

(x) Graduação () Pós-Graduação

**PREVISÃO DE SÉRIES TEMPORAIS NA PRODUTIVIDADE AGRÍCOLA
BRASILEIRA: uma abordagem com machine learning**

Anderson Oliveira Brito
Universidade Regional do Cariri (URCA)
anderson.oliveira@urca.br

Maria Avyla Batista da Silva
Universidade Regional do Cariri (URCA)
avyla.batista@urca.br

Manoel Alexandre de Lucena
Universidade Federal do Cariri (UFCA)
manoel.alexandre@aluno.ufca.edu.br

RESUMO

A agricultura é peça-chave na dinâmica econômica brasileira, uma vez que confere ao setor competitividade, inovação e possibilidades de inversões em outras áreas. Neste sentido, a produtividade total dos fatores está diretamente atrelada à competitividade apresentada pelo setor. Logo, o seu aprimoramento e crescimento são essenciais tanto para o desenvolvimento agrícola, como para o crescimento continuado do setor. Diante de tamanha relevância, a literatura tem considerado realizar previsões de produtividade como forma de nortear políticas públicas e tomadas de decisão dos produtores. Nesse sentido, o presente trabalho objetivou, por meio da utilização do modelo tradicional de previsão de séries temporais *Autoregressive Integrated Moving Average* (ARIMA) e dois modelos de *machine learning*, *Support Vector Regression* (SVR) e *Decision Tree Regression* (DTR), prever as séries temporais de produtividade agrícola brasileira no recorte temporal de 2012 a 2023, em trimestres. Os principais resultados obtidos apontam que o SRV e o DTR obtiveram melhor desempenho na previsão da produtividade agrícola brasileira, sendo que o primeiro apresentou menor erro. Ademais, a projeção para o ano de 2024 revela que a produtividade deve manter a mesma tendência dos anos anteriores, com pico no primeiro trimestre, seguido por uma redução nos demais.

Palavras-chave: Previsão; Produtividade agrícola; Séries temporais; *Machine learning*.

1 INTRODUÇÃO

A agricultura desempenha um papel crucial na economia brasileira, impulsionando a inovação e o dinamismo tanto no comércio internacional quanto interno, além de influenciar investimentos em outras áreas econômicas (Freitas, 2014). Assim, Klein e Luna (2020) destacam que a competitividade do setor agrícola está diretamente ligada à sua produtividade. Assim, fica evidente que a agricultura não apenas impulsiona a economia, mas também desempenha um papel crucial na busca por maior eficiência e no desenvolvimento de práticas inovadoras para garantir o crescimento contínuo e a competitividade do setor. Ademais, para incrementar a análise da produtividade agrícola, a literatura tem incorporado a utilização do aprendizado de máquina (*machine learning*) para prever, a partir de um conjunto de dados, comportamentos, essencialmente como meio para otimizar a tomada de decisões e instauração de políticas públicas mais eficientes (Sousa *et al.*, 2020).

Em vista do exposto, o objetivo deste trabalho consiste em modelar e prever a produtividade agrícola do Brasil no período de 2012 a 2023, trimestralmente, utilizando-se de modelos de séries temporais, notadamente, algoritmos de *machine learning*. Para tal, utilizou-se uma série extraída do Observatório da Produtividade Regis Bonelli, da Fundação Getúlio Vargas (FGV). Inicialmente, a série histórica foi dividida em dois conjuntos, sendo 75% destinado ao treinamento dos modelos e 25% para testar o desempenho relativo das previsões. No conjunto de treinamento, aplicou-se uma transformação de escala, em que de cada observação, subtrai-se a média e se divide pelo desvio padrão. Em seguida, foram treinados três modelos: *Autoregressive Integrated Moving Average* (ARIMA), *Support Vector Regression* (SVR) e *Decision Tree Regression* (DTR). Sendo o primeiro um modelo clássico de previsão de séries temporais e dois últimos de *machine learning*. Após a modelagem, realizou-se a previsão considerando o horizonte do conjunto de teste. Em seguida, aplica-se a transformação reversa de escala, levando as previsões para a escala da variável modelada. Por fim, a raiz do erro quadrático médio (*Root Mean Squared Error* - RMSE) é calculado para avaliar o desempenho dos modelos em termos de erros de previsão. Ressalta-se que todas as previsões, gráficos e tabelas apresentadas neste trabalho foram geradas utilizando a linguagem Python.

