.

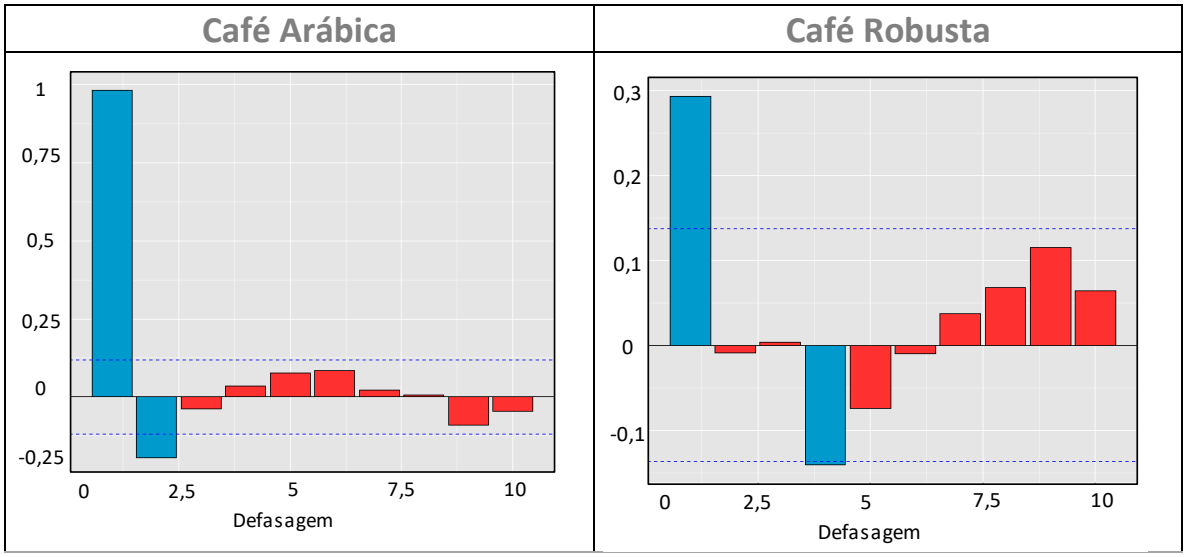



Figura 2 – Gráfico da função de autocorrelação parcial (FACP) para as séries temporais diferenciadas de preços dos cafés arábica e robusta.
Fonte: Dados da pesquisa (2019)

Em relação ao último, como as séries necessitaram de apenas uma diferenciação para torná-las estacionárias, o parâmetro *d* foi definido como unitário [OSHODI et al. 2017]. Já os parâmetros *p* e *q* foram escolhidos utilizando os mesmos procedimentos do AR, mas, neste caso, os coeficientes MA foram definidos como até 4 entradas. Ao final, escolheu-se ARIMA (1,1,1) para café Arábica e ARIMA (3,1,1) para café Robusta. Note que o estimador de máxima verossimilhança foi utilizado para ajuste.

Para a previsão dos dados utilizando ambas as redes neurais, foi variada empiricamente a quantidade de neurônios na camada intermediaria começando com 3 neurônios, com incremento de 5, até 50 neurônios. Para cada variação, as redes foram executadas trinta vezes. Para o café Arábica as redes utilizaram os atrasos 1 e 2, enquanto para a serie Robusta foram utilizados de 1 a 4 atrasos.

Com os parâmetros dos modelos definidos, fez-se a previsão de preços praticados das séries de café para 1 passo à frente, referentes aos 24 meses dos anos de 2017 e 2018. A Tabela 1 mostra os resultados computacionais obtidos dos conjuntos de testes pelo Erro Quadrático Médio (MSE). Nela, os valores encontrados pelas redes neurais são referentes a melhor das 30 simulações. Observa-se que o teste de Friedman foi aplicado e os p-valores encontrados foram muito próximos de zero. Isso garante que mudar o preditor leva a resultados estatisticamente distintos.

