

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

DEPARTAMENTO DE ECONOMIA
BACHARELADO EM CIÊNCIAS ECONÔMICAS

LUIZ FELIPE DIAS DE LIMA

ANÁLISE DA PREVISIBILIDADE DO PREÇO SPOT DO MILHO NA
DETERMINAÇÃO DO PREÇO FUTURO: UM ESTUDO UTILIZANDO RANDOM
FOREST.

Recife

2025

LUIZ FELIPE DIAS DE LIMA

ANÁLISE DA PREVISIBILIDADE DO PREÇO SPOT DO MILHO NA
DETERMINAÇÃO DO PREÇO FUTURO: UM ESTUDO UTILIZANDO RANDOM
FOREST.

Monografia apresentada à coordenação do curso de Bacharelado em Ciências Econômicas, sob a orientação do professora **Dr^a. Gisleia Benini Duarte** submetido à Universidade Federal Rural de Pernambuco como parte dos requisitos para a obtenção do título de Bacharel em Ciências Econômicas, de acordo com as exigências.

Recife

2025

Dados Internacionais de Catalogação na Publicação
Sistema Integrado de Bibliotecas da UFRPE
Bibliotecário(a): Auxiliadora Cunha – CRB-4 1134

L732a Lima, Luiz Felipe Dias de.
Análise da previsibilidade do preço spot do milho na
determinação do preço futuro: um estudo utilizando
Random Forest / Luiz Felipe Dias de Lima. – Recife, 2025.
44 f.; il.

Orientador(a): Gisleia Benini Duarte.

Trabalho de Conclusão de Curso (Graduação) –
Universidade Federal Rural de Pernambuco, Bacharelado
em Ciências Econômicas, Recife, BR-PE, 2025.

Inclui referências e apêndice(s).

1. Random Forest. 2. Milho. 3. Milho - Preço . I. Duarte,
Gisleia Benini, orient. II. Título

CDD 330

LUIZ FELIPE DIAS DE LIMA

ANÁLISE DA PREVISIBILIDADE DO PREÇO SPOT DO MILHO NA
DETERMINAÇÃO DO PREÇO FUTURO: UM ESTUDO UTILIZANDO RANDOM
FOREST.

Monografia apresentada à coordenação do curso de Bacharelado em Ciências Econômicas, sob a orientação do professora Dr^a. Gisleia Benini Duarte submetido à Universidade Federal Rural de Pernambuco como parte dos requisitos para a obtenção do título de Bacharel em Ciências Econômicas, de acordo com as exigências.

Aprovado em: 21/07/2025

BANCA EXAMINADORA

Prof^a. Dr^a. GISLEIA BENINI DUARTE (Orientadora)
Universidade Federal Rural de Pernambuco

Prof. Dr. ANDRE DE SOUZA MELO (Examinadora interna)
Universidade Federal Rural de Pernambuco

Prof^a. Dr^a. JANAÍNA DA SILVA ALVES (Examinador externo)
Universidade Federal Rural do Rio Grande do Norte

Dedico este trabalho à minha amada mãe,minha eterna parceira de jornada.

Maria Ângela,por iluminar meus dias, sejam de desafios ou conquistas, com sua energia vibrante e alegria inesgotável

AGRADECIMENTOS

A Jornada até aqui foi marcada por desafios, aprendizados e conquistas que não seriam possíveis sem o apoio de pessoas especiais. Em primeiro lugar agradeço aos meus familiares, especialmente minha mãe Fernanda, minha avó Ângela e minha irmã Ana Clara. Aos meus amigos, em especial Ítalo, Johan, Michell, Erickson e Tiago, e a cada colega que esteve presente em alguma etapa desse processo, oferecendo palavras de encorajamento, conselhos, risadas e, principalmente, companhia nos momentos mais intensos.

Agradeço também a minha orientadora Gisleia Duarte, por toda a orientação, paciência e incentivo ao longo deste trabalho. Sua dedicação foi essencial para que este projeto se concretizasse.

“Se não puder voar, corra. Se não puder correr, ande. Se não puder andar, rasteje, mas continue em frente de qualquer jeito”

Martin Luther King

RESUMO

Este estudo investigou a relevância do preço do contrato futuro de milho como variável preditora do preço spot da commodity para o período de 2018 a 2020 e de 2022 a 2024, com periodicidade diária e assim para as demais variáveis. Para tanto, adotou-se como metodologia o algoritmo Random Forest, considerando como variáveis explicativas a cotação do dólar, o preço futuro da soja e o próprio preço presente (spot) do milho. O principal objetivo foi avaliar se o preço atual do milho constitui um bom predito para o comportamento do mercado futuro. Dessa forma o Random Forest demonstrou alto desempenho na previsão do contrato futuro do milho, indicando boa capacidade de generalização a partir do preço spot, além disso demonstrando que a cotação do dólar é uma variável importante no comportamento do preço futuro do milho.

Palavras-chave: Random Forest; Mercado futuro do milho; preço presente .

ABSTRACT

This study investigated the relevance of the corn spot price as a predictor for its future contract price, using daily data for the periods of 2018-2020 and 2022-2024. The Random Forest algorithm was adopted as the methodology, considering the dollar exchange rate, soybean future prices, and the corn spot price itself as explanatory variables. The main objective was to assess whether the current price of corn is a strong predictor for the behavior of the futures market. The Random Forest model demonstrated high performance in forecasting the corn future contract, indicating a good generalization capacity from the spot price. Additionally, the findings show that the dollar exchange rate is an important variable in the behavior of the future corn price.

Keywords: Random Forest; corn futures market; spot price

LISTA DE ILUSTRAÇÕES

Quadro 1	18
Figura 1 — Fluxo de variáveis para o modelo Random Forest	22
Tabela 1 — Resultados do teste Feature Importance para o milho futuro	26
Tabela 2 — Resultados dos testes de todas as modelagens utilizadas	28
Figura 2 — Resultados do teste Importância de Variáveis para o milho futuro	31
Figura 3 — Serie de Preço Futuro do Milho 2018 a 2024 – excluindo período pandêmico	33
Figura 4 — Serie de Preço à vista do Milho 2018 a 2024 – excluindo período pandêmico	33
Figura 5 — Serie Preço à vista do Milho 2022 a 2024	34
Figura 6 — Serie de Preço à vista do Milho 2018 a 2020	34
Figura 7 — Serie de Preço Futuro da Soja 2018 a 2024 – excluindo período pandêmico	35
Figura 8 — Serie de Preço Futuro da Soja 2022 a 2024	35
Figura 9 — Serie do Preço Futuro da Soja 2018 a 2020	36
Figura 10 — Serie da Cotação do Dólar 2018 a 2024 – excluindo período pandêmico	36
Figura 11 — Serie de Cotação do Dólar 2018 a 2020	37
Figura 12 — Serie de Cotação do Dólar 2022 a 2024	37

LISTA DE ABREVIATURAS E SIGLAS

MAE	Mean Absolut Error
RMSE	Root Mean Squared Error
R2	Coeficiente de Determinação
IBGE	Instituto Brasileiro de Geografia e Estatística
B3	Brasil, Bolsa, Balcão
BM&F	Bolsa de Mercadorias e Futuros
BOVESPA	Bolsa de Valores do Estado de São Paulo
CCM	Contrato Futuro de Milho
CBOT	Chicago Board of Trade
SJC	Contrato Futuro de Soja
CVM	Comissão de Valores Mobiliários

LISTA DE SÍMBOLOS

SUMÁRIO

1	INTRODUÇÃO	12
2	REFERENCIAL TEÓRICO	14
2.1	MERCADO FUTURO E COMMODITIES AGRICOLAS	14
2.2	CONTRATO FUTURO E PRECIFICAÇÃO DO MILHO	15
2.3	PREVISÃO DE COMMODITIES AGRICOLAS	19
3	METODOLOGIA	21
3.1	RANDOM FOREST	21
3.2	BASE DE DADOS E PRÉ-PROCESSAMENTO	23
3.3	MÉTRICAS DE AVALIAÇÃO	24
3.4	IMPORTÂNCIA DAS VÁRIAVEIS (FEATURE IMPORTANCE)	24
4	ANÁLISE DOS RESULTADOS	26
4.1	IMPORTÂNCIA DO PREÇO SPOT DO MILHO	26
4.2	IMPORTÂNCIA DA COTAÇÃO DO DÓLAR	26
4.3	IMPORTÂNCIA DO PREÇO FUTURO DA SOJA	27
4.4	DESEMPENHO DO MODELO RANDOM FOREST	27
5	CONCLUSÃO	31
6	APÊNDICE	33
	REFERÊNCIAS	42

1 INTRODUÇÃO

A Predição de preços é um campo bastante complexo independente do mercado em que estamos se referindo, e assim quando trazemos esse assunto para o mercado de *commodities* agrícolas, tornamos essa tarefa ainda mais complexa, mas é de grande relevância para diversos agentes econômicos, incluindo produtores, consumidores, *traders* e formuladores de políticas públicas. A volatilidade desses preços pode impactar a rentabilidade e a estabilidade de produtos e mercados. Assim, a capacidade de captar tendências futuras nos preços com base no tempo e variáveis correlacionadas torna uma ferramenta fundamental para tomadas de decisões. (Silva, 2024)

Para (Michellin et al., 2017) o crescimento da economia mundial traz consigo um aumento considerável da demanda por proteína animal que por sua vez tem como maior base o milho na ração de todas as criações de animais para consumo. Aproximadamente 80% do milho produzido no país é consumido nessa forma sendo um número bastante elevado quando nos referimos a avicultura e suinocultura, além disso o cereal é matéria prima para a produção de etanol sendo uma fonte de combustível para economia global.

