# Sales Forecasting Using Time Series Models with Exogenous Logistics Variables

**Ricaelle N. T. Pontes**[1], **Priscilla K. M. V. Azevêdo**[1]

[1]Universidade Federal do Agreste de Pernambuco (UFAPE)
Garanhuns – PE – Brasil

ricaelle.nascimento@ufape.edu.br, priscilla.azevedo@ufape.edu.br

***Abstract.*** *Sales and Operations Planning (S&OP) aims to balance supply and demand. However, traditional univariate models rely solely on historical sales data, disregarding logistical fluctuations. This study proposes the use of the multivariate predictive model (SARIMAX), incorporating inventory pressure and budget targets as exogenous covariates with logistical impact. Using out-of-sample validation on real manufacturing data, seven models were compared. The SARIMAX model statistically outperformed classical methods (endogenous variables), achieving a Volumetric Accuracy of 98.13% with an overall mean MAPE of 2.85%, reaching 2.2% in medium-term forecasting simulations. The results indicate that stochastic learning combined with the incorporation of exogenous logistical variables contributes to more accurate projections, enabling cost reduction in S&OP.*

***Resumo.*** *O Planejamento de Vendas e Operações (S&OP) busca balancear a oferta e a demanda. Contudo, modelos univariados tradicionais baseiam-se apenas no histórico de vendas, ignorando flutuações logísticas. Este estudo propõe o uso do modelo preditivo multivariado (SARIMAX), integrando a pressão de estoque e metas orçamentárias como covariáveis exógenas de impacto logístico. Utilizando validação out-of-sample em dados reais de manufatura, compararam-se sete modelos. O modelo SARIMAX superou estatisticamente os métodos clássicos (variáveis endógenas), alcançando uma Acurácia Volumétrica de 98,13% com MAPE médio global de 2,85%, chegando a 2,2% em simulações de previsões de médio prazo. Conclui-se que o aprendizado estocástico aliado à incorporação de variáveis exógenas logísticas contribui para as projeções, viabilizando a redução de custos no S&OP.*

## 1. Introdução

O Planejamento de Vendas e Operações (*Sales and Operations Planning*– S&OP) é fundamental para a sustentabilidade financeira e a eficiência operacional das organizações, ao alinhar decisões estratégicas aos processos operacionais. Seu principal desafio é equilibrar continuamente demanda e oferta, envolvendo capacidade produtiva e estoques. Desequilíbrios nesse processo podem gerar tanto excesso de estoque e impactos no fluxo de caixa quanto perda de vendas e redução do nível de serviço ao cliente [Wallace 2001]. Neste cenário, a previsão de vendas constitui o principal insumo do S&OP.

Para estas previsões, modelos estatísticos baseados em dados históricos são amplamente utilizados [Slack et al. 2002]. Contudo, eles não capturam plenamente

dinâmicas complexas da cadeia de suprimentos. A correta modelagem de tendência e sazonalidade é, portanto, essencial para se obter uma acurácia preditiva, especialmente em cenários dinâmicos nos quais tais padrões se modificam ao longo do tempo [Heizer and Render 2011, Hyndman and Athanasopoulos 2021, Saadeddin 2024]. Assim, abordagens baseadas em análise avançada, integrada à variáveis externas, pode permitir previsões mais robustas e o ajuste proativo da capacidade operacional. Neste contexto, este trabalho avalia o desempenho de sete algoritmos preditivos com modelagem de séries temporais sobre uma base de dados reais, com o faturamento histórico de 10 anos (2016 - 2025) e a combinação desta base com variáveis exógenas logísticas, como as pressões de estoque nos centros de distribuição e na rede de distribuidores.

Para a avaliação dos resultados, foram discutidas as parametrizações definidas, o treinamento, bem como a acurácia e os erros dos diferentes algoritmos, com o objetivo de identificar um modelo capaz de incorporar informações das variáveis da operação logística e explorar sua relação preditiva com os resultados de vendas, além de identificar e prever os padrões cíclicos de sazonalidade, contribuindo para a melhoria das previsões. Ao fim da avaliação, observou-se que a abordagem multivariada, com as variáveis propostas, consolidou o modelo SARIMAX. Este superou estatisticamente os métodos univariados clássicos, alcançando uma Acurácia Volumétrica de 98,13% e reduzindo o Erro Percentual Absoluto Médio (MAPE) a 2,85%, indicando que o uso da defasagem de estoque apresenta relevância preditiva para a identificação antecipada de possíveis restrições logísticas e para o planejamento diante das flutuações no S&OP.

O restante do texto está organizado da seguinte forma. A Seção 2 apresenta o embasamento teórico acerca de séries temporais e algoritmos preditivos, assim como os principais trabalhos relacionados. Na Seção 3 é apresentada a metodologia adotada. Os resultados encontrados são discutidos na Seção 4. Por fim, a Seção 5 sintetiza as conclusões, limitações e direcionamentos para trabalhos futuros.

