

Ainda Estou Aqui Para Prever: uma comparação entre SARIMA e SARIMAX para vendas sazonais de aves no Brasil

I'm Still Here to Forecast: a comparison between SARIMA and SARIMAX for seasonal poultry sales in Brazil

Todavía Estoy Aquí Para Predecir: una comparación entre SARIMA y SARIMAX para las ventas estacionales de aves en Brasil

Como citar:

Lucas, Edimilson C.; Yoshikuni, Adilson C.; Agustini, Carlos A. D. & Silva, Vinicius A. B. (2025). Ainda estou aqui para prever: uma comparação entre SARIMA e SARIMAX para vendas sazonais de aves no Brasil. Revista Gestão & Tecnologia, vol. 25, nº 3, p:98-118

Edimilson Costa Lucas, Professor do Departamento de Ciências Atuariais da EPPEN/UNIFESP.

<https://orcid.org/0000-0003-0533-6067>

Adilson Carlos Yoshikuni, Professor do Mestrado e Doutorado Profissional em Controladoria, Finanças e Tecnologias de Gestão Empresariais.

<https://orcid.org/0000-0003-4611-6933>

Carlos Alberto Di Agustini, Professor pesquisador da Strong Business School e Professor da USCS

<https://orcid.org/0000-0002-3075-5370>

Vinicius Augusto Brunassi Silva, Professor e Pesquisador da FECAP.

<https://orcid.org/0000-0002-1299-321X>

"Os autores declaram não haver qualquer conflito de interesse de natureza pessoal ou corporativa, em relação ao tema, processo e resultado da pesquisa".

Editor Científico: José Edson Lara
Organização Comitê Científico
Double Blind Review pelo SEER/OJS
Recebido em 07/02/2025, Aprovado em 20/06/2025



Resumo

Objetivo: Essa pesquisa tem por objetivo comparar a aplicação dos modelos econométricos SARIMA e SARIMAX na previsão de demanda de produtos sazonais, utilizando como base real de dados as vendas mensais de aves congeladas de duas marcas concorrentes no Brasil.

Metodologia: A metodologia empregada baseia-se na modelagem de séries temporais com os modelos SARIMA e SARIMAX. A abordagem adotou o procedimento de separação dos dados em conjuntos de treinamento e teste, técnica amplamente aplicada em estudos empíricos com foco em previsão fora da amostra, para uma averiguação da robustez dos modelos, em linha com as fundamentações das técnicas de *machine learning*.

Originalidade/Relevância: A originalidade do trabalho reside na aplicação desses modelos à previsão de demanda de aves natalinas de uma empresa brasileira, um nicho pouco explorado na literatura nacional, além de incorporar uma análise com variável exógena de concorrência direta. O estudo também contribui ao aproximar práticas tradicionais de séries temporais à lógica preditiva do aprendizado de máquina.

Principais resultados: Os resultados demonstraram desempenho superior do modelo SARIMA, com menor erro de previsão e maior capacidade de generalização, mesmo sem o uso de variáveis externas.

Contribuições teóricas/metodológicas: A pesquisa reforça a aplicabilidade de modelos sazonais em mercados com forte ciclicidade e mostra que, em certos contextos, a simplicidade estatística pode superar a complexidade multivariada, mesmo em ambientes voláteis e competitivos.

Palavras-chave: Séries Temporais, SARIMA, SARIMAX, Sazonalidade.

Abstract

Objective: This study aims to compare the application of SARIMA and SARIMAX econometric models in forecasting the demand for seasonal products, using real monthly sales data of frozen poultry from two competing brands in Brazil.

Methodology: The methodology is based on time series modeling using SARIMA and SARIMAX models. The approach involved splitting the dataset into training and testing sets, a procedure widely adopted in empirical studies focused on out-of-sample forecasting, to assess model robustness, in line with the principles of machine learning validation techniques.

Originality/Relevance: The originality of this research lies in the application of these models to forecast holiday-season poultry demand in a Brazilian company—a niche rarely explored in national literature—while also incorporating an analysis with an exogenous variable representing direct market competition. The study contributes by bridging traditional time series practices with predictive logic inspired by machine learning.

Main Results: The results demonstrated superior performance of the SARIMA model, with lower forecasting errors and greater generalization capability, even without the use of external variables.

Theoretical/Methodological Contributions: The research reinforces the applicability of seasonal models in markets characterized by strong cyclicity and shows that, in certain

contexts, statistical simplicity can outperform multivariate complexity, even in volatile and competitive environments.

Keywords: Time Series, SARIMA, SARIMAX, Seasonality.

Resumen

Propósito: Esta investigación tiene como objetivo comparar la aplicación de los modelos econométricos SARIMA y SARIMAX en la previsión de la demanda de productos estacionales, utilizando como base datos reales de ventas mensuales de aves congeladas de dos marcas competidoras en Brasil.

Metodología: La metodología se basa en la modelización de series temporales mediante los modelos SARIMA y SARIMAX. La estrategia adoptada incluyó la división de los datos en conjuntos de entrenamiento y prueba, una técnica ampliamente utilizada en estudios empíricos con énfasis en la predicción fuera de la muestra, con el fin de verificar la robustez de los modelos, en consonancia con los fundamentos de las técnicas de aprendizaje automático.

Originalidad/Relevancia: La originalidad del estudio reside en la aplicación de estos modelos a la previsión de la demanda de aves navideñas de una empresa brasileña, un nicho poco explorado en la literatura nacional, además de incorporar una variable exógena que representa la competencia directa en el mercado. El estudio también contribuye al aproximar prácticas tradicionales de series temporales a la lógica predictiva del aprendizaje automático.

Principales Resultados: Los resultados mostraron un rendimiento superior del modelo SARIMA, con menor error de previsión y mayor capacidad de generalización, incluso sin el uso de variables externas.

Contribuciones Teóricas/Metodológicas: La investigación refuerza la aplicabilidad de modelos estacionales en mercados con fuerte ciclicidad y demuestra que, en ciertos contextos, la simplicidad estadística puede superar la complejidad multivariada, incluso en entornos volátiles y competitivos.

Palabras clave: Series Temporales, SARIMA, SARIMAX, Estacionalidad.