2 DISCUSSÃO E ANÁLISE DOS DADOS

Na Tabela 1 é apresentado o RMSE do conjunto de teste dos modelos empregados na previsão da produtividade agrícola no Brasil. Em particular, observa-se que o ARIMA obteve um erro maior em relação aos algoritmos de aprendizado de máquina. Desta forma,

relativamente, o modelo com melhor desempenho é a SVR, com RMSE igual a 1,0979 reais por hora. A DTR apresenta desempenho inferior, com RMSE igual a 2,6802 reais por hora.

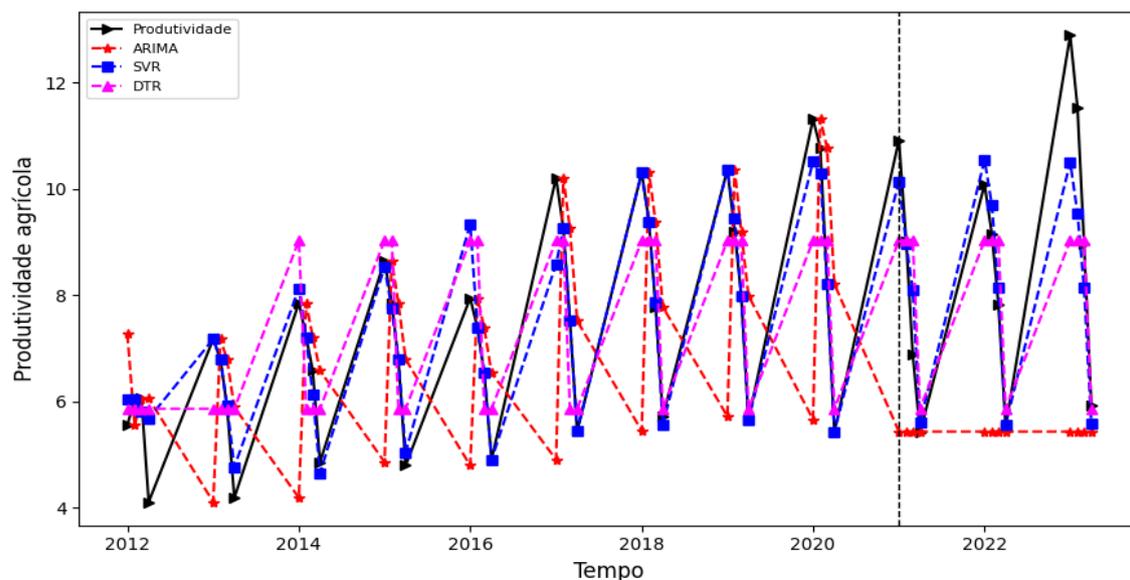
Tabela 1: RMSE das previsões no conjunto de teste dos modelos considerados na previsão da produtividade agrícola no Brasil (menor RMSE em negrito)

Métrica	ARIMA	SVR	DTR
RMSE	15,9537	1,0979	2,6802

Fonte: Elaborado pelos autores

Na Figura 1 são exibidos os gráficos das previsões dos modelos, considerando os conjuntos de treinamento e teste. No caso do ARIMA, por ser do tipo ARIMA (0,1,0) a previsão no conjunto de teste é constante e, dessa forma, não acompanhou a série. Por outro lado, a SVR e DTR conseguem acompanhar melhor a série de produtividade. Contudo, nitidamente, observa-se que a SVR conseguiu modelar melhor os picos da série. Esta percepção, portanto, ratifica o RMSE do erro deste modelo apresentado na Tabela 1. Este resultado se alinha com a literatura, em que algoritmos de aprendizado de máquina, têm sobressaído na previsão de produtividade agrícola (Tabar *et al.*, 2022).

Figura 1: Previsão da produtividade agrícola no Brasil no período de 2012 a 2023, trimestralmente



Nota: a linha vertical pontilhada separa o conjunto de treinamento do teste.