Já segundo Silva et al.(2017), o mercado futuro do milho vem ano a ano ganhando destaque na participação na produção nacional nesse insumo um aumento de quase 10 vezes nessa participação do início do século para primeira década, o que mostra o tanto que esse mercado movimentava a economia brasileira, no ano de 2023, o Brasil produziu cerca de 131 milhões de toneladas de milho (IBGE, 2022), onde 34 milhões de toneladas foram exportadas de janeiro a setembro, um recorde histórico, atualmente o Brasil é um dos 3 maiores produtores mundiais da *commodity* ficando atrás apenas de EUA e China, o produto também se configura entre as primeiras posições no rol de produtos exportados (Ministério da Agricultura e pecuária, 2024).

O referido trabalho tem como objetivo examinar se o preço *spot* (preço atual ou presente) do milho é um importante preditor do preço futuro da *commodity* através da aplicação do modelo de aprendizagem de máquina Random Forest. A motivação para este estudo surge na necessidade de aprimorar técnicas de predição de preço no setor agrícola oferecendo aos agentes de mercado mais uma ferramenta precisa e confiável para tomada de decisões além disso validarmos o peso do preço *spot* como

preditor do preço futuro, a demonstração do uso de aprendizagem de máquinas no problema em questão, vai permitir extrair padrões e relações que podem não ser evidentes em abordagens tradicionais.

A relevância desse estudo, tem como maior intuito a contribuição para o aprimoramento das estratégias de previsão de preços no mercado de commodities agrícolas, que tem como característica, a elevada volatilidade. Penso que ao identificar variáveis com grande poder preditivo para o preço futuro do milho, o trabalho oferece subsídios importantes para agentes do agronegócio, como produtores, tradings e investidores, no planejamento de vendas, definição de contratos e estratégias de hedge. Além disso, a utilização de técnicas de aprendizado de máquina, como o Random Forest, reforça o potencial de aplicação de métodos computacionais na análise de séries econômicas, promovendo maior precisão nas tomadas de decisão e no gerenciamento de riscos de mercado.

A priori, foi realizada uma revisão da literatura sobre o tema de previsão de *commodities*, apresentando os principais estudos, metodologias e resultados que têm sido encontrados na literatura. A primeira seção traz o mercado futuro com o enfoque nas *commodities* agrícolas dentro dos principais estudos sobre economia.

A segunda seção trata sobre os principais pontos da pesquisa: Mercado futuro, contrato futuro e precificação do milho a modelos utilizados na predição de *commodities* no qual fundamenta se a eficiência e casos utilizados, fala também sobre o Random Forest e sua aplicação em séries temporais.

A terceira seção refere-se a metodologia da pesquisa, a descrição geral, discorre sobre o modelo escolhido, a base de dados, e as considerações sobre a amostra, bem como as variáveis escolhidas.

A última seção trabalha com os resultados obtidos através dos testes estáticos estimados na linguagem de programação *Python*, por meio dos ajustes diários informados pela B3 (Brasil, Bolsa, Balcão) e dos preços à vista extraídos da CEPEA, junto com outras variáveis que foram utilizadas como variáveis externas, como SJC(contrato futuro da soja) e dólar. Foi feito um tratamento dos dados, que os tornam em séries temporais e após isso foram transformados em base de treinamento e de teste, após isso os dados foram modelados com e sem as variáveis externas ao modelo, para assim fornecer a comparação e, por fim, a conclusão.

2 REFERENCIAL TEÓRICO

2.1 MERCADO FUTURO E COMMODITIES AGRICOLAS

De acordo com Corrêa (2005), a visão que temos sobre o surgimento da ideia de mercado futuro, nasceu no berço da civilização, na Grécia antiga, as evidências remontam que Tales de Mileto (624 – 546 a.c), filósofo e matemático, foi o primeiro a adotar o conceito de opções e derivativos, na lenda o grego, ainda no inverno se antecipou que no verão teria uma boa colheita de azeitona, assim alugando as prensas para esmagar azeitonas no período de safra e assim lucrando.

Assim para CVM(2022), o mercado futuro pode ser entendido como uma melhoria do mercado a termo onde são negociados contratos com liquidação em datas futuras em que o preço permanece fixo, estipulado em contrato. Dessa forma os investidores podem comprar e vender os contratos de forma de mitigar riscos.

Já (Corrêa, 2017), afirma que na efetivação de uma transação há dois momentos, o de efeito de celebração, onde se firma o negócio e o de execução, quando a mercadoria é repassada., No mercado à vista (*spot*), esses efeitos são realizados em momentos simultâneos, já no mercado a termo esses efeitos acontecem em momentos diferentes, isso sendo notoriamente comum no mercado agrícola, em que os produtores realizam as negociações dos contratos ainda no período do plantio. Nesse caso, negocia-se parte da safra, tendo a possibilidade de riscos para ambas as partes, pois se houver quebra de safra ou se a mesma dobrar de produção há uma relação de ganho ou perda, caso o preço suba ou desça. Portanto, cada parte assume esses riscos, mas o produtor ainda deve efetuar a entrega do produto no preço combinado.

Segundo (Michelin et al., 2012) o mercado de *commodities* agrícolas tem como principais características a alta volatilidade e de ser influenciado por variáveis como oferta, demanda, sazonalidade, clima e políticas, além disso esse mercado movimenta a economia global sendo um pilar crucial para o desenvolvimento de políticas e a economia das nações.

Já para (Michelotti & Siqueira, 2019) aumento da participação de investidores institucionais no mercado futuro de *commodities* tem mudado a dinâmica da formação de preços. Antes utilizados predominantemente como instrumentos de hedge por

produtores e compradores, os contratos futuros tornaram-se ativos financeiros negociados por fundos de investimento, bancos e especuladores. Esse fenômeno é conhecido como financeirização das *commodities* agrícolas, onde os preços são cada vez mais influenciados por fatores financeiros e menos pelos fundamentos da oferta e demanda

No Brasil ainda pontua (Corrêa, 2005), os estudos do mercado de opções se deram na década de 1970, se iniciando na bolsa do Rio, mas sendo implementado primeiramente pela Bovespa na década de 1980. Assim, o mercado de opções foi popularizado com a criação da BM&F, que visava criar um produto para gestão de riscos., Dessa forma, os produtores teriam suas colheitas protegidas por *hedging*, podendo captar crédito a taxas menores, e os investidores teriam uma boa opção de rentabilizar.

Historicamente, a precificação de *commodities* agrícolas esteve ligada a fatores fundamentais como clima, produção, estoques e consumo global. O modelo tradicional sugere que uma quebra de safra, por exemplo, resulta em aumento dos preços devido à menor oferta. Entretanto, estudos indicam que essa relação tem sido cada vez mais alterada por mecanismos financeiros. Com a globalização e a interconectividade dos mercados, os preços das *commodities* passaram a ser influenciados não apenas por condições físicas da produção, mas também por movimentos de capital em mercados de derivativos e oscilações cambiais (Michelotti & Siqueira, 2019).

Estudos mostram que, desde os anos 2000, a correlação entre o volume de contratos futuros e os preços das *commodities* agrícolas tem aumentado, sugerindo uma maior influência da especulação no comportamento dos preços. Esse efeito pode amplificar oscilações naturais do mercado, criando bolhas especulativas que se descolam das condições reais da produção e do consumo (Barbosa, 2022).

2.2 CONTRATO FUTURO E PRECIFICAÇÃO DO MILHO

Mesmo o mercado futuro tenha sido construído com o objetivo de auxiliar na negociação e na proteção contra riscos, seu modelo foi alterado no início do século XXI, com a entrada de importantes instituições bancárias, fundos de investimentos e especuladores, houve um aumento considerável de agentes que não participam da

produção e nem do consumo das commodities, mas que buscam o lucro no balanço do mercado, assim destaca (Michelotti & Siqueira, 2019)

Para Barbosa (2022), as especulações têm dois efeitos: por um lado, podem aumentar a liquidez e melhorar a formação dos preços; por outro, têm o efeito de amplificar a volatilidade e descolar os preços futuros dos principais fundamentos de oferta e demanda. Dessa forma, em momentos de crise, reforça-se ainda mais essa dinâmica, na qual especuladores podem direcionar capital para determinados ativos, provocando um aumento artificial de preços, mesmo sem alterações na oferta ou demanda.