Introdução

A previsão de demanda com base em séries temporais representa uma prática essencial no planejamento estratégico das organizações, especialmente em setores nos quais a variabilidade da demanda afeta diretamente os níveis de produção, estoque e atendimento ao cliente. Ao permitir a antecipação de padrões futuros de consumo, uma previsão eficaz contribui para a eficiência operacional, a redução de custos logísticos e o aumento da competitividade (Khalid, 2024; Bertaglia, 2003; Slack, Chambers & Johnston, 1997).

Não obstante, a obtenção de previsões precisas é desafiadora, uma vez que os dados de vendas apresentam características complexas, como sazonalidade, tendências não lineares e

flutuações externas imprevisíveis. Métodos tradicionais de suavização, como a Média Móvel Exponencialmente Ponderada, podem ser eficazes em ambientes mais estáveis, mas tendem a falhar quando aplicados a séries temporais com estruturas sazonais acentuadas ou impactadas por fatores exógenos (Zhou et al., 2021; Tubino, 2000).

Nesse contexto, os modelos da classe Box-Jenkins, como o ARIMA (*Autoregressive Integrated Moving Average*), têm sido amplamente utilizados por sua capacidade de modelar dependências temporais. Todavia, quando há padrões sazonais e a influência de variáveis externas, modelos como SARIMA (*Seasonal ARIMA*) e SARIMAX (*Seasonal ARIMA with Exogenous Variables*) se tornam alternativas metodológicas mais apropriadas (Tamura, 2013; Alencar, 2022). O SARIMA é eficaz na modelagem de sazonalidade explícita, enquanto o SARIMAX permite a incorporação de variáveis exógenas, o que pode enriquecer o processo preditivo (Abirami, 2024; Kulkarni & Rane, 2020; Handayani, Maslim & Mudjihartono, 2020).

Apesar da comprovada eficácia desses modelos em diversas aplicações empíricas, seu uso ainda é restrito no contexto empresarial brasileiro, sobretudo em mercados sazonais como o de alimentos congelados. Dessa forma, nesse estudo, realiza-se uma análise comparativa entre os modelos SARIMA e SARIMAX na previsão de vendas mensais de aves congeladas de duas marcas reais concorrentes, com foco em períodos natalinos.

O diferencial metodológico da pesquisa reside na adoção de uma abordagem baseada em previsão fora da amostra (*out-of-sample*) (Hawinkel, Waegeman & Maere, 2023), por meio da divisão dos dados em conjuntos de treinamento e teste. Esse procedimento, comum em estudos empíricos de validação de modelos, permite avaliar a capacidade de generalização dos métodos testados — aspecto fundamental em ambientes de alta variabilidade (Lucas, Mendes-da-Silva & Lyons, 2017).

Vale a pena ressaltar que, além da escolha apropriada do modelo econométrico, um aspecto central em estudos de previsão é a forma como os modelos são validados. A simples capacidade de um modelo ajustar-se bem aos dados históricos (ajuste *in-sample*) não garante sua utilidade preditiva em dados futuros. Por isso, diversos autores destacam a importância da validação fora da amostra (*out-of-sample validation*) como uma estratégia essencial para testar a capacidade de generalização do modelo (Hyndman & Athanasopoulos, 2020), em linha com

as metodologias aplicadas em aprendizado de máquina. Essa abordagem consiste em dividir a série temporal em duas partes: uma usada para estimar os parâmetros do modelo (conjunto de treinamento) e outra utilizada exclusivamente para avaliação da previsão (conjunto de teste). O uso dessa técnica reduz o risco de superajuste (*overfitting*), situação em que o modelo capta ruídos específicos da amostra histórica e perde poder de generalização para novos dados. O fenômeno do *overfitting* é particularmente preocupante quando se utilizam modelos com muitas variáveis explicativas ou estruturas excessivamente flexíveis, como pode ocorrer em aplicações do modelo SARIMAX. Nesses casos, embora o modelo possa apresentar ótimo desempenho no ajuste aos dados de treinamento, seu desempenho fora da amostra pode deteriorar-se significativamente.

Nesse sentido, ganha destaque o papel dos critérios de seleção de modelos, como o Critério de Informação de Akaike (AIC), que penaliza modelos excessivamente complexos e privilegia aqueles com melhor equilíbrio entre ajuste e parcimônia. A parcimônia, por sua vez, é um princípio estatístico que favorece modelos mais simples sempre que eles forem capazes de capturar adequadamente os padrões dos dados. Segundo Box et al. (2015), modelos com menor número de parâmetros tendem a ser mais robustos, especialmente quando se lida com séries curtas ou altamente sazonais. Assim, a escolha entre SARIMA e SARIMAX não deve se basear apenas em sua sofisticação teórica, mas na adequação empírica ao contexto analisado, no comportamento dos resíduos, nas métricas de erro e na estabilidade das previsões fora da amostra.

Os resultados indicam que, embora o SARIMAX incorpore uma variável exógena (vendas do concorrente), o modelo SARIMA apresentou desempenho superior em termos de acurácia e estabilidade preditiva (Lee & Bang, 2024). Tais achados reforçam a relevância da modelagem econométrica clássica com ajustes sazonais adequados, oferecendo contribuições tanto para a literatura acadêmica quanto para a prática gerencial em ambientes com forte sazonalidade e competição de mercado.

Referencial teórico

A previsão de demanda é um dos pilares da gestão operacional e estratégica nas organizações modernas. Quando baseada em séries temporais, a previsão permite identificar padrões recorrentes, antecipar variações sazonais e projetar tendências futuras, tornando-se essencial para empresas que desejam manter níveis ótimos de estoque, reduzir perdas e melhorar seu nível de serviço (Liu et al., 2021; Bertaglia, 2003; Slack et al., 1997).

Entretanto, a previsão em ambientes reais envolve desafios consideráveis. As séries de vendas podem conter ruídos, comportamentos sazonais irregulares, *outliers* e efeitos causados por fatores exógenos como clima, políticas promocionais e ações da concorrência. Por esse motivo, a literatura tem buscado métodos que transcendam as abordagens tradicionais e que ofereçam maior poder de explicação e previsão (Zhang et al., 2021; Tubino, 2000).

