REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

PREVISÃO DA TAXA DE DESEMPREGO NO BRASIL COM MODELO ESTATÍSTICO SARIMA UNIVARIADO E COMPONENTE SAZONAL: APLICAÇÃO AOS DADOS DA PNAD CONTÍNUA (2012-2023)

UNEMPLOYMENT RATE FORECASTING IN BRAZIL USING A UNIVARIATE SEASONAL SARIMA MODEL: EVIDENCE FROM PNAD CONTÍNUA (2012–2023)

Ciências Sociais Aplicadas • 20/01/2026

REGISTRO DOI: 10.5281/zenodo.18309808

Gabriel Diego de Carvalho[1]

RESUMO

O desemprego representa um dos principais desafios socioeconômicos do Brasil, impactando diretamente o bem-estar da população e a formulação de políticas públicas. Este estudo avalia a capacidade preditiva do modelo SARIMA *(Seasonal Autoregressive Integrated Moving Average)* na previsão da taxa de desemprego, com base em dados da PNAD Contínua (IBGE) entre 2012 e 2023. A metodologia emprega abordagem quantitativa e aplicação da modelagem SARIMA à série trimestral, interpolada para frequência mensal via *spline* cúbico. O modelo selecionado — SARIMA (1,1,0)(2,0,0)[12] — apresentou bom ajuste (AIC = -108,39), resíduos não autocorrelacionados e desempenho satisfatório (MAPE = 4,1%, RMSE = 0,27). As previsões mensais para 2024 indicam trajetória de queda da taxa de desemprego, variando de 7,59% em fevereiro para 6,85% em janeiro de 2025. A comparação com dados reais confirma a precisão do modelo em capturar a tendência geral, apesar de erros pontuais em períodos de choques conjunturais. A pesquisa se destaca por validar as estimativas com dados observados e reforça o potencial da modelagem SARIMA como ferramenta confiável para análises preditivas de curto prazo, subsidiando a formulação de políticas públicas baseadas em evidências econômicas.

Palavras-chave: desemprego; previsão; modelo SARIMA; séries temporais; mercado de trabalho.

ABSTRACT

Unemployment is one of the key socioeconomic challenges in Brazil, directly affecting the population's well-being and the design of public policies. This study evaluates the predictive performance of the SARIMA (Seasonal Autoregressive Integrated Moving Average) model in forecasting the unemployment rate, using PNAD Contínua (IBGE) data from 2012 to 2023. A quantitative approach was adopted,

applying the SARIMA model to the quarterly unemployment rate series, which was interpolated to monthly frequency using cubic spline. The selected model — SARIMA (1,1,0)(2,0,0)[12] — showed a good fit (AIC = -108.39), non-autocorrelated residuals, and satisfactory accuracy (MAPE = 4.1%, RMSE = 0.27). Monthly forecasts for 2024 indicated a downward trend, with unemployment decreasing from 7.59% in February to 6.85% in January 2025. Comparisons with real data confirmed the model's ability to capture general trends, despite isolated errors during periods of economic shocks. The study contributes to the national literature by validating forecasts with actual observed data and underscores the SARIMA model's potential as a reliable tool for short-term predictive analysis, supporting evidence-based public policy formulation.
Keywords: unemployment; forecasting; SARIMA model; time series; labor market.

1. INTRODUÇÃO

O desemprego é um dos principais desafios macroeconômicos enfrentados por países em desenvolvimento e desenvolvidos, afetando não apenas o desempenho econômico agregado, mas também o bem-estar da população, a estabilidade social e as perspectivas de crescimento sustentável (Paul & Moser, 2009; OIT, 2022). No Brasil, a taxa de desemprego tem oscilado de forma significativa nas últimas décadas, refletindo vulnerabilidades estruturais do mercado de trabalho, além de forte sensibilidade a choques cíclicos, como a recessão de 2014-2016 e os efeitos devastadores da pandemia da COVID-19 (IBGE, 2021; IPEA, 2023). No terceiro trimestre de 2021, por exemplo, a taxa atingiu 14,7%, revelando a intensidade da crise e a limitação das políticas de contenção adotadas até então (IBGE, 2021).

A análise do desemprego envolve múltiplas dimensões, refletidas em diferentes categorias consolidadas na literatura econômica, como o desemprego friccional, estrutural, cíclico, sazonal e tecnológico. A coexistência desses tipos evidencia a complexidade do fenômeno e reforça a necessidade de abordagens analíticas capazes de captar suas variações ao longo do tempo e em distintos contextos econômicos e institucionais.

Nos últimos anos, o desemprego juvenil e a precarização das relações de trabalho ganharam destaque na literatura e nas políticas públicas. A informalidade crescente, impulsionada por formas de contratação como a pejotização[2] e a "uberização[3]", reflete um processo contínuo de desproteção social (Araújo et al., 2022; Silva, Courseuil & Costa, 2020). Esses fenômenos, ainda que frequentemente invisibilizados nas estatísticas convencionais, ampliam a desigualdade e reduzem o potencial de inserção produtiva, sobretudo entre jovens, mulheres e trabalhadores com baixa escolaridade (IPEA, 2023).

Diante desse cenário desafiador, surge uma questão central: em que medida é possível antecipar a trajetória da taxa de desemprego brasileira com ferramentas estatísticas robustas e adequadas à natureza da série? A resposta a essa pergunta se mostra crucial para a formulação de políticas públicas mais eficazes, capazes de mitigar os efeitos do desemprego e promover uma inserção produtiva sustentável.

Este estudo tem como objetivo avaliar a capacidade preditiva do modelo SARIMA (*Seasonal Autoregressive Integrated Moving Average*) na estimação da taxa de desemprego trimestral no Brasil entre 2012 e 2023, com projeções para o ano de 2024. O modelo

SARIMA é uma extensão do clássico ARIMA que incorpora componentes sazonais explícitos, tornando-se particularmente útil para séries temporais com padrões periódicos, como é o caso da taxa de desemprego brasileira (Box & Jenkins, 1976; Enders, 2015). A escolha desse modelo fundamenta-se na hipótese de que a taxa de desemprego apresenta componentes sazonais, estruturais e aleatórios que podem ser estatisticamente modelados para produzir previsões confiáveis.

Do ponto de vista metodológico, este trabalho adota uma abordagem quantitativa e aplica a metodologia de Box & Jenkins (1976) à série temporal da taxa de desemprego nacional, conforme divulgada pela PNAD Contínua do Instituto Brasileiro de Geografia e Estatística (IBGE). A modelagem é composta pelas etapas de identificação da ordem do modelo, estimação dos parâmetros, diagnóstico dos resíduos e previsão dos valores futuros. Os critérios de seleção incluem o Critério de Informação de *Akaike* (AIC) e o Critério de Informação Bayesiano (BIC), enquanto os testes de estacionariedade aplicados são o *Augmented Dickey-Fuller* (ADF) e o KPSS. A validação dos resíduos ocorre por meio de testes de autocorrelação (Ljung-Box), normalidade (Shapiro-Wilk) e heterocedasticidade (ARCH-LM). As métricas de avaliação do desempenho preditivo incluem o Erro Médio Absoluto Percentual (MAPE) e o Erro Quadrático Médio (RMSE).

Para garantir a robustez empírica da análise, todas as etapas foram realizadas com o suporte do software *RStudio*, utilizando o pacote *forecast*, amplamente reconhecido na literatura econométrica por sua precisão em modelagem de séries temporais (Hyndman & Athanasopoulos, 2021). A série analisada contempla um período de intensas transformações macroeconômicas e políticas, incluindo

crises, reformas trabalhistas e avanços tecnológicos, permitindo uma avaliação crítica do modelo SARIMA em cenários de instabilidade.

Além disso, os resultados preditivos obtidos são comparados com os dados efetivos do ano de 2024, permitindo testar a precisão do modelo frente à realidade observada. A análise também considera a inclusão futura de variáveis exógenas como PIB, inflação e taxa de juros para melhorar o desempenho preditivo do modelo, abrindo caminho para a incorporação de modelos ARIMAX ou até abordagens não-lineares como redes neurais artificiais (Makridakis, Spiliotis & Assimakopoulos, 2018).

Embora o SARIMA seja uma ferramenta poderosa para séries com padrão sazonal, ele difere de outros modelos como o ARIMA (que não considera sazonalidade explícita), os modelos VAR (que consideram múltiplas séries simultâneas) e as redes neurais (que capturam não linearidades e interações complexas, porém com maior demanda computacional e menor interpretabilidade) (Zhang, Patuwo & Hu, 1998; Enders, 2015). Dessa forma, a escolha do SARIMA busca o equilíbrio entre precisão, transparência e capacidade de replicação.

Por fim, este artigo está estruturado da seguinte maneira: após esta introdução, a Seção 2 apresenta o referencial teórico, abordando os principais conceitos sobre o desemprego e os modelos de previsão utilizados na literatura. A Seção 3 detalha a metodologia adotada, desde a coleta de dados até os procedimentos de estimação e validação. A Seção 4 apresenta e analisa os resultados empíricos obtidos, incluindo gráficos, testes estatísticos e comparações com os dados reais observados. A Seção 5 discute as conclusões e implicações do estudo, bem como suas limitações e recomendações

para futuras pesquisas. Por fim, a Seção 6 reúne as referências bibliográficas utilizadas ao longo do trabalho.

2. REFERENCIAL TEÓRICO

A compreensão do desemprego exige uma abordagem multifacetada que articule definições conceituais, interpretações teóricas e instrumentos empíricos. No contexto econômico, o desemprego refere-se à parcela da população economicamente ativa que, embora disponível e em busca de trabalho, não encontra ocupação formal ou informal (IBGE, 2023). Trata-se, portanto, de uma variável crítica para a avaliação do desempenho econômico de um país e de seus efeitos sociais colaterais.

Entre os conceitos fundamentais para a análise da desocupação, destacam-se os diferentes tipos de desemprego, cuja categorização permite interpretações mais precisas das dinâmicas do mercado de trabalho. O desemprego friccional resulta de transições naturais entre empregos, como mudanças de ocupação ou ingresso no mercado. O desemprego estrutural, por sua vez, decorre do descompasso entre as qualificações dos trabalhadores e as exigências do sistema produtivo, sendo agravado por fatores como baixa escolaridade, avanço tecnológico e mudanças no perfil ocupacional (Carleial, 2015; Antunes, 2010). O desemprego cíclico está diretamente relacionado aos ciclos econômicos, elevando-se em momentos de recessão e reduzindo-se nas fases de crescimento, como evidenciado nas crises de 2008 e 2015–2016 (Marasca et al., 2017). O desemprego sazonal surge de oscilações previsíveis na demanda por trabalho, como nos setores agrícola e de varejo. Já o desemprego tecnológico tem se tornado mais relevante nas últimas décadas, dado o crescimento da automação e da digitalização das

tarefas produtivas, fenômeno com impactos crescentes no Brasil e no mundo (Jansen & Uexkull, 2010).

Esses diferentes tipos não são mutuamente excludentes. Ao contrário, muitas vezes coexistem e interagem entre si, o que torna a previsão do desemprego uma tarefa complexa e desafiadora, especialmente em economias emergentes, onde fatores estruturais, informais e institucionais têm peso significativo.

Nesse contexto, torna-se fundamental recorrer a modelos estatísticos que permitam decompor e prever o comportamento dessa variável ao longo do tempo. O modelo SARIMA (*Seasonal Autoregressive Integrated Moving Average*), desenvolvido como uma extensão do modelo ARIMA clássico de Box e Jenkins (1976), é amplamente reconhecido por sua capacidade de lidar com séries temporais que apresentam tendência, ruído e, especialmente, sazonalidade. O modelo permite identificar e estimar os componentes autorregressivos, de médias móveis e de diferenciação, tanto em sua forma não sazonal quanto sazonal, sendo especialmente útil para séries mensais ou trimestrais, como a da taxa de desemprego no Brasil.

A teoria econômica que embasa a modelagem da desocupação tem raízes tanto na tradição neoclássica, que entende o desemprego como resultado de rigidezes institucionais e imperfeições nos mercados, quanto nas abordagens keynesianas, que argumentam que o desemprego involuntário pode persistir devido à insuficiência de demanda agregada (Keynes, 1936). A teoria da busca por emprego, desenvolvida por Diamond, Mortensen e Pissarides (1994), introduz uma perspectiva microeconômica, explicando o desemprego como um fenômeno de fricção no emparelhamento

entre trabalhadores e empregadores. Essa teoria enfatiza a importância de fatores como assimetria de informação, custos de busca e tempo de transição, contribuindo para explicar o desemprego mesmo em mercados aparentemente equilibrados.

No campo aplicado, estudos empíricos têm explorado diferentes modelos para a previsão da taxa de desemprego. Cavalcante e Silva (2011) demonstraram a eficácia de modelos ARIMA para captar oscilações conjunturais no Brasil, mas apontaram limitações quanto à sazonalidade. Marasca et al. (2017) testaram modelos SARIMA sobre séries mensais da taxa de desocupação e identificaram melhor desempenho preditivo nos períodos pós-crise. No plano internacional, Hyndman e Athanasopoulos (2021) confirmam que, apesar da ascensão de técnicas mais complexas, como redes neurais artificiais (RNA), os modelos SARIMA ainda são altamente competitivos pela sua interpretabilidade e estabilidade em séries econômicas.

Embora estudos como os de Zhang, Patuwo e Hu (1998) e Makridakis et al. (2018) sugiram que métodos baseados em aprendizado de máquina superem os modelos tradicionais em alguns contextos, essas abordagens enfrentam desafios de transparência, replicabilidade e risco de overfitting[4]. A modelagem SARIMA, por sua vez, oferece uma estrutura mais transparente e ajustável, especialmente relevante para organismos públicos e formuladores de políticas que necessitam de previsões fundamentadas e auditáveis.

No entanto, apesar da consolidação do modelo SARIMA na literatura econométrica internacional, ainda existem lacunas importantes na aplicação empírica dessa abordagem ao mercado de trabalho

brasileiro. Grande parte das pesquisas nacionais limita-se à análise descritiva da desocupação ou à aplicação de modelos ARIMA simples, sem considerar componentes sazonais ou realizar comparações com os dados efetivamente observados. Além disso, poucos estudos avaliam a precisão das previsões ao longo do tempo, nem exploram a sensibilidade do modelo a choques exógenos como reformas institucionais, crises sanitárias ou flutuações do mercado internacional. O presente trabalho busca preencher essa lacuna ao aplicar o modelo SARIMA à série histórica da taxa de desemprego trimestral no Brasil, comparando suas previsões com os valores efetivamente observados em 2024.

Ao articular essas contribuições teóricas e empíricas, este estudo posiciona-se na fronteira entre a teoria econômica do desemprego e a modelagem estatística aplicada, propondo uma análise fundamentada e atualizada da capacidade preditiva do modelo SARIMA. O objetivo é oferecer subsídios técnicos e metodológicos para a melhoria da previsão econômica, com implicações diretas sobre a formulação de políticas públicas voltadas à redução da desocupação e ao fortalecimento do mercado de trabalho.

A partir das abordagens macroeconômicas discutidas, observa-se que o comportamento da taxa de desemprego no Brasil está fortemente condicionado a fatores sazonais, cíclicos e estruturais — como os ciclos de demanda agregada (Keynes, 1936), os choques institucionais e a fricção no mercado de trabalho (Diamond, Mortensen & Pissarides, 1994). Nesse contexto, a escolha do modelo SARIMA se mostra apropriada por permitir a captura estatística de padrões de recorrência e tendência, sem sacrificar a interpretabilidade essencial para a aplicação em formulação de políticas. Ao incorporar explicitamente a sazonalidade, o SARIMA

dialoga com as dinâmicas cíclicas apontadas pela literatura e oferece uma ferramenta capaz de traduzir complexidades macroeconômicas em previsões objetivas e auditáveis.

A modelagem SARIMA assume papel central na análise empírica da taxa de desemprego brasileira por sua capacidade de capturar padrões sazonais e persistentes ao longo do tempo. Como evidenciado pelas teorias de Kaynes (1936), o desemprego pode se manter elevado mesmo em situações de equilíbrio, refletindo insuficiências de demanda agregada que não são imediatamente corrigidas pelo mercado. Nesse contexto, a modelagem SARIMA permite explorar essa persistência estatística, ajustando componentes sazonais e estruturais que influenciam os ciclos do desemprego.

Adicionalmente, ao incorporar termos autorregressivos e de médias móveis em múltiplas dimensões (sazonal e não sazonal), o modelo possibilita previsões mais robustas e fundamentadas — especialmente importantes em economias emergentes, mercadas por elevada volatilidade e sensibilidade a fatores externos. A compatibilidade dessa abordagem com os fundamentos apresentados por Friedman (1968), que discute a existência de uma taxa natural de desemprego, reforça seu uso como ferramenta analítica que transita entre os planos teórico e aplicado. Assim, o modelo SARIMA contribui para uma compreensão mais acurada da dinâmica do mercado de trabalho, subsidiando decisões estratégicas e formulações de políticas públicas orientadas pela evidência.

3. METODOLOGIA

A presente pesquisa adota uma abordagem quantitativa, com foco na modelagem de séries temporais univariadas, por meio da aplicação do modelo SARIMA *(Seasonal Autoregressive Integrated Moving Average*), visando à previsão da taxa de desemprego no Brasil. A metodologia segue o procedimento clássico de Box e Jenkins (1976), que compreende as etapas de identificação do modelo, estimação dos parâmetros, diagnóstico dos resíduos e previsão. O estudo foi operacionalizado no ambiente R, versão 4.3.2, utilizando o pacote *forecast*, amplamente consagrado para a análise de séries temporais.

A variável de interesse é a taxa de desemprego brasileira, originalmente disponibilizada em periodicidade trimestral pela Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua) do IBGE, no intervalo de 2012 a 2023. Para possibilitar a geração de previsões mensais, os dados foram interpolados por *spline cúbico*[5], mantendo a estrutura original da série. O modelo foi ajustado com base na série interpolada, assumindo frequência mensal (s = 12) para fins de identificação e estimativa dos componentes sazonais.

A interpolação cúbica foi adotada para transformar a série trimestral da PNAD Contínua em frequência mensal, de modo a explorar a estrutura sazonal mais rica do modelo SARIMA. Essa abordagem permite maior granularidade na previsão e segue práticas comuns na literatura quando se busca compatibilidade com modelos sazonais mensais. Embora essa transformação introduza valores não observados diretamente, ela mantém a coerência temporal da série e preserva a tendência de longo prazo e a sazonalidade média, sendo adequada para fins de previsão, conforme discutido em Hyndman & Athanasopoulos (2021).

Formalmente, o modelo SARIMA pode ser representado como:

$$SARIMA(p,d,q)(P,D,Q)s$$

onde:

- p é a ordem do componente autorregressivo (AR) não sazonal;
- d é o grau de diferenciação não sazonal;
- q é a ordem do componente de médias móveis (MA) não sazonal;
- P,D,Q são os respectivos parâmetros sazonais;
- s representa a sazonalidade (neste caso, s=12).

A expressão geral do modelo é dada por:

$$\Phi P(Bs)\phi p(B)(1-B)d(1-Bs)DYt = \Theta Q(Bs)\theta q(B)\varepsilon t$$

em que:

- B é o operador de defasagem (BYt=Yt−1);
- ϕp(B) e θq(B) são polinômios não sazonais de ordem p e q;
- ΦP(Bs) e ΘQ(Bs) são polinômios sazonais de ordem P e Q;
- εt é o termo de erro aleatório (ruído branco).

A identificação da estrutura do modelo foi realizada por meio da análise gráfica da série e das funções de autocorrelação (ACF) e autocorrelação parcial (PACF), incluindo seus componentes sazonais.

Para verificar a estacionariedade, aplicaram-se os testes ADF (Augmented Dickey-Fuller) e KPSS (Kwiatkowski-Phillips-Schmidt-Shin). O ADF testa a hipótese nula de presença de raiz unitária e é representado pela seguinte equação:

$$\Delta Yt = \alpha + \beta t + \gamma Yt - 1 + i = 1 \sum p \delta i \Delta Yt - i + \varepsilon t$$

Após os testes indicarem não estacionariedade da série original, uma diferenciação foi aplicada (d=1d = 1d=1), alcançando estacionariedade conforme verificado pelo ADF (p-valor = 0,01) e pelo KPSS (p-valor > 0,1).

A estimação dos parâmetros foi realizada pelo método da máxima verossimilhança, utilizando a função auto.arima() do pacote forecast, que testa diversas combinações de ordens e seleciona aquela com menor Critério de Informação de Akaike (AIC) e Critério de Informação Bayesiano (BIC). Essas métricas são calculadas pelas expressões:

$$AIC = -2\ln(L) + 2k$$
$$BIC = -2\ln(L) + k\ln(n)$$

em que:

- L é o valor da função de verossimilhança;
- k é o número de parâmetros estimados;
- n é o tamanho da amostra.

O modelo escolhido foi SARIMA (1,1,0)(2,0,0)[12], com os seguintes coeficientes estimados:

- AR (1) = 0,8071

- SAR (12) = 0,3327

- SAR(24) = 0,4604

O diagnóstico dos resíduos foi conduzido com base em três testes: (i) **teste de Ljung-Box** para autocorrelação serial dos resíduos, (ii) **teste de Shapiro-Wilk** para verificação de normalidade, e (iii) **teste ARCH-LM** para heterocedasticidade.

O teste de Shapiro-Wilk tem a forma:

$$W = \frac{\left(\Sigma\, a_i\, X_{(i)}\right)^2}{\Sigma\,(X_i - \bar{X})^2}$$

em que:

- X(i): são os valores ordenados da amostra;

- ai: são coeficientes calculados com base na média e variância esperadas sob normalidade.

A avaliação da acurácia do modelo foi feita com base em duas métricas de erro amplamente utilizadas:

- Erro Médio Absoluto Percentual (MAPE):

$$MAPE = \frac{100}{n} * \Sigma \left|\frac{(Yt - \hat{Y}t)}{Yt}\right|$$

- Raiz do Erro Quadrático Médio (RMSE):

$$RMSE = \sqrt{(\,(1/n) * \Sigma\,(Yt - \hat{Y}t)\text{^}2\,)}$$

Os valores obtidos foram MAPE = 4,1% e RMSE = 0,27, indicando bom desempenho preditivo do modelo.

Adicionalmente, considerou-se a possibilidade de aplicar modelos alternativos de previsão como o ARIMA tradicional (sem componente sazonal), modelos ARIMAX com inclusão de variáveis exógenas e abordagens baseadas em rede neurais artificiais. No entanto, o ARIMA simples apresentou desempenho inferior em testes preliminares, com resíduos autocorrelacionacionados e métricas de erro menos satisfatórias. Os modelos ARIMAX foram descartados neste primeiro momento por limitações de disponibilidade de variáveis explicativas mensais compatíveis com a série interpolada. Quanto às redes neurais, embora ofereçam potencial de acurácia superior em certos contextos, enfrentam desafios de interpretabilidade e replicabilidade, especialmente em aplicações voltadas a políticas públicas. Dessa forma, optou-se pelo modelo SARIMA por seu equilíbrio entre robustez estatística, transparência metodológica e capacidade de capturar a sazonalidade explícita de série.

A previsão foi realizada para o período de fevereiro de 2024 a janeiro de 2025, com resultados mensais obtidos por meio da função forecast() aplicada ao modelo ajustado. Embora o modelo tenha sido estimado com base em dados interpolados, as previsões foram validadas comparando-se as médias trimestrais previstas com os valores reais divulgados pelo IBGE em 2024.

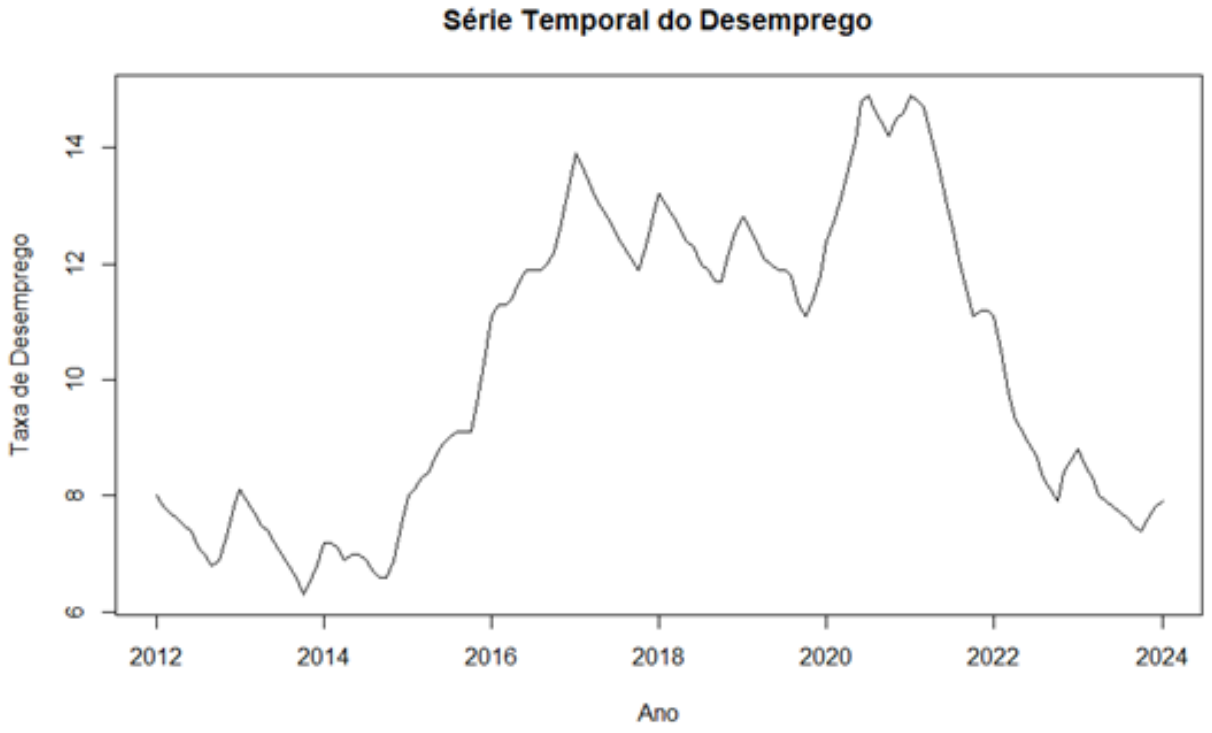
Por fim, reconhece-se que o modelo SARIMA, por ser univariado, apresenta limitações no que tange à sensibilidade a choques exógenos e políticas econômicas específicas. Eventos como estímulos fiscais, mudanças institucionais ou choques de oferta podem impactar o desemprego de forma que não seja captada pela estrutura histórica da série. A ausência de variáveis exógenas limita

sua capacidade explicativa, ainda que mantenha forte desempenho preditivo de curto prazo em contextos de relativa estabilidade.

4. RESULTADOS E DISCUSSÃO

A aplicação do modelo SARIMA (1,1,0)(2,0,0)[12] à série trimestral da taxa de desemprego no Brasil permitiu uma análise abrangente da estrutura temporal da variável e sua previsão para o período de fevereiro de 2024 a janeiro de 2025. Inicialmente, a série original, apresentada no Gráfico 1, evidencia um comportamento cíclico com tendência decrescente após o pico ocorrido entre 2016 e 2017. Destaca-se o impacto da pandemia em 2020, seguido de uma recuperação progressiva nos anos seguintes.

Gráfico 1 – Série Original da Taxa de Desemprego no Brasil (2012–2023)

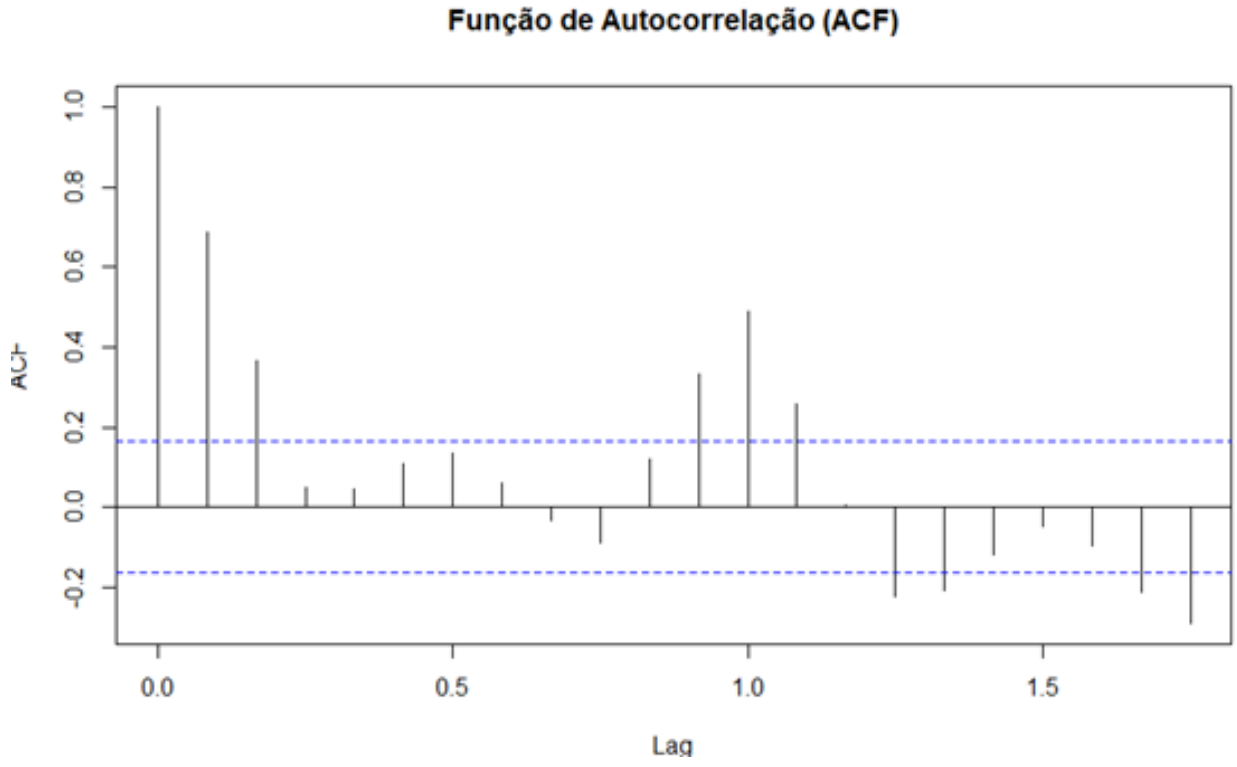


Fonte: Elaborado pelo autor através do Software RStudio.

O comportamento visual da série indicou a necessidade de diferenciação para garantir estacionariedade, especialmente por apresentar tendência de longo prazo e picos periódicos. Para confirmar isso, foram aplicados os testes ADF e KPSS. O ADF, com p-valor = 0,8965, não rejeitou a hipótese de presença de raiz unitária, enquanto o KPSS rejeitou a hipótese nula de estacionariedade (p < 0,01). Após a primeira diferenciação, os testes indicaram estacionariedade adequada.

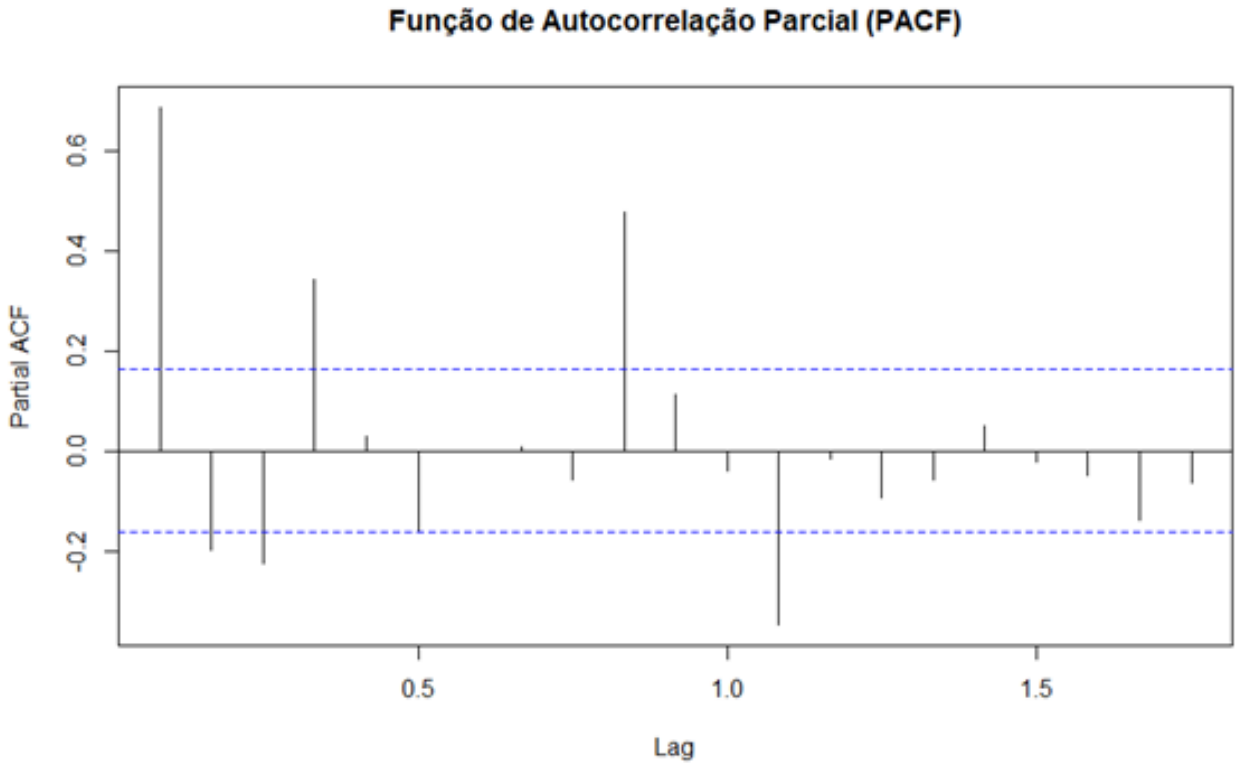
Na etapa de identificação do modelo, as funções de autocorrelação e autocorrelação parcial foram analisadas. O Gráfico 2, correspondente à FAC (ACF), mostra autocorrelações significativas até defasagens longas, indicando persistência da série. Já o Gráfico 3, referente à FACP (PACF), mostra corte acentuado na defasagem 1, sugerindo a presença de um componente autorregressivo não sazonal.

Gráfico 2 – Função de Autocorrelação (FAC)



Fonte: Elaborado pelo autor através do Software RStudio.

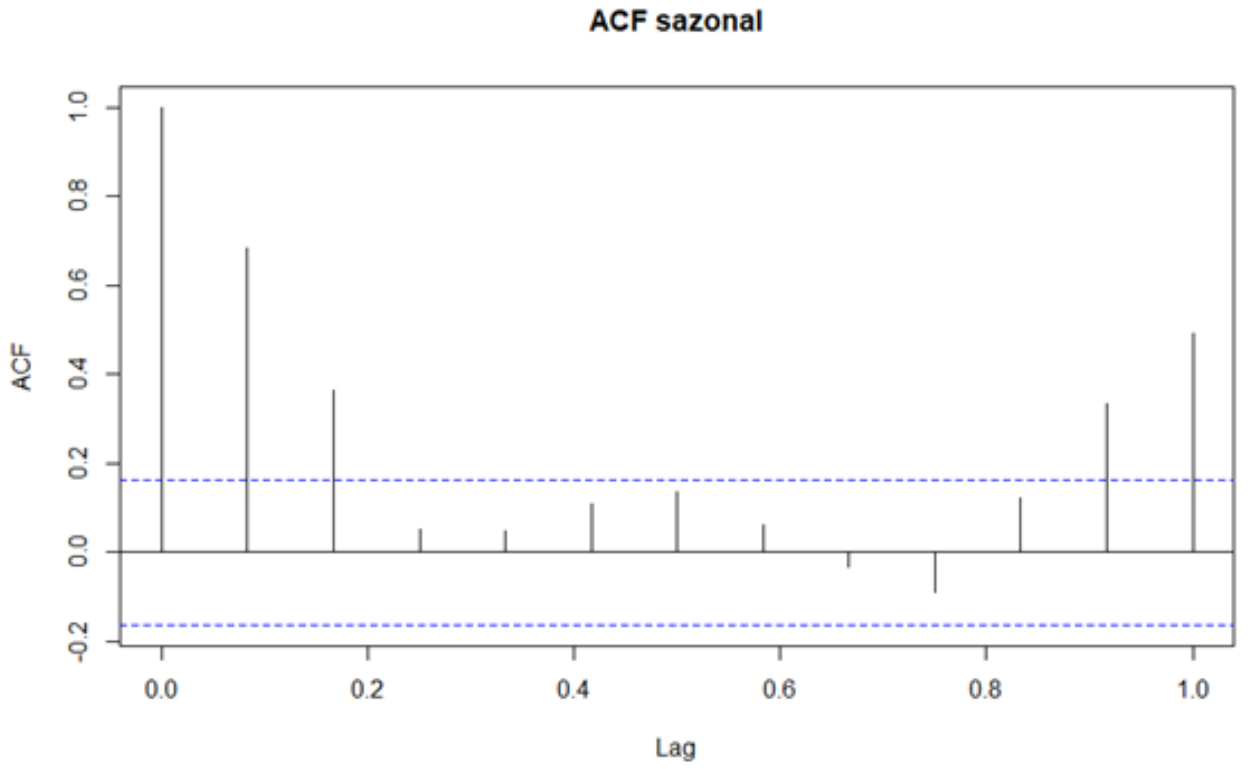
Gráfico 3 – Função de Autocorrelação Parcial (FACP)



Fonte: Elaborado pelo autor através do Software RStudio.

Além disso, o Gráfico 4 – PACF Sazonal evidencia picos marcantes nas defasagens múltiplas de 12, o que justifica a inclusão de dois componentes autorregressivos sazonais no modelo. Esses resultados embasaram a escolha da estrutura SARIMA (1,1,0)(2,0,0)[12].

Gráfico 4 – PACF Sazonal