Um mercado eficiente segundo (Fama, 1970), reflete todas as informações que estão disponíveis, mas os contratos futuros tendem a responder a mudanças de oferta e demanda, podendo assim sofrer resultado de movimentos especulativos.

Contratos futuros permitem que os produtores fixem os preços dos insumos antes mesmo de colheitas, criações ou produção, dessa forma podendo se proteger dos impactos da volatilidade no momento da comercialização, isso é o pilar da criação dos mercados futuros, o hedge (mecanismo de proteção), para estabilizar os preços e dar segurança na operação de produtores e industriais (Gomes, 2002).

Em mercados muito especulativos, segundo (Barbosa, 2022) os contratos futuros tendem a ser menos eficientes como um mecanismo de proteção, por conta da alta volatilidade, o que dificulta a tomada de decisão. Embora os derivativos sejam úteis para suavizar as oscilações no curto prazo, eles não conseguem evitar ciclos no longo prazo e os preços tendem a ser influenciados por fatores macroeconômicos.

De acordo com (Michelotti & Siqueira, 2019) a inclusão de contratos de milho cada vez mais em cestas de derivativos criou um vínculo entre mercados agrícolas e financeiros de forma que oscilações de mercados acionários e taxas de juros passam a impactar diretamente o preço do milho futuro, tornando o mercado cada vez mais influenciado por mercados externos e tornando o preço cada vez mais imprevisível.

Ademais, entre as commodities, o milho se destaca por possuir um dos mercados mais sensíveis aos efeitos da financeirização., pois é negociado em números elevados em mercado futuro. A Chicago Board of Trade (CBOT) exerce papel central na precificação do milho, servindo como base para os contratos futuros e influenciando diretamente os preços praticados no mercado físico.

Assim os impactos da financeirização nesse mercado é o aumento da volatilidade dos preços. Investidores quando ingressam nesse mercado, operam sem considerar variáveis produtivas, criando flutuações repentinas nos preços. Dificultando a cadeia produtiva, que enfrentam maior incerteza ao realizar planejamentos de safra e de comercialização da *comodity*.

No Brasil a precificação de contratos futuros está altamente correlacionada a taxa de câmbio, destaca (Barbosa, 2022) a maior parte da safra produzida nacionalmente tem destino para países estrangeiros, logo a moeda local desvalorizada tende a fazer com que o milho nacional fique mais competitivo no mercado internacional, elevando os nossos preços. E ainda políticas de intervenção governamental, como a regulação de estoque e de preços mínimos podem influenciar a formação de preços do mercado futuro, tornando menos previsível do que o mercado americano.

O Brasil é um dos maiores produtores e exportadores de milho, A B3 – Brasil, Bolsa, Balcão¹ trabalha com contratos futuros funcionando similar a bolsa de Chicago (CBOT), uma das maiores operadoras desse mercado no mundo. Porém, ambas possuem suas diferenças, a bolsa brasileira enfrenta problemas relacionados a menor liquidez e maior influência externas como políticas agrícolas nacionais e logísticas dos produtores, problemas que já foram contornados pela consolidada bolsa americana (Michelotti & Siqueira, 2019).

OBS: essa seção como um todo precisa ser muito revisada. Ela está truncada, sem fluidez. Favor corrigir para a versão final.

No Brasil, os contratos futuros de milho são negociados na **B3**, e apresentam as características conforme o quadro 1:

¹B3 é uma das principais empresas de infraestrutura de mercado financeiro no mundo, com atuação em ambiente de bolsa e de balcão. Sociedade de capital aberto – cujas ações (B3SA3) são negociadas no Novo Mercado –, a Companhia integra os índices Ibovespa, IBrX-50, IBrX e Itag, entre outros. Reúne ainda tradição de inovação em produtos e tecnologia e é uma das maiores em valor de mercado, com posição global de destaque no setor de bolsas.” Disponível em: https://www.b3.com.br/pt_br/b3/institucional/quem-somos/. Acesso em: 21 jul. 2025.

Quadro 1

CARACTERÍSTICAS DO PRODUTO	
Objeto de negociação	Milho em grão a granel, com odor e aspectos normais, duro ou semiduro e amarelo.
Código de negociação	CCM
Padrão de qualidade	Milho a granel com 14% de umidade e até 6% de ardidos.
Tamanho do contrato	450 sacas de 60kg.
Meses de vencimento:	janeiro, março, maio, julho, setembro e novembro.
Cotação	Reais por saca, com duas casas decimais.

Fonte: O autor (2025)

Fonte: elaboração própria adequada da B3.

Assim afirma (Gomes, 2002) que os preços do milho podem ser sofrer alterações por diversas formas tanto com a imposição de tarifas e subsídios do governo, como alterações na taxa de câmbio afetando diretamente a curva da oferta. Por outro lado, a demanda crescente do milho para produção de combustível e para alimentação de animais coloca cada vez mais uma complexidade maior na precificação da commodity do milho.

De acordo com (Caldarelli & Bacchi, 2012), estudos indicam que além de fatores tradicionais relacionados a precificação do milho, variáveis como taxa de câmbio e o preço da soja podem exercer uma influência significativa sobre o milho no mercado a termo. Essa relação entre milho e soja pode ser explicada por uma relação de substituição na demanda entre esses contratos ou de complementaridade da oferta. Assim, variações no preço da soja influenciam as expectativas do mercado e, conseqüentemente, impactam o preço do milho.

Outro ponto a se destacar segundo (Gomes, 2002) afirma que o grande número de especuladores no mercado do milho tende a comprometer a predição do preço

spot na determinação do preço futuro. Assim se o preço futuro não só refletir as expectativas dos agentes racionais, mas também as oscilações de capital especulativo, a relação entre essas duas variáveis pode ser menos assertiva. Portanto, o uso de metodologias mais avançadas de predição podem captar padrões complexos do mercado de milho, conforme explanado nessa seção.

OBS: como disse, revisar, corrigir e melhorar toda essa seção!

2.3 PREVISÃO DE COMMODITIES AGRICOLAS

Como afirmam Pinto et al. (2008), a previsibilidade de preços de commodities agrícolas representa um desafio complexo, uma vez que envolve uma ampla gama de variáveis, como fatores climáticos, oferta, demanda, políticas governamentais e oscilações no mercado internacional. A metodologia mais frequentemente utilizada nesse contexto são os modelos estatísticos tradicionais, especialmente o modelo AutoRegressive Integrated Moving Average (ARIMA), que continua sendo amplamente empregado por sua capacidade de capturar padrões temporais e gerar previsões relativamente precisas no curto prazo. No entanto, esse modelo apresenta limitações, entre as quais se destaca a dificuldade de incorporar fatores externos inesperados.

Para diminuir algumas dessas limitações, modelos híbridos são frequentemente utilizados, os quais capturam melhor a volatilidade dos preços das *commodities* agrícolas. Um dos exemplos de tais modelos é a combinação do ARIMA com GARCH (*Generalized AutoRegressive Conditional Heteroskedasticity*), que são muito úteis em mercados com grande flutuação nos preços, como observado no trabalho de Lima et al. (2010).

Com início de estudos em técnicas de aprendizado de máquina, o avanço em metodologias de predição vem sendo cada vez mais uma alternativa eficaz, na previsão de preços de *commodities*. Métodos como *Random Forest* (RF) e *Long Short-Term Memory* (LSTM), são exemplos de modelos promissores no que diz respeito à predição. Além desses, os modelos baseados em Redes Neurais Artificiais (RNA) tem sido cada vez mais estudados e demonstrado ser eficientes por sua capacidade de capturar padrões não lineares e relações complexas entre variáveis preditoras. O LSTM por sua vez, indica poder superar modelos estatísticos tradicionais na predição

de preços de *commodities*, devido à sua habilidade em lidar com dependências temporais de longo prazo (Tavares & Quadrelli, 2022).

Ademais, os modelos ARFIMA (*AutoRegressive Fractionally Integrated Moving Average*) chegam como alternativa de modelos ARIMA, podendo utilizar diferenciação fracionária para capturar dependências de longo prazo em séries temporais robustas para *commodities* agrícolas, especialmente em períodos de alta volatilidade (Lima et al., 2007).

Dessa forma, o modelo a ser utilizado varia conforme a especificidade das séries temporais analisadas e seus respectivos contextos. Com base na literatura, observa-se que essa seleção varia conforme as características dos dados e os objetivos da análise. , Por exemplo, os métodos ARIMA e ARFIMA são utilizados em previsão de séries temporais de curto prazo, sendo adequados para séries estacionárias ou transformadas em estacionárias., Modelos da família GARCH, por sua vez, são utilizados em modelagens, cujas séries apresentam alta volatilidade. Já os modelos de aprendizado de máquinas como *Random Forest* e LSTM, são utilizados no contexto de previsão de longo prazo, sendo especialmente eficazes quando combinados com dados exógenos. A utilização de modelos híbridos, combinando modelos estatísticos e inteligência artificial, tem sido uma alternativa futura para aprimorar a previsão de preços de *commodities* agrícolas.

