

ANÁLISE E PREVISÃO DOS PREÇOS DO LEITE NO BRASIL: UMA ABORDAGEM BASEADA EM MODELOS SARIMA E ARIMAX

ANALYSIS AND FORECASTING OF MILK PRICES IN BRAZIL: AN APPROACH BASED ON SARIMA AND ARIMAX MODELS

Patrícia Taynara de Jesus das Chagas

Zootecnista. Mestranda no Programa de Pós-Graduação em Economia Rural – UFC.
patriciataynara@alu.ufc.br.

Vitor Hugo Miro Couto Silva

Doutor em Economia. Professor do Programa de Pós-Graduação em Economia Rural – UFC.
vitormiro@ufc.br

Grupo de Trabalho (GT): GT03. Mercados agrícolas e cadeias produtivas.

Resumo

O Brasil figura entre os maiores produtores mundiais de leite bovino, e o comportamento de seus preços reflete a influência de fatores como tendência, sazonalidade e custos dos principais insumos. Este estudo tem como objetivo analisar e prever o comportamento do preço real do leite no Brasil, aplicando modelos econométricos de regressão linear múltipla e de séries temporais. Entre as variáveis exógenas consideradas estão os preços do milho, da soja e o tamanho do rebanho, além de componentes de tendência e efeitos sazonais. Os resultados indicam que o preço do milho exerce influência significativa sobre o preço do leite, refletindo a transmissão de custos na cadeia produtiva. A análise mostrou ainda que a série apresenta estacionariedade em torno de uma tendência determinística, sendo o modelo ARIMA (5,0,0) com tendência o que apresentou melhor desempenho preditivo, segundo os critérios de informação e métricas de erro. As previsões para o curto prazo indicam manutenção dos níveis atuais com leve tendência de alta, sugerindo estabilidade no mercado leiteiro e confirmando a utilidade dos modelos de séries temporais para o monitoramento e planejamento da atividade.

Palavras-chave: preço do leite, séries temporais, previsão.

Abstract

Brazil ranks among the world's largest producers of bovine milk, and the behavior of its prices reflects the influence of factors such as trend, seasonality, and the cost of key inputs. This study aims to analyze and forecast the behavior of the real milk price in Brazil by applying multiple linear regression and time series models. The exogenous variables considered include the prices of corn and soybean, herd size, as well as trend and seasonal components. The results indicate that the price of corn exerts a significant influence on milk prices, reflecting cost transmission along the production chain. The analysis also shows that the series is stationary around a deterministic trend, with the ARIMA (5,0,0) model with trend presenting the best predictive performance according to information criteria and forecast accuracy metrics. Short-term forecasts indicate relative stability with a slight upward trend in milk prices, suggesting a stable market environment and confirming the usefulness of time series models for monitoring and planning in the dairy sector.

Keywords: milk price; time series; forecasting.

1. Introdução

A produção de leite ocupa posição central na agropecuária brasileira, tanto pela sua relevância econômica quanto pelo papel social que desempenha na geração de renda e empregos. A atividade está presente em mais de um milhão de propriedades rurais distribuídas por todo o território nacional, sendo responsável por cerca de três milhões de postos de trabalho apenas na produção primária (Vilela; Leite; Resende, 2002; Rocha; Carvalho; Resende, 2020).

Segundo a Pesquisa da Pecuária Municipal (PPM) do Instituto Brasileiro de Geografia e Estatística (IBGE, 2024), 5.491 dos 5.570 municípios brasileiros registraram produção de leite em 2023, totalizando 35,4 bilhões de litros e aproximadamente R\$ 80 bilhões em valor de produção, o que evidencia a importância da pecuária leiteira no agronegócio nacional. No cenário internacional, dados da FAO (2025) indicam que o Brasil ocupa a quarta posição mundial em volume de produção de leite.

O comportamento dos preços do leite tem impacto direto sobre a rentabilidade dos produtores e sobre a competitividade da cadeia produtiva. As variações no preço refletem fatores estruturais e conjunturais, como custos de alimentação, clima, rebanho e dinâmica de mercado. Diante desse contexto, compreender a dinâmica dos preços do leite é fundamental para subsidiar decisões de planejamento, comercialização e formulação de políticas públicas.

Assim, o presente estudo tem como objetivo analisar e prever o comportamento do preço real do leite no Brasil, considerando variáveis de custo de produção (preços do milho e da soja), tamanho do rebanho, tendência temporal e efeitos sazonais. Essas variáveis foram consideradas de forma a testar as seguintes hipóteses de pesquisa: (i) os preços dos insumos — especialmente milho e soja — exercem influência positiva e significativa sobre o preço do leite pago ao produtor; (ii) o preço real do leite apresenta uma tendência de crescimento no longo prazo, associada à elevação dos custos de produção e à dinâmica de mercado; e (iii) existem padrões sazonais sistemáticos que afetam os preços ao longo do ano, refletindo variações climáticas e produtivas.

Para alcançar esse objetivo, o estudo adota abordagens econométricas complementares, combinando modelos de regressão linear múltipla, estimados por Mínimos Quadrados Ordinários (MQO), e modelos de séries temporais, em particular o *AutoRegressive Integrated Moving Average* (ARIMA) e derivações como o *Seasonal ARIMA* (SARIMA) e as extensões que incorporam variáveis exógenas (ARIMAX e SARIMAX). Essa estratégia analítica permite avaliar simultaneamente os determinantes econômicos de curto e longo prazo e a capacidade preditiva das diferentes especificações, oferecendo uma visão abrangente sobre a dinâmica e o monitoramento do mercado leiteiro brasileiro.

O artigo está estruturado da seguinte forma: além desta introdução, a seção 2 apresenta a revisão de literatura sobre os principais determinantes econômicos e temporais dos preços do leite; a seção 3 descreve a base de dados e os procedimentos metodológicos utilizados; a seção 4 discute os resultados empíricos e as previsões obtidas; e, por fim, a seção 5 apresenta as conclusões e sugestões para pesquisas futuras.

2. Revisão de Literatura

2.1. Fatores Econômicos na Produção Leiteira

O setor leiteiro no Brasil exerce uma função essencial na economia e na garantia do abastecimento alimentar, sendo moldado por uma variedade de fatores econômicos e sociais. A trajetória dos preços do leite, considerando valores nominais e reais, evidencia essas influências, destacando os efeitos da inflação e de outras variáveis macroeconômicas (Vilela *et al.*, 2016).

A pecuária leiteira brasileira tem apresentado crescimento expressivo nas últimas décadas, acompanhando a expansão da produção agropecuária e o aumento do consumo de derivados lácteos. De acordo com a Companhia Nacional de Abastecimento (CONAB, 2025), o mercado nacional iniciou 2024 com elevação da produção e demanda relativamente estável. No quarto trimestre de 2024, a captação formal de leite atingiu 6,75 bilhões de litros — um aumento de 4,1% em relação ao mesmo período do ano anterior e 7,3% em relação ao trimestre anterior. Esses resultados refletem condições climáticas favoráveis e preços mais atrativos aos produtores, fatores que impulsionaram a produção.

Entretanto, o setor enfrenta fortes desafios relacionados à volatilidade de preços e à gestão de risco (Siqueira; Silva; Aguiar, 2008). Nas principais bacias leiteiras do Brasil monitoradas pelo Centro de Estudos Avançados em Economia Aplicada (Cepea), o leite cru é entregue diariamente aos laticínios, enquanto o pagamento aos produtores ocorre mensalmente, no mês seguinte à captação, o que aumenta a incerteza e dificulta o planejamento financeiro (Grigol, 2023). Essa particularidade gera incertezas e dificulta o planejamento dos pecuaristas, que precisam lidar também com questões sazonais não só quanto ao preço, mas também à produção (Alves; Sousa; Ervilha, 2014).

