



# Previsão do Preço do Arroz: Um Estudo com Modelos Econométricos e Variáveis Exógenas

Lucas Valle Mielke<sup>1</sup>  
ICMC-USP  
Paulino Ribeiro Villas-Boas<sup>2</sup>  
Embrapa Instrumentação  
ICMC-USP

## 1 Introdução

O arroz (*Oryza sativa* L.) é um dos cereais mais cultivados e consumidos no mundo, desempenhando um papel crucial na alimentação de aproximadamente 2,5 bilhões de pessoas. Em 2022, a produção global de arroz atingiu cerca de 776 milhões de toneladas, com o Brasil contribuindo com aproximadamente 11 milhões de toneladas, posicionando-se como um dos principais produtores fora da Ásia [1].

Como ocorre com toda commodity agrícola, o preço do arroz é influenciado por diversos fatores, incluindo a dinâmica de oferta e demanda, as condições climáticas e as políticas governamentais. Essa complexa interação resulta em volatilidade que impacta tanto produtores como consumidores, especialmente as famílias de baixa renda, mais vulneráveis a flutuações nos preços dos alimentos. A volatilidade dos preços é particularmente agravada ao considerarmos o ciclo de produção do arroz, que abrange até 5 meses entre plantio e colheita, expondo produtores a incertezas prolongadas e torna as decisões de plantio, tomadas meses antes da colheita, extremamente desafiadoras.

Diante desse cenário, este trabalho propõe desenvolver um modelo preditivo baseado em técnicas de aprendizado de máquina para antecipar o preço do arroz com antecedência de cinco meses. A proposta utiliza dados históricos e considera variáveis relacionadas à oferta, como os preços de fertilizantes e de commodities concorrentes por terra cultivável, além de variáveis de demanda, como a renda média da população. Essa abordagem visa fornecer percepções que possam auxiliar produtores e formuladores de políticas na tomada de decisões mais informadas, contribuindo assim para melhor estabilidade do mercado de arroz.

---

<sup>1</sup>lucas.mielke@alumni.usp.br

<sup>2</sup>paulino.villas-boas@embrapa.br

## 2 Metodologia

A etapa inicial consistiu na identificação das variáveis com impacto econômico direto no preço do arroz, refletindo as forças de oferta e demanda. A literatura aponta que fatores como a taxa de câmbio [2], os custos de insumos [3] e a competição com outras commodities [3] são determinantes importantes para a oferta do arroz, enquanto a demanda tende a ter alguma relação com a renda média da população [4]. Com base nessas evidências, selecionamos variáveis representativas de ambos os lados: o custo médio de fertilizantes, os preços médios da soja e do milho, e a taxa de câmbio como indicadores da oferta, e a renda média da população como indicador da demanda.

A variável dependente deste estudo é o preço à vista da saca de 50 kg de arroz no Rio Grande do Sul obtido do Centro de Estudos Avançados em Economia Aplicada (CEPEA/USP), que disponibiliza informações sobre preços em reais e dólares. As variáveis independentes dos preços de soja e milho, e taxa de câmbio também foram provenientes do CEPEA. Para representar os custos de insumos, utilizamos o preço médio mensal da tonelada de fertilizantes NPK, disponível na base de dados da Companhia Nacional de Abastecimento (CONAB). A demanda é representada pela renda média mensal da população, obtida por meio da Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE). A série temporal gerada abrange o período de junho de 2012 a maio de 2024, com os dados ajustados para frequência mensal e deflacionados pelo Índice de Preços ao Consumidor Amplo (IPCA). Além disso, aplicamos diferenciação nas variáveis para analisar as variações mensais.

Para prever o preço do arroz com cinco meses de antecedência, adotamos um procedimento de defasagem direta, que consiste em aplicar um deslocamento temporal nas variáveis independentes do período  $t$  para prever a variável dependente no período  $t+5$  permitindo realizar previsões diretas com base nos dados observados.