Entre os métodos tradicionais, a Média Móvel Exponencialmente Ponderada continua a ser amplamente utilizada, especialmente por sua simplicidade de implementação. No entanto, sua eficácia é limitada quando confrontada com séries que apresentam padrões de longo prazo, ciclos ou sazonalidade acentuada (Sukparungsee et al., 2020). Nesse sentido, métodos mais sofisticados, como os modelos ARIMA, vêm ganhando destaque desde a década de 1970, a partir das contribuições de Box e Jenkins (Oh & Seong, 2024; Bianchi et al., 1998).

O modelo ARIMA permite capturar a autocorrelação presente nas séries, tornando-o adequado para séries estacionárias ou que podem ser tornadas estacionárias por diferenciação. Contudo, quando a série apresenta sazonalidade ou sofre influência de variáveis externas, torna-se necessário o uso de modelos mais abrangentes, como o SARIMA e o SARIMAX (Manigandan et al., 2021; Adli, 2020).

O SARIMA incorpora explicitamente os componentes sazonais, sendo especialmente eficaz em séries com padrões cíclicos recorrentes. Já o SARIMAX, ao adicionar variáveis exógenas ao modelo, expande sua capacidade explicativa e preditiva (Elshewey et al., 2022). Ao considerar variáveis externas - como ações da concorrência, indicadores econômicos ou variações climáticas - o modelo passa a capturar efeitos que não são observáveis diretamente na estrutura autoregressiva da série principal (Wahyudi & Febriani, 2024; Tamura, 2013).

Embora ainda pouco explorados no contexto de previsão de vendas no Brasil, alguns estudos demonstram a aplicabilidade e o potencial desses modelos. Lizzi (2012), por exemplo, utilizou SARIMA e SARIMAX para prever casos de dengue, incorporando variáveis climáticas. Walter (2013) aplicou SARIMA na previsão da frota de motocicletas, evidenciando sua robustez para séries de vendas. No contexto agrícola, Divisekara (2020) utilizou SARIMA para prever os preços de lentilhas, mostrando a utilidade do modelo na formulação de decisões de produção e comercialização.

Além disso, estudos comparativos, como o de previsão do consumo de gás natural nos Estados Unidos, demonstraram que, em determinadas circunstâncias, o modelo SARIMAX supera o SARIMA em precisão, especialmente quando variáveis exógenas relevantes são incluídas de forma adequada assim como feito em Manigandan et al. (2021) e Broni-Bediako et al. (2024).

Ao trazer esse arcabouço para o contexto do varejo brasileiro, especialmente em cadeias alimentares com forte sazonalidade e competição intensa entre marcas, a escolha entre SARIMA e SARIMAX torna-se uma decisão crítica. A inclusão de variáveis exógenas, como o desempenho da concorrência, políticas de preço e promoções, pode ser decisiva para capturar a real dinâmica da demanda (Serrano et al., 2024; Wahyudi & Febriani, 2024; Alharbi & Csala, 2022; Rocha, 2014).

Enfatiza-se que é importante não negligenciar o papel do julgamento humano. Fatores qualitativos, como expectativas de mercado, lançamentos de produtos e campanhas promocionais, frequentemente escapam aos métodos quantitativos, sendo melhor percebidos por analistas experientes. A combinação de técnicas estatísticas robustas com a experiência do gestor pode resultar em previsões mais acuradas e decisões mais precisas (Elamin & Fukushige, 2018; Wanke, 2006).

A previsão de demanda com base em séries temporais tem ocupado posição central no campo da administração de empresas onde as decisões estratégicas em torno de estoques, compras, produção e distribuição são profundamente impactadas pela capacidade da empresa de antecipar o comportamento da demanda de forma confiável (Bertaglia, 2003; Slack, Chambers & Johnston, 1997).

Ao longo do tempo, diversas abordagens foram desenvolvidas para lidar com esse desafio. Métodos tradicionais, como a Média Móvel Exponencialmente Ponderada, ainda são amplamente utilizados por sua simplicidade e rapidez de implementação. No entanto, esses métodos se mostram limitados quando confrontados com padrões temporais mais complexos, como ciclos irregulares, sazonalidades múltiplas ou variações abruptas causadas por fatores externos (Tubino, 2000; Wanke, 2006).

Nesse sentido, os modelos da família ARIMA (*Autoregressive Integrated Moving Average*), propostos por Box e Jenkins em 1976, representam um avanço significativo ao permitir a modelagem explícita de dependências temporais e estruturas não estacionárias. Esses modelos assumem que o comportamento presente de uma série pode ser explicado por seus próprios valores passados, bem como por erros aleatórios igualmente autocorrelacionados. No entanto, o ARIMA tradicional não é sensível à presença de sazonalidade ou à influência de fatores exógenos, o que limita sua aplicabilidade em determinados contextos (Alencar, 2022).

Como evolução natural dessa abordagem, o modelo SARIMA (*Seasonal ARIMA*) foi desenvolvido para incorporar padrões sazonais explícitos nas séries temporais. Ao adicionar componentes específicos para captar flutuações periódicas, como aquelas observadas em ciclos anuais, trimestrais ou mensais, o SARIMA amplia significativamente o escopo das aplicações do ARIMA, tornando-se uma ferramenta adequada para contextos onde a sazonalidade desempenha papel determinante no comportamento da demanda (Nau, 2014; Box et al., 2015).

De forma mais abrangente, o modelo SARIMAX (*Seasonal ARIMA with Exogenous Variables*) inclui a possibilidade de introduzir variáveis externas, enriquecendo o processo preditivo ao considerar fatores que não estão diretamente refletidos na série temporal original. Esse tipo de modelagem multivariada é particularmente útil em cenários nos quais elementos como preços da concorrência, campanhas promocionais ou indicadores macroeconômicos influenciam a demanda (Adli, 2020; Tamura, 2013).

Apesar de sua relevância teórica, os modelos SARIMA e SARIMAX ainda são pouco explorados em estudos aplicados à previsão de vendas no contexto brasileiro. Em uma das poucas investigações nacionais, Lizzi (2012) demonstrou a utilidade do SARIMAX na previsão de casos de dengue, utilizando variáveis climáticas como entradas exógenas. De forma

semelhante, Walter (2013) aplicou o SARIMA à previsão da demanda por motocicletas no Brasil, revelando sua capacidade de capturar sazonalidades de maneira eficaz.