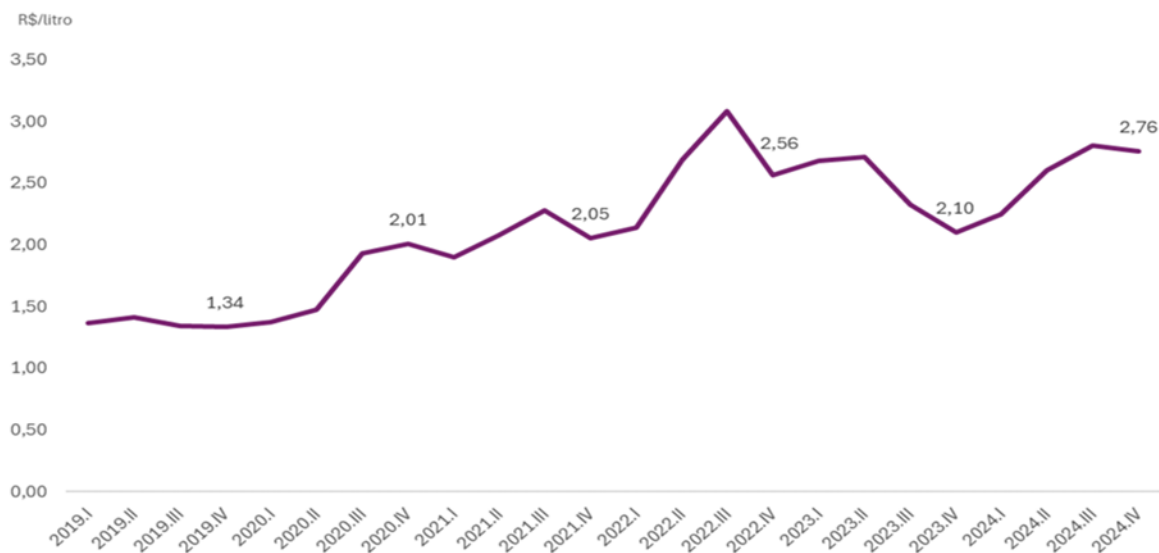
Estudos sobre a formação de preços do leite no Brasil indicam forte integração espacial entre os mercados regionais. No estudo de Carvalho *et al.* (2013), foi analisada a transmissão de preços entre mercados utilizando os valores pagos ao produtor nos estados de Minas Gerais, Rio Grande do Sul, Paraná, São Paulo e Goiás. Constatou-se que Minas Gerais apresenta o maior poder explicativo sobre os preços nos demais estados. Esse impacto pode ser explicado

devido ao fato de o estado de Minas Gerais ser o maior produtor de leite, com 24,9% da produção nacional (IBGE, 2025).

Avaliando o comportamento dos preços do leite e a influência de alguns fatores sobre os preços, o artigo de Simões, Nicholson e Carvalho (2023), revelou que o preço nominal médio anual pago aos produtores no Brasil apresentou um aumento de R\$ 0,93 por litro (194%) entre 2005 e 2019. No mesmo período, o preço real, ajustado pelo Índice Geral de Preços – Disponibilidade Interna (IGP-DI) para dezembro de 2019, cresceu R\$ 0,36 por litro (34%).

A Pesquisa Trimestral do Leite feita pelo IBGE mostrou que no quarto trimestre de 2024, o preço médio do litro de leite cru pago aos produtores foi de R\$ 2,76, representando um aumento de 31,4% em relação ao mesmo período do ano anterior. Contudo, comparado ao valor médio registrado no terceiro trimestre de 2024, houve uma redução de 1,4% (IBGE, 2025). O gráfico da Figura 1, a seguir, apresenta a trajetória recente do preço médio do leite cru pago ao produtor, onde é possível notar uma evolução no preço pago aos produtores entre os anos de 2019 e 2024.

Figura 1 - Evolução do preço líquido médio do leite cru pago ao produtor (R\$/l) - média trimestral no período 2019-2024.



Fonte: IBGE (2025).

2.2. Custo de insumos

Os custos com alimentação representam a maior parcela das despesas operacionais nas propriedades leiteiras (Hoffmann *et al.*, 1987). Nesse contexto, soja e milho são os principais insumos utilizados, desempenhando papel central na rentabilidade da atividade. A soja, normalmente processada na forma de farelo, constitui suplemento proteico de alto valor

nutricional para a alimentação animal e representa um subproduto de elevada rentabilidade para a indústria de óleo. Seu preço no mercado varia conforme o índice de proteína, que se situa entre 43% e 48% (Fernandes *et al.*, 2009; Boschiero, 2023).

O consumo interno de farelo de soja no Brasil tem crescido de forma contínua, atingindo aproximadamente 19,25 milhões de toneladas na safra 2023/24 (EMBRAPA, 2025). Já o milho é o principal insumo energético das dietas para vacas leiteiras, com amido entre 65% e 72% da matéria seca, sendo diretamente associado ao incremento na produção de leite (Moura, 2013). Seu consumo para fins pecuários alcançou 56,3 milhões de toneladas na safra 2024/25 (ABIMILHO, 2025).

As flutuações nos preços desses grãos afetam diretamente os custos de produção e, consequentemente, o preço do leite. Conforme Neiva (2021), aumentos no preço do milho e da soja reduzem as margens de lucro dos produtores e tendem a elevar os preços pagos ao produtor e ao consumidor. Simões, Nicholson e Carvalho (2023) observaram que um acréscimo de 10% no preço da ração resulta em elevação média de 0,8% no preço do leite, confirmando a transmissão de custos ao longo da cadeia. Informação que revela que os aumentos no preço da ração são absorvidos, em parte, pelo produtor, comprimindo suas margens de lucro.

2.3. Variáveis Temporais

Além dos fatores de custo, a dinâmica temporal dos preços é um aspecto crucial na pecuária leiteira. O setor é fortemente influenciado por padrões sazonais associados ao regime climático tropical brasileiro, caracterizado por um inverno seco e verão chuvoso. Essa variação sazonal impacta diretamente os produtores de leite, resultando em uma queda de receita durante a entressafra. Nesse período, há redução na produção de leite e aumento nos custos de produção (Junqueira; Zoccal; Miranda, 2013).

A literatura sobre séries temporais em economia agrícola enfatiza que a decomposição dos preços em tendência e sazonalidade é essencial para compreender sua evolução e formular previsões precisas. A tendência capta movimentos de longo prazo, como mudanças estruturais de oferta e demanda, enquanto a sazonalidade identifica padrões recorrentes de curto prazo relacionados a fatores climáticos e produtivos.

Esses elementos justificam a aplicação de modelos econométricos de séries temporais, como ARIMA e SARIMA, que permitem representar a dependência temporal e os efeitos sazonais de forma sistemática. Assim, a revisão da literatura evidencia que o comportamento do preço do leite resulta da interação entre custos de insumos, variações sazonais e tendências estruturais de mercado, fatores que orientam as hipóteses e a modelagem adotadas neste estudo.

3. Metodologia

O estudo utiliza dados mensais de janeiro de 2005 a dezembro de 2024, provenientes do Cepea, IBGE e Fundação Getúlio Vargas (FGV), todos convertidos em valores reais pelo IGP-DI. As variáveis de interesse são o preço real do leite, o preço real do milho, o preço real da soja e o rebanho leiteiro. Todas foram transformadas em logaritmos para estabilizar a variância e permitir a interpretação dos coeficientes como elasticidades.