## 2. Planejamento de Vendas e Operações (S&OP)

Para que uma cadeia de suprimentos funcione com eficiência, é fundamental manter o alinhamento contínuo entre a demanda do mercado e a capacidade de oferta da empresa [Wallace 2001]. Quando esse alinhamento falha e a informação não flui com clareza entre os elos da cadeia, as empresas tendem a aumentar seus estoques de forma preventiva para se proteger contra possíveis flutuações [Coelho et al. 2009]. Esse descompasso, em que o processamento isolado de previsões transforma pequenas oscilações nas vendas da ponta, em grandes distorções nos pedidos enviados às fábricas, é conhecido como efeito chicote [Coelho et al. 2009]. Sem uma visão integrada do S&OP, gestores confundem ruídos logísticos com tendências reais, resultando em acúmulo ou rupturas de estoques.

Para superar essa vulnerabilidade e garantir decisões de S&OP ágeis e seguras, faz-se necessário substituir avaliações puramente intuitivas por infraestruturas de ciência de dados. É neste contexto que se torna indispensável a correta distinção e aplicação de séries temporais e algoritmos preditivos. Ao utilizar algoritmos preditivos avançados para processar as séries temporais de vendas em conjunto com informações de estoque, o sistema atua como um filtro analítico, atenuando distorções do efeito chicote e entregando previsões mais otimizadas, fornecendo ao S&OP o suporte necessário para uma tomada de decisão orientada a dados.

## 2.1. Séries Temporais para Previsão de Vendas

Previsão de vendas apoia-se fortemente na estruturação dos dados históricos. Uma série temporal é definida como qualquer conjunto de observações ordenadas sequencialmente no tempo, permitindo que as organizações mapeiem a história dos acontecimentos passados para entender o comportamento futuro do mercado [Morettin and Toloi 2006].

Geralmente, uma série temporal corporativa pode ser decomposta em padrões fundamentais: tendência, sazonalidade, ciclos e variações aleatórias ou ruídos [Heizer and Render 2011]. A tendência representa o movimento gradual de crescimento ou declínio dos dados a longo prazo, enquanto a sazonalidade engloba padrões cíclicos que se repetem em intervalos fixos e regulares ao longo do ano [Hyndman and Athanasopoulos 2021]. A compreensão correta da estrutura desses dados é o primeiro passo para qualquer estudo preditivo, pois ignorar as flutuações sazonais invariavelmente resulta em conclusões inadequadas para o S&OP. Enquanto a série temporal representa os dados, os algoritmos preditivos são as ferramentas matemáticas aplicadas sobre eles para estimar as vendas futuras.

## 2.2. Algoritmos Preditivos e Variáveis Exógenas para Previsão de Vendas

Na literatura, observa-se uma ampla variedade de métodos de previsão, que vão desde técnicas clássicas de suavização até algoritmos mais avançados. Entre os métodos tradicionais, destacam-se os de suavização exponencial (como os modelos de *Holt* e *Holt-Winters*), que se baseiam em atribuir pesos decrescentes às observações passadas, priorizando o histórico recente para atualizar a projeção da tendência e as extrações sazonais aditivas ou multiplicativas [Almada-Lobo 2007, Silva Junior and Costa 2014]. Outra abordagem robusta é a decomposição STL (*Seasonal and Trend using Loess*), que utiliza regressões locais não lineares para lidar com sazonalidades complexas que mudam dinamicamente com o tempo [Hyndman and Athanasopoulos 2021].

Avançando para as técnicas mais robustas, destacam-se os algoritmos baseados em processos estocásticos. Na estatística, um modelo estocástico é aquele que incorpora elementos de aleatoriedade e probabilidade em sua formulação, em vez de regras puramente fixas. O modelo ARIMA (*AutoRegressive Integrated Moving Average*), por exemplo, é um modelo estocástico que modela a autocorrelação interna dos dados a partir da dependência de valores passados, aplicando diferenciações matemáticas para alcançar a estacionariedade da série [Saadeddin 2024].Contudo, grande parte dos modelos clássicos e estocásticos baseiam-se em análises univariadas, isto é, projetam o futuro olhando apenas para a própria variável alvo (o histórico de vendas). Em ambientes de S&OP sujeitos ao efeito chicote, essa abordagem torna o algoritmo ”cego”a fatores de atrito da cadeia de suprimentos [Coelho et al. 2009].

É neste ponto que a engenharia de variáveis exógenas se faz necessária. A evolução multivariada do ARIMA, conhecida como SARIMAX (*Seasonal AutoRegressive Integrated Moving Average with eXogenous factors*), permite incorporar covariáveis externas como regressores, possibilitando explorar sua associação com a série de vendas e ampliar a capacidade preditiva do modelo. No contexto logístico, a introdução da pressão de estoque no centro de distribuição como uma variável exógena defasada no tempo, pode servir para ensinar matematicamente ao modelo que as sobras ou faltas de produto no mês passado restringem ou impulsionam o potencial de venda do mês atual.

Esse embasamento justifica a necessidade de realizar experimentações comparativas entre modelos puros (sem a integração de exógenas) e abordagens multivariadas.