No cenário internacional, Divisekara (2020) utilizou o SARIMA para prever preços futuros de lentilhas com base em dados semanais coletados ao longo de nove anos. O modelo foi capaz de captar oscilações sazonais e fornecer suporte à tomada de decisão por parte de produtores e consumidores. Na mesma linha, Manigandan et al. (2021) compararam o desempenho de SARIMA e SARIMAX na previsão de produção e consumo de gás natural nos Estados Unidos, concluindo que o SARIMAX obteve melhor desempenho em função da inclusão de variáveis explicativas relevantes.

A partir dessas evidências, percebe-se que tanto o SARIMA quanto o SARIMAX podem ser valiosos em contextos de previsão de vendas, desde que as características das séries e a disponibilidade de variáveis externas sejam consideradas. No ambiente varejista brasileiro, por exemplo, produtos com forte apelo sazonal, como aves natalinas, estão sujeitos não apenas a ciclos regulares de demanda, mas também à concorrência direta entre marcas, campanhas promocionais e flutuações no poder de compra do consumidor.

Além das variáveis quantificáveis, a literatura também aponta para a importância da dimensão subjetiva na previsão de demanda. Conforme argumenta Wanke (2006), o julgamento humano, baseado na experiência do gestor, no conhecimento do mercado e na percepção de fatores intangíveis, desempenha papel complementar às técnicas quantitativas. Promoções específicas, reações inesperadas do mercado ou mudanças abruptas no comportamento dos consumidores são exemplos de situações que, muitas vezes, escapam ao alcance dos modelos estatísticos, por mais sofisticados que sejam.

Dessa forma, o presente estudo se insere em um espaço relevante da literatura ao aplicar e comparar os modelos SARIMA e SARIMAX na previsão de vendas sazonais de aves congeladas, incorporando uma variável exógena de concorrência direta no modelo SARIMAX. A análise contribui para uma compreensão mais ampla da aplicabilidade desses modelos no setor varejista e ressalta a importância da modelagem adequada às características específicas de cada contexto de negócio.

Metodologia

Conforme apresentado, esse estudo adota uma abordagem quantitativa aplicada à modelagem de séries temporais com o objetivo de comparar o desempenho preditivo dos modelos SARIMA e SARIMAX na previsão de vendas sazonais. O foco recai sobre uma série mensal de vendas de aves congeladas, produto tipicamente comercializado em maior volume no período natalino. A pesquisa fundamenta-se em dados reais disponibilizados por uma empresa brasileira do setor alimentício.

Estrutura dos modelos

O modelo ARIMA é considerado para prever os valores futuros de uma função linear com valores passados e erros aleatórios, sendo um dos métodos mais comuns utilizados na análise de séries temporais (Manigandan, 2021). O modelo ARIMA consiste nos termos modelo autorregressivo (AR), integração (I) e média móvel (MA) (Ediger, 2017). Como aditivo, o ARIMAX propôs a expansão do modelo ARIMA (Bierens, 1987). Neste método, são aplicadas variáveis independentes adicionais além do método ARIMA. O método ARIMAX pode ser especificado da seguinte forma (Jalalkamali & Moradi, 2015):

$$y_t = \alpha + \sum_{i=1}^m \frac{\gamma_i(G)}{\varphi(G)} GLX^{it} + e_t \quad (1)$$

$$e_t = \frac{\gamma_i(G)}{\varphi(G)} a_t \quad (2)$$

onde $\varphi(G)$ denota os coeficientes AR, $\gamma_i(G)$ representa os coeficientes MA, l_i indica o grau de defasagem, e_t mostra a análise de erros na regressão e a_t a média zero nos termos de erro da série temporal.

Já o modelo SARIMA $(p, d, q) * (P, D, Q)^S$ é um complemento do ARIMA que foi desenvolvido para aprimorar o desempenho da média móvel integrada auto-regressiva na modelagem de séries sazonais. A forma geral do método ARIMA sazonal é expressa como (Nau, 2014):

$$\varphi_p(G)\phi_p(G^{sazonal})(1-G)^d(1-G^{sazonal})^D X_t = \gamma_q(G)\omega_Q(G^{sazonal})e_t \quad (3)$$

As expressões $\varphi_p(G)$ e $\gamma_q(G)$ indicam a ordem dos polinômios característicos dos componentes de autorregressão não sazonais (AR) e das médias móveis não sazonais (MA). $\phi_p(G^{sazonal})$ e $\omega_Q(G^{sazonal})$ representam os polinômios de autorregressão sazonais (SAR) e as médias móveis sazonais (SMA), respectivamente. $(1 - G)$ e $(G^{sazonal})$ indicam os componentes de diferenciação para as séries temporais não sazonais e sazonais, respectivamente. Os termos d e D representam as ordens de diferenciação ordinária do ARIMA não sazonal e as ordens de diferenciação sazonal do ARIMA sazonal da série, respectivamente (Box, 2015). Além disso, X_t representa o valor observado no tempo t , e_t indica os termos de erro da previsão, s é o período do padrão sazonal (por exemplo, $s = 12$ para uma série mensal) e G denota os termos do operador de defasagem.

Os componentes do SARIMA podem ser escritos como:

$$AR: \varphi_p(G) = 1 - \varphi_1(G) - \varphi_2 G^2 - \varphi_3 G^3 - \dots - \varphi_p G^p \quad (4)$$

$$MA: \gamma_q(G) = 1 - \gamma_1(G) - \gamma_2 G^2 - \gamma_3 G^3 - \dots - \gamma_q G^q \quad (5)$$

$$SAR: \phi_p(G^{sazonal}) = 1 - \phi_1 G^{sazonal} - \phi_2 G^{2sazonal} - \phi_3 G^{3sazonal} - \dots - \phi_p G^{Psazonal} \quad (6)$$

$$SMA: \omega_Q(G^{sazonal}) = 1 - \omega_1 G^{sazonal} - \omega_2 G^{2sazonal} - \omega_3 G^{3sazonal} - \dots - \omega_Q G^{Qsazonal} \quad (7)$$

Finalizando os modelos econométricos, temos o modelo SARIMAX, que é um avanço do ARIMA sazonal (SARIMA) com variáveis externas (X) na busca de aprimorar seu desempenho, chamado SARIMAX $(p, d, q) * (P, D, Q)$. O modelo pode ser escrito da seguinte forma (Tarsitano, 2017) apresentado em (8):

$$\varphi_p(G)\phi_p(G^{sazonal})(1 - G)^d(1 - G^{sazonal})^D X_t = \alpha_k y_{k,t} + \gamma_q(G)\omega_Q(G^{sazonal})e_t \quad (8)$$

onde $y_{k,t}$ são as observações correspondentes da k -ésima representada como o número de variáveis exógenas no tempo t , e α_k representa o valor do coeficiente de correlação da k -ésima variável de entrada exógena (X).