3.1. Modelos de regressão linear

Como etapa inicial da análise, foram estimados modelos de regressão linear múltipla pelo método dos Mínimos Quadrados Ordinários (MQO), com o objetivo de identificar os principais determinantes do preço real do leite. As especificações consideraram diferentes combinações de variáveis, incluindo tendência temporal, componentes sazonais, preços de insumos (milho e soja) e o tamanho do rebanho leiteiro.

A variável de tendência (*tend*) foi construída de forma crescente ao longo do tempo, representando possíveis padrões de evolução ou declínio do preço real do leite. Por sua vez, os efeitos sazonais foram incorporados por meio de um conjunto de variáveis *dummy* mensais, que capturam variações sistemáticas associadas a períodos específicos do ano (como entressafra e safra).

O modelo geral estimado pode ser representado da seguinte forma:

$$\ln P_{leite} = \beta_0 + \beta_1 \cdot tend + \beta_2 \cdot \ln P_{soja} + \beta_3 \cdot \ln P_{milho} + \beta_4 \cdot \ln Reb + \sum_{j=2}^{12} \gamma_j \cdot D_j + \epsilon_t \quad [1]$$

Nesta expressão $\ln P_{leite}$ é o preço do leite (em escala de logaritmo natural); *tend* representa a tendência temporal; $\ln P_{soja}$ e $\ln P_{milho}$ correspondem aos preços reais da soja e do milho, respectivamente; $\ln Reb$ indica o tamanho do rebanho leiteiro; os termos D_j são variáveis *dummy* mensais (com janeiro como categoria base), que assumem valor 1 quando o mês é o *j*-ésimo e 0 caso contrário; ϵ_t é o termo de erro aleatório, assumido com média zero e variância constante.

A partir desse modelo, dado por [1] foram testadas diferentes especificações, de modo a avaliar a relevância das variáveis explicativas e dos efeitos sazonais. Diferentes testes foram aplicados. O teste F foi utilizado para comparar os modelos. A normalidade dos resíduos foi avaliada por meio do teste de Shapiro-Wilk. E a verificação de autocorrelação dos resíduos foi

realizada por meio do teste de Durbin-Watson, enquanto a heterocedasticidade foi analisada utilizando o teste de Breusch-Pagan.

3.2. Modelagem de Séries Temporais

O modelo ARIMA (p, d, q) é uma combinação de componentes autoregressivos (AR) e de médias móveis (MA), aplicados a uma série que foi diferenciada d vezes para torná-la estacionária. A forma geral do modelo ARIMA (p, d, q) pode ser representada pela equação:

$$(1 - B)^d Y_t = c + \sum_{i=1}^p \varphi_i (1 - B)^d Y_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad [2]$$

Nesta expressão, Y_t é o valor da série no tempo t ; B é o operador de defasagem, tal que $B Y_t = Y_{t-1}$; d é a ordem de diferenciação necessária para atingir a estacionariedade; p é a ordem do componente autoregressivo (AR), que captura a dependência de um valor em relação a suas próprias defasagens; q é a ordem do componente de Médias Móveis (MA), que modela a dependência do valor corrente em relação aos erros de previsão passados; φ_i e θ_j são os parâmetros a serem estimados para os componentes AR e MA, respectivamente; ε_t é um ruído branco, um termo de erro aleatório com média zero e variância constante.

A metodologia de Box-Jenkins para ajustar um modelo ARIMA segue um processo iterativo de três etapas:

1. Identificação: Análise das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) para determinar os valores candidatos de p , d e q .
2. Estimação: Uso de métodos como a Máxima Verossimilhança para estimar os parâmetros φ_i e θ_j do modelo.
3. Diagnóstico: Verificação da adequação do modelo, analisando-se se os resíduos resultantes se comportam como um ruído branco. Neste caso, o teste de Ljung-Box é comumente empregado.

Quando uma série temporal exibe sazonalidade com período s ($s = 12$ para dados mensais), o modelo ARIMA padrão é insuficiente. O modelo SARIMA, denotado por ARIMA $(p, d, q)(P, D, Q)_s$, incorpora termos sazonais análogos aos componentes não sazonais. A estrutura do modelo SARIMA é dada por:

$$\Phi_P(B^s) \varphi_p(B) (1 - B^s)^D (1 - B)^d Y_t = \theta_Q(B^s) \theta_q(B) \varepsilon_t \quad [3]$$

Nesta expressão (p, d, q) são as ordens não sazonais (conforme definido no ARIMA); $(P, D, Q)_s$ são as ordens sazonais; $\Phi_P(B^s)$ é o polinômio autoregressivo sazonal de ordem P ;

$\theta_Q(B^S)$ é o polinômio de médias móveis sazonal de ordem Q ; $(1 - B^S)^D$ é o operador de diferenciação sazonal de ordem D .

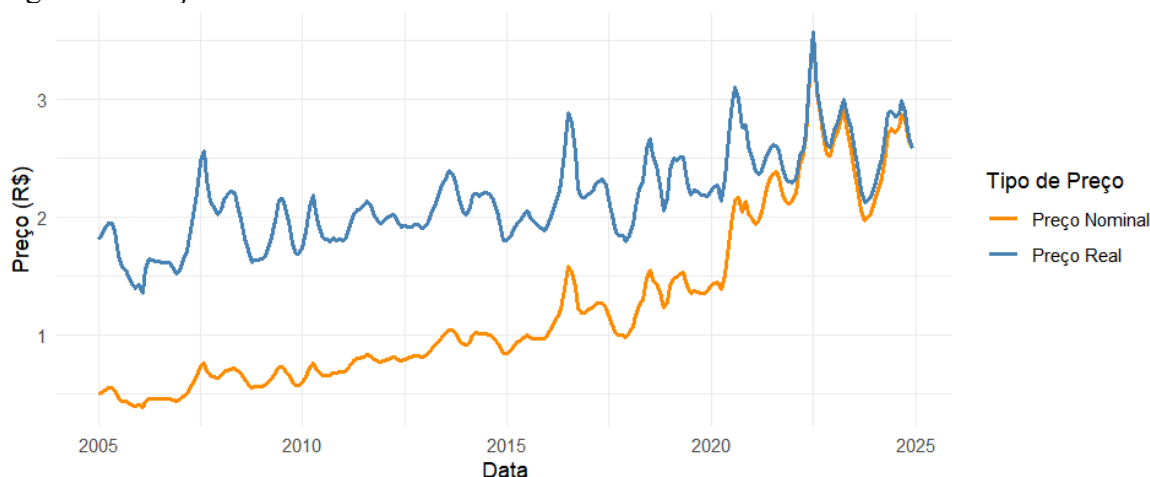
Em termos práticos, o modelo SARIMA não apenas modela a correlação entre um ponto e seus predecessores imediatos (componente não sazonal), mas também a correlação entre um ponto e seus valores no mesmo período sazonal do ano (ou ciclo) anterior (componente sazonal).

4. Resultados e Discussão

4.1 Análise do Preço Real do Leite no Brasil

Para uma compreensão mais ampla, um gráfico comparativo (Figura 2) entre os preços nominal e real do leite no Brasil foi elaborado, abrangendo o período de 2005 a 2025, o que permite uma análise detalhada das transformações ocorridas ao longo dos anos.

Figura 2 - Preço Nominal e Real do Leite.