3 METODOLOGIA

A metodologia deste estudo foi estruturada coletando ajustes diários dos preços spot do milho no Brasil, cotação do dólar e SJC - contratos futuros da soja e do milho negociados na bolsa no período de 2018 a 2020 e 2022 a 2024, excluindo o período em que houve a pandemia que poderia ocasionar outliers na série. A fonte dos dados foi Centro de Estudos Avançados em Economia Aplicada (CEPEA) para os preços spot e a B3 para os contratos futuros através de acesso histórico e diário de ajuste de pregões.

Segundo MultiPy-Project (2025) a análise dos dados foi realizada através do software Python, que vem cada vez mais sendo utilizado em problemas similares, permite a execução de uma ampla gama de testes estatísticos e é amplamente utilizado em pesquisas na área econômica.

3.1 RANDOM FOREST

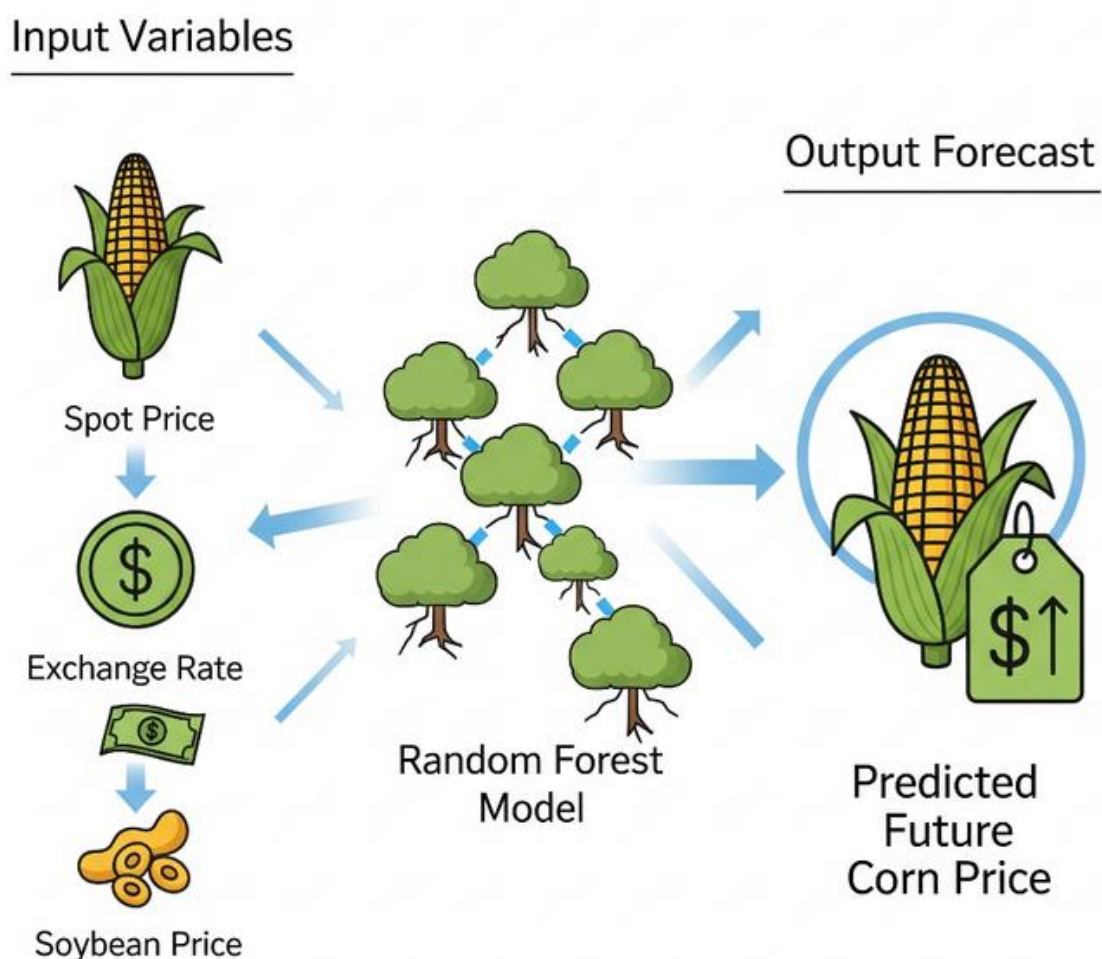
O Random Forest é um modelo de aprendizado de máquina que se baseia nas técnicas de ensemble, onde um conjunto de árvores de decisões simples são combinadas para produzir uma previsão mais precisa e robusta. Essas árvores são constituídas individualmente a partir de subconjuntos aleatórios de dados. Assim, esse é o algoritmo desenvolvido para solucionar problemas de sobreajustes frequentemente observados em árvores de decisão isoladas.

A ideia de combinar modelos distintos para gerar melhores previsões foi formalmente introduzida por (Bates & Granger, 1969), que demonstraram que a combinação de previsões a partir de diferentes abordagens estatísticas poderia resultar em uma redução do erro quadrático médio. Esse resultado, na época contraintuitivo, serviu como base conceitual para a criação de modelos do tipo ensemble. Atualmente, essa abordagem se tornou natural para a comunidade de aprendizado de máquina, sendo o Random Forest um dos exemplos mais bem-sucedidos de aplicação prática desse princípio.

Segundo (Géron, 2019), uma Random Forest é formada a partir do treinamento de várias árvores de decisão sobre subconjuntos aleatórios dos dados, tanto no que se refere às amostras quanto às variáveis. Em vez de cada árvore procurar a melhor

divisão entre todas as variáveis disponíveis, o algoritmo restringe a busca a um subconjunto aleatório dessas variáveis, o que aumenta a diversidade entre as árvores do modelo. Essa aleatoriedade adicional, embora introduza um viés maior em cada árvore individual, reduz a correlação entre elas, o que diminui a variância do modelo final e melhora sua capacidade de generalização.

Figura 1 — Fluxo de variáveis para o modelo Random Forest



Fonte: Elaborado pelo autor (2025), com auxílio da ferramenta Gemini (Google).

Além de oferecer boa performance preditiva, o Random Forest é altamente valorizado por sua capacidade de interpretação. Um dos grandes diferenciais do

algoritmo é sua habilidade de calcular a importância relativa das variáveis explicativas, com base na redução da impureza observada em cada divisão dos nós das árvores. Cada vez que uma variável é usada para dividir um nó, o algoritmo verifica quanto essa divisão contribuiu para reduzir a impureza (no caso da regressão, geralmente medida pelo erro quadrático). Essas contribuições são então ponderadas pela quantidade de amostras que passaram por esse nó, e agregadas ao longo de todas as árvores da floresta. O resultado é um ranking das variáveis mais relevantes para o modelo preditivo.

(Nielsen, 2019) complementa essa análise ao destacar que o Random Forest representa a aplicação prática da “sabedoria da multidão”. Mesmo que nenhuma árvore de decisão isolada seja particularmente eficaz, a combinação de muitas delas — cada uma captando diferentes aspectos da estrutura dos dados — pode gerar uma previsão mais precisa e confiável do que qualquer modelo individual. A força da Random Forest, portanto, está na diversidade de suas árvores, o que permite reduzir o erro geral do modelo.

3.2 BASE DE DADOS E PRÉ-PROCESSAMENTO

Os dados utilizados no estudo foram coletados de fontes confiáveis do mercado de commodities agrícolas, abrangendo informações sobre:

Preço spot do milho (mercado à vista);

Preço futuro do milho (contratos negociados na B3);

Preço futuro da soja (como variável comparativa);

Cotação do dólar (para capturar efeitos do mercado internacional).

Antes da aplicação dos modelos preditivos, foram realizadas etapas de pré-processamento dos dados, linhas com dados incompletos foram removidas utilizando método estatísticos. A base foi dividida em 80% para treinamento e 20% para teste. A escolha da divisão foi baseada no (Duarte & Sobral, 2023). moderado.

A configuração do modelo incluiu:

Número de árvores ($n_{estimators}$): 200; para modelo de regressão, (Breimer, 2001) sugere que não é necessário testar valores para número de árvores desde que esse número seja alto. A rotina Carter do R utiliza 500 como padrão.

Máximo de variáveis por nó (`max_features = mtry`), onde `mtry` varia de 2 a 4, dependendo do conjunto de dados (número de covariáveis dividido por três) `random_state = 11` para garantir a reprodutibilidade dos resultados.

3.3 MÉTRICAS DE AVALIAÇÃO

A performance dos modelos fora verificadas utilizando três métricas principais: RMSE, MAE e R^2 .

3.3.1 Root Mean Squared Error (RMSE)

O Erro Quadrático Médio da Raiz (RMSE) mede a diferença entre os valores previstos e os valores reais. Essa métrica penaliza erros maiores de maneira mais intensa, sendo útil para avaliar a qualidade do modelo. Quanto menor o RMSE, melhor a precisão do modelo.