O modelo SARIMAX desenvolve a capacidade do modelo SARIMA pela integração de informações contextuais, como atividades econômicas, feriados regionais, clima, vendas da

concorrência, e outros parâmetros que podem impactar diretamente a demanda por determinado produto.

Para avaliar os dois métodos, a medida utilizada para prever os modelos neste estudo é o erro quadrático médio (RMSE) e o erro percentual médio absoluto (MAPE%), (Armstrong, 1992) mostrados nas equações (9) e (10):

$$MAPE\% = \frac{\sum_{t=1}^n \frac{|E_t - F_t|}{E_t}}{n} \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (E_t - F_t)^2} \quad (10)$$

onde E_t e F_t representam, respectivamente, os valores reais e os valores previstos no tempo t e n são os valores estimados.

Apresentação e discussão dos resultados

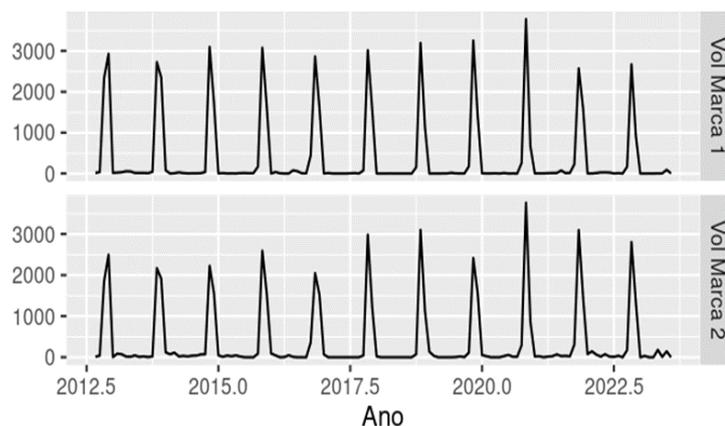
A base de dados utilizada nesta pesquisa foi fornecida por uma grande empresa do setor alimentício no Brasil, contendo registros mensais de vendas de aves congeladas de duas marcas distintas, com forte concentração no período natalino. As duas marcas pertencem ao mesmo grupo empresarial, mas concorrem diretamente nos pontos de venda. A série abrange o período de setembro de 2012 a agosto de 2023, totalizando 123 observações mensais. A marca 1 representa a variável de interesse a ser prevista, enquanto a marca 2 foi utilizada como variável exógena no modelo SARIMAX.

As estatísticas descritivas das duas séries indicam forte presença de sazonalidade, com concentração das vendas nos meses de novembro e dezembro. A mediana da série da Marca 1 é de 10,5 kg, com média de 402,1 kg, enquanto a Marca 2 apresenta mediana de 24 kg e média de 378,1 kg. O valor máximo de vendas para ambas as marcas foi registrado em novembro de 2021, ultrapassando 3700 kg. Esses dados confirmam o comportamento altamente concentrado da demanda.

Tabela 1*Estatísticas descritivas dos dados*

Métricas	Marca 1	Marca 2
Min.	0,00	0,00
1º Qu.	2,00	0,00
Mediana	10,50	24,00
Média	402,11	378,10
3º Qu.	73,75	97,00
Máx.	3779,00	3760,00

A série temporal da Marca 1 foi representada graficamente junto à da Marca 2, permitindo visualização clara da correlação entre os ciclos de vendas, sendo apresentado na Figura 2.

Figura 1*Histórico Mensal de Vendas de Aves Natalinas das Marcas 1 e 2*

Com base nessa estrutura, os modelos SARIMA e SARIMAX foram testados com o objetivo de comparar sua performance preditiva. A base foi dividida em dois subconjuntos: o conjunto de treinamento (setembro de 2012 a setembro de 2022, com 121 observações) e o conjunto de teste (outubro de 2022 a agosto de 2023, com 11 observações). Essa divisão permitiu avaliar a capacidade de generalização dos modelos por meio de previsão fora da amostra (*out-of-sample*).

Antes da modelagem, foi aplicada a verificação da estacionariedade utilizando o teste KPSS. Para ambas as séries, o valor-p reportado foi 0,10, indicando que não há evidência para

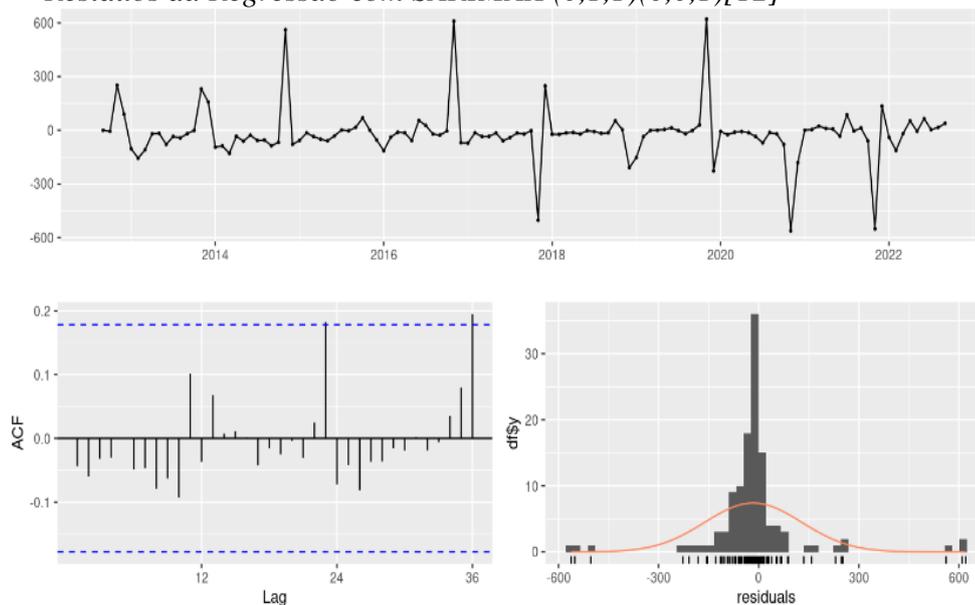
rejeitar a hipótese nula de estacionariedade. Dessa forma, foi possível prosseguir com a modelagem sem a necessidade de transformações adicionais.