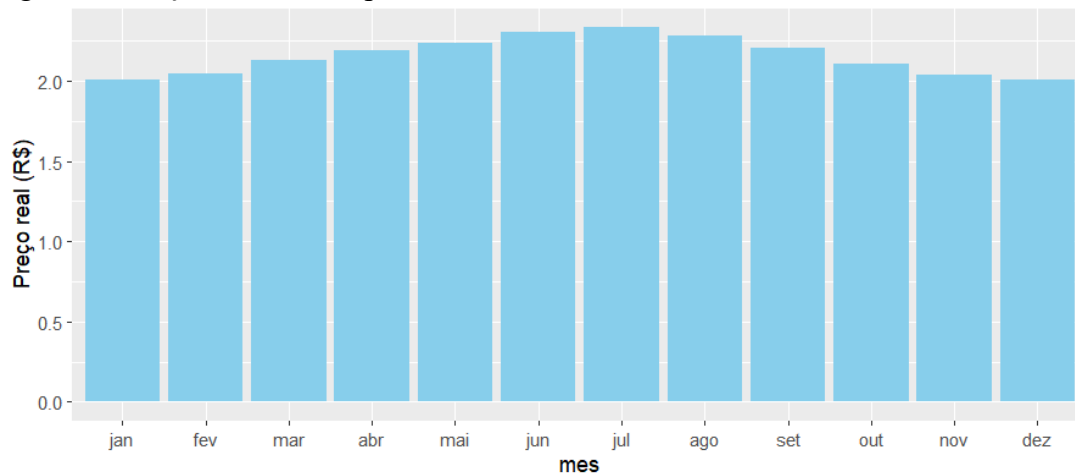
Fonte: Dados da Pesquisa.

A linha laranja, demonstra uma tendência de crescimento ao longo do período. A linha azul apresenta variações mais marcadas e valores geralmente mais altos que os do preço nominal, refletindo os ajustes pela inflação e outros fatores econômicos. Esses elementos, combinados, indicam que, embora o preço nominal do leite tenha aumentado de maneira constante, o preço real demonstra uma maior volatilidade, mas também uma tendência de crescimento sistemático.

O gráfico da Figura 3 mostra a distribuição média do preço real do leite em diferentes meses do ano. Os dados apresentados indicam que o período de abril a setembro, com destaque para os meses de maio a julho, corresponde a uma fase de maior valorização nos preços do leite.

Esse comportamento pode estar relacionado tanto à sazonalidade na produção como o período de entressafra, quanto a alterações na demanda.

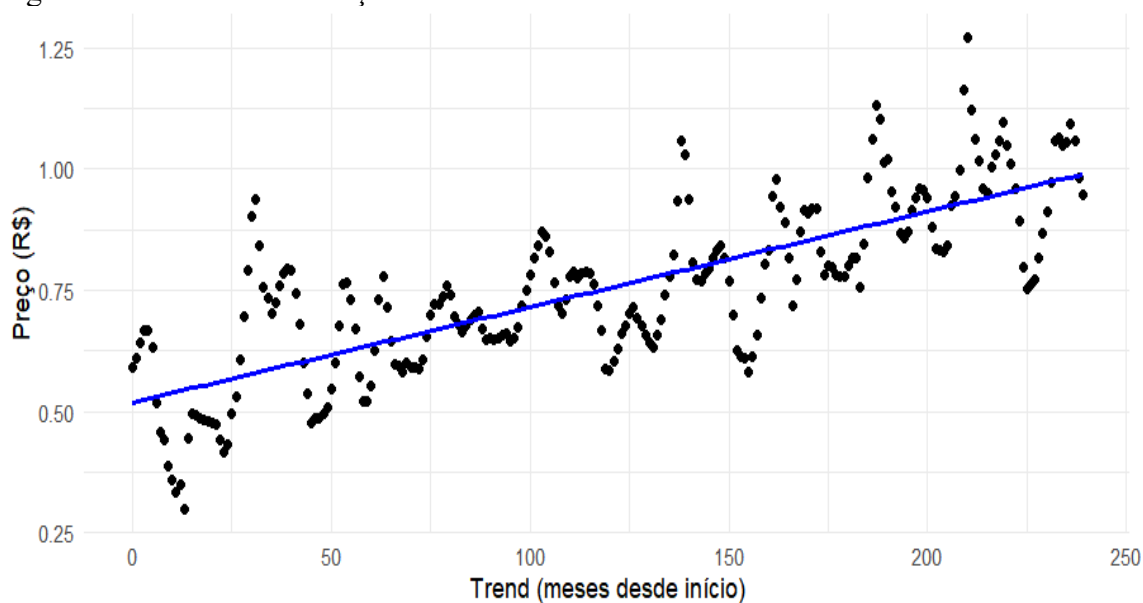
Figura 3 - Preço Médio Real por Mês.



Fonte: Dados da Pesquisa.

Na análise inicial foi estimado um modelo de regressão considerado apenas a variável de tendência (*tend*). No gráfico da Figura 4 temos o modelo ajustado aos dados. Os resultados sugerem que o preço real do leite no Brasil segue uma tendência ascendente ao longo do tempo, com variações sazonais significativas.

Figura 4 - Tendência do Preço Nominal do Leite.



Fonte: Dados da Pesquisa.

4.2. Modelagem com modelos de regressão linear

Conforme apresentado na metodologia, foram estimadas diferentes especificações do modelo de regressão linear. A Tabela 1 apresenta, de forma sintética, os resultados das estimações por Mínimos Quadrados Ordinários (MQO) para diferentes especificações do modelo de preço real do leite. De modo geral, observa-se que a inclusão de componentes sazonais e variáveis de custo (milho e soja) melhora o ajuste dos modelos, conforme indicado pelo aumento do R^2 ajustado e pela redução dos critérios de informação (AIC e BIC).

O coeficiente de tendência foi estatisticamente significativo em todas as especificações do modelo, indicando um resultado bastante robusto. Em todos os modelos, tem-se uma evidência de crescimento médio de aproximadamente 0,2% ao mês no preço real do leite. Em termos anuais, isso dá cerca de 2,4% ao ano ($((1,002)^{12} - 1 \approx 2,4\%)$).

As *dummies* para o comportamento sazonal também se mostram relevantes, em que os coeficientes para os meses março–setembro apresentam efeitos positivos e significativos, com picos entre maio–julho ($\approx 0,13$ – $0,16$ em log), ou seja, ~ 14 – 17% acima de janeiro (mês-base). No fim do ano (out–dez) os efeitos são menores e, em alguns casos, não significativos ou ligeiramente negativos (nov/dez).

O preço do milho quando incluído, apresenta um coeficiente entre 0,18 e 0,22, sugerindo elasticidade de 0,18–0,22, ou seja, um aumento de 10% no preço do milho associa-se a aumentos entre 1,8–2,2% no preço real do leite, *ceteris paribus*. As evidências nos modelos estimados indicam que o milho é insumo mais robusto na transmissão de custos.

O preço da soja se mostrou significativo apenas em algumas especificações, mas perde significância em outras quando combinada com o preço do milho. Ou seja, o efeito é positivo, porém não se pode afirmar que seja robusto (possível colinearidade entre insumos e/ou sobreposição com sazonalidade).

Por sua vez, a variável que representa o tamanho do rebanho não se mostrou significativa nas especificações testadas. Talvez por ser uma variável com variações anuais (sem mudanças ao longo dos meses), tende a não explicar flutuações mensais de preço.