Fonte: Elaborado pelo autores

Considerando o melhor modelo obtido, neste caso, a SVR, realizou-se a projeção da

produtividade agrícola para o ano de 2024, ou seja, para os quatro trimestres. Os resultados indicam que a produtividade deve seguir a mesma trajetória dos anos anteriores, com um pico no primeiro trimestre (10,53 reais por hora) seguido por uma trajetória decrescente, alcançando 5,56 reais por hora no quarto trimestre. Esse resultado também mostra que a produtividade agrícola tende a crescer, embora apresente oscilações ao longo do ano, o que a literatura corrobora (Antunes; Stege, 2020).

4 CONCLUSÕES

Os resultados apontaram para o melhor desempenho da SVR , seguida pela DTR, enquanto o ARIMA apresentou um desempenho inferior, com um RMSE mais elevado. A análise do gráfico de previsões, revelou que o ARIMA não conseguiu acompanhar adequadamente a série devido à sua previsão constante, enquanto o SVR e a DTR mostraram uma capacidade superior de modelar a série de produtividade, com destaque para o primeiro, com melhor representação dos picos da série. Com base no melhor modelo (SVR), foi feita uma projeção da produtividade agrícola para 2024. Os resultados indicaram uma tendência esperada de seguir o padrão dos anos anteriores, com um pico no primeiro trimestre seguido por uma redução ao longo do ano. Apesar dos desafios históricos enfrentados pelo setor agrícola brasileiro, como a concentração de investimentos em centros industriais no passado, as mudanças estruturais adotadas ao longo das últimas décadas tornaram o setor significativamente eficiente e capaz de contribuir de maneira efetiva para o crescimento econômico do país. Este estudo fornece *insights* valiosos sobre a previsão da produtividade agrícola no Brasil, elucidando a importância de modelos de séries temporais e algoritmos de *machine learning* para esse fim.

REFERÊNCIAS

ANTUNES, L.; STEGE, A. Análise espacial da convergência da produtividade agrícola brasileira (1990-2018). *In: Encontro de Economia da Região Sul*, 23., 2020. **Anais [...]**. Porto Alegre: ANPEC SUL, 2020.

FREITAS, R. E. Produtividade agrícola no Brasil. *In: NEGRI, F.; CAVALCANTE, L. R. Produtividade no Brasil: desempenho e determinantes*. Brasília: ABDI, p. 373-401, 2014.

KLEIN, H. S.; LUNA, F. V. **Alimentando o mundo: o surgimento da moderna economia agrícola no Brasil**. Editora FGV, 2020.

KHAN, A. S.; SILVA, L. M. R. Assistência técnica, eficiência na utilização dos fatores de produção e da produtividade diferencial em propriedades rurais. **Revista de Economia e**

Sociologia Rural, Brasília, v. 35, n. 2, p. 95-114, 1997.

PEROBELLI, F. S.; ALMEIDA, E. S.; ALVIM, M. I. S. A.; FERREIRA, P. G.C.
Produtividade do setor agrícola brasileiro (1991-2003): uma análise espacial. **Nova economia**, v. 17, p. 65-91, 2007.

SOUZA, C. C.; OLIVEIRA, A. R.; BEZERRA, W. A. O.; CÁCERES, K. G.; VERONEZ, A. C. C. S. Previsão de produtividade da agropecuária em assentamentos rurais de Mato Grosso do Sul utilizando inteligência artificial. Encontro Internacional de Gestão, Desenvolvimento e Inovação (EIGEDIN), v. 4, n. 1, 2020. **Anais [...]**. EIGEDIN, 30 out. 2020.

TABAR, M.; LEE, D.; HUGHES, D. P.; YADAV, A. Mitigating Low Agricultural Productivity of Smallholder Farms in Africa: Time-Series Forecasting for Environmental Stressors. **Proceedings of the AAAI Conference on Artificial Intelligence**, [S. l.], v. 36, n. 11, p. 12608-12614, 2022. DOI: 10.1609/aaai.v36i11.21534. Disponível em: <https://ojs.aaai.org/index.php/AAAI/article/view/21534>. Acesso em: 1 abr. 2024.