Tabela 1 – Erro Quadrático Médio (MSE) das previsões

| MSE | SE | AR | ARIMA | ELM | MLP |
|---|---|---|---|---|---|
| **Arábica** | 181,55 | 172,82 | 172,61 | 132,74 | 161,27 |
| **Robusta** | 274,80 | 288,63 | 364,31 | 247,07 | 226,93 |

Fonte: Dados da pesquisa (2019)

Considerando o desempenho geral dos modelos, o ELM obteve melhores resultados para previsão da série de café Arábica, enquanto que o MLP obteve resultados mais precisos em relação ao Robusta.

Como esperado, os modelos lineares (SE, AR e ARIMA), atingiram resultados inferiores de previsão para as duas séries analisadas, se comparado as Redes Neurais. Como as séries eram bastante variáveis, as RNAs por serem técnicas não-lineares e com grande capacidade de mapeamento, possuem maior capacidade para ajustar os dados previstos aos valores reais. Além disso, a capacidade de generalização das mesmas é fator preponderante para seus bons resultados [Silva et al. 2018a].

Entretanto, vale ressaltar que os modelos lineares são simples e rápidos em sua implementação e execução, podendo apresentar resultados equivalentes exigindo menor esforço, se comparado com as RNAs.

A Figura 3 mostra cada previsão feita com os modelos descritos.

## 5. Conclusão

Esse estudo investigou a aplicação de Redes Neurais Perceptron Multicamadas (MLP) e Máquina de Aprendizado Extremo (ELM) comparado aos modelos lineares de Suavização Exponencial (SE), Autorregresssivo e Auto Regressivo Integrado de Médias Móveis (ARIMA), para a previsão dos preços praticados de café Arábica e Robusta.

A MLP e ELM são bastante semelhantes, porém a segunda possui um custo computacional mais baixo, pois sua camada intermediária não é ajustada e seu treinamento é baseado em um método determinístico.

Para todos os modelos considerou a previsão dos últimos 24 meses de ambas as séries, com horizontes de previsão de 1 passos a frente. Os parâmetros de entrada para os modelos foram definidos conforme as caracteristicas das séries e seus atrasos e com base na função de autocorrelação parcial. O número de neurônio das redes foi definido empiricamente.

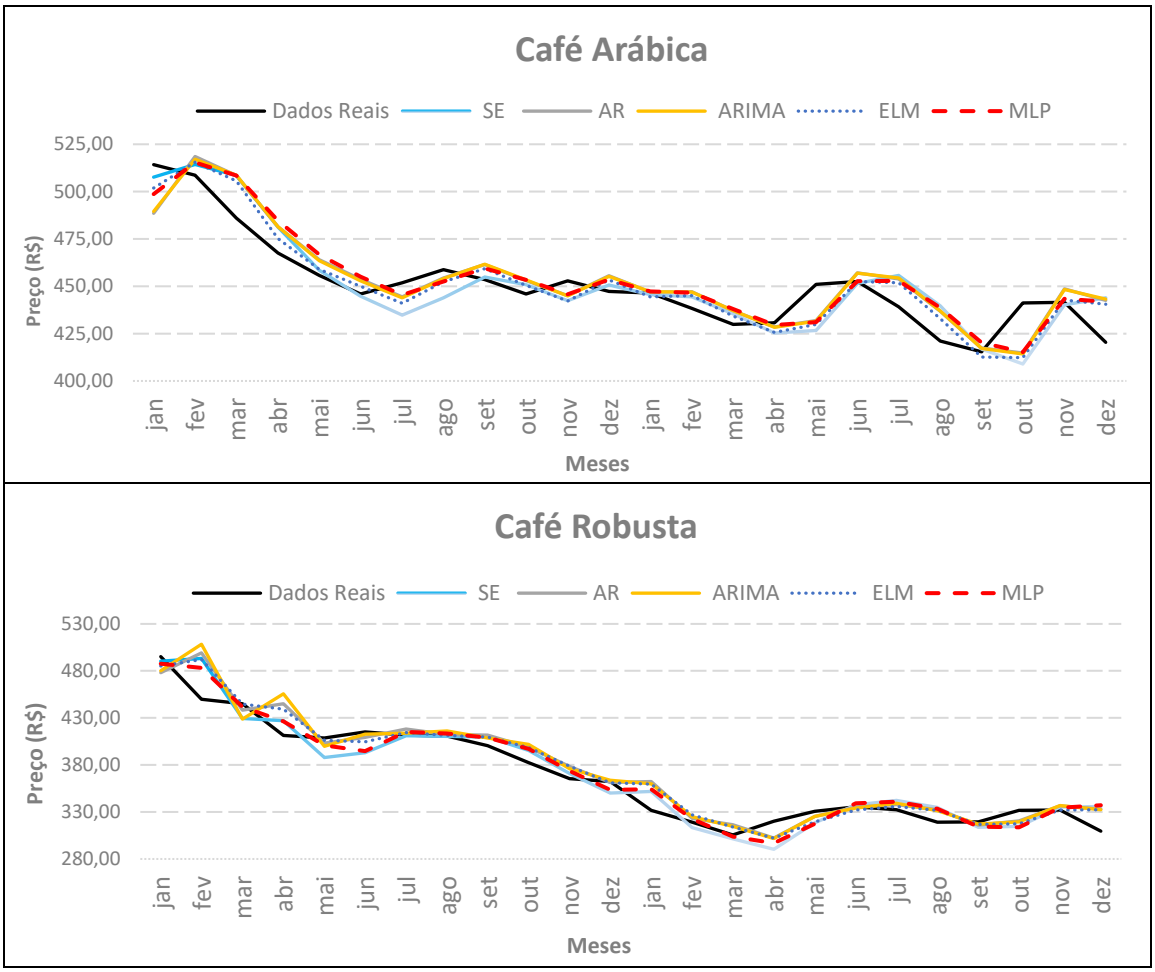



Figura 3 – Gráfico da comparação dos dados reais com os dados previstos para 1 passo a frente referentes aos 24 meses dos anos de 2017 e 2018, utilizando os modelos SE, AR, ARIMA, ELM e MLP.
Fonte: Dados da pesquisa (2019)

Os modelos lineares possuem maior simplicidade em sua aplicação, porém, como os dados analisados eram bastante variáveis, as RNAs apresentaram resultados mais precisos de previsão para ambas as séries de dados analisadas.

Trabalhos futuros poderão ser realizados utilizando outras arquiteturas de redes neurais e diferentes dados de entrada. Ademais, faz-se necessário aplicar outros otimizadores para os modelos ARIMA.

## Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), projeto Universal nº 405580/2018-5, pelo apoio financeiro ao projeto de pesquisa.

## Referências

Ayyanathan, N.; Kannammal, A. (2014). Share price time series forecasting for effective supply chain information exchange. *International Journal of Logistics Systems and Management.* 18: 139-158.

Bouzada, M. A. C. (2012). *Aprendendo Decomposição Clássica: Tutorial Para Um Método De Análise De Séries Temporais. Tecnologias De Administração E Contabilidade*. 2: 1-18.

Box, G. E. P ; Jenkins, G. M.; Reinsel, G. C. (1994). Times series analysis: forecasting and control, 3ª Ed. San Francisco: Holden-Day.

Castro, L. N. (2006). Fundamentals of natural computing: Basic concepts, algorithms and applications. Boca Raton: Chapman& Hall.

CEPEA - Centro de Estudos Avançados em Economia Aplicada (ESALQ/USP). Disponível em: < http://cepea.esalq.usp.br>. Acesso em: 4 fev. 2019.

Cuaresma, C. J.; Hlouskova, J.; Obersteiner, M. (2018). Fundamentals, speculation or macroeconomic conditions? Modelling and forecasting Arabica coffee prices. *European Review of Agricultural Economics*. 45: 583–615.

David, S. A.; Trevisan, L. R.; Lopes, A. M.; Machado, J. A. T.; Inácio Junior, C. M. C. (2017). Dynamics of commodities prices: Integer and fractional models. *Fundamental Informatica*. 151: 389–408.

Fauziah, F. N.; Gunaryati, A. (2017). Comparison forecasting with double exponential smoothing and artificial neural network to predict the price of sugar. *International Journal of Simulation: Systems, Science and Technology*. 18: 1-8.

Fousekis, P. (2017). Price co- movement and the hedger's value-at-risk in the futures markets for coffee. *Agricultural Economics Review*. 18: 35–47.

Haykin, S. (1999). Neural networks: A comprehensive foundation. 2nd ed. Upper Saddle River, NJ: Prentice-Hall.

Huang, G. B.; Zhu Q. Y.; Siew C. K. (2004). Extreme learning machine: A new learning scheme of feedforward neural networks. In Proceedings of *international joint conference on neural networks*, 985–990.

Huang, G. B.; Zhu, Q. Y.; Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing.* 70: 489–501.

Hyndman, R. J.; Athanasopoulos, G. (2018). Forecasting: principles and practice. Melbourne, Australia.

ICO - International Coffee Organization. coffe Brazil. Disponível em: <http://www.ico.org/>. Acesso em: 4 abr. 2019.

Meneghini, M.; Anzanello, M.; Kahmann, A.; Tortorella, G. L. (2018). Ajuste de previsão de demanda quantitativa com base em fatores qualitativos: estudo de caso em um restaurante fast food. *Sistemas & Gestão*. 13: 68-80.

Miche, Y. A.; Sorjamaa, P.; Bas, O.; Simula, C.; Jutten, and Lendasse A. (2010) OP-ELM: Optimally pruned extreme learning machine". *IEEE Transactions on Neural Networks*. 158–162.

Milas, C.; Otero, J.; Panagiotidis, T. (2004). Forecasting the spot spices of various coffee types using linear and non-linear error correction models. *International Journal of Finance and Economics*. 9: 227–288.

Montgomery, D.; Johnson, L.; Gardiner, J. (1990). Forecasting and Time Series Analysis. New York: McGraw-Hill.

Oshodi, O. S.; Ejohwomu, O. A.; Famakin, I. O.; Cortez, P. (2017). Comparing univariate techniques for tender price index forecasting: Box-Jenkins and neural network model. *Construction Economics and Building*. 17: 109–123.

Pradkhan, E. (2017). Financial activity in agricultural futures markets: evidence from quantile regressions. *Australian Journal of Agricultural and Resource Economics.* 61: 610–625.

Ramyar, S.; Kianfar, F. (2019). Forecasting Crude Oil Prices: A Comparison Between Artificial Neural Networks and Vector Autoregressive Models. *Computational Economics*. 53: 743–761.

Silva, C. A. G. (2018). Previsão do preço da commodity café arábica: Uma aplicação da Metodologia. *Espacios*. 39: 1-18.

Silva, N.; Tay, I.; Okida, S.; Stevan Jr., S. L.; Siqueira, H. (2018a). Neural Networks for Predicting Prices of Sugarcane Derivatives. *Sugar Tech*. DOI: 10.1007/s12355-018-0648-5.

Siqueira, H., L.; Boccato, I.; Luna, R.; Attux, and C. Lyra. (2018). Performance analysis of unorganized machines in streamflow forecasting of Brazilian plants. *Applied Soft Computing.* 68: 494–506

Taylor, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*. 19: 715– 725.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science.* 6: 324-342.

Yashavanth, B. S.; Singh, K. N.; Paul, A.; Paul, R. K. (2017). Forecasting prices of coffee seeds using Vector Autoregressive Time Series Model. *Indian Journal of Agricultural Sciences*. 87: 754-758.

Zhang G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 50: 159–175.