Os resultados dos modelos lineares indicam que a inclusão de variáveis explicativas adicionais, como os preços do milho, da soja e o tamanho do rebanho, melhora substancialmente o ajuste em relação às especificações mais simples. Os coeficientes de determinação ajustados (R^2 ajust.) aumentam de 0,60 no modelo básico para cerca de 0,73 nas versões mais completas (Modelos 6 a 8), evidenciando que essas variáveis contribuem para explicar a variação do logaritmo do preço real do leite.

Os critérios de informação (AIC e BIC) também apontam para melhor desempenho das especificações intermediárias — especialmente os Modelos 7 e 8 — que equilibram poder

explicativo e parcimônia. Em todos os casos, o teste F global é altamente significativo, confirmando que os conjuntos de regressores são estatisticamente relevantes.

Por sua vez, os testes de diagnóstico revelam problemas típicos de séries temporais. O teste de Durbin–Watson apresentou valores muito baixos ($\approx 0,2$), com p-valores nulos, indicando forte autocorrelação positiva dos resíduos. Já o teste de Breusch–Pagan mostrou evidência de heterocedasticidade em boa parte dos modelos ($p < 0,05$), sugerindo variância não constante dos erros.

Tabela 1. Resultados dos modelos de regressão linear estimados.

Variável	Modelo1	Modelo2	Modelo3	Modelo4	Modelo5	Modelo6	Modelo7	Modelo8	Modelo9
Intercepto	0,515 (0,014)***	0,452 (0,025)***	-0,449 (0,139)**	-0,536 (0,202)**	0,306 (1,244)	-1,414 (1,263)	-0,663 (0,196)***	-1,896 (1,185)	0,876 (1,186)
Tendência (trend)	0,002 (0,000)***	0,002 (0,000)***	0,002 (0,000)***	0,002 (0,000)***	0,002 (0,000)***	0,002 (0,000)***	0,002 (0,000)***	0,002 (0,000)***	0,001 (0,000)***
ln(Preço do milho)			0,213 (0,032)***			0,188 (0,044)***	0,177 (0,040)***	0,220 (0,033)***	
ln(Preço da soja)				0,209 (0,042)***		0,062(0,056)	0,078 (0,050)		0,223 (0,044)***
ln(Rebanho)					0,009 (0,073)	0,046 (0,076)		0,084 (0,068)	-0,087 (0,072)
Mês 02		0,020 (0,032)	0,024 (0,029)	0,026 (0,030)	0,020 (0,032)	0,025 (0,029)	0,025 (0,029)	0,024(0,029)	0,027 (0,030)
Mês 03		0,061 (0,032)	0,063 (0,029)*	0,067 (0,030)*	0,061 (0,032)	0,064 (0,029)*	0,065 (0,029)*	0,062 (0,029)*	0,068(0,030)*
Mês 04		0,089 (0,032)**	0,104 (0,029)***	0,098 (0,030)**	0,089 (0,032)**	0,105 (0,029)***	0,105 (0,029)***	0,104 (0,029)***	0,099 (0,030)**
Mês 05		0,108 (0,032)***	0,128 (0,029)***	0,113 (0,030)***	0,107 (0,032)***	0,127 (0,029)***	0,126 (0,029)***	0,128 (0,029)***	0,114 (0,030)***
Mês 06		0,131 (0,032)***	0,154 (0,029)***	0,133 (0,030)***	0,131 (0,032)***	0,151 (0,029)***	0,151 (0,029)***	0,154 (0,029)***	0,134 (0,030)***
Mês 07		0,137 (0,032)***	0,164 (0,029)***	0,136 (0,030)***	0,137 (0,032)***	0,160 (0,030)***	0,159 (0,030)***	0,164 (0,029)***	0,137 (0,030)***
Mês 08		0,113 (0,032)***	0,136 (0,029)***	0,110 (0,030)***	0,113 (0,032)***	0,132 (0,030)***	0,131 (0,029)***	0,136 (0,029)***	0,111 (0,030)***
Mês 09		0,076 (0,032)*	0,094 (0,029)**	0,070 (0,030)*	0,075 (0,032)*	0,089 (0,029)**	0,088 (0,029)**	0,093 (0,029)**	0,071 (0,030)*
Mês 10		0,028 (0,032)	0,041 (0,029)	0,024 (0,030)	0,028 (0,032)	0,037 (0,029)	0,037 (0,029)	0,040 (0,029)	0,025 (0,030)
Mês 11		-0,002 (0,032)	0,006 (0,029)	-0,007 (0,030)	-0,002 (0,032)	0,002 (0,029)	0,003 (0,029)	0,005 (0,029)	-0,006 (0,030)
Mês 12		-0,020 (0,032)	-0,015 (0,029)	-0,022 (0,030)	-0,020 (0,032)	-0,017 (0,029)	-0,017 (0,029)	-0,017 (0,029)	-0,020 (0,030)
Estatísticas do modelo									
R ²	0,603	0,693	0,742	0,723	0,693	0,745	0,745	0,744	0,725
R ² ajust.	0,601	0,677	0,728	0,707	0,675	0,728	0,729	0,728	0,708
AIC	-367,819	-407,577	-447,531	-430,114	-405,591	-446,440	-448,052	-447,138	-429,666
BIC	-357,377	-358,848	-395,321	-377,904	-353,382	-387,269	-392,362	-391,448	-373,975
F (global)	361,594	42,726	50,087	45,364	39,269	43,731	46,962	46,722	42,313
p-valor F	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Durbin-Watson	0,204	0,182	0,245	0,219	0,182	0,250	0,249	0,248	0,222
p-valor DW	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Breusch-Pagan	1,229	21,059	29,447	21,784	37,863	39,657	29,537	39,264	31,949
p-valor BP	0,268	0,050	0,006	0,059	0,000	0,001	0,009	0,000	0,004

Fonte: Dados da pesquisa.

Valores de erro padrão entre parênteses. Níveis de significância: 0 ‘****’ 0.001 ‘***’ 0.01 ‘**’ 0.05 ‘.’ 0.1 ‘.’ 1.

4.3. Modelagem com modelos de séries temporais

Os resultados obtidos com a estimação dos modelos por regressão linear indicam que, embora eles expliquem bem o comportamento médio da série, suas inferências devem ser interpretadas com cautela. O padrão de autocorrelação e heterocedasticidade sugere a necessidade de modelos dinâmicos mais adequados, como os modelos ARIMA, SARIMA e ARIMAX, que tratam explicitamente as dependências temporais e sazonais da série.