3.3.2 Mean Absolute Error (MAE)

O Erro Médio Absoluto (MAE) calcula a média dos erros absolutos entre as previsões e os valores reais, fornecendo uma medida da precisão global do modelo. Um MAE baixo indica que, em média, erros da previsão são menores.

3.3.3 Coeficiente de Determinação (R^2)

O R^2 mede o quanto do comportamento dos preços futuros do milho pode ser explicado pelo modelo. Um R^2 próximo de 1 indica que o modelo explica bem a variação.

3.4 IMPORTÂNCIA DAS VÁRIAVEIS (FEATURE IMPORTANCE)

O modelo usado leva em consideração a importância das variáveis preditivas, ou seja, quais variáveis extraídas do conjunto de dados possuem maior impacto na previsão do preço futuro do milho. A estimativa da importância das variáveis é realizada com base na diminuição da impureza dos nós dentro das árvores de decisão que compõem o Random Forest.

Para o contexto das árvores randomizadas, (Breiman, 2001) introduziu o conceito de diminuição ponderada da impureza sobre todas as árvores no modelo. Isso significa que, para cada variável utilizada em uma divisão, o impacto na previsão

é calculado considerando todas as árvores do Random Forest, resultando na seguinte equação:

onde:

representa a ponderação das amostras que atingem o nó t ;

indica a variável utilizada para dividir o nó.

Os valores de importância variam entre 0 e 1, e a soma de todas as importâncias das variáveis no modelo deve ser igual a 1. Essa análise permite identificar quais variáveis possuem maior impacto na previsão dos preços futuros do milho.

4 ANÁLISE DOS RESULTADOS

A análise da predição do preço spot do milho na determinação do seu preço futuro foi feita utilizando o modelo Random Forest, foi usado diferentes combinações de variáveis explicativas. As previsões consideraram inicialmente apenas o preço spot do milho, seguido pela inclusão do preço futuro da soja e da cotação do dólar. Os resultados permitiram uma análise detalhada da importância de cada variável na formação dos preços futuros para o modelo em questão.

Tabela 1 — Resultados do teste Feature Importance para o milho futuro

Modelos	Milho Spot	Soja	Futuro	Dólar
Milho Spot	1,0			
Milho Spot e Soja	0,989745	0,010255		
Milho Spot, Soja e Dólar	0,537179	0,007396		0,455424

Fonte: O autor (2025)

Fonte: Elaboração própria, através do ambiente Jupyter.

4.1 IMPORTÂNCIA DO PREÇO SPOT DO MILHO

Em todas três aplicações testadas, o preço à vista do milho mostrou ser a variável mais significativa para previsão da variável explicada, apresentando uma relevância ao modelo chegando próxima ao 100% (Como mostrado na Tabela 1) nas modelagens realizadas. Desta forma confirmando a hipótese inicial do estudo de que o mercado futuro do milho é bastante influenciado pelos preços do mercado spot, refletindo as informações disponíveis no presente e antecipando tendências futuras.

4.2 IMPORTÂNCIA DA COTAÇÃO DO DÓLAR

A adição da cotação da moeda do dólar no modelo indicou um impacto expressivo na predição do preço futuro do milho, se destacando como a segunda variável mais relevante. Sua importância na previsão ficou em aproximadamente 45%

(Como mostrado na Tabela 1). Demonstrando o peso da taxa de câmbio na precificação de commodities agrícolas.

O resultado é bem coerente com a realidade do mercado, já que o milho é uma das principais commodities exportados pelo país, quando o real está enfrentando uma desvalorização frente ao dólar, os preços tendem a subir tornando as exportações mais competitivas e alavancando o valor dos contratos futuros. Esse efeito corrobora a necessidade de que o câmbio é um fator fundamental ao se analisar a formação de preços no mercado de milho.

4.3 IMPORTÂNCIA DO PREÇO FUTURO DA SOJA

Mesmo que a soja seja uma commodity que se relacionada com milho no agronegócio, a sua inclusão no modelo demonstrou baixa importância, com uma significância inferior a 1% (Como mostrado na Tabela 1). Indicando que apesar de que os mercados de soja e de milho apresentem dinâmicas econômicas semelhantes, o preço do mercado futuro da soja não é um fator relevante para formação do preço futuro do milho.

Essa baixa influencia no preço futuro da soja pode ter relação a diversos fatores, dentre políticas de comercialização especificações, diferenças em ciclos produtivos e padrões diferentes de exportação e de consumo. Esse descobrimento sugere que, para predição de preços futuros do milho, o eixo deve ser mantido no preço à vista da própria commodity e em variáveis macroeconômicas.

4.4 DESEMPENHO DO MODELO RANDOM FOREST

As modelagens feitas possibilitaram que fosse realizada uma avaliação detalhada da predição do preço *spot* do milho para definição do preço futuro. Ademais, para além dos resultados da inserção de variáveis adicionais, para o preço futuro da soja e a cotação do dólar.

A tabela a seguir define valores da MAE, RMSE e R^2 para cada um dos contextos de teste.

Tabela 2 — Resultados dos testes de todas as modelagens utilizadas

	MAE	RMSE	R ²
Milho Spot	R\$ 2,055	R\$ 3,078	0,98068
Milho Spot e Soja	R\$ 1,367	R\$ 2,127	0.99006
Milho Spot, Soja e Dólar	R\$ 1,164	R\$ 1,871	0.99006

Fonte: O autor (2025)

Fonte: Elaboração própria, através do ambiente Jupyter.

A primeira testagem teve como sua única variável preditora o preço do milho spot, assim atingindo um R² de 0,98068, que aponta uma interessante definição de modelo considerando à série de preços futuros. Contudo, é possível que a margem de erro das previsões pode persistir baixa em função do RMSE de R\$ 3,078 e o MAE de R\$ 2,055, reverberando assim no ponto de vista que outras variáveis poderiam complementar a eficiência do modelo.

Desse modo, as conclusões do modelo apontam que a hipótese de que os preços do mercado *spot* são reflexo do mercado futuro do milho são verdadeiras, reafirmando assim a correlação esperada entre dois mercados. Porém, a eficiência do modelo ainda poderia evoluir considerando fatores complementares que influenciam o preço do milho.

Para o segundo teste, adicionamos a variável preço futuro da soja, que trouxe resultados que reduziram os erros RMSE e MAE para R\$ 2,127 e R\$ 1,367, de maneira respectiva. O coeficiente do R² aumentou para 0,99006, indicando um crescimento na capacidade de explicar o modelo.

No terceiro teste foi adicionada a cotação do dólar ao grupo de variáveis preditoras, indicando assim um resultado mais eficiente. O RMSE caiu para R\$ 1,871 e o MAE caiu para R\$ 1,164, mantendo assim o R² em 0,99006. Em outras palavras, o dólar melhorou a previsão consideravelmente.

Os impactos do dólar que tendem a ser variáveis, são justificados pelo fato de o milho brasileiro ser amplamente exportado, tornando suas cotações sensíveis às variações cambiais (Figura 10) e interferem diretamente nos preços do mercado interno. Com isso, os preços do milho tendem a subir em razão da competitividade

das exportações, como notamos nas (Figura 3) e (Figura 4), tornando o preço do dólar fundamental para a modelagem preditiva.

O comparativo entre as três situações testadas indica que a inserção de algumas variáveis adicionais fez melhorar a precisão do modelo, com destaque para a variável do dólar. De modo geral, o modelo que se baseia no *spot* do milho indicou um bom desempenho, mas sua precisão sendo melhorada com a adição de informações macroeconômicas.

Dessa forma, os ganhos obtidos neste estudo seguem em linha com o que é proposto na literatura de Fama (1970), que mercados refletem, em boa parte, as informações disponíveis, o que é evidenciado pela grande explicabilidade do modelo somente com o preço *spot*. Todavia, a inclusão de variáveis externas, como o preço futuro da soja e a cotação do dólar, ocasionou em uma melhoria no modelo, indicando que apesar dos resultados, existe ineficiência relativa e informações complementares não plenamente incorporadas nos preços futuros.

Segundo French (1986), em muitos casos, os preços futuros não são previsores perfeitos dos preços à vista futuros. Há evidências de que os preços futuros incluem prêmios de risco e outros fatores que podem distorcer sua capacidade de prever os preços à vista com precisão. O estudo já corrobora com essa linha ao comprovar que apenas o preço *spot* do milho, de forma isolada já demonstra uma base sólida de previsão, mas sua acurácia se amplifica quando colocamos outras variáveis que afetam a formação de preços no mercado futuro, como a taxa de câmbio, que é essencial, no caso do milho brasileiro, devido a influência das exportações.

