

Uso do modelo ARIMA para previsão do preço do milho pago aos produtores no estado de Mato Grosso

Rodolfo B. Z. da Silva,¹ Fábria F. F. C. Aires²

UFMT, Várzea Grande, MT

André Krindges,³ Moiseis Ceconello⁴

UFMT, Cuiabá, MT

Adauto F. Bueno⁵

O estado de Mato Grosso é o maior produtor de milho do país, sendo que sua produção vem aumentando ano após ano. Na última safra (2021/2022), o estado foi responsável pela produção total de 41.620,10 mil toneladas, a partir de 6.547,40 mil hectares (ha) destinadas ao plantio, resultando em uma produtividade de 6.357 kg/ha [1]. Diante do exposto, prever o preço pago por este produto aos produtores rurais através de modelos de séries temporais (conjuntos de observações de uma variável coletadas em intervalos de tempo regulares) é de fundamental importância para que eles possam estimar seus ganhos e realizar o planejamento das safras. Este artigo teve como objetivo realizar previsões futuras do preço médio mensal pago aos produtores no estado de Mato Grosso a partir do modelo ARIMA (*Auto Regressive Integrated Moving Average*).

Para atingir este objetivo, foram utilizados dados históricos mensais (de janeiro de 2019 a fevereiro de 2023) do preço médio do milho, disponibilizados pela Companhia Nacional de Abastecimento (CONAB) [1].

Todos os procedimentos de análise de dados foram realizados com auxílio do software R. Primeiramente, foram obtidas as estatísticas descritivas da série temporal estudada. Em seguida, a base de dados original com 50 observações foi dividida em duas porções: (i) treinamento (38 observações) para estimação do parâmetro do modelo; e (ii) teste (12 observações) para validação e análise do desempenho apresentado através dos erros de previsão - *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE) e *Mean Absolute Percentage Error* (MAPE).

Para identificar o modelo ARIMA apropriado para prever a série temporal analisada, foi utilizado um algoritmo para seleção automática de modelos ARIMA desenvolvido em [2] e que está presente no pacote *forecast* (função *auto.arima*) do software R.

Após a identificação do modelo, foram realizadas previsões para fora da amostra (4 meses), ou seja, para os meses de março, abril, maio e junho de 2023. O modelo ARIMA (p, d, q) é formado pelos componentes autoregressivo de ordem p - $AR(p)$ e de média móvel de ordem q - $MA(q)$, com a série original diferenciada d vezes, sendo representado pela Equação 1 [3].

$$w_t = c + \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_p \epsilon_{t-p} \quad (1)$$

Em que: $w_t = \Delta^d v_t = v_t - v_{t-d}$ As estatísticas descritivas para série temporal estudada durante o período de janeiro de 2019 a fevereiro de 2023, foram as seguintes: Valor Mínimo - R\$ 21,44; 1º Quartil - R\$ 31,46; Mediana - R\$ 61,90; Média - R\$ 52,15; 3º Quartil - R\$ 68,75; e Valor Máximo

¹rodolfo.silva@ufmt.br

²fabia_aires@hotmail.com

³krindges@gmail.com

⁴moiseis@gmail.com

⁵adauto.bueno@unemat.br

– R\$ 78,99. O modelo ARIMA identificado automaticamente pelo software R, foi o ARIMA (1,1,1). Os parâmetros estimados para este modelo, os critérios AIC (*Akaike's Information Criteria*) e BIC (*Bayesian Information Criteria*), bem como as medidas de acurácia obtidas para a porção de dados de teste (12 últimas observações), são apresentadas na Tabela 1.

Tabela 1: Parâmetros do modelo ARIMA(1,1,1) e medidas de acurácia para a porção de dados de teste (12 últimas observações).

Modelo	Parâmetros	Erro Padrão	AIC	BIC	MAE	RMSE	MAPE
ARIMA(1,1,1)	$\phi_1 = -0,63099$	0,15033	281,98	287,66	2,98	3,95	5,74
	$\theta_1 = 0,96827$	0,10914					

A Figura 1 apresenta a série temporal do preço do milho pago ao produtor mato-grossense dividida em dados de treinamento e de teste, os valores ajustados que foram previstos pelo modelo ARIMA (1,1,1) na porção de dados de treinamento, na porção de dados de teste e para os 4 períodos fora da amostra (março, abril, maio e junho de 2023).

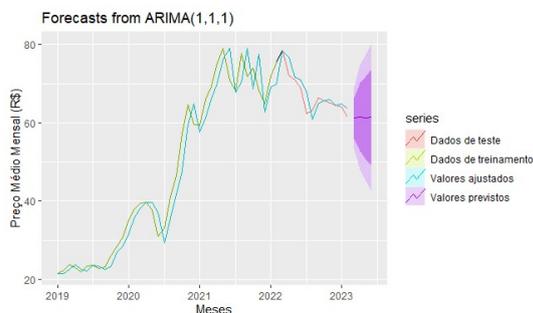


Figura 1: Série temporal estudada (dividida em dados de treinamento e teste), valores ajustados aos dados de treinamento, valores previstos para a porção de dados de teste e para fora da amostra pelo modelo ARIMA (1,1,1)

Na Figura 1, a área sombreada pela cor roxa para o período fora da amostra, delimita os intervalos de confiança de 95% para as previsões obtidas. Os valores previstos pelo modelo ARIMA (1,1,1) para o período fora da amostra foram os seguintes: março/2023 – R\$ 61,17; abril/2023 – R\$ 61,41; maio/2023 – R\$ 61,26; e junho/2023 – R\$ 61,35.

Agradecimentos

Agradecemos à FAPEMAT pelo apoio financeiro.

Referências

- [1] Conab. **Companhia Nacional de Abastecimento**. Online. Acessado em 204/04/2023, <https://www.conab.gov.br/info-agro/safras/serie-historica-das-safras/itemlist/category/910-Milho>.
- [2] R. J. Hyndman e Y Khandakar. “Automatic Time Series Forecasting: The forecast Package for R.” Em: **Journal of Statistical Software** 27 (2008), pp. 1–22.
- [3] S. G. Makridakis, S. C. Wheelwright e R. J. Hyndman. **Forecasting: Methods and Applications**. 3a. ed. New York: John Wiley & Sons, 1998.