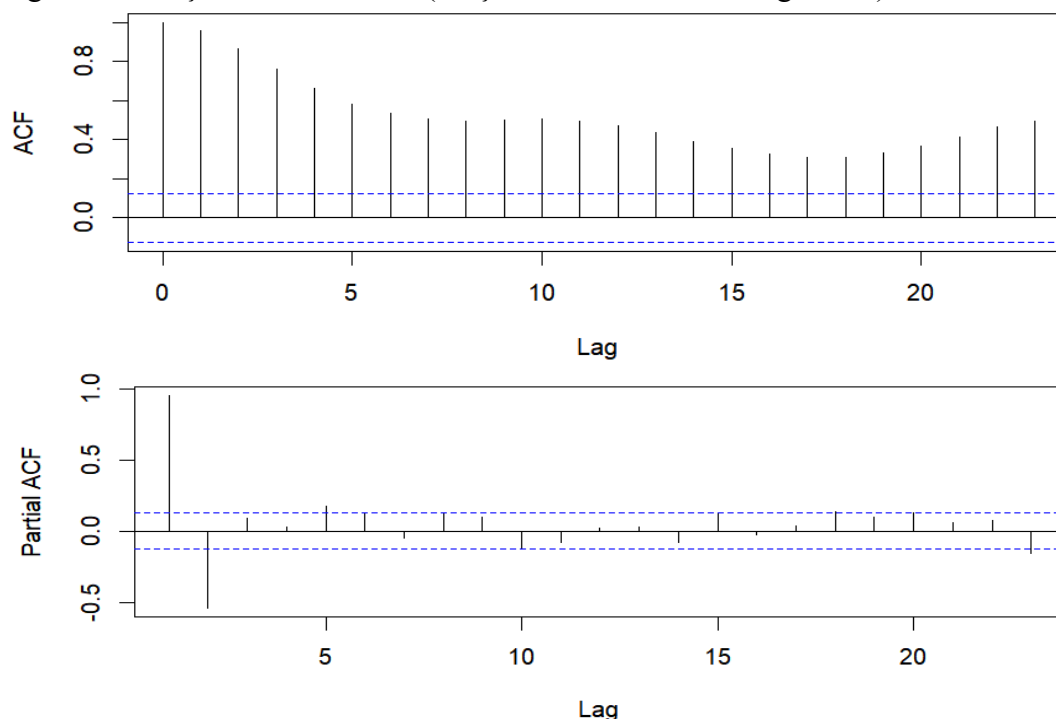
Os testes de raiz unitária indicaram que o logaritmo do preço do leite apresenta comportamento estacionário em torno de uma tendência determinística. O teste de Dickey-Fuller aumentado (ADF) rejeitou a hipótese nula de raiz unitária ($p = 0,01$), enquanto o teste de Kwiatkowski-Phillips-Schmidt-Shin (KPSS) rejeitou a hipótese de estacionariedade em nível ($p < 0,01$), mas não rejeitou a hipótese de estacionariedade em torno de uma tendência determinística ($p \geq 0,10$). De forma complementar, os resultados do teste ADF com especificações alternativas reforçam essa evidência: na versão com tendência determinística (type = “trend”), a hipótese nula de raiz unitária foi rejeitada ($\tau_3 = -4,79$), ao passo que na versão com apenas intercepto (type = “drift”) não houve rejeição ($\tau_2 = -2,33$).

Com base nestes resultados foi especificado um modelo SARIMA com $d = 0$, incluindo a tendência como regressor determinístico (xreg) e tratando a sazonalidade por meio dos componentes sazonais $(P, D, Q)_{12}$. O teste de OCSB foi empregado para confirmar esta escolha¹.

A análise gráfica das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série temporal do logaritmo do preço real do leite destaca aspectos importantes da dependência temporal e estacionariedade dos dados. No caso da série em análise, os gráficos na Figura 5 indicam uma ACF que apresenta um decaimento lento enquanto a PACF apresenta um corte abrupto após a primeira defasagem. Este comportamento revela evidências de um modelo autorregressivo.

¹ O teste de OCSB (*Osborn-Chui-Smith-Birchenhall test*) verifica se há uma raiz unitária no operador de defasagem sazonal $(1 - L^s)$, onde s é o período (por exemplo, 12 meses).

Figura 5 - Funções ACF e PACF (Preço Real do Leite, em logaritmo).



Fonte: Dados da Pesquisa.

4.4. Resultados do modelo ARIMA

Foram testadas diferentes especificações dos modelos. Aqui são apresentados, de forma resumida, os resultados dos modelos com os melhores resultados.

A comparação entre diferentes especificações do modelo ARIMA evidencia resultados consistentes com o diagnóstico preliminar de tendência estacionária da série do preço do leite (em log). Entre as especificações avaliadas, os modelos ARIMA (5,0,0) e ARIMA (2,0,2), ambos com tendência determinística, apresentaram os melhores desempenhos segundo os critérios de informação (AIC e BIC), com valores de AIC em torno de -879 e -878, respectivamente (Tabela 2). Esses resultados indicam que a série foi adequadamente modelada sem diferenciação ($d = 0$), desde que a tendência determinística seja incluída como regressor.

Adotando como critério os menores erros de previsão, o modelo SARIMA(1, 1, 1)(1,1,1)₁₂ apresentou o melhor desempenho entre todas as especificações testadas, registrando o menor RMSE (0,0358) e MAPE (3,47%). Esse resultado indica um melhor desempenho preditivo de curto prazo. Por sua vez, especificações que incluíram uma variável exógena – o preço do milho no presente caso – não resultaram em melhorias significativas.

Tabela 2 - Comparação de Modelos ARIMA para Previsão do Preço do Leite

Especificação	AIC	BIC	RMSE	MAPE	LogLik
ARIMA(5,0,0)(0,0,0)[12] + trend	-879,53	-851,68	0,0372	3,80	447,76
ARIMA(2,0,2)(0,0,0)[12] + trend	-878,17	-853,8	0,0375	3,80	446,08
ARIMA(2,0,2)(1,0,0)[12] + trend	-876,45	-848,6	0,0375	3,81	446,22
ARIMA(2,1,1) + trend + milho	-865,08	-847,69	0,0386	3,80	437,54
ARIMA(2,1,1)(0,0,0)[12] + trend	-863,95	-850,04	0,0389	3,85	435,98
ARIMA(1,1,1)(0,0,0)[12] + trend	-840,23	-826,32	0,0409	4,10	424,11
ARIMA(1,1,1)(1,1,1)[12] + trend	-805,46	-788,34	0,0358	3,47	407,73

Fonte: Dados da Pesquisa.

Embora o modelo SARIMA(1, 1, 1)(1,1,1)₁₂ tenha apresentado as menores medidas de RMSE e MAPE, optou-se pelo ARIMA(5,0,0) com tendência determinística devido à sua superioridade nos critérios de informação AIC e BIC, que penalizam a complexidade e são mais adequados para inferência. O modelo ARIMA (5,0,0) ajustado ao logaritmo do preço do leite, com inclusão de uma tendência determinística (Tabela 3), revelou uma estrutura autorregressiva de ordem elevada, indicando forte persistência temporal. O termo de tendência (“drift”), é positivo e altamente significativo (0,002; $p < 0,001$), coerente com os modelos de regressão apresentados anteriormente.

Tabela 3 - Resultados da Estimação ARIMA(5,0,0) com Componente de Tendência

Parâmetro	Coefficiente	Erro Padrão	Valor-z	p-valor
AR1	1,5069***	0,0640	23,5480	0,0000
AR2	-0,7955***	0,1154	-6,8950	0,0000
AR3	0,2993*	0,1248	2,3990	0,0164
AR4	-0,2794*	0,1152	-2,4250	0,0153
AR5	0,1076·	0,0642	1,6750	0,0939
intercepto	0,5172***	0,0295	17,5240	0,0000
drift	0,002***	0,0002	8,8250	0,0000

Fonte: Dados da Pesquisa.