O modelo SARIMAX foi estimado considerando a série da Marca 1 como variável dependente e a série da Marca 2 como variável exógena. A função `auto.arima` do pacote `forecast` do `software R`, foi utilizada para identificar a melhor especificação com base no menor AIC. O modelo selecionado foi SARIMAX (0,1,1)(0,0,1)[12], com AIC de 1549,61.

As métricas obtidas no conjunto de treinamento foram: MAPE de 19,67% e RMSE de 394,73. Tais valores indicam um desempenho razoável, com previsão acertando cerca de 80% das vendas em termos absolutos.

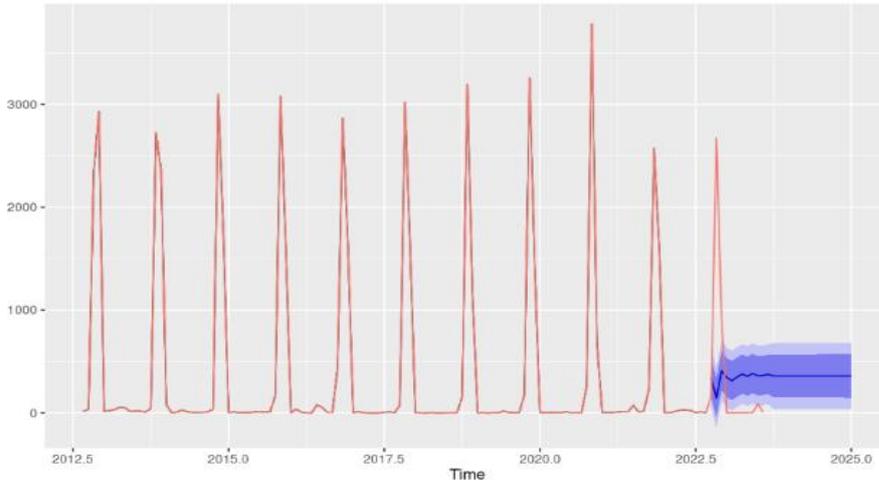
Figura 2

Resíduos da Regressão com SARIMAX (0,1,1)(0,0,1)[12]



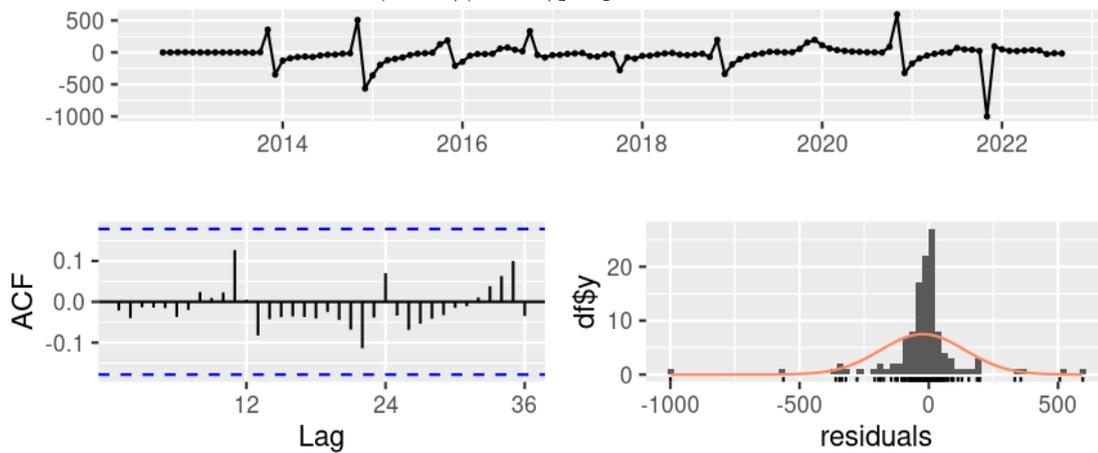
Conforme a Figura 2, a análise dos resíduos por meio da função de autocorrelação (ACF) revelou que quase todas as defasagens estavam dentro dos limites aceitáveis. O teste de Ljung-Box confirmou esse diagnóstico, com p-valor de 0,94, indicando que os resíduos podem ser tratados como ruído branco.

Figura 3
Previsão com SARIMAX (0,1,1)(0,0,1)[12]



Contudo, de acordo com a Figura 4, no conjunto de teste, o modelo SARIMAX mostrou limitação. As previsões geradas para o período de outubro de 2022 a agosto de 2023 apresentaram desvio significativo dos valores reais e não conseguiram capturar o padrão de sazonalidade típico da série, especialmente nos meses que antecedem o Natal.

Figura 4
Resíduos do modelo SARIMA (0,0,1)(1, 1,0)[12]



A modelagem SARIMA foi realizada de forma univariada, considerando apenas os dados da Marca 1. Novamente, foi utilizada a função `auto.arima` com base no AIC. O modelo

selecionado foi SARIMA (0,0,1)(1,1,0)[12], com AIC de 1440,65 — valor inferior ao modelo anterior.

Além disso, o desempenho do SARIMA no conjunto de treinamento foi superior ao SARIMAX: MAPE de 8,32% e RMSE de 164,73. Isso indica previsões muito próximas dos valores reais, com mais de 90% de acerto em termos absolutos.

A verificação dos resíduos também foi satisfatória. Todas as autocorrelações encontravam-se dentro dos intervalos esperados, e o teste de Ljung-Box apresentou p-valor de 0,99, confirmando que os resíduos podem ser considerados ruído branco.

Mais importante ainda, no conjunto de teste, o modelo SARIMA conseguiu prever com precisão os padrões de sazonalidade observados nos anos anteriores, especialmente os picos de vendas nos meses de fim de ano, conforme apresentado na Figura 5.