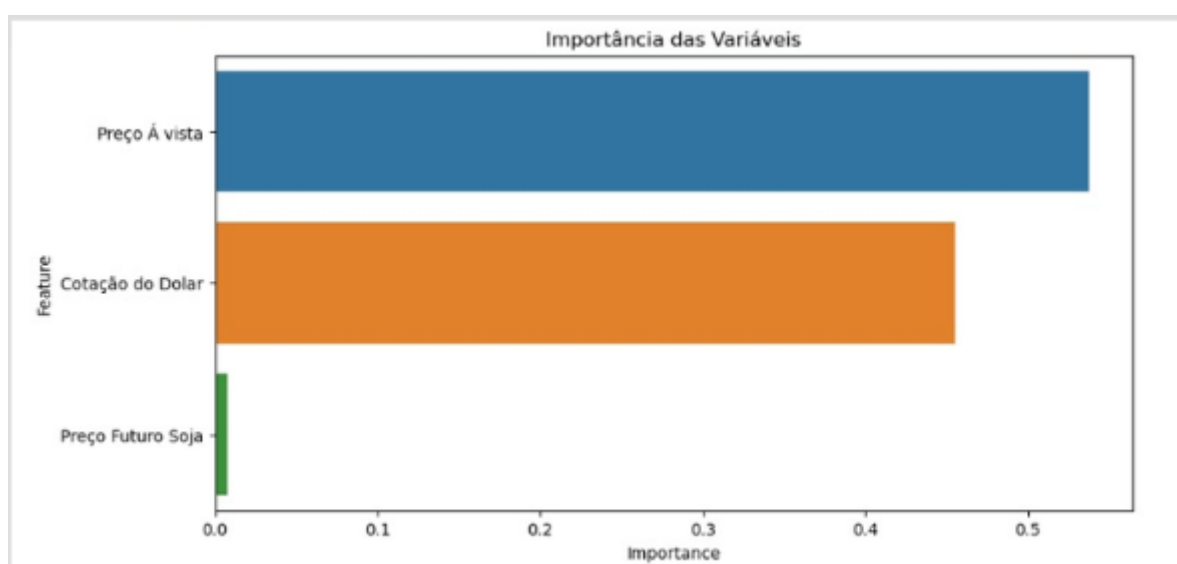
Dessa forma, apesar dos resultados não invalidarem a existência de um mercado eficiente, eles sugerem que há espaço para ganhos preditivos a partir do uso de modelos avançados de aprendizado de máquina, o que corrobora com o fato de que temos imperfeições e assimetrias de informações no mercado futuro do milho. Isso reforça a teoria que a eficiência dos mercados, conforme proposta por Fama (1970), pode não ser absoluta em ambientes complexos como os mercados de *comodities* agrícolas, nos quais variáveis macroeconômicas e sazonais exercem influência significativa e nem sempre perfeitamente precificada pelos contratos futuros.

Tal resultado reforça a relevância da consideração de um número significativo de variáveis para a previsão dos preços futuros do milho, principalmente os fatores que influenciam a dinâmica de demanda e oferta.

5 CONCLUSÃO

O estudo em questão teve como objetivo avaliar a previsibilidade do preço à vista do milho na determinação do seu preço futuro, através do modelo de aprendizagem de máquina *Random Forest*. A avaliação foi conduzida considerando diferentes variáveis explicativas, sendo o preço *spot* do milho, o preço futuro da soja e a cotação do dólar, permitindo analisar a robustez do modelo e a significância de cada variável na formação dos preços futuros.

Figura 2 — Resultados do teste Importância de Variáveis para o milho futuro



Fonte: Elaboração própria, através do ambiente Jupyter.

Os resultados obtidos mostram que o preço *spot* do milho é a variável mais relevante na predição do seu preço futuro (Figura 2), corroborando a hipótese inicial do estudo em questão. Em todos os modelos testados, o preço *spot* apresentou a maior significância preditiva, indicando que o mercado futuro reflete, boa parte, as movimentações do mercado a vista. Essa descoberta reforça a lógica em que preços futuros tendem a incorporar informações disponíveis no presente, tornando um indicativo das expectativas do mercado.

Da mesma forma, a inclusão da cotação do dólar como variável explicativa mostrou um impacto relevante na predição do preço futuro do milho. Isso pode ser exemplificado pela forte dependência do setor agrícola brasileiro ao mercado internacional, já que boa parte da produção interna é exportada.

Assim, variações no câmbio influenciam diretamente a competitividade do milho brasileiro no mercado externo, afetando as expectativas de preços futuros.

Em contrapartida, a variável preço futuro da soja demonstrou baixa importância preditiva, indicando que, embora as duas *commodities* compartilhem uma relação econômica no setor agrícola, podendo até ser bens substitutos nas cestas dos investidores. O impacto direto da soja na precificação do milho futuro não foi tão expressivo quanto o de outras variáveis, conforme dados dessa pesquisa. Isso indica que, para a predição do preço futuro do milho, o preço *spot* da própria *commodity* é o principal fator a ser considerado, seguido por variáveis macroeconômicas como o câmbio.

A análise das métricas de desempenho do modelo demonstrou altos valores de R^2 e baixos erros médios (MAE e RMSE), confirmando a eficácia do *Random Forest* na previsão dos preços futuros do milho. Isso evidencia que técnicas de aprendizado de máquina são ferramentas úteis para analisar a dinâmica de precificação no mercado de *commodities*, capturando padrões complexos e aprimorando a tomada de decisão para agentes do setor agrícola.

Diante desses achados, este estudo contribui para a discussão sobre previsibilidade de preços de *commodities*, destacando que a modelagem preditiva pode ser uma ferramenta valiosa para produtores, investidores e formuladores de políticas públicas. A identificação do preço *spot* como principal variável explicativa reforça a necessidade de acompanhamento contínuo do mercado à vista para a antecipação de tendências no mercado futuro.

Para estudos futuros, recomenda-se explorar outras técnicas avançadas de modelagem, como XGBoost, LSTM, Redes Neurais e Modelos Híbridos, além de ampliar a base de dados para verificar a estabilidade dos resultados em períodos de maior volatilidade da série (Figura 3). Do mesmo modo, a inclusão de variáveis climáticas, variáveis macroeconômicas, como contrato do petróleo e indicadores de oferta e demanda podem enriquecer a análise e fornecer *insights* adicionais sobre a formação dos preços futuros do milho.

6 APÊNDICE

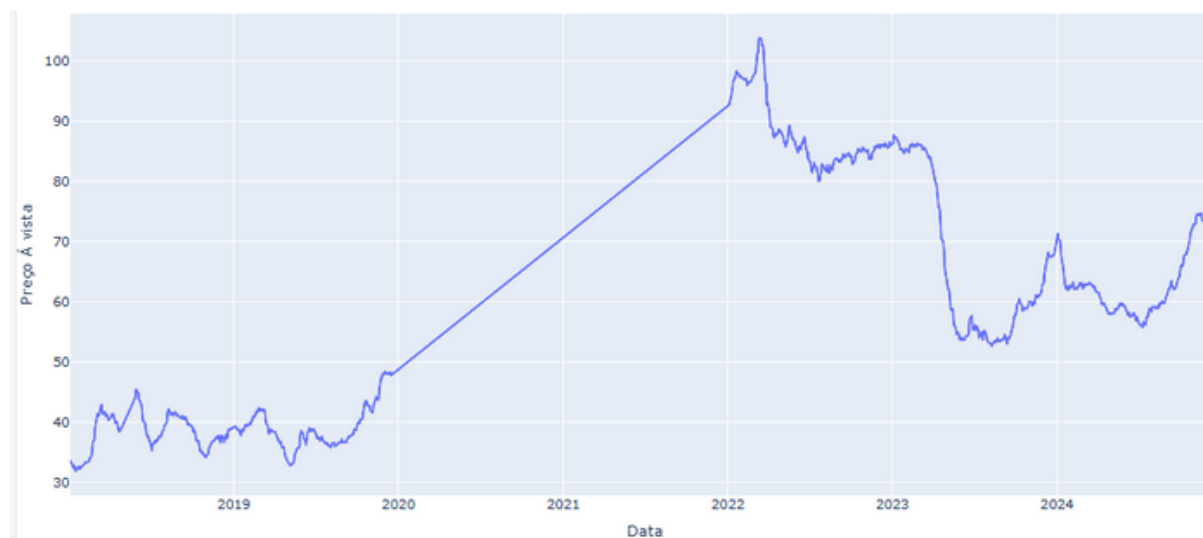
Gráficos que mostram o comportamento da variável explicada e variáveis explicativas

Figura 3 — Serie de Preço Futuro do Milho 2018 a 2024 – excluindo período pandêmico