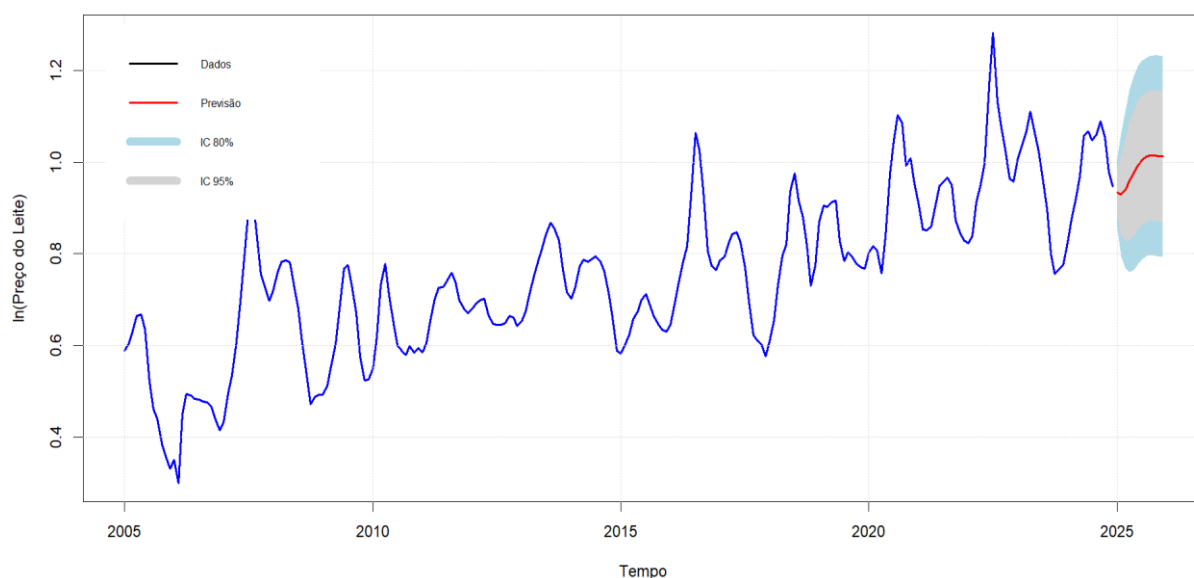
Níveis de significância: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1.

Os coeficientes autorregressivos até a ordem 4 são estatisticamente significantes e descrevem uma forte inércia de curto prazo, sugerindo que o preço reage a choques passados com ajustamentos alternadamente positivos e negativos até estabilizar-se em torno da tendência.

De forma mais direta, tem-se que os choques no preço do leite têm efeitos temporários, dissipando-se gradualmente em aproximadamente quatro meses.

O gráfico da Figura 6 apresenta as previsões do modelo ARIMA ajustado ao logaritmo do preço real do leite, com horizonte de projeção de 12 meses. As previsões (linha vermelha) seguem uma trajetória levemente ascendente, refletindo a tendência determinística positiva identificada nos modelos anteriores. As faixas de confianças de 80% (azul claro) e 95% (cinza) indicam o intervalo esperado de variação dos preços futuros. Em termos práticos, o modelo projeta manutenção dos níveis atuais de preço com leve tendência de alta, compatível com o padrão histórico de recuperação gradual após quedas sazonais.

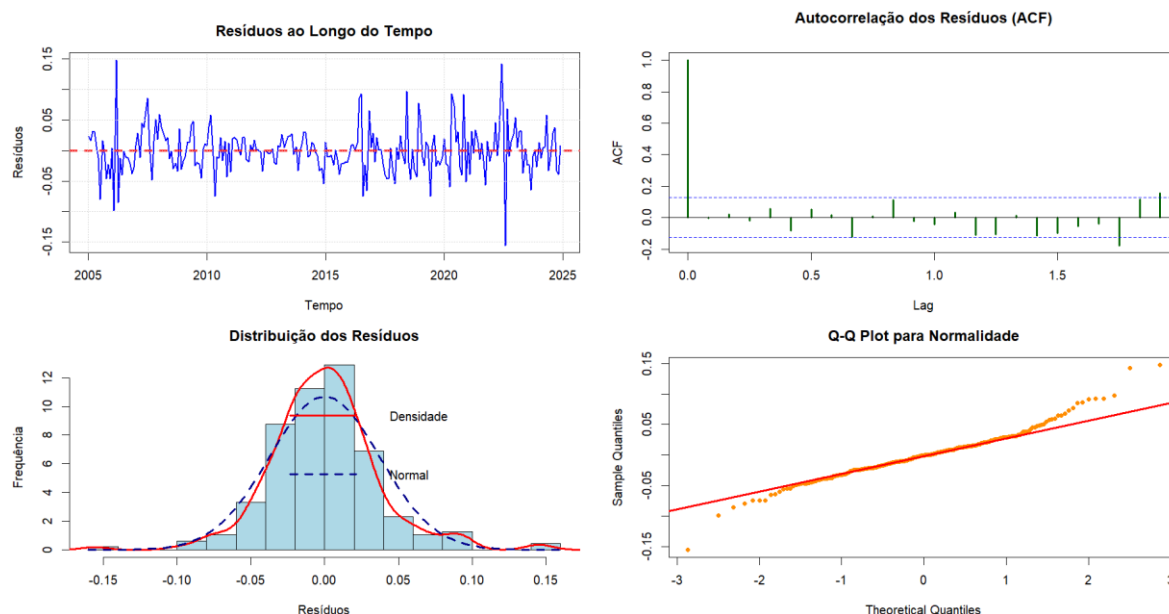
Figura 6 - Previsão do preço real do leite (em log).



Fonte: Dados da Pesquisa.

A avaliação dos resíduos (Figura 7) indica que o modelo não captura integralmente a dependência temporal. O teste de Ljung-Box ($Q=43,9$; $p=0,0078$) rejeita a hipótese de ruído branco, e a ACF mostra autocorrelação remanescente, sobretudo na primeira defasagem. Além disso, o teste de Shapiro-Wilk ($W=0,963$; $p<0,001$) e o Q-Q plot evidenciam desvios da normalidade, com caudas mais pesadas. Esses resultados sugerem a necessidade de ajustes na especificação. Refinamentos futuros nessa modelagem podem tornar os resíduos compatíveis com ruído branco e melhorar a confiabilidade da inferência e das previsões.

Figura 7 - Diagnóstico dos resíduos do modelo



Fonte: Dados da Pesquisa.

5. Considerações Finais

O presente artigo apresentou uma análise do comportamento do preço real do leite no Brasil iniciou-se com a aplicação de modelos de regressão linear múltipla, serem modelos bastante intuitivos e interpretáveis. Buscando aprimorar a capacidade explicativa e preditiva da análise, foram posteriormente empregados modelos de séries temporais, seguindo a metodologia Box-Jenkins, incluindo especificações ARIMA, SARIMA e ARIMAX.

Os resultados obtidos com a aplicação de modelos de regressão linear apresentaram evidências de que o comportamento do preço real do leite no Brasil é determinado por fatores de custo (especialmente o preço do milho), por padrões sazonais e de tendência de longo prazo. Os modelos de regressão linear confirmaram uma trajetória ascendente média de cerca de 2,4% ao ano e significância de fatores sazonais, captados por *dummies* para os meses do ano.

Por sua vez, modelos de séries temporais da família ARIMA foram aplicados com a pretensão de aprimorar a análise. Entre as diferentes especificações avaliadas, o modelo ARIMA (5,0,0) com tendência determinística apresentou o melhor desempenho segundo os critérios AIC e BIC. O termo de tendência mostrou-se altamente significativo e coerente com a hipótese de crescimento gradual dos preços reais. Por outro lado, os diagnósticos de resíduos indicaram autocorrelação remanescente e leve desvio da normalidade, sugerindo espaço para aprimoramentos.

As previsões obtidas projetam estabilidade e leve elevação nos preços do leite, compatível com o comportamento histórico da série. Esses resultados reforçam o potencial dos modelos ARIMA para fins de monitoramento e planejamento de curto prazo no setor leiteiro.