### 2.3. Trabalhos Relacionados

Trabalhos anteriores têm explorado técnicas de séries temporais para previsão em diferentes domínios. Por exemplo, o estudo sobre previsão de consumo de energia elétrica na Região Sul do Brasil utiliza modelos clássicos, como Suavização Exponencial de *Holt* e ARIMA, obtendo bom desempenho em cenários de curto prazo, porém restringindo-se a uma abordagem univariada baseada exclusivamente em dados históricos [Silva et al. 2020]. De forma complementar, abordagens mais recentes, como a previsão de eventos hidrológicos por meio da combinação de gráficos de recorrência e redes neurais convolucionais, demonstram o potencial de métodos de aprendizado profundo na captura de padrões complexos em séries temporais, [Oliveira et al. 2023]. Apesar desses avanços, tais trabalhos não incorporam explicitamente variáveis exógenas com interpretação operacional. Nesse contexto, o presente estudo diferencia-se ao propor um modelo SARIMAX que integra fatores logísticos do S&OP, como pressão de estoque e metas orçamentárias, permitindo capturar relações e aprimorar a acurácia preditiva em um cenário real de manufatura, com validação *out-of-sample* e comparação entre múltiplas abordagens.

No contexto específico de previsão em negócios corporativos,[Khaw et al. 2025] apresenta um estudo com séries temporais multivariados, incorporando variáveis exógenas, como indicadores de tendência de busca *online*. Entretanto, evidências mostraram que a simples inclusão dessas variáveis não garantiu ganhos significativos de desempenho. Nesse contexto, o presente trabalho diferencia-se ao incorporar variáveis exógenas associadas aos processos logísticos no ambiente de S&OP, como pressão de estoque e metas orçamentárias, buscando capturar relações mais diretamente associadas à dinâmica operacional da demanda, além de realizar uma comparação entre vários modelos distintos.

## 3. Metodologia de Otimização da Previsão de Vendas

Suponha um cenário de uma fábrica de calçados que precisa planejar a sua produção para o ano seguinte. O gestor de S&OP tem a série histórica de vendas e as oscilações de estoque do ano anterior apresentadas na Tabela 1. A grosso modo, o gestor pode ser levado a destacar apenas o viés de crescimento nas vendas e um pico sazonal no final do ano (Novembro e Dezembro). Contudo, nota-se também que, nos meses de pico, o estoque zerou. O objetivo da modelagem preditiva e dos algoritmos que são avaliados neste trabalho não é apenas traçar uma linha estatística, mas responder a uma pergunta de negócios crítica: *"Considerando o histórico de vendas e as limitações de estoque, qual deve ser o volume exato de produção para os meses de janeiro, novembro e dezembro do próximo ano, de forma a garantir o máximo de atendimento sem gerar excessos?"*

| **Mês** | Jan | Fev | Mar | Abr | Mai | Jun | Jul | Ago | Set | Out | Nov | Dez |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Vendas** | 10 | 12 | 15 | 14 | 18 | 16 | 15 | 19 | 22 | 25 | 35 | 40 |
| **Estoque** | 5 | 6 | 4 | 7 | 5 | 8 | 9 | 6 | 4 | 2 | 0 | 0 |

**Tabela 1. Cenário Motivacional de Vendas e Estoque de Calçados.**

Considerando esta problemática, este trabalho iniciou com uma revisão da literatura, que permitiu mapear o estado da arte dos algoritmos preditivos e identificar o

problema logístico do efeito chicote. Em seguida, foi realizado um estudo empírico com base em simulações sobre dados corporativos reais, visando avaliar as soluções propostas. A Figura 1 apresenta o fluxo de trabalho em suas outras cinco etapas, desde os ajustes nas bases de dados até a avaliação final de desempenho dos algoritmos. Ao longo desta seção, cada uma destas etapas será apresentada em mais detalhes. Para nortear a solução deste desafio, o estudo foi estruturado com base nas seguintes Questões de Pesquisa (QPs):

- **QP1:** A integração de variáveis logísticas exógenas (como metas de orçamento e pressões de estoque) em modelos preditivos multivariados proporciona uma acurácia significativamente superior na previsão de vendas, quando comparado com métodos univariados tradicionais?
- **QP2:** É possível modelar matematicamente o impacto do efeito chicote utilizando o estado do estoque da rede do mês anterior como um regressor exógeno de controle para mitigar erros nas projeções futuras?
- **QP3:** Do ponto de vista do S&OP, a adoção de um modelo estocástico otimizado reduz as taxas de erro a patamares que viabilizem o balanceamento seguro entre custos operacionais e atendimento ao cliente?

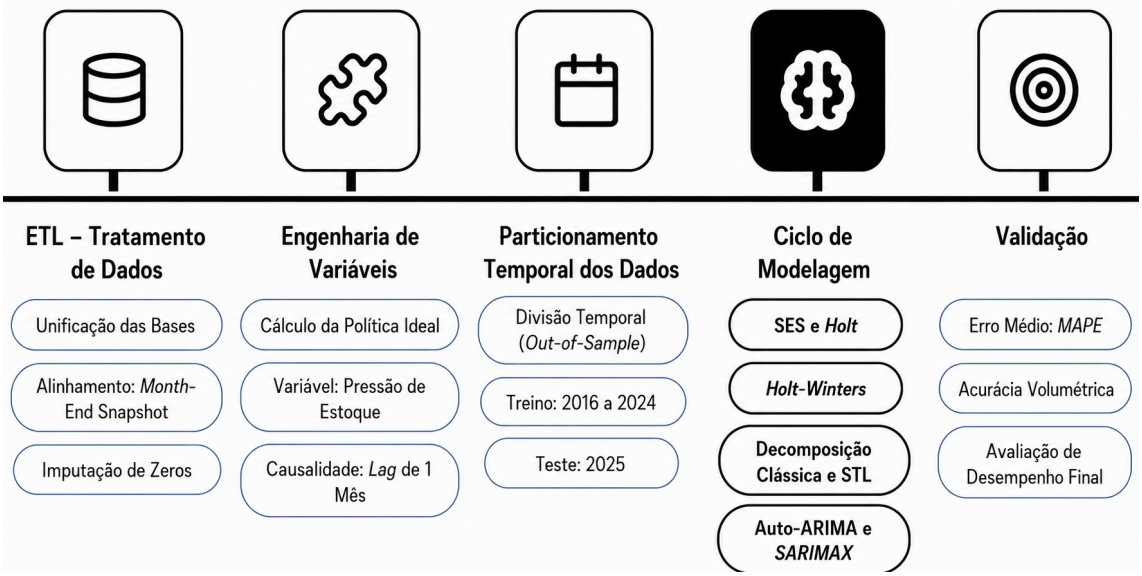



**Figura 1. Fluxograma das etapas metodológicas da pesquisa.**

### 3.1. Extração, Transformação e Carga de Dados

O processo de Extração, Transformação e Carga (*Extract, Transform, Load* - ETL), destacado na Figura 1, foi realizado sobre três bases de dados principais [1]: (i) Histórico de vendas e metas de orçamento (formato de planilha, periodicidade mensal), (ii) Registros logísticos do Centro de Distribuição (CD) e (iii) Histórico logístico da rede de distribuidores (formato CSV, periodicidade diária, a partir de 2022).

Filtraram-se os atributos originais, mantendo apenas colunas essenciais para a modelagem de volume e restrições logísticas; dados secundários foram descartados. O

[1]Repositório com todos os experimentos e dados - https://github.com/ricaellenascimento/Sales-Forecasting.git

principal desafio técnico na consolidação consistiu na diferença de granularidade temporal. Enquanto as vendas eram contabilizadas no fechamento do mês, a movimentação de estoque ocorria diariamente. Para solucionar essa discrepância, aplicou-se a técnica de *Month-End Snapshot* [Kimball and Ross 2013], assim, foi isolada a posição física do estoque correspondente ao último dia de cada mês, alinhando-a ao volume de faturamento daquele mesmo período.

Adicionalmente, como o histórico de vendas remonta a 2016 e os controles da rede iniciaram-se apenas em 2022, aplicou-se uma técnica de imputação de dados ausentes (*missing data*). Para o período "cego"(2016 a 2021), o algoritmo imputou o valor numérico zero (estado de equilíbrio logístico neutro) para as pressões de estoque, o que garantiu a integridade matricial contínua necessária para que os algoritmos autorregressivos processassem o longo prazo sem falhas.

Após estas etapas, obteve-se uma matriz com 132 linhas (representando as observações mensais ininterruptas de janeiro de 2016 a dezembro de 2025 - *DateTime*), e outros quatro atributos:

- **`Venda_Real`** (Numérico Inteiro): Atua como a variável alvo (*target*) para todos os algoritmos. Representa o volume real de itens faturados no mês.
- **`Orçamento_Planejado`** (Numérico Inteiro): Representa a meta comercial estipulada pela empresa no início do ano. Atua como um balizador interno, não sendo considerada uma variável externa do mercado, mas sim uma diretriz corporativa.
- **`Pressao_De_Estoque`** (Numérico Contínuo): Variável exógena que representa o desvio absoluto entre o estoque físico real retido e a sua respectiva política ideal de cobertura.
- **`Pressao_Rede`** (Numérico Contínuo): Variável exógena que representa o desvio entre o estoque real presente na ponta (distribuidores) e a política ideal, em que valores positivos indicam excessos, e negativos apontam rupturas.

Cabe ressaltar que a variável `Venda_Real` foi utilizada ativamente em todos os sete algoritmos experimentados, enquanto as variáveis exógenas de pressão logística e a `Meta_Orcamento` foram direcionadas especificamente para a etapa multivariada deste estudo (modelo SARIMAX), para avaliar o impacto preditivo do efeito chicote.

### 3.2. Engenharia de Variáveis

A modelagem multivariada incluiu variáveis exógenas ligadas ao efeito chicote, com destaque para a variável `Pressao_De_Estoque` (Figura 1), criada para representar a saúde da cadeia de suprimentos. Essa variável foi modelada como o desvio absoluto entre o estoque real e a política ideal ($PI_t$) de cobertura mensal ($t$) estabelecida pela empresa, definida pela média entre os estoques máximo ($Emax_t$) e mínimo permitidos ($Emin_t$) (Equação 1).

$$PI_t = \frac{(Emax_t + Emin_t)}{2} \tag{1}$$

Assim, a Pressão de Estoque ($PE_t$) foi definida como a diferença entre o Estoque Real do período ($ER_t$) e a Política Ideal ($PI_t$) (Equação 2). Neste contexto, $PE_t > 0$,

representa excesso de inventário, enquanto $PE_t < 0$, sinaliza risco de ruptura e desabastecimento.

$$PE_t = ER_t - PI_t \tag{2}$$

Para capturar a associação logística no contexto do S&OP e subsidiar a **QP2**, aplicou-se um deslocamento temporal de um mês (`shift(1)`) na variável de pressão da rede de distribuidores, gerando a `Pressao_Rede_Anterior`. Esse procedimento permitiu ao modelo associar o excesso ou a falta de estoque do período anterior à capacidade de faturamento do mês atual, traduzindo as restrições logísticas do efeito chicote em um preditor matemático direto.

### 3.3. Particionamento Temporal de Dados

Para responder às questões de pesquisa **QP1** e **QP3**, os experimentos compararam métodos univariados tradicionais ao modelo multivariado proposto (Figura 1). A validação foi conduzida por meio de segmentação cronológica da série temporal, simulando um cenário real de S&OP e avaliando a capacidade de generalização dos modelos em previsões *out-of-sample*. Para este fim, os blocos foram particionados da seguinte forma:

- **Conjunto de Treinamento (2016–2024):** composto por 108 meses de dados históricos utilizados no aprendizado dos algoritmos, incluindo tendências, sazonalidades e uma quebra estrutural a partir de 2020 (decorrente de anomalias e choques econômicos globais do período), incorporada como teste de estresse contra *overfitting*.
- **Conjunto de Teste (janeiro 2025 até março de 2026):** formado por 15 meses isolados e não utilizados no treinamento, garantindo uma validação *out-of-sample* representativa da capacidade preditiva dos modelos.

### 3.4. Ciclo de Modelagem

O ciclo de modelagem da experimentação envolveu a parametrização, o treinamento e a avaliação comparativa de sete algoritmos preditivos, partindo de métodos de suavização clássica até algoritmos estocásticos (Figura 1). Para a escolha destes algoritmos, foram considerados os métodos univariados mais recorrentes na literatura em cenários semelhantes ao avaliado neste trabalho. Após isto, foi destacado o que apresentou melhor resultado. Com isto, seria possível ressaltar com firmeza se os métodos mais simples poderiam sanar o problema ou se de fato era necessário um modelo mais robusto.

Na problemática deste trabalho, o modelo de Suavização Exponencial Simples (SES) [Morettin and Toloi 2006, Silva Junior and Costa 2014], utilizado como modelo base (*baseline*), assume que a série de vendas é localmente estacionária. Ele prevê o futuro baseando-se em uma média ponderada, onde os dados de faturamento mais recentes possuem maior peso. O algoritmo otimiza o parâmetro de suavização Alfa ($\alpha$) para calcular o volume de vendas previsto ($\hat{V}_{t+1}$), onde $V_t$ é o volume real de vendas observado no mês atual $t$, e $\hat{V}_t$ é a previsão de vendas feita anteriormente para este mesmo mês. (Equação 3).

$$\hat{V}_{t+1} = \alpha V_t + (1 - \alpha)\hat{V}_t \tag{3}$$

Para o modelo de suavização *Holt*, o Nível base de vendas ($N_t$) é controlado pelo parâmetro $\alpha$ (Equação 4), e a Tendência de vendas ($T_t$) é controlada pelo parâmetro $\beta$ (Equação 5). A previsão do volume de vendas para $h$ meses à frente é dada por: $\hat{V}_{t+h} = N_t + hT_t$. Desta forma, o algoritmo entende se as vendas da empresa estão em um viés de alta ou de baixa e projeta essa inclinação logarítmica para o futuro.

$$N_t = \alpha V_t + (1 - \alpha)(N_{t-1} + T_{t-1}) \tag{4}$$

$$T_t = \beta(N_t - N_{t-1}) + (1 - \beta)T_{t-1} \tag{5}$$

O modelo *Holt-Winters* foi utilizado com o objetivo de identificar se determinados meses do ano possuem um comportamento sistematicamente superior ou inferior à média de vendas (como ocorre em picos de alta temporada). Ele adiciona o componente de Sazonalidade ($Saz_t$), controlado pelo parâmetro Gama ($\gamma$) . Adotou-se a abordagem Aditiva, cuja previsão final consolida os três fatores projetados ao longo do ciclo sazonal ($s = 12$ meses). Na Equação 6, o volume projetado ($\hat{V}_{t+h}$) deixa de ser uma fórmula isolada e passa a ser o resultado prático da soma do patamar atual de faturamento da empresa ($N_t$), somado à velocidade em que essas vendas estão crescendo ($hT_t$), e corrigido pelo fator de oscilação sazonal esperado para aquele mês específico do calendário ($Saz_{t+h-s}$).

$$\hat{V}_{t+h} = N_t + hT_t + Saz_{t+h-s} \tag{6}$$

Para o modelo de decomposição clássica [Hyndman and Athanasopoulos 2021], foi considerado que a venda observada ($y_t$) é a soma da Tendência $T_t$( extraída via médias móveis), da Sazonalidade constante ($S_t$) e do Resíduo aleatório ($R_t$) (Equação 7). De forma complementar, com o modelo STL foram geradas previsões projetando-se as extrações suavizadas de tendência e sazonalidade isoladamente.

$$y_t = T_t + S_t + R_t \tag{7}$$

Por fim, foram avaliados os modelos Auto-ARIMA e SARIMAX [Saadeddin 2024]. Considerando o cenário motivacional de calçados (Tabela 1), o objetivo do uso do SARIMAX é prever as vendas de Janeiro considerando dados além da tendência matemática de alta. Assim, a Equação 8, considera que o volume de vendas ($V_t$) é explicado pelos componentes temporais tradicionais, somados a uma regressão do vetor de variáveis exógenas ($X_t$). Neste estudo, o vetor $X_t$ é instanciado com as variáveis: `Meta_Orcamento`, `Pressao_De_Estoque` e `Pressao_Rede_Anterior`. Matematicamente, a matriz de regressão pondera o impacto logístico, ensinando ao modelo que o desvio de estoque (ruptura ou excesso) do mês passado atua como um limitador da venda do mês atual, entregando ao S&OP uma projeção muito mais realista e acurada.

$$V_t = \beta X_t + \phi(L)V_{t-1} + \theta(L)\epsilon_t \tag{8}$$

### 3.5. Validação

Para a avaliação de desempenho e o ranqueamento dos algoritmos, foram utilizadas duas métricas (Figura 1), o MAPE (*Mean Absolute Percentage Error* - Erro Percentual Absoluto Médio) e a Acurácia Volumétrica (AV).

O MAPE (Equação 9) mede a precisão estatística da previsão mês a mês [Silva Junior and Costa 2014]. Por utilizar valores absolutos, ele impede que erros de superestimativa anulem erros de subestimativa, fornecendo o desvio médio real em escala percentual. Na Equação 9, $n$ é o número de períodos avaliados (15 meses do conjunto de teste), $y_t$ é o volume real de vendas no tempo $t$, e $\hat{y}_t$ é o volume previsto pelo algoritmo para o mesmo período.

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \tag{9}$$

No contexto de S&OP, prever o volume total a ser movimentado na cadeia durante o ano fiscal é tão vital quanto acertar o total mensal, pois dita a estratégia de contratação de fretes e dimensionamento fabril. Seu cálculo afere a distância percentual entre a soma do volume previsto ($V_{prev}$) e a soma do volume real ($V_{real}$) executado em 15 meses (Equação 10).

$$AV = 100\% - \left( \frac{|V_{prev} - V_{real}|}{V_{real}} \times 100\% \right) \tag{10}$$

## 4. Resultados e Discussão

Os algoritmos foram avaliados individualmente no teste. A Tabela 2 ranqueia as abordagens por maior AV e menor MAPE nos 15 meses avaliados. A utilização conjunta destas duas métricas, conforme estruturado na metodologia, garante que o algoritmo selecionado seja estatisticamente robusto tanto no dimensionamento agregado para táticas de suprimentos no longo prazo (AV) quanto na precisão mês a mês (MAPE).

**Tabela 2. Desempenho dos Modelos Preditivos (Visão de 15 meses)**

| **Modelo / Algoritmo** | **MAPE** | **Acurácia** ($AV$) |
|---|---|---|
| SARIMAX | 2,85% | 98,13% |
| Holt (Tendência) | 4,90% | 95,27% |
| Auto-ARIMA | 5,08% | 94,88% |
| SES (*Baseline*) | 5,21% | 94,73% |
| Decomposição Clássica | 5,93% | 94,59% |
| Decomposição STL | 6,33% | 94,41% |
| Holt-Winters | 7,01% | 92,99% |
| Planejamento Empresa (Meta) | 10,15% | 89,77% |

Portanto, a conquista do modelo SARIMAX ao obter, de forma simultânea, a maior AV (98,13%) e o menor MAPE (2,85%) indica para a questão de pesquisa **QP1**, que a integração de variáveis logísticas pode trazer uma superioridade da previsão, não apenas no volume tático acumulado da cadeia, mas também na precisão mês a mês. Destaca-se que as estimativas da empresa (Orçamento_Planejado), sem modelos preditivos, apresentaram o maior erro e acurácia inferior ao esperado no ambiente corporativo (cerca de 97%).

### 4.1. Análise de Resiliência e Antecipação

Embora a avaliação do erro para o mês imediatamente seguinte (M-1), apresentados na Tabela 2, seja fundamental para o balizamento tático de curto prazo, o processo de S&OP em indústrias de manufatura pode exigir uma visão estratégica em um horizonte temporal mais amplo. No fluxo produtivo real, o horizonte de três meses (M-3) pode ser considerado um período crítico para o planejamento de suprimentos, aquisição de matérias-primas importadas (insumos) e contratação de capacidade de fretes . Modelos imprecisos para o trimestre seguinte inviabilizam a proteção da cadeia.

Para avaliar a resiliência estocástica do modelo proposto, foi conduzida uma análise de simulação de janela móvel travando os dados com 90 dias de antecedência (M-3). Por exemplo, para prever março, o modelo usa os dados reais até dezembro (sem considerar janeiro e fevereiro). No passo seguinte, para prever abril, ele usa os dados reais até janeiro (desconsiderando fevereiro e março). O objetivo deste experimento foi verificar comparativamente o desempenho de previsões de venda mês a mês considerando um tempo mais alongado (três passos à frente ”no escuro”.) e mais curto, da melhor configuração descoberta neste trabalho, e destes com a previsão da empresa, sem o uso de modelos ( `Orçamento_Planejado`).

Na Figura 2 são apresentados o total de vendas do orçamento planejado, das previsões geradas pelo SARIMAX nos horizontes M-1 e M-3 e as vendas reais realizadas ( `Venda_Real`).

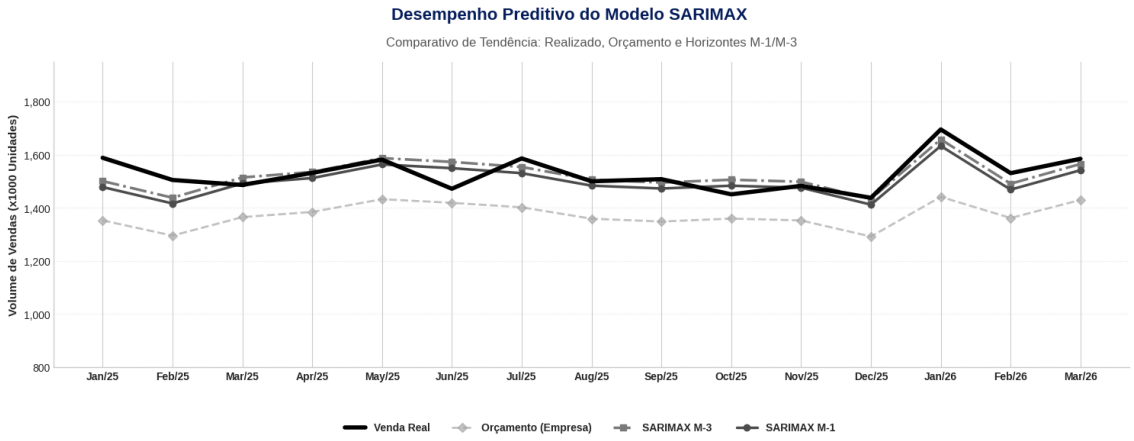



**Figura 2. Comparativo de previsões de vendas (M-1 e M-3), orçamento e vendas reais (15 meses).**

Os resultados mostram que o modelo SARIMAX (M-3) manteve previsões próximas das vendas reais, com MAPE de 2,2% e AV de 99,6%, próximo ao resultado obtido pelo SARIMAX (M-1) (Tabela 2). O modelo M-3 reduziu a incerteza logística em aproximadamente 120 mil unidades mensais, frente ao planejamento manual da empresa, proporcionando boa previsibilidade ao processo de S&OP, auxiliando negociações com fornecedores, otimização de estoques e redução de rupturas e excessos.

### 4.2. Respostas às Questões de Pesquisa

A análise comparativa das métricas de erro estatístico e do comportamento algorítmico nas simulações de curto (M-1) e médio prazo (M-3) fornecem os subsídios práticos e

teóricos necessários para responder às Questões de Pesquisa (QPs) formuladas no início deste estudo.

Em resposta à **QP1** (sobre a superioridade de modelos preditivos e variáveis exógenas logísticas), os resultados indicaram que o modelo multivariado SARIMAX superou significativamente todas as abordagens univariadas, reduzindo o MAPE para 2.2%. Em contraste, seu predecessor estocástico puramente endógeno (Auto-ARIMA) estacionou em 14,41% e os modelos clássicos de suavização (como o *Holt-Winters*) apresentaram erros de até 20,02%. Isso indica que, em ambientes corporativos complexos, o histórico de vendas puro pode ser insuficiente para explicar e projetar o faturamento futuro.

Com relação à **QP2** (modelagem preditiva do efeito chicote), os experimentos indicaram que a utilização da variável `Pressao_De_Estoque` com defasagem temporal (*lag* de 1 mês da rede) atua como um adequado regressor de associação logística. A arquitetura multivariada do modelo SARIMAX validou matematicamente o que a literatura de S&OP defende na prática logística: o desbalanceamento de inventário do mês anterior (sobra ou escassez) age como um restritor físico que limita ou impulsiona o potencial de vendas do mês atual. O modelo SARIMAX, com as variáveis modeladas neste trabalho, foi o único capaz de mapear essa correlação e antecipar os vales decorrentes de desabastecimentos reais.

Por fim, no que tange à **QP3** (impacto gerencial e viabilidade do balanceamento no S&OP), a Acurácia Volumétrica de 98,13% alcançada pelo modelo usado neste trabalho, com as variáveis exógenas propostas, pode atender aos rigorosos padrões de excelência exigidos pelas indústrias de manufatura. A constatação empírica de que a arquitetura preserva sua precisão direcional mesmo no horizonte de antecedência trimestral ($M-3$) indica que a adoção desta proposta não apenas mitiga falhas no nível de serviço ao cliente, mas fornece aos executivos da área de suprimentos o tempo de resposta necessário para o dimensionamento econômico de frotas de frete e o balanceamento seguro dos estoques na cadeia.

## 5. Conclusões

Este trabalho buscou mitigar distorções do efeito chicote no processo de S&OP por meio da avaliação de modelos de previsão de vendas baseados em séries temporais e variáveis exógenas corporativas. Os resultados evidenciaram a relevância da incorporação de informações logísticas às projeções de demanda, com potencial para apoiar decisões de planejamento e otimização de custos.

A abordagem multivariada baseada no SARIMAX, com as variáveis *pressão de estoque* e *pressão de rede*, superou os modelos univariados de suavização e o ARIMA, alcançando MAPE médio global de 2,85% e Acurácia Volumétrica de 98,13%. No horizonte trimestral (M-3), apresentou MAPE de 2,2%, indicando desempenho consistente em previsões de médio prazo. O estudo também disponibiliza uma base de dados real utilizada nos experimentos, contribuindo para pesquisas futuras.

Apesar das contribuições alcançadas, reconhece-se que as validações apresentam limitações inerentes ao escopo, uma vez que os experimentos se concentraram em uma base de dados corporativa de um único setor e utilizaram um número restrito de variáveis logísticas de atrito. Portanto, sugere-se como trabalhos futuros: (i) a expansão da modelagem multivariada para outras categorias de produtos e unidades de negócios; (ii) a

inclusão de variáveis macroeconômicas (como inflação e juros) para buscar a captura de oscilações de mercado de longo prazo; e (iii) o desenvolvimento de estudos comparativos (tempo de processamento, acurácia e proteção dos dados) entre o algoritmo destacado neste trabalho e algoritmos de aprendizagem de máquina recentes.

## Referências


Almada-Lobo, B. (2007). Métodos estatísticos de previsão: Amortecimento exponencial. Material técnico/Apresentação.

Coelho, L. C., Follmann, N., and Rodriguez, C. M. T. (2009). O impacto do compartilhamento de informações na redução do efeito chicote na cadeia de abastecimento. *Gestão & Produção*, 16(4):571–583.

Heizer, J. and Render, B. (2011). *Operations Management*. Prentice Hall, New Jersey, 10 edition.

Hyndman, R. J. and Athanasopoulos, G. (2021). *Previsão: Princípios e prática*. OTexts, Melbourne, Australia, 3 edition.

Khaw, B., Irwanto, R., Yunis, R., and Elly, E. (2025). Analisis time series dan perancangan dashboard untuk memprediksi penjualan dengan metode prophet dan sarimax. *Jurnal SIFO Mikroskil*, 26(2):207–222.

Kimball, R. and Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Wiley, Indianapolis, 3 edition.

Morettin, P. A. and Toloi, C. M. C. (2006). *Análise de séries temporais*. Blucher, São Paulo, 2 edition.

Oliveira, F. S., Costa, L. R., and Almeida, P. H. (2023). Flood forecasting using recurrence plots and convolutional neural networks. In *Anais do Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, Brasil. SBC.

Saadeddin, Z. (2024). Arima para previsão de séries temporais: Um guia completo. Acessado em: 08 abr. 2026.

Silva, J. A., Pereira, M. L., and Santos, R. C. (2020). Previsão da demanda de energia elétrica na região sul do brasil utilizando séries temporais. In *Anais do Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, Brasil. SBC.

Silva Junior, L. J. d. and Costa, F. P. (2014). Aplicação dos métodos de amortecimento exponencial para previsão de demanda de clientes em oficina mecânica automotiva. In *Anais do IV Congresso Brasileiro de Engenharia de Produção (ConBRepro)*, Ponta Grossa, PR.

Slack, N., Chambers, S., and Johnston, R. (2002). *Administração da Produção*. Atlas, São Paulo, 2 edition.

Wallace, T. F. (2001). *Planejamento de vendas e operações: guia prático*. IMAM, São Paulo.