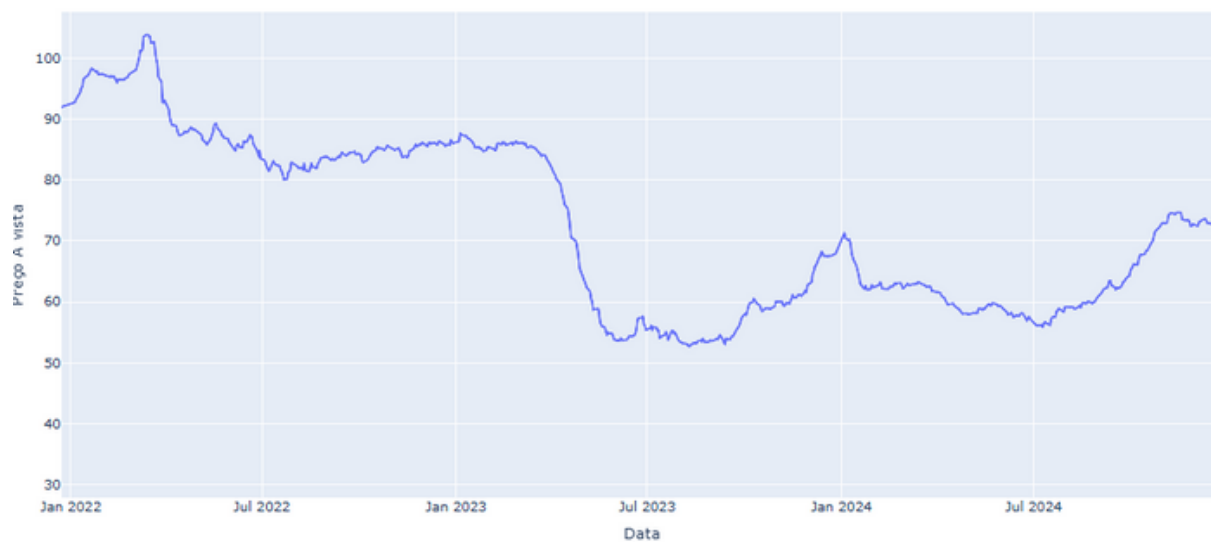
Fonte: O autor (2025)

Figura 4 — Serie de Preço à vista do Milho 2018 a 2024 – excluindo período pandêmico



Fonte: O autor (2025)

Figura 5 — Serie Preço à vista do Milho 2022 a 2024



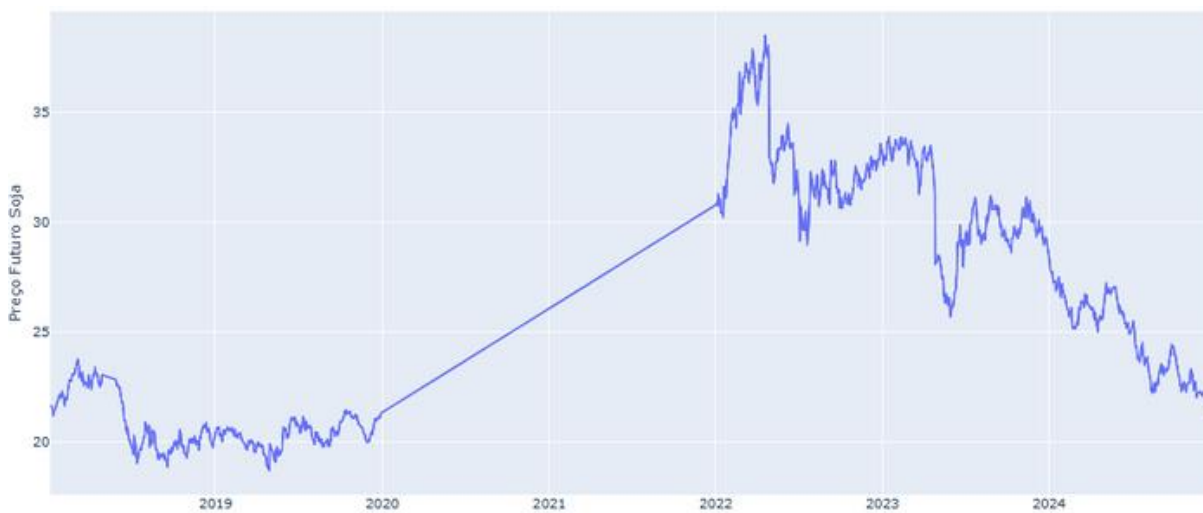
Fonte: O autor (2025)

Figura 6 — Serie de Preço à vista do Milho 2018 a 2020



Fonte: O autor (2025)

Figura 7 — Serie de Preço Futuro da Soja 2018 a 2024 – excluindo período pandêmico



Fonte: O autor (2025)

Figura 8 — Serie de Preço Futuro da Soja 2022 a 2024



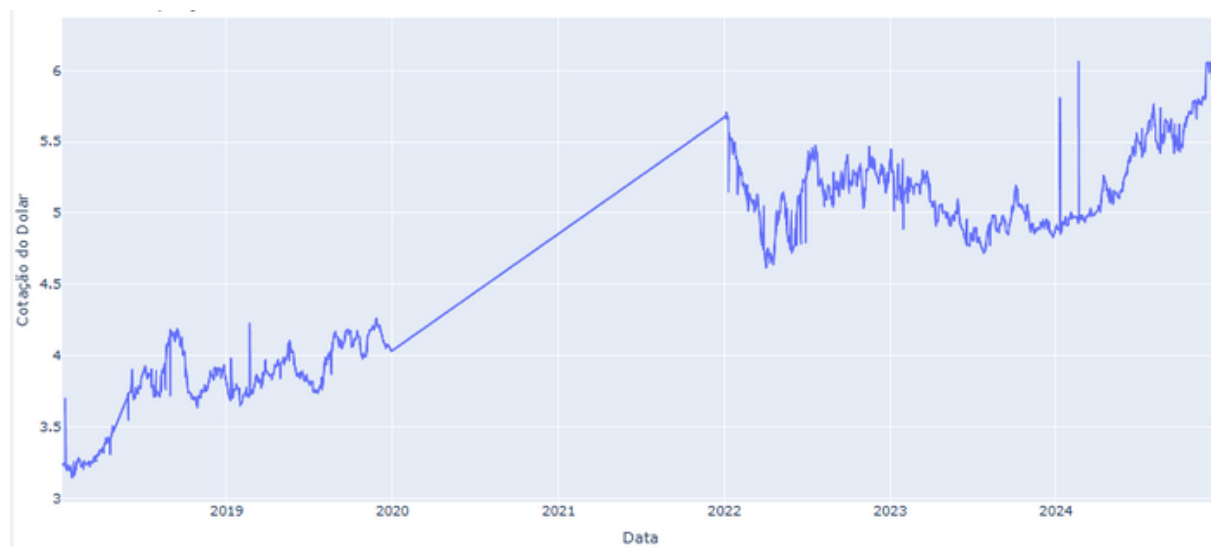
Fonte: O autor (2025)

Figura 9 — Serie do Preço Futuro da Soja 2018 a 2020



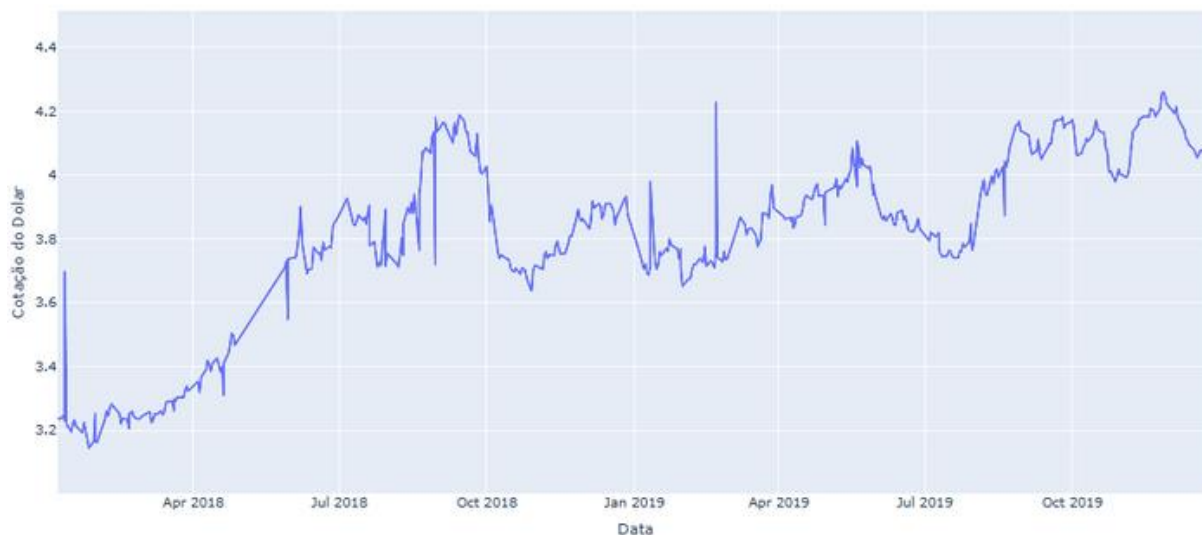
Fonte: O autor (2025)

Figura 10 — Serie da Cotação do Dólar 2018 a 2024 – excluindo período pandêmico



Fonte: O autor (2025)