Referências

ALVES, Frederick Fagundes; SOUSA, Lucas Vitor de Carvalho; ERVILHA, Gabriel Teixeira. **Planejamento e previsão do preço do leite em Minas Gerais: análise empírica com base no modelo X12-Arima**. *Revista de Economia e Agronegócio*, v. 12, p. 115–134, 2014.

ABIMILHO – ASSOCIAÇÃO BRASILEIRA DOS PRODUTORES DE MILHO. **Estatísticas**. Disponível em: <<https://www.abimilho.com.br/estatisticas>>. Acesso em: 1 abr. 2025.

BOSCHIERO, Beatriz Nastaro. **Subprodutos da soja**. *Agroadvance*, 2023. Atualizado em 2025. Disponível em: <<https://agroadvance.com.br/blog-subprodutos-da-soja/#:~:text=Dados%20do%20fechamento%20do%20ano,Figura%201>>. Acesso em: 1 abr. 2025.

CARVALHO, Bruno Henrique Picon de; ROSADO, Patrícia Lopes; SHIKI, Simone de Faria Narciso; CRUZ, Aline Cristina da. **Integração intra e inter-mercado: o caso dos preços do leite e derivados no Brasil**. São João del-Rei: Universidade Federal de São João del-Rei, 2013.

CEPEA – CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA. **Indicador da soja**. Disponível em: <<https://www.cepea.esalq.usp.br/br/indicador/soja.aspx>>. Acesso em: 01 abr. 2025a.

CEPEA – CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA. **Indicador do milho**. Disponível em: <<https://www.cepea.esalq.usp.br/br/indicador/milho.aspx>>. Acesso em: 01 abr. 2025b.

CONAB – COMPANHIA NACIONAL DE ABASTECIMENTO. **Análise mensal: leite e derivados**. jan. 2025. Disponível em: <<https://www.conab.gov.br/info-agro/analises-do-mercado-agropecuário-e-extrativista/analises-do-mercado/historico-mensal-de-leite/item/25659-analise-mensal-do-leite-janeiro-de-2024>>. Acesso em: 15 mar. 2025.

EMBRAPA – EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Dados econômicos da soja**. Disponível em: <<https://www.embrapa.br/web/portal/soja/cultivos/soja1/dados-economicos>>. Acesso em: 1 abr. 2025.

FAO – FOOD AND AGRICULTURE ORGANIZATION. **FAOSTAT**. Disponível em: <<https://www.fao.org/faostat/en/#data/QCL/visualize>>. Acesso em: 10 mar. 2025.

FERNANDES, Juliano José de Resende et al. **Farelo de soja em substituição à ureia em dietas para bovinos de corte em crescimento**. *Ciência Animal Brasileira*, Goiânia, v. 10, n.

2, p. 373–378, 2009. Disponível em: <<https://revistas.ufg.br/vet/article/view/6583>>. Acesso em: 2 abr. 2025.

FGV – FUNDAÇÃO GETULIO VARGAS. **FGVDados** – Consulta de séries históricas econômicas. Rio de Janeiro: FGV IBRE, 2025. Disponível em: <https://extra-ibre.fgv.br/IBRE/sitefgvdados/Consulta.aspx>. Acesso em: 21 mar. 2025.

GRIGOL, Natália Salaro. **Ajustes metodológicos na pesquisa do preço do leite cru do Cepea garantem evolução conjunta com cadeia leiteira**. *Cepea*, 9 jan. 2023. Disponível em: <<https://www.cepea.esalq.usp.br/br/opinioao-cepea/ajustes-metodologicos-na-pesquisa-do-preco-do-leite-cru-do-cepea-garantem-evolucao-conjunta-com-cadeia-leiteira.aspx>>. Acesso em: 2 abr. 2025.

HOFFMANN, Rodolfo et al. **Administração da empresa agrícola**. 5. ed. São Paulo: Pioneira, 1987.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Indicadores IBGE: estatística da produção pecuária**. Rio de Janeiro: IBGE, out.–dez. 2024. Publicado em: 18 mar. 2025. Disponível em: <<https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=72380>>. Acesso em: 1 abr. 2025.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Valor da produção da pecuária e aquicultura chega a R\$ 122,4 bilhões em 2023**. Agência de Notícias IBGE, 18 out. 2024 Disponível em: < <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/41352-valor-da-producao-da-pecuaria-e-aquicultura-chega-a-r-122-4-bilhoes-em-2023> >. Acesso em: 7 mar. 2025.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Tabela 74 - Produção de origem animal, por tipo de produto**. Pesquisa da Pecuária Municipal. SIDRA – Sistema IBGE de Recuperação Automática. Disponível em: <https://sidra.ibge.gov.br/Tabela/74>. Acesso em: 21 mar. 2025.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Tabela 94 - Vacas ordenhadas**. Pesquisa da Pecuária Municipal. SIDRA – Sistema IBGE de Recuperação Automática. Disponível em: <https://sidra.ibge.gov.br/tabela/94>. Acesso em: 21 out. 2025.

JUNQUEIRA, Rafael Villela Bastos; ZOCCAL, Rosangela; MIRANDA, João Eustáquio Cabral de. **Análise da sazonalidade na produção de leite no Brasil**. *Engormix*, 2013. Disponível em: <https://pt.engormix.com/pecuaria-leite/qualidade-leite/analise-sazonalidade-producao-leite_a38244/>. Acesso em: 14 mar. 2025.

MOURA, André Morais. **Milho diferindo no processamento para vacas leiteiras em pastejo**. 2013. Tese (Mestrado em Zootecnia – Área de concentração: Produção Animal) – Universidade Federal de Minas Gerais, Belo Horizonte.

NEIVA, Rubens Antônio. **Pecuária de leite vive incertezas quanto ao preço de insumos**. *Embrapa Gado de Leite*, 2021. Disponível em: <<https://www.embrapa.br/busca-de-noticias/-/noticia/63586475/pecuaria-de-leite-vive-incertezas-quanto-ao-preco-de-insumos>>. Acesso em: 2 abr. 2025.

ROCHA, Denis Teixeira da; CARVALHO, Glauco Rodrigues; RESENDE, João Cesar de. **Cadeia produtiva do leite no Brasil: produção primária**. Juiz de Fora: Embrapa Gado de Leite, 2020. 15 p. (Circular Técnica, 123).

SIMÕES, Andre Rozemberg Peixoto; NICHOLSON, Charles Frederick; CARVALHO, Glauco Rodrigues. **Components of farm milk price behavior in Brazil from 2005 to 2020**. *Revista Brasileira de Zootecnia*, v. 52, e20220116, 2023. DOI: <<https://doi.org/10.37496/rbz5220220116>>.

SIQUEIRA, Kennya Bestriz; SILVA, Carlos Arthur Barbosa da; AGUIAR, Danilo Rolim Dias de. **Viability of introducing milk futures contracts in Brazil: a multiple criteria decision analysis**. *Agribusiness*, v. 24, p. 491–509, 2008. DOI: <<https://doi.org/10.1002/agr.20175>>.

VILELA, Duarte; LEITE, José Luiz Bellini; RESENDE, João Cesar de. **Políticas para o leite no Brasil: passado, presente e futuro**. *Simpósio sobre Sustentabilidade da Pecuária Leiteira na Região Sul do Brasil*, v. 1, p. 1–26, 2002.

VILELA, Darcy et al. **Pecuária de leite no Brasil: cenários e avanços tecnológicos**. Brasília, DF: Embrapa, 2016.