Figura 5
Previsão com SARIMA (0,0,1)(1,1,0)[12]



A comparação entre os modelos SARIMA e SARIMAX evidencia a superioridade do modelo univariado em todos os critérios avaliados. Além de apresentar menor AIC, o SARIMA demonstrou menor erro absoluto (MAPE) e menor erro quadrático médio (RMSE), tanto no treinamento quanto na previsão fora da amostra. A Tabela 2 consolida tais achados.

Tabela 2*Comparação entre as métricas dos modelos SARIMA e SARIMAX*

Métricas	SARIMAX	SARIMA
Ordem do Modelo	$(0, 1, 1)(0, 0, 1)[12]$	$(0, 0, 1)(1, 1, 0)[12]$
Critério AIC	1549,61	1440,65
MAPE	0,1967	0,0832
RMSE	394,73	164,73
Teste de Ljung - Box (p)	0,94	0,99

A comparação gráfica das previsões dos dois modelos reforça essas conclusões. O SARIMAX não conseguiu acompanhar adequadamente as variações da demanda no conjunto de teste, ao passo que o SARIMA demonstrou estabilidade e precisão em suas estimativas.

Conclusão

O presente estudo teve por objetivo avaliar comparativamente o desempenho dos modelos estatísticos SARIMA e SARIMAX na previsão de demanda de vendas sazonais, utilizando como base séries temporais mensais de vendas de aves congeladas comercializadas no Brasil, com foco nos períodos que antecedem o Natal. A inclusão da série de vendas de uma marca concorrente como variável exógena permitiu explorar a hipótese de que fatores competitivos poderiam melhorar a acurácia preditiva dos modelos.

Os resultados revelaram que o modelo SARIMA, apesar de sua estrutura univariada e relativa simplicidade, superou o SARIMAX em todos os critérios de desempenho avaliados: menor erro percentual absoluto médio (MAPE), menor raiz do erro quadrático médio (RMSE) e maior adequação estatística dos resíduos (Perez-Guerra et al., 2023). Mais importante ainda, o modelo SARIMA demonstrou excelente capacidade de generalização, mantendo sua acurácia preditiva mesmo quando aplicado ao conjunto de dados fora da amostra — aspecto fundamental para sua aplicabilidade em ambientes reais de decisão.

A análise dos resíduos indicou que ambos os modelos atendiam ao pressuposto de ruído branco, mas apenas o SARIMA foi capaz de capturar adequadamente o padrão de sazonalidade típico do setor, refletindo sua robustez na modelagem de séries temporais com ciclos bem definidos.

Do ponto de vista prático, os resultados obtidos reforçam a importância da escolha criteriosa dos modelos de previsão, sugerindo que, em contextos nos quais a sazonalidade é dominante, modelos mais parcimoniosos como o SARIMA podem oferecer desempenho superior, mesmo sem o uso de variáveis exógenas (Perez-Guerra et al., 2023). Essa constatação é especialmente relevante para empresas que operam em mercados com alto grau de previsibilidade sazonal e que buscam soluções preditivas confiáveis para o planejamento de estoques e estratégias de produção.

No âmbito acadêmico, a pesquisa contribui para a literatura ao comparar empiricamente dois modelos consolidados (Manigandan et al. 2021), aplicando técnicas de validação fora da amostra e explorando um setor ainda pouco abordado em estudos com séries temporais: o mercado de alimentos sazonais no Brasil. A abordagem adotada pode ser replicada e estendida a outros produtos e setores, incentivando a adoção de modelos estatísticos robustos no apoio à tomada de decisão baseada em dados.

Como agenda para pesquisas futuras, sugere-se explorar o uso de abordagens híbridas que combinem modelos estatísticos com componentes de inteligência artificial, bem como testar a inclusão de outras variáveis exógenas, como indicadores macroeconômicos, datas promocionais e clima, que possam ter influência mais significativa sobre o comportamento da demanda.

Referências

- Abirami, S. (2024). Sales prediction based on sarimax time series algorithm. *International Scientific Journal of Engineering and Management*. <https://doi.org/10.55041/isjem01493>.
- Adli, K. A. (2020). Forecasting steel prices using ARIMAX model: A case study of Turkey. *The International Journal of Business Management and Technology*, 4(5), 62–68.
- Alencar, J. F. de. (2022). Seleção de modelo de previsão de demanda agregada para série temporal no setor industrial de tintas e vernizes (Trabalho de Conclusão de Curso). Universidade Federal de Pernambuco.
- Alharbi, F., & Csala, D. (2022). A Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX) Forecasting Model-Based Time Series Approach. *Inventions*. <https://doi.org/10.3390/inventions7040094>.
- Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8, 69–80.