Com a base de dados devidamente tratada e ajustada, aplicamos os modelos econométricos de Regressão Linear, Regressão de Cume e o modelo de séries temporais SARIMAX a cada conjunto de dados. Adicionalmente, utilizamos um modelo SARIMA, sem defasagem e sem inclusão de variáveis exógenas, como modelo de controle, para comparar a eficácia da adição de variáveis exógenas e do procedimento de defasagem. O conjunto de dados foi dividido em nove períodos de oito anos, dos quais os primeiros sete anos e sete meses foram destinados ao treinamento dos modelos, enquanto os cinco meses finais foram reservados para avaliar a capacidade preditiva.

## 3 Resultados

A análise exploratória revelou uma tendência geral de crescimento nos preços do arroz, milho e soja ao longo do período estudado. No entanto, essa tendência foi marcada por oscilações significativas, especialmente na segunda metade de 2020, quando a crise da COVID-19 impactou as cadeias de suprimento global. Paralelamente, os preços dos fertilizantes também aumentaram, impulsionados pelo conflito entre Rússia e Ucrânia, que afetou a oferta desses insumos. Em contraste, a renda da população apresentou uma trajetória constante de desvalorização, encerrando o período em pouco mais de 50% do seu valor inicial. A Figura 1 ilustra essas movimentações.

Análises de estacionariedade revelaram que, inicialmente, nenhuma das séries era estacionária; no entanto, todas se tornaram estacionárias após a diferenciação, contribuindo com a eficiência dos

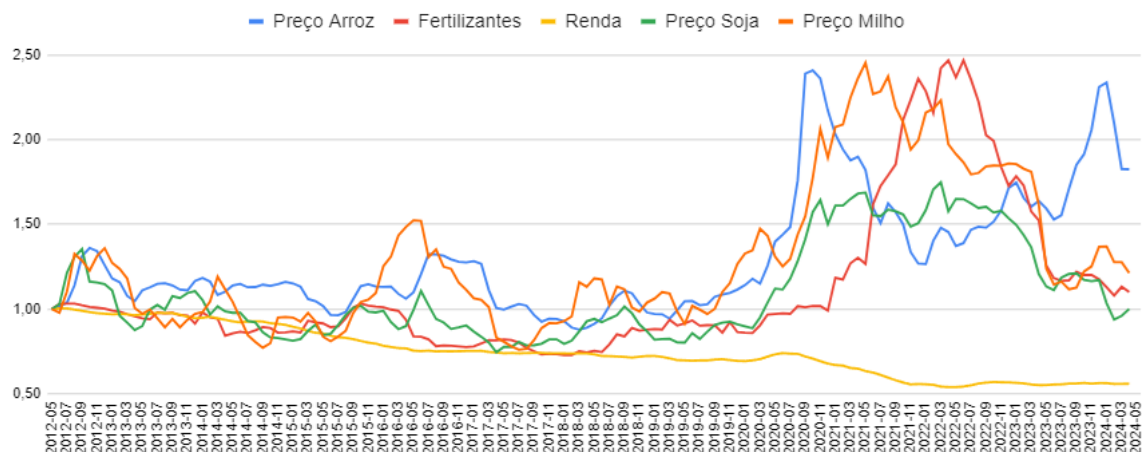


Figura 1: Evolução dos Preços do Arroz, Milho, Soja e Fertilizantes, e da Renda da População (2012-2024)

modelos de regressão. Também foi avaliada a correlação de todas as variáveis atestando ausência de multicolinearidade, destacando uma correlação praticamente nula entre a renda e o preço do arroz. Este resultado reforça a relação inelástica mencionada na literatura [4].

Por fim, a Tabela 1 apresenta a matriz de Erro Quadrático Médio (RMSE) dos três modelos testados com a base de dados defasada, além do modelo de controle SARIMA. Os três primeiros modelos atingiram um RMSE em torno de 2,9, com leve superioridade para o SARIMAX, ajustado para SARIMAX(0,0,1)(0,0,0)[12], que considera apenas um atraso e uma sazonalidade anual de 12 meses. Observou-se uma variação significativa no desempenho dos modelos em três períodos, entre 2013 e 2020, além dos períodos mais recentes, sugerindo a influência de choques macroeconômicos nas previsões. O modelo de controle SARIMA apresentou erros significativamente mais altos na maioria dos períodos, destacando a vantagem da inclusão de variáveis exógenas e dos ajustes de defasagem feitos na base de dados, o que reforça a eficácia das abordagens econométricas aplicadas neste estudo.

