

# ANÁLISE DO ÍNDICE DE PREÇOS AO CONSUMIDOR AMPLO (IPCA) POR MEIO DE COMBINAÇÕES DE PREVISÕES

*Data de aceite: 26/01/2024*

**Glaucio Jorge Ferreira Rosa**

Universidade Federal de Santa Maria

**Diego Brenner Dos Reis**

Universidade Federal de Santa Maria

**Cleber Bisognin**

Universidade Federal de Santa Maria

**RESUMO:** O Índice de Preços ao Consumidor Amplo (IPCA) possui o objetivo de medir a taxa de inflação com base em uma cesta de produtos e serviços comercializados no varejo, que representam o consumo pessoal das famílias. Permanecer atualizado acerca dos índices econômicos é de suma importância para compreender as tendências e a situação da economia brasileira. Esses índices oferecem dados relevantes sobre a inflação, o desempenho do setor produtivo, a variação dos preços e a renda das famílias, entre outros aspectos essenciais. Estar informado sobre tais métricas possibilita uma visão abrangente do cenário econômico do país. Dada a importância do IPCA, o objetivo deste trabalho é prever os futuros valores desse índice utilizando combinações de previsões. Os modelos

base utilizados são o SARIMA, os Modelos de Suavização Exponencial e Holt-Winters, totalizando seis modelos ajustados. Para a obtenção de previsões mais precisas, foram aplicados métodos de combinações de previsões, incluindo a média aritmética simples, a variância mínima e a regressões linear (OLS) e robusta (LAD). Para avaliar a acurácia das previsões e alcançar o objetivo proposto, foram utilizadas as seguintes medidas de desempenho: RSME, MAPE e MAE. Com base nas medidas de acurácia três combinações foram selecionadas, para as quais foram calculadas as previsões para o ano de 2023: combinação por mínimos quadrados e robusta utilizando seis modelos e a regressão robusta utilizando os modelos SARIMA, Holt-Winters Aditivo e Holt-Winters Multiplicativo com fator de amortecimento.

**PALAVRAS-CHAVE:** Índice de Preços ao Consumidor Amplo – IPCA; Combinação de Previsões; Modelagem.

**ABSTRACT:** The Índice de Preços ao Consumidor Amplo (IPCA) aims to measure the inflation rate based on a basket of goods and services traded at retail, representing the personal consumption of households. Staying updated on economic indicators is

of utmost importance to understand the trends and the situation of the Brazilian economy. These indicators provide relevant data on inflation, the performance of the productive sector, price variations, and household income, among other essential aspects. Being informed about such metrics enables a comprehensive view of the country's economic landscape. Given the importance of this index, the objective of this study is to forecast future values of the CPI using forecast combinations. The base models used are SARIMA, Exponential Smoothing Models, and Holt-Winters, totaling six adjusted models. To obtain more accurate forecasts, forecast combination methods were applied, including simple arithmetic mean, minimum variance, and linear regressions (OLS) and robust regressions (LAD). To evaluate the accuracy of the forecasts and achieve the proposed objective, the following performance measures were used: RMSE, MAPE, and MAE. Based on the accuracy measures, three combinations were selected, for which forecasts were calculated for the year 2023: the combination using least squares and robust regression with six models, and the robust regression using the SARIMA, Additive Holt-Winters, and Multiplicative Holt-Winters models with damping factor.

**KEYWORDS:** Índice de Preços ao Consumidor Amplo – IPCA; Forecast *Combination*; *Modeling*.

## 1 | INTRODUÇÃO

O Índice Nacional de Preços ao Consumidor Amplo (IPCA) desempenha um papel crucial como um dos principais indicadores empregados pelo Banco Central do Brasil para mensurar a inflação no país. Mensalmente, o Instituto Brasileiro de Geografia e Estatística (IBGE) é responsável pelo cálculo do IPCA, que reflete a variação dos preços de um conjunto de bens e serviços consumidos pelas famílias brasileiras. O índice ganha destaque ao servir como referência para o sistema de metas de inflação implementado pelo Banco Central do Brasil. O objetivo primordial é manter a inflação dentro de uma faixa estabelecida pelo governo, atualmente fixada em 3,75% ao ano, com uma margem de tolerância de 1,5 pontos percentuais, seja para mais ou para menos. O IPCA é composto por nove grupos distintos, que são: alimentação e bebidas, artigos de residência, comunicação, despesas pessoais, educação, habitação, saúde e cuidados pessoais, transportes e vestuário. Cada um desses grupos possui um peso específico no cálculo do índice, levando em consideração sua relevância no orçamento das famílias brasileiras.

O impacto do IPCA na vida dos brasileiros é abrangente. Quando a inflação está em alta, os preços dos produtos e serviços tendem a aumentar, resultando em uma redução do poder de compra das pessoas e afetando a economia como um todo. Por essa razão, torna-se fundamental que o governo e os agentes econômicos acompanhem de perto a evolução do IPCA. De fato, os índices econômicos desempenham um papel significativo na economia de um país, pois fornecem informações valiosas sobre as condições e tendências do mercado. Esses indicadores abrangem diversas áreas e setores, oferecendo uma visão abrangente do desempenho econômico em diferentes aspectos. Como resultado, eles exercem uma influência relevante nas decisões dos agentes econômicos, incluindo

consumidores, empresas e governos.

Portanto, os índices econômicos desempenham um papel vital na análise e no funcionamento da economia, sendo ferramentas essenciais para entender a conjuntura econômica atual. Desta forma, a busca por melhorias é essencial para garantir o sucesso e a eficiência das atividades em diversas áreas. Nesse contexto, as técnicas de previsão desempenham um papel fundamental ao auxiliar a tomada de decisões em atividades que requerem planejamento e avaliação constantes. As previsões permitem antecipar tendências futuras e possíveis cenários, possibilitando que gestores e tomadores de decisão se preparem para enfrentar os desafios que surgirão. Através do uso de métodos de previsão, é possível identificar oportunidades e ameaças, o que é especialmente valioso em um ambiente de negócios altamente competitivo. Além disso, as técnicas de previsão ajudam a reduzir a incerteza presente em diversas áreas de interesse, proporcionando uma base sólida para o planejamento estratégico.

Neste sentido, Carvalho et al. (2016) indicam as previsões têm o poder de fornecer informações cruciais que apoiam o desenvolvimento de estratégias eficientes e ajudam a direcionar recursos de forma mais efetiva. Isso se traduz em vantagens competitivas para empresas, instituições e organizações, ao permitir que se adaptem melhor às mudanças do mercado e às demandas do ambiente globalizado.

De fato, a utilização de diversas técnicas de previsão pode aumentar a abrangência das informações disponíveis, contribuindo para prognósticos mais precisos e acurados. Uma abordagem frequentemente bem-sucedida é a combinação de previsões, que consiste em unir diferentes técnicas de previsão individuais para gerar uma previsão combinada. Segundo Costantini e Pappalardo (2010), esse método busca capturar os diversos fatores que influenciam cada técnica de previsão individual utilizada como base para a obtenção da previsão combinada. Ao combinar várias previsões, é possível aproveitar os pontos fortes de cada método, compensando eventuais limitações de um único modelo. Estudos, como os de Clemem (1989), Costantini e Pappalardo (2010) e Muniroh et al. (2017), relatam resultados positivos para a aplicação de combinações de previsões. Muitas vezes, as combinações avaliadas demonstram maior acurácia em comparação com as previsões individuais. Esse aumento na precisão pode ser atribuído à capacidade de incorporar diferentes perspectivas e fontes de informação, tornando o processo de previsão mais robusto e confiável. Assim, a combinação de previsões é uma estratégia valiosa que tem se mostrado eficiente em melhorar a qualidade das previsões, proporcionando uma visão mais completa e aprimorada do comportamento futuro das séries temporais ou variáveis estudadas. Isso é especialmente relevante em cenários em que a precisão das previsões é crucial para tomadas de decisão estratégicas e planejamento.

Pelo apresentado, o objetivo deste trabalho é calcular previsões acuradas para o IPCA utilizando combinações de previsões. Para o cálculo das previsões a serem combinadas por diversos métodos serão utilizados os modelos SARIMA, Holt-Winters e Suavização

Exponencial (ETS). Tais modelos são amplamente utilizados nas mais diversas áreas para realizar previsões de séries temporais com as características de tendência (determinística e/ou estocástica) e/ou sazonalidade. As técnicas de combinação de previsões utilizadas são média aritmética simples (MAS), variância mínima, regressão por mínimos quadrados ordinários (MQO) e regressão robusta (LAD).

## 2 | TÉCNICAS DE PREVISÃO

Nesta seção são apresentados os modelos usados para realizar a análise e previsão dos dados mensais do IPCA de 1995 a 2022, a saber: modelos Holt Winters, SARIMA( $p, d, q$ ) $\times$ ( $P, D, Q$ ) $_s$  e Suavização Exponencial, além de quatro métodos para realizar combinações de previsões: variância mínima, por regressão linear e robusta e média aritmética. A seguir são definidos os processos SARIMA( $p, d, q$ ) $\times$ ( $P, D, Q$ ) $_s$  propostos por Box; Jenkins (1976).

**Definição 1:** Considere o processo estocástico  $\{X_t\}_{t \in \mathbb{Z}}$  satisfazendo a Equação (1)

$$\Phi(B^s)\phi(B)(1 - B^s)^D(1 - B)^d(X_t - \mu) = \theta(B)\theta(B^s)\varepsilon_t, \quad (1)$$

em que  $\mu$  é a média do processo,  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  é o processo ruído branco,  $s \in \mathbb{N}$  é a sazonalidade,  $B$  é o operador defasagem ou de retardo, isto é,  $B^l(X_t) = X_{t-l}$  e  $B^{sl}(X_t) = X_{t-sl}$ , para  $j \in \mathbb{N}$ ,  $\phi(\cdot)$ ,  $\theta(\cdot)$ ,  $\Phi(\cdot)$  e  $\Theta(\cdot)$  são os polinômios de ordem  $p$ ,  $q$ ,  $P$  e  $Q$ , respectivamente, definidos pelas Equações (2) e (3).

$$\phi(z) = \sum_{l=1}^p (-\phi_l)z^l, \quad \theta(z) = \sum_{m=1}^q (-\theta_m)z^m, \quad (2)$$

$$\Phi(z) = \sum_{r=1}^P (-\Phi_r)z^r, \quad \Theta(z) = \sum_{l=1}^Q (-\Theta_l)z^l, \quad (3)$$

com  $\phi_l$ ,  $1 \leq l \leq p$ ,  $\theta_m$ ,  $1 \leq m \leq q$ ,  $\Phi_r$ ,  $1 \leq r \leq P$  e  $\Theta_l$ ,  $1 \leq l \leq Q$ , constantes reais e  $\phi_0 = \Phi_0 = -1 = \theta_0 = \Theta_0$ . Então  $\{X_t\}_{t \in \mathbb{Z}}$  é um processo sazonal autorregressivo integrado de média móvel de ordem  $(p, d, q) \times (P, D, Q)_s$  com sazonalidade  $s$ , denotado por SARIMA( $p, d, q$ ) $\times$ ( $P, D, Q$ ) $_s$ , onde  $d$  e  $D$  são, respectivamente, o grau de diferenciação e o grau de diferenciação sazonal.

A equação (4) seguir apresenta a previsão de um processo SARIMA ( $p, d, q$ ) $\times$ ( $P, D, Q$ ) $_s$  invertível,

$$\hat{X}_t(h) = \sum_{j \geq 1} \pi_j \hat{X}_t(j - h), \quad (4)$$

onde  $h \geq 1$  e  $\{\pi_j\}_{j \geq 0}$ , são os coeficientes da representação autoregressiva infinita do processo. Pode-se perceber que  $\hat{X}_t(h)$  pode ser expressa com uma soma ponderada das observações presentes e passadas  $(X_t)_{t \leq n}$ , com  $n$  é o tamanho da série temporal.

Os modelos de suavização exponencial destacam-se entre outras técnicas de

previsão devido à sua simplicidade, facilidade de ajuste e boa acurácia (PELLEGRINI, 2000). Abaixo estão descritos os modelos de Holt-Winters para efeitos sazonais aditivos e multiplicativos. Para mais informações detalhadas, consultar Makridakis et al. (1998) e Morettin; Toloi (2006).

**Definição 2:** Para os Modelos de Holt-Winters para efeitos sazonais aditivos com fator de amortecimento para a tendência, considera-se que tal série é formada pela soma do nível, tendência, um fator sazonal e um erro aleatório. As estimativas do nível da série no tempo  $t$ , denotado por  $L_t$ , da tendência, denotada por  $T_t$  e do fator sazonal, denotado por  $S_t$ , são dadas, respectivamente, pelas Equações de (5) a (7).

$$L_t = \alpha(X_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + \phi T_{t-1}) \quad (5)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)\phi T_{t-1} \quad (6)$$

$$S_t = \gamma(X_t - L_t) + (1 - \gamma)S_{t-s}, \quad (7)$$

em que  $\alpha$  é o coeficiente de ponderação exponencial do nível ( $0 \leq \alpha \leq 1$ ),  $\beta$  é o coeficiente de ponderação exponencial da tendência ( $0 \leq \beta \leq 1$ ) e  $\gamma$  é o coeficiente de ponderação exponencial do fator sazonal ( $0 \leq \gamma \leq 1$ ) e  $0 \leq \phi \leq 1$  fator de amortecimento.

As previsões  $h$  passos à frente, são dadas pela expressão dada pela Equação (8)

$$\hat{X}_t(h) = L_t + (\phi + \phi^2 + \dots + \phi^h)T_t + S_{t+h-s(k+1)}, \quad (8)$$

em que  $k$  é a parte inteira de  $\frac{(h-1)}{s}$  o que garante que as estimativas dos índices sazonais utilizados para a previsão venham do último ano da amostra. A notação  $\hat{X}_t(h)$  indica a previsão de origem  $t$  e horizonte  $h \geq 1$ .

**Observação 1:** Para os Modelos de Holt-Winters para efeitos sazonais aditivos sem fator de amortecimento para a tendência, considere  $\phi=1$  na Definição 2.

Na definição a seguir apresentamos os modelos de Holt-Winters para efeito sazonal multiplicativos com amortecimento.

**Definição 3:** Para os Modelos de Holt-Winters para efeitos sazonais multiplicativos com fator de amortecimento para a tendência, considera-se que tal série é formada por uma tendência aditiva, por um fator sazonal multiplicativo e um erro aleatório. As estimativas do nível da série no tempo  $t$ , denotado por  $L_t$ , da tendência, denotada por  $T_t$  e do fator sazonal, denotado por  $S_t$ , são dadas, respectivamente, pelas Equações de (9) a (11).

$$L_t = \alpha \frac{X_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + \phi T_{t-1}) \quad (9)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)\phi T_{t-1} \quad (10)$$

$$S_t = \gamma \frac{X_t}{L_t} + (1 - \gamma)S_{t-s}, \quad (11)$$

em que  $\alpha$  é o coeficiente de ponderação exponencial do nível ( $0 \leq \alpha \leq 1$ ),  $\beta$  é o coeficiente de ponderação exponencial da tendência ( $0 \leq \beta \leq 1$ ) e  $\gamma$  é o coeficiente de ponderação exponencial do fator sazonal ( $0 \leq \gamma \leq 1$ ) e  $0 \leq \phi \leq 1$  fator de amortecimento.

As previsões  $h$  passos à frente, são dadas conforme a Equação (12)

$$\hat{X}_t(h) = (L_t + (\phi + \phi^2 + \dots + \phi^h)T_t)S_{t+h-s(k+1)}. \quad (12)$$

A notação  $\hat{X}_t(h)$  indica a previsão de origem  $t$  e horizonte  $h \geq 1$ .

**Observação 2:** Para os Modelos de Holt-Winters para efeitos sazonais multiplicativos sem fator de amortecimento para a tendência, considere  $\phi=1$  na Definição 3.

Neste trabalho, também empregamos os Modelos de Suavização Exponencial - ETS ( $M, A_d, N$ ), ou seja, com ruído multiplicativo, tendência aditiva amortecida. As definições desses modelos são apresentadas a seguir.

**Definição 4:** Nos Modelos de Suavização Exponencial – ETS ( $M, A_d, N$ ), (considere-se que tal série é formada por erros multiplicativos e tendência aditiva amortecida, os quais são definidos a seguir, como apresentado pela Equação (13).

$$\hat{X}_t(t) = (l_{t-1} + \phi b_{t-1}) \quad (13)$$

As estimativas do nível da série no tempo  $t$ , denotado por  $l_t$  e da tendência, denotada por  $b_t$ , são dadas, respectivamente, pelas Equações de (14) e (15).

$$l_t = (l_{t-1} + \phi b_{t-1})(1 + \alpha \varepsilon_t) \quad (14)$$

$$b_t = \beta(l_{t-1} + \phi b_{t-1})\varepsilon_t + \phi b_{t-1} \quad (15)$$

em que  $\{\varepsilon_t\}_t=1$  é considerado um erro independente e identicamente distribuído com média zero e variância constante,  $\alpha$ , ( $0 \leq \alpha \leq 1$ ), é o coeficiente de ponderação exponencial do nível,  $\beta$  é o coeficiente de ponderação exponencial da tendência ( $0 \leq \beta \leq 1$ ) e  $\phi$  é o parâmetro de amortecimento.

As previsões  $h$  passos à frente, são definidas conforme a Equação (17).

$$\hat{X}_t(h) = (l_{t+h-1} + \phi b_{t+h-1}). \quad (17)$$

A notação  $\hat{X}_t(h)$  indica a previsão de origem  $t$  e horizonte  $h \geq 1$ . Maiores detalhes sobre a Taxonomia dos Modelos de Modelos de Suavização Exponencial – ETS ver Hyndman et al. (2020).

A proposta metodológica de Bates e Granger (1969) para realizar previsões por meio da combinação de modelos envolve a formulação de uma combinação linear de previsões não viciadas. Em outras palavras, diferentes modelos geram previsões que são combinadas de forma ponderada, com os pesos sendo selecionados de maneira a garantir que as previsões não apresentem viés ou tendenciosidade.

O objetivo dessa abordagem é aproveitar as vantagens de cada modelo individual e criar uma previsão mais robusta e precisa, por meio da combinação ponderada de suas previsões. Ao fazer isso, busca-se reduzir a variabilidade e melhorar a acurácia geral das previsões.

A combinação de modelos pode ser uma estratégia útil para aumentar a confiabilidade das previsões, uma vez que diferentes modelos podem capturar diferentes aspectos dos

dados e podem ter desempenhos variados em diferentes cenários. Ao combinar suas previsões, é possível mitigar as fraquezas de cada modelo individual e obter uma previsão mais confiável e geralmente mais precisa.

A representação matemática da combinação de previsões é expressa por:

$$f_c = w_1 f_1 + w_2 f_2 + \dots + w_k f_k, \quad (18)$$

onde  $w_i$  denota o peso da previsão  $i$ ,  $w_i \geq 0$ ,  $i = 1, \dots, k$ , com  $\sum_{i=1}^k w_i = 1$  para  $k = 2, \dots, 6$ .

O primeiro método de combinação de previsões e também o mais usual entre os métodos é o de média aritmética simples.

A combinação de previsão pelo método da média aritmética simples (MAS) é o método cuja soma das combinações divididas por suas respectivas quantidades  $K$ , a expressão da média é dada por:

$$f_{MS} = \frac{f_1 + \dots + f_k}{k}. \quad (19)$$

em que  $f_i$  e  $f_k$  são as previsões individuais da variável de interesse e  $k=2, \dots, 6$  são as combinações dos modelos possíveis.

O segundo método descrito trata da combinação pela variância mínima. Com a proposta de minimizar a variância da previsão combinada  $\sigma_c^2$ . Esse procedimento de redução da variância é então denominado de método da Variância Mínima.

Considerando  $k$  previsões de uma variável de interesse  $Z_t$ ,  $w$  o vetor de pesos, deve ser determinado por:

$$w = \frac{\sum^{-1} \mathbf{1}}{\mathbf{1}' \sum^{-1} \mathbf{1}}. \quad (20)$$

onde,  $\sum$  representa a matriz de covariância amostral dos erros e  $\mathbf{1}$  é o vetor coluna de dimensão  $k$ .

A ideia de utilizar modelos de regressão para combinar previsões foi proposta por Crane e Crotty (1967) e amplamente divulgada por Granger e Ramanathan (1984).

A previsão combinada é uma função linear do previsões individuais onde os pesos são determinados utilizando um modelo de regressão das previsões individuais sobre os dados reais, como dado pela Equação (21).

$$y_t = \beta_0 + \beta_1 f_1 + \dots + \beta_k f_k + \varepsilon_t. \quad (21)$$

onde,  $\beta_0, \beta_1, \dots, \beta_k$  representam os coeficientes de regressão,  $f_1, \dots, f_k$  representam o  $k$ -ésimo nível da variável independente e  $k=2, \dots, 6$  são as combinações das previsões dos modelos.

A combinação de Previsão pelo método de Regressão de Mínimos Quadrados (MQO) é dada por

$$f_R = \hat{\beta}_0 + \hat{\beta}_1 f_1 + \dots + \hat{\beta}_k f_k. \quad (22)$$

A estimação de  $\beta = (\beta_0, \beta_1, \dots, \beta_k)'$  realizada pelo procedimento de mínimos quadrados ordinários (MQO), resulta em um valor estimado  $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)' \in \mathbb{R}^{n+1}$ , que possibilita a redução da função perda,

$$L_1(\beta) = \sum_{t=1}^n r_t^2. \quad (22)$$

sendo os resíduos  $r_t$  escritos pela expressão a seguir:

$$r_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 f_{t1} - \dots - \hat{\beta}_k f_{tk}. \quad (23)$$

A combinação de previsões pelo método de Regressão Robusta, é uma generalização da regressão linear e normalmente é aplicada em casos que as observações apresentam *outliers*, alta assimetria e heteroscedasticidade.

Um caso particular da Regressão Robusta é quando minimizamos os desvios absolutos. A estimação dos coeficientes do modelo da Equação (21) pelo método de regressão pelo desvio absoluto mínimo (LAD) é realizado minimizando a norma  $L_1$  dos resíduos, expressa por:

$$L_2(\beta) = \sum_{t=1}^n |r_t|. \quad (24)$$

em que os resíduos  $r_t$  são dados pela equação (23).

Segundo Makridakis et al. (1998), o termo acurácia diz respeito à capacidade do modelo (ou da combinação) em reproduzir os dados já conhecidos, ou seja, a qualidade do ajuste. No entanto, para determinar qual técnica de previsão - seja individual ou combinada - é a mais adequada, é crucial obter medidas de acurácia. Dentre as existentes, as mais utilizadas são: o (MAE) Erro Médio Absoluto de Previsão o MAPE (percentual médio absoluto de erro) e o RMSE (a raiz quadrada da média dos erros quadráticos).

### 3 | PROCEDIMENTOS METODOLÓGICOS

A pesquisa em questão é de natureza aplicada, pois tem como objetivo resolver problemas específicos e foi conduzida para gerar conhecimentos sobre os dados do mensais do IPCA de 1995 a 2022. Os dados históricos do IPCA foram obtidos do Banco de Dados do IBGE - Instituto Brasileiro de Geografia e Estatística (<http://www.ibge.gov.br>), cujo objetivo principal é medir a variação média dos preços de uma cesta de produtos e serviços. O índice IPCA também é chamado de taxa de inflação do Brasil, ainda que este seja apenas um dos indicadores para sua mensuração. Foram utilizadas 348 observações mensais disponíveis, de janeiro de 1995 a dezembro de 2022. Os dados foram encontrados pelos autores e acessados em 13/01/2023. O uso deste período se deve ao fato de marcar



o início do Real como moeda oficial brasileira e os índices registrados passaram a ser mensurados com a atual moeda como referência.

Após a etapa de análise preliminar, foram ajustados os modelos SARIMA, Holt-Winters e ETS. Para a estimação dos parâmetros dos processos SARIMA  $(p,d,q) \times (P,D,Q)_s$ , a função *auto.arima* foi utilizada. Para os modelos de Holt-Winters foi utilizado a função *hw*. Para a estimação dos parâmetros dos modelos ETS, utilizou-se a função *ets*.

Após definir os modelos e estimar seus parâmetros, foram aplicadas as técnicas de combinação de variância mínima, média aritmética e por regressão no software R. Essas combinações foram obtidas através da junção de dois modelos individuais, seguida de três, quatro, cinco e, por fim, seis modelos.

A avaliação da precisão das técnicas para realizar as previsões no período de Jan/1995 a Dez/2022 foi realizada utilizando três medidas: MAE (erro médio absoluto de previsão), RMSE (raiz quadrada da média dos erros quadráticos) e MAPE (média dos erros percentuais absolutos). Essas métricas foram calculadas utilizando a rotina *accuracy*. Todas as funções utilizadas neste trabalho são do software R.

## 4 | RESULTADOS E DISCUSSÕES

Estamos interessados em encontrar uma combinação de previsões mais acurada para prever a série temporal do IPCA mensal para 2023, baseando-se no período de janeiro de 1995 a dezembro de 2022. A Figura 1 apresenta o gráfico dos números índices mensais do IPCA no período de 1995 a 2022 e sua respectiva função de autocorrelação amostral. Analisando o gráfico da série temporal e da função de autocorrelação amostral (ver Figura 1) nos leva a concluir que a série temporal é não estacionária.

Para testar a estacionariedade da série, foram aplicados os seguintes testes de raiz unitária: Dickey-Fuller (ADF) e Phillips-Perron (PP), cujas hipóteses são:  $H_0$ : Há raiz unitária, a série não é estacionária versus  $H_1$ : Não há raiz unitária, a série é estacionária. Ambos os testes resultaram em um valor-p = 0,99, indicando que a série temporal possui raiz unitária, sendo, portanto, não estacionária. Realizou-se a análise de tendência determinística nos dados, por meio do teste de Cox-Stuart, cujas hipóteses são:  $H_0$ : Sem tendência determinística e  $H_1$ : Com tendência determinística. O resultado do teste teve um valor-p < 0.001, rejeitando, dessa forma a hipótese nula e indicando haver tendência determinística.

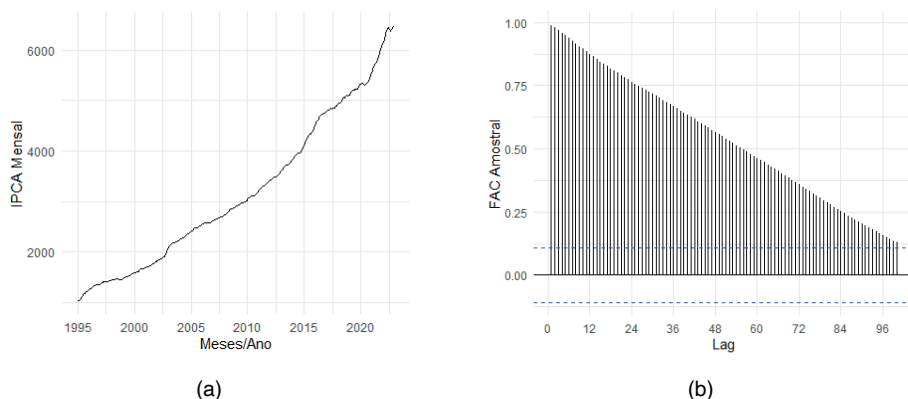


FIGURA 1 – Índice Nacional de Preços ao Consumidor Amplo (IPCA), de jan. de 1995 a dez. de 2022: (a) gráfico da série temporal; (b) função de autocorrelação amostral.

Para análise da sazonalidade da série temporal, foram aplicados os testes de Kruskal Wallis e Friedman, de modo a testar:  $H_0$ : Não há sazonalidade versus  $H_1$ : há sazonalidade. Ambos os testes apresentaram valor-p < 0.001, ou seja, menores que o índice de significância ( $\alpha=0,05$ ). Com isso rejeita-se a hipótese de ausência de sazonalidade.

Dado que a série temporal em questão é não estacionária, apresentando sazonalidade, e considerando que o objetivo deste estudo é encontrar a combinação de previsões mais acurada para prever os futuros valores do IPCA, adotaremos diferentes modelos que levam em consideração sazonalidade, tendência estocástica e determinística. Tais modelos foram definidos na Seção 2. A seguir são apresentados os modelos que foram ajustadas à Série Temporal do IPCA.

**Modelo 1** - Modelo de Suavização Exponencial Holt-Winters (Aditivo): as estimativas para os parâmetros do modelo encontram-se na Tabela 1. Vale ressaltar que no modelo aditivo, as variâncias sazonais são aproximadamente constantes ao longo da série. Para este modelo, a soma do quadrado dos erros (SSE) foi igual à 13,8015.

Parâmetro	$\alpha$	$\beta$	$\gamma$
Estimativa	0,9999	0,1338	0,0001

TABELA 1 - Resultado do ajuste do Modelo de Suavização Exponencial Holt-Winters (Aditivo).

**Modelo 2** - Modelo de Suavização Exponencial Holt-Winters Multiplicativo: as estimativas para os parâmetros do modelo encontram-se na Tabela 2. Vale ressaltar que no modelo multiplicativo, a amplitude da variação sazonal aumenta com o acréscimo no nível médio da série. Para este modelo, obteve-se a  $SSE = 0,0108$ .

Parâmetro	$\alpha$	$\beta$	$\gamma$
Estimativa	0,2321	0,0974	0,0234

TABELA 2 - Resultado do ajuste do Modelo de Suavização Exponencial Holt-Winters (Multiplicativo).

**Modelo 3** - Modelo de Suavização Exponencial Holt-Winters Aditivo com fator de amortecimento: as estimativas para os parâmetros do modelo encontram-se na Tabela 3. Para este modelo, obteve-se a SSE= 13,8173.

Parâmetro	$\alpha$	$\beta$	$\gamma$	$\phi$
Estimativa	0,9897	0,1326	0,0103	0,9800

TABELA 3 - Resultado do ajuste do Modelo de Suavização Exponencial Holt-Winters (Aditivo com fator de amortecimento).

**Modelo 4** - Modelo de Suavização Exponencial Holt-Winters Multiplicativo com fator de amortecimento: as estimativas para os parâmetros do modelo encontram-se na Tabela 4. Neste modelo, obteve-se a SSE= 0,0045.

Parâmetro	$\alpha$	$\beta$	$\gamma$	$\phi$
Estimativa	0,9608	0,1312	0,0388	0,9775

TABELA 4 - Resultado do ajuste do Modelo de Suavização Exponencial Holt-Winters (Multiplicativo com fator de amortecimento).

**Modelo 5** - Modelo SARIMA  $(1,1,0) \times (0,0,1)_{12}$  com drift: as estimativas dos coeficientes encontram-se na Tabela 5. Este modelo leva em consideração a tendência determinística (drift). Pelo Teste de Ljung-Box [valor-p=0,3512], os resíduos do modelo são não correlacionados, sendo assim, o modelo proposto é adequado aos dados.

Parâmetros	Estimativa	Erro Padrão	Estat. Z	P(> z )
$\phi_1$	0,641613	0,042398	15,1330	<0,0001
$\phi_2$	0,128108	0,060300	2,1245	0,0336
drift	16,388111	2,018999	8,1169	< 0,0001

Log Verossimilhança: -1303,91; AIC = 2615,82  
 Teste Resíduos (Box-Pierce): df=10 – valor-p=0,3512

TABELA 5 - Resultado do ajuste do Modelo SARIMA  $(1,1,0) \times (0,0,1)_{12}$ .

**Modelo 6** - Modelo ETS(M,A<sub>d</sub>,N): as estimativas dos coeficientes encontram-se na Tabela 6. Para este modelo, obteve-se uma variância dos resíduos estimada igual a  $\hat{\sigma}^2=0,0035$ , e Critério de Informação de Akaike AIC= 3505,656.

Parâmetro	$\alpha$	$\beta$	$\phi$
Estimativa	0,9999	0,8151	0,9142

TABELA 6 - Resultado do ajuste do Modelo ETS(M,A<sub>d</sub>,N).

Após o ajuste dos modelos e cálculo das predições, passou-se ao cálculo das combinações de previsão e medidas de acurácia (ver Tabela 7).

Combinação	Modelos	RMSE	MAE	MAPE
Média	1 5	11,9144	7,7213	0,2504
	1 5 6	11,7190	7,5038	0,2411
	1 2 5 6	11,8493	7,6037	0,2481
	1 2 4 5 6	12,0270	7,7729	0,2562
	1 2 3 4 5 6	18,8072	12,7653	0,4568
MQO	1 5	11,5861	7,3662	0,2346
	1 2 5	11,5615	7,4076	0,2372
	1 2 5 6	11,5457	7,4136	0,2376
	1 2 3 5 6	11,5113	7,4123	0,2377
	1 2 3 4 5 6	<b>11,5106</b>	7,4105	0,2375
VM	3 6	11,8124	7,5439	0,2425
	1 3 6	11,6500	7,3702	0,2328
	1 3 5 6	11,7161	7,4932	0,2409
	1 2 3 5 6	11,8868	7,6069	0,2486
	1 2 3 4 5 6	12,0510	7,7494	0,2553
LAD	1 5	11,6127	7,3255	0,2307
	1 4 5	11,6213	7,3246	<b>0,2305</b>
	1 3 5 6	11,5968	7,3088	0,2309
	1 3 4 5 6	11,5873	7,3081	0,2310
	1 2 3 4 5 6	11,6044	<b>7,3074</b>	0,2311

TABELA 7 - Medidas de acurácia das combinações de previsão para a série temporal do Índice Nacional de Preços ao Consumidor Amplo (IPCA).

A combinação de previsão por regressão MQO com todos os modelos apresentou menor RMSE, a combinação de previsão por regressão Robusta utilizando todos os modelos apresentou menor MAE e utilizando os modelos 1, 4 e 5 apresentou menor MAPE. Na Tabela 8 encontra-se os coeficientes das combinações pelos métodos de regressão, utilizando o método dos Mínimos Quadrados Ordinários e Robusta LAD selecionados pelas menores medidas de acurácia após a análise da Tabela 7. A Tabela 9 apresenta os valores previsto da série temporal do Índice Nacional de Preços ao Consumidor Amplo (IPCA), para o período de janeiro a dezembro de 2023, utilizando as combinações de previsão selecionadas pelas medidas de acurácia. A Figura 2 apresenta a predição e previsão série temporal do Índice Nacional de Preços ao Consumidor Amplo (IPCA) utilizado os métodos de combinações com menores medidas de acurácia conforme a Tabela 7.

Medida de Acurácia	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
RMSE	0,0093	0,4893	-0,2372	0,0622	-0,0426	0,3841	0,3447
MAE	-1,5709	0,7108	0,0438	0,0496	-0,0364	0,1077	0,1251
MAPE	-1,8587	0,8319	0,0324	0,1364			

TABELA 8 - Modelos de combinação de previsão: parâmetros das combinações de previsão selecionadas pelas menores medidas de acurácia conforme a Tabela 7.

Mês/ Ano	Dados Reais	Combinação de Previsões		
		Regressão MQO com Seis Modelos	Regressão LAD com Seis Modelos	Regressão LAD com Três Modelos
Jan/23	6508,4	6510,32	6509,81	6510,21
Fev/23	6563,07	6544,83	6543,54	6545,18
Mar/23	6609,67	6578,48	6577,90	6581,75
Abr/23	6649,99	6604,69	6607,49	6612,74
Mai/23	6665,28	6624,91	6631,71	6638,77
Jun/23	6659,95	6643,51	6656,12	6664,80
Jul/23		6655,04	6673,33	6682,02
Ago/23		6666,36	6690,77	6700,99
Set/23		6680,65	6709,71	6721,53
Out/23		6700,68	6734,88	6748,09
Nov/23		6718,38	6758,32	6773,74
Dez/23		6741,22	6784,98	6802,43

TABELA 9 – Previsão da série temporal do Índice Nacional de Preços ao Consumidor Amplo (IPCA), para o período de janeiro a dezembro de 2023.

## 51 CONCLUSÃO

Este trabalho teve como objetivo prever o Índice Nacional de Preços ao Consumidor Amplo (IPCA). Para isto foi utilizado uma série temporal mensal no período entre janeiro de 1995 a dezembro de 2022.

Inicialmente foi analisado a estacionariedade e a componente sazonal da série temporal. A seguir, foram ajustados seis modelos individuais com base no período citado. Os modelos ajustados foram Holt-Winters (aditivo e multiplicativo / com e sem fator de amortecimento), Suavização exponencial (ETS) e SARIMA. Com base nos seis modelos ajustados foram obtidas as combinações das previsões para as quatro técnicas em estudo: média aritmética simples; regressão via MQO; regressão LAD e por variância mínima.

Foram analisadas as medidas de acurácia das combinações de previsões. O método por regressão linear utilizando as previsões de todos os modelos apresentou menor RMSE. A técnica de combinação por regressão robusta utilizando as previsões de todos os modelos apresentou menor MAE. Já as previsões dos modelos 1, 4 e 5 apresentaram o menor MAPE.

Todas as previsões calculadas para o período de janeiro a junho 2023 ficaram pouco abaixo dos valores reais do IPCA (ver Tabela 9). Além disso, a taxa de crescimento apresenta uma leve desaceleração como pode ser percebido na Figura 2. Este resultado é corroborado pela Carta de Conjuntura (2023): “a projeção do Grupo de Conjuntura para o Índice Nacional de Preços ao Consumidor (INPC) em 2023 também foi revista, recuando de 5,5% para 4,9%.”

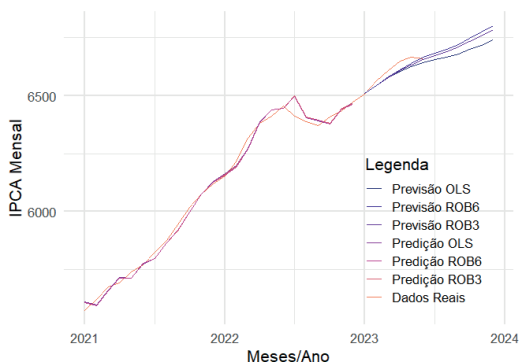


FIGURA 2 - Predição e previsão da série temporal do Índice Nacional de Preços ao Consumidor Amplo (IPCA), para o período de janeiro a dezembro de 2023. **Notação:** Predição/Previsão OLS: Predição/Previsão utilizando combinação de previsões através do método dos mínimos quadrados. Predição/Previsão ROB3: Predição/Previsão utilizando combinação de previsões através do método robusto LAD com três modelos ajustados. Predição/Previsão ROB6: Predição/Previsão utilizando combinação de previsões através do método robusto LAD com os seis modelos ajustados.

## REFERÊNCIAS

BATES, J. M.; GRANGER, C. W. J. The Combining of Forecasts. **Operational Research Quarterly**, v.20, n.4, 1969, p. 451-468.

BOX, G. E.; JENKINS, G. M. **Time series analysis, control, and forecasting**. San Francisco, CA: Holden Day, 1976.

COSTANTINI, M.; PAPPALARDO, C. A hierarchical procedure for the combination of forecasts. **International Journal of Forecasting**, v. 26 n.4, p.725–743, 2010.

CARVALHO, R. L. S.; NASCIMENTO, B. I. S.; QUERINO, C. A. S.; SILVA, M. J. G.; DELGADO, A. R. S. Comportamento das séries temporais de temperatura do ar, umidade e precipitação pluviométrica no município de Ariquemes (Rondônia-Brasil). **Revista Brasileira de Climatologia**, Ano 12 v.18, p.123-1442, jan-jun 2016.

CLEMEN, R.T. Combining Forecasts: A review and annotated bibliography. **International Journal of Forecasting**, v.5, n.4, p.559-583, 1989.

CRANE, D. B.; J. R. CROTTY. A two-stage forecasting model: Exponential smoothing and multiple regression. **Management Science**, v.13 n.8, p. 501–507, 1967.

GRANGER, C. W.; R, RAMANATHAN. Improved methods of combining forecasts. **Journal of Forecasting**, v.3, n.2, p.97–204, 1984.

IPEA, **Carta de Conjuntura**. n. 59, Nota de Conjuntura 33, Brasília, 2023. Disponível em: <http://tiny.cc/y2b9vz>. Acessado em: 24 de julho de 2023.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. **Forecasting**: methods and applications. John Wiley & Sons, 3<sup>rd</sup> edition, 1998.

MORETTIN, P. A.; TOLOI, C. **Análise de séries temporais**. Ed. Blücher. 2006.

MUNIROH, M. F.; ISMAIL, N.; LAZIM, M. A Combination of Forecasts with an Application to Unemployment Rate. **Pertanika Journal Science & Technology**. v.25, n.3, p. 787 – 796, 2017.

PELLEGRINI, F. R. **Metodologia para implementação de sistemas de previsão de demanda**. Dissertação (Mestrado em Engenharia de Produção). Programa de Pós-Graduação em Engenharia de Produção. Universidade Federal do Rio Grande do Sul. Porto Alegre, 2000.