- Bertaglia, P. R. (2003). *Logística e gerenciamento da cadeia de abastecimento*. São Paulo: Saraiva.
- Bianchi, L., Jarrett, J., & Hanumara, R. (1998). Improving forecasting for telemarketing centers by ARIMA modeling with intervention. *International Journal of Forecasting*, 14, 497-504. [https://doi.org/10.1016/S0169-2070\(98\)00037-5](https://doi.org/10.1016/S0169-2070(98)00037-5).
- Bierens, H. J. (1987). Armax model specification testing, with an application to unemployment in the Netherlands. *Journal of Econometrics*, 35, 161–190.
- Box, G. E., Jenkins, G. M., & MacGregor, J. F. (1974). Some recent advances in forecasting and control. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 23, 158–179.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. Hoboken, NJ: John Wiley & Sons.
- Broni-Bediako, E., Buabeng, A., & Allotey, P. (2024). Predicting Ghana's Daily Natural Gas Consumption Using Time Series Models. *Petroleum Science and Engineering*. <https://doi.org/10.11648/j.pse.20240801.14>.
- Ediger, V. S., & Akar, S. (2007). ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, 35, 1701–1708.
- Elamin, N., & Fukushige, M. (2018). Modeling and forecasting hourly electricity demand by SARIMAX with interactions. *Energy*. <https://doi.org/10.1016/J.ENERGY.2018.09.157>.
- Elshewey, A., Shams, M., Elhady, A., Shohieb, S., Abdelhamid, A., Ibrahim, A., & Tarek, Z. (2022). A Novel WD-SARIMAX Model for Temperature Forecasting Using Daily Delhi Climate Dataset. *Sustainability*. <https://doi.org/10.3390/su15010757>.
- Formigoni, W. C., Olga, M., Henning, E., Moro, G., & Wayne, R. S. (2013). Aplicação de um modelo SARIMA na previsão de vendas de motocicletas. *Exacta*, 11(1), 77–88.
- Handayani, N., Maslim, M., & Mudjihartono, P. (2020). Forecasting of Catfish Sales by Time Series Using the SARIMA method. *Journal of Biomedical Informatics*, 11, 83. <https://doi.org/10.24002/jbi.v11i2.3535>.
- Hawinkel, S., Waegeman, W., & Maere, S. (2023). Out-of-Sample R2: Estimation and Inference. *The American Statistician*, 78, 15 - 25. <https://doi.org/10.1080/00031305.2023.2216252>.
- Herrera, R. L., Petropoulos, F., Safari, A., & Davallou, M. (2019). Forecast: Forecasting Functions for Time Series and Linear Models. R Foundation for Statistical Computing. <https://cran.r-project.org/web/packages/forecast/index.html>
- Hyndman, R. J., & Athanasopoulos, G. (2020). Checkresiduals: Check Residuals from Fitted Time Series Models. R Foundation for Statistical Computing. <https://cran.r-project.org/web/packages/checkresiduals/index.html>
- Jalalkamali, A., Moradi, M., & Moradi, N. (2015). Application of several artificial intelligence models and ARIMAX model for forecasting drought using the Standardized Precipitation Index. *International Journal of Environmental Science and Technology*, 12, 1201–1210.
- Khalid, O. (2024). Short-term and long-term product demand forecasting with time series models. *Journal of Trends in Financial and Economics*. <https://doi.org/10.61784/jtfe3022>.
- Kulkarni, R., & Rane, M. (2020). Pattern Recognition - Product Sales Analysis Using SARIMA Model in Time Series Forecasting. .

- Kwiatkowski, D., Morettin, P. A., & Singer, J. D. (1992). Tseries: Time Series Analysis and Computational Finance. R Foundation for Statistical Computing. <https://cran.r-project.org/web/packages/tseries/index.html>
- Lee, G., & Bang, J. (2024). Forecasting Container Throughput of Singapore Port Considering Various Exogenous Variables Based on SARIMAX Models. *Forecasting*. <https://doi.org/10.3390/forecast6030038>.
- Liu, Z., Zhu, Z., Gao, J., & Xu, C. (2021). Forecast Methods for Time Series Data: A Survey. *IEEE Access*, 9, 91896-91912. <https://doi.org/10.1109/ACCESS.2021.3091162>.
- Lucas, E. C., Mendes-Da-Silva, W., & Lyons, A. C. (2017). Gender differences in attitudes towards driving and demand for private Insurance: Evidence from middle class drivers. *Transportation research part F: traffic psychology and behaviour*, 47, 72-85.
- Manigandan, P., Alam, M. S., Alharthi, M., Khan, U., Alagirisamy, K., Pachiyappan, D., & Rehman, A. (2021). Forecasting natural gas production and consumption in United States—Evidence from SARIMA and SARIMAX models. *Energies*, 14, 6021.
- Oh, J., & Seong, B. (2024). Forecasting with a combined model of ETS and ARIMA. *Communications for Statistical Applications and Methods*. <https://doi.org/10.29220/csam.2024.31.1.143>.
- Perez-Guerra, U., Macedo, R., Manrique, Y., Condori, E., Gonzáles, H., Fernández, E., Luque, N., Pérez-Durand, M., & García-Herreros, M. (2023). Seasonal autoregressive integrated moving average (SARIMA) time-series model for milk production forecasting in pasture-based dairy cows in the Andean highlands. *PLOS ONE*, 18. <https://doi.org/10.1371/journal.pone.0288849>.
- Rocha, I. B. M. S. (2014). Impacto da concorrência na performance do retalhista (Dissertação de Mestrado). Universidade Católica Portuguesa – Porto.
- Serrano, A., Rodrigues, G., Martins, P., Saiki, G., Filho, G., Gonçalves, V., & De Oliveira Albuquerque, R. (2024). Statistical Comparison of Time Series Models for Forecasting Brazilian Monthly Energy Demand Using Economic, Industrial, and Climatic Exogenous Variables. *Applied Sciences*. <https://doi.org/10.3390/app14135846>.
- Slack, N., Chambers, S., & Johnston, R. (2002). Administração da produção (2ª ed., M. T. C. Oliveira & F. Alher, Trad.). São Paulo: Atlas.
- Sukparungsee, S., Areepong, Y., & Taboran, R. (2020). Exponentially weighted moving average—Moving average charts for monitoring the process mean. *PLoS ONE*, 15. <https://doi.org/10.1371/journal.pone.0228208>.
- Tamura, L. K. (2013). Séries temporais com variáveis exógenas e gráficos de controle como ferramentas de decisão no mercado financeiro. (Dissertação de Mestrado). São Paulo.
- Tarsitano, A., & Amerise, I. L. (2017). Short-term load forecasting using a two-stage SARIMAX model. *Energy*, 133, 108–114.
- Tubino, D. F. (2000). Manual de planejamento e controle da produção (2ª ed.). São Paulo: Atlas.
- Wahyudi, A., & Febriani, F. (2024). Time-Series Forecasting of Particulate Organic Carbon on the Sunda Shelf: Comparative Performance of the SARIMA and SARIMAX Models. *Regional Studies in Marine Science*. <https://doi.org/10.1016/j.rsma.2024.103863>.
- Wanke, R., & Julianelli, L. (2006). Previsão de vendas: Processos organizacionais e métodos quantitativos e qualitativos (1ª ed.). São Paulo: Atlas.

- Zhang, C., Tian, Y., & Fan, Z. (2021). Forecasting sales using online review and search engine data: A method based on PCA–DSFOA–BPNN. *International Journal of Forecasting*. <https://doi.org/10.1016/j.ijforecast.2021.07.010>.
- Zhou, W., Jiang, R., Ding, S., Cheng, Y., Li, Y., & Tao, H. (2021). A novel grey prediction model for seasonal time series. *Knowl. Based Syst.*, 229, 107363. <https://doi.org/10.1016/j.knosys.2021.107363>.