Tabela 1: Resultados dos Modelos de Previsão (2012-2024)

Modelo	SARIMAX	Ridge Regression	Linear Regressão	Controle
2012-06 a 2020-05	1,84	1,90	1,92	0,88
2013-01 a 2020-12	7,72	7,91	7,95	6,66
2013-06 a 2021-05	1,42	1,45	1,44	2,06
2014-01 a 2021-12	2,03	1,73	1,85	5,02
2014-06 a 2022-05	1,95	2,99	2,88	2,91
2015-01 a 2022-12	1,62	1,03	1,13	6,02
2015-06 a 2023-05	1,42	1,70	1,73	6,13
2016-01 a 2023-12	3,57	3,08	3,09	3,15
2016-06 a 2024-05	4,10	4,21	4,20	7,39
Total	2,85	2,89	2,91	4,47

## 4 Conclusões

Os resultados deste estudo confirmam a viabilidade de modelos econométricos para prever o preço do arroz com cinco meses de antecedência, demonstrando a eficácia da incorporação de variáveis exógenas relacionadas à oferta e demanda, bem como o uso do procedimento de defasagem na base de dados. A análise exploratória revelou que, após a deflação pelo IPCA, os preços agrícolas aumentaram a taxas superiores à inflação, enquanto a renda da população declinou, o que reflete um agravamento das condições socioeconômicas no Brasil. As variações nos preços do arroz, milho e soja foram diretamente impactadas pela crise da COVID-19, enquanto o aumento dos preços dos fertilizantes foi fortemente influenciado pelo conflito entre Rússia e Ucrânia.

Os modelos econométricos aplicados apresentaram resultados robustos, com a Regressão Linear e a Regressão de Cume mostrando consistência, enquanto o modelo SARIMAX se destacou por sua maior precisão preditiva. Contudo, erros de previsão aumentaram em períodos específicos, sugerindo que choques macroeconômicos afetam a acurácia dos modelos. Isso destaca a importância de adaptações contínuas na modelagem e na incorporação de fatores adicionais no futuro.

Este trabalho contribui para a literatura ao destacar a importância de incorporar variáveis exógenas e ao utilizar métodos menos explorados na previsão de preços agrícolas com antecedência de vários meses, preenchendo uma lacuna nos estudos sobre previsões de commodities a médio e longo prazo. Futuras pesquisas poderiam ampliar essa análise, investigando a lucratividade do produtor rural e integrando uma visão mais abrangente que relacione preços agrícolas, custos de insumos e renda familiar. Isso permitiria uma compreensão mais prática e holística do setor agrícola, fornecendo subsídios importantes para decisões estratégicas.

## Referências

- [1] Food and Agriculture Organization of the United Nations. *Food and agriculture data*. FAO, 2024. Disponível em: <https://www.fao.org/faostat>. Acesso em 10/10/2024.
- [2] Barros, G. S. C., Spolador, H. F. S., Bacchi, M. R. P. *Choques de Oferta e Demanda e o Crescimento da Agricultura Brasileira*. Disponível em: <https://periodicos.fgv.br/rbe/article/view/1030>. Acesso em 09/10/2024.
- [3] Kreter, A. C., Souza Júnior, J. R. C., Teixeira, W. S., Castro, N. R. *Mercados e Preços Agropecuários*. Carta Conjuntura 60. IPEA. Brasília, 2023. Disponível em: [https://www.ipea.gov.br/cartadeconjuntura/wp-content/uploads/2023/09/230905\\_cc\\_60\\_nota\\_20.pdf](https://www.ipea.gov.br/cartadeconjuntura/wp-content/uploads/2023/09/230905_cc_60_nota_20.pdf). Acesso em 09/10/2024.
- [4] Alves, L. R. A., Barros, G. S. C. *Oferta e Demanda de Arroz no Brasil: demanda interna fraca e exportação restrita pressionam área*. CEPEA. Piracicaba, 2024. Disponível em: <https://www.cepea.org.br/br/opiniao-cepea/demanda-fraca-e-exportacao-reduzida-limitam-producao-brasileira-de-arroz.aspx>. Acesso em 10/10/2024.