Figura 11 — Serie de Cotação do Dólar 2018 a 2020



Fonte: O autor (2025)

Figura 12 — Serie de Cotação do Dólar 2022 a 2024



Fonte: O autor (2025)

Script do passo a passo

```

import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_absolute_error, mean_squared_error,
r2_score

import matplotlib.pyplot as plt
import seaborn as sns

# Definição da semente para reprodutibilidade
# Converter todas as colunas para numérico corretamente
# Define vírgula como separador decimal
# Remove colunas indesejadas
# Converte para float
# Remove linhas com valores NaN antes do treinamento
np.random.seed(11)
def carregar_dados(nome_arquivo):
    base = pd.read_csv(nome_arquivo, decimal=",")
    base.drop(columns=['Data', 'Unnamed: 0'], inplace=True, errors='ignore')

    for col in base.columns:
        base[col] = base[col].astype(str).str.replace(",", ".").astype(float)

    base.dropna(inplace=True)
    return base

# Treinamento do modelo
# Cálculo das métricas
# Importância das variáveis
def treinar_random_forest(base, mtry):
    X = base.drop(columns=['Preço Futuro'])
    y = base['Preço Futuro']

```

```

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=11)
rf = RandomForestRegressor(n_estimators=200, max_features=mtry,
random_state=11)
rf.fit(X_train, y_train)

y_pred = rf.predict(X_test)

metrics = {
    'MAE': mean_absolute_error(y_test, y_pred),
    'RMSE': np.sqrt(mean_squared_error(y_test, y_pred)),
    'R²': r2_score(y_test, y_pred)
}

feature_importance = pd.DataFrame({'Feature': X.columns, 'Importance':
rf.feature_importances_})
feature_importance = feature_importance.sort_values(by='Importance',
ascending=False)

return rf, metrics, feature_importance, X_test, y_test, y_pred
# Gráfico de Importância das Variáveis
def plotar_importancia(feature_importance):
    plt.figure(figsize=(10, 5))
    sns.barplot(x=feature_importance['Importance'],
y=feature_importance['Feature'])
    plt.title('Importância das Variáveis')
    plt.show()
# Lista de datasets para processar
datasets = [
    ("CCM - Milho.csv", 2),
    ("CCM - Milho e SJC - Soja.csv", 3),

```

```

("CCM - Milho, SJC - Soja e dolar.csv", 4)
]
# Loop para treinar o modelo em cada dataset
for dataset, mtry in datasets:
    print(f"\nProcessando {dataset}...")
    base = carregar_dados(dataset)
    rf, metrics, feature_importance, X_test, y_test, y_pred =
treinar_random_forest(base, mtry)

    print("Colunas do treinamento:", list(rf.feature_names_in_))
    print("Colunas do X_test:", list(X_test.columns))
    print("Métricas de desempenho:", metrics)
    print("Importância das Variáveis:\n", feature_importance)

    plotar_importancia(feature_importance)

# Criar DataFrame com resultados reais e previstos
# Simulação de datas
# Gráfico interativo comparando real e previsão
df_resultado = pd.DataFrame({
    "Data": pd.date_range(start="2024-01-01", periods=len(y_test)),
    "Preço Á vista": X_test.iloc[:, 0].values,
    "Preço Futuro Real": y_test.values,
    "Preço Futuro Previsto": y_pred
})

import plotly.express as px
fig = px.line(df_resultado,
              x="Data",
              y=["Preço Futuro Real", "Preço Futuro Previsto"],
              title=f"Previsão X Real - {dataset}",
              labels={"value": "Preço", "variable": "Legenda"})

```

fig.show()

REFERÊNCIAS

- . In: NIELSEN, Aileen. **Análise prática de séries temporais: predição com estatística e aprendizado de máquina**. Rio de Janeiro: Alta Books, 2021. 252–265 p.
- B3. **Brasil, Bolsa, Balcão**. Disponível em: <http://www.b3.com.br/>. Acesso em: 12 nov. 2024.
- BARBOSA , Francisco Benedito da Costa. FORMAÇÃO DOS PREÇOS AGRÍCOLAS: FUNDAMENTOS OU ATIVIDADE ESPECULATIVA? . **IPADES**.
- BATES, John M; GRANGER, Clive W. J. The combination of forecasts. **Operational Research Quarterly**, v. 20, n. 4. 451–468 p, 1969.
- BREIMAN, Leo. Random forests. **Machine Learning**, v. 45, n. 1. 5–32 p, 2001. DOI: 10.1023/A:1010933404324.
- CALDARELLI, Carlos Eduardo; BACCHI, Mírian Rumenos Piedade. Fatores de influência no preço do milho no Brasil. **Texto para Discussão**, Brasília, DF, n. 39, 2010.
- CORRÊA, Arnaldo Luiz. **Derivativos Agrícolas**. 3 ed. Editora Comunnicar, v. 3, f. 199, 2017. 397 p. cap. 2.
- DUARTE, Gisleia Benini; SOBRAL, Marcos Felipe Falcão. CONSEQUENCES OF VIOLENCE ON MATHEMATICS LEARNING IN BRAZILIAN PUBLIC SCHOOL STUDENTS. **Manuscrito não publicado**..
- FAMA, Eugene F. Efficient Capital Markets: A Review of Theory and Empirical Work. **The Journal of Finance**, Nova York, v. 25, n. 2. 383–417 p, 1970.
- FRENCH, Kenneth R.. Detecting spot price forecasts in futures prices. **The Journal of Business**, Chicago, v. 59, n. 2, pt. 2. S39–S54 p, 1986.
- GOMES, Marcos Faria. . **Formação de Preços de Commodities no Brasil**. São Paulo, 2002. 52 p Dissertação (Mestrado em Economia de Empresas) - Fundação Getúlio Vargas, São Paulo, 2002.
- GÉRON, Aurélien. **Mãos À Obra**: Aprendizado de Máquina com Scikit-Learn E TensorFlow. Rio de Janeiro: Alta Books, v. 2, f. 288, 2018. 576 p.

LIMA, Fabiano Guasti *et al.* Previsão de preços de commodities com modelos ARIMA-GARCH e redes neurais com ondaletas: Velhas tecnologias – novos resultados. **Revista de Administração (R.Adm.)**, São Paulo, v. 45, n. 2. 188–202 p, 2010. abr./maio/jun..

LIMA, Ricardo Chaves; GÓIS, Marcos Roberto; ULISES, Charles. Previsão de preços futuros de commodities agrícolas com diferenciações inteira e fracionária, e erros heteroscedásticos. **Revista de Economia e Relações Internacionais (RER)**, Rio de Janeiro, v. 45, n. 3. 621–644 p, 2007. Impressa em agosto de 2007.

MICHELIN , Fernanda Piccinin; SILVA , Fabiano Mello da; RUPPENTHAL , Janis Elisa. ANÁLISE DA EFICIÊNCIA DO MERCADO FUTURO DE MILHO NO BRASIL. **Ciências Sociais Aplicadas em Revista**, v. 12 , n. 22, p. 89-105, 2012 .

MICHELOTTI , Fernando; SIQUEIRA , Hipólita . FINANCEIRIZAÇÃO DAS COMMODITIES AGRÍCOLAS E ECONOMIA DO AGRONEGÓCIO NO BRASIL. **Semestre Económico** , v. 22 , n. 50 , p. 87-106, 2019 .

MINISTÉRIO da Agricultura e Pecuária. Secretaria de Comércio e Relações Internacionais.. Brasília, 2024. 16 p. Exportações Brasileiras Milho.

MULTIPY-PROJECT. PyEcon. Disponível em: <https://pyecon.org/>. Acesso em: 8 abr. 2025.

PINTO, Pablo Aurélio Lacerda de Almeida *et al.* APLICAÇÃO DO MODELO ARIMA À PREVISÃO DO PREÇO DAS COMMODITIES AGRÍCOLAS BRASILEIRAS. **Congresso da Sociedade Brasileira de Economia, Administração e Sociologia Rural, 46**, Rio Branco - Acre, 2008.

SILVA, Felipe Morelli da *et al.* LIQUIDEZ DOS CONTRATOS FUTUROS DE MILHO NEGOCIADOS NA BM&FBOVESPA. **BARA Revista de Administração e Negócios da Amazônia**, v. 9, n. 1, p. 26-44, 2017.

SILVA, Marcos V. N. **Análise da eficiência do mercado futuro e spot de milho no contexto regional dos anos de 2018 a 2022: Os casos de São Paulo, Pernambuco e Bahia.** Universidade Federal Rural de Pernambuco. Recife, 2024 Trabalho de Conclusão de Curso, Recife, 2024.

TAVARES, Ubiratan da Silva; QUADRELLI, Giovane. ANÁLISE COMPARATIVA ENTRE OS MODELOS ARIMA E LSTM NA PREVISÃO DE CURTO PRAZO DA DEMANDA DE POTÊNCIA ATIVA. **Revista de Engenharia e Tecnologia**, v. 14, n. 1, p. 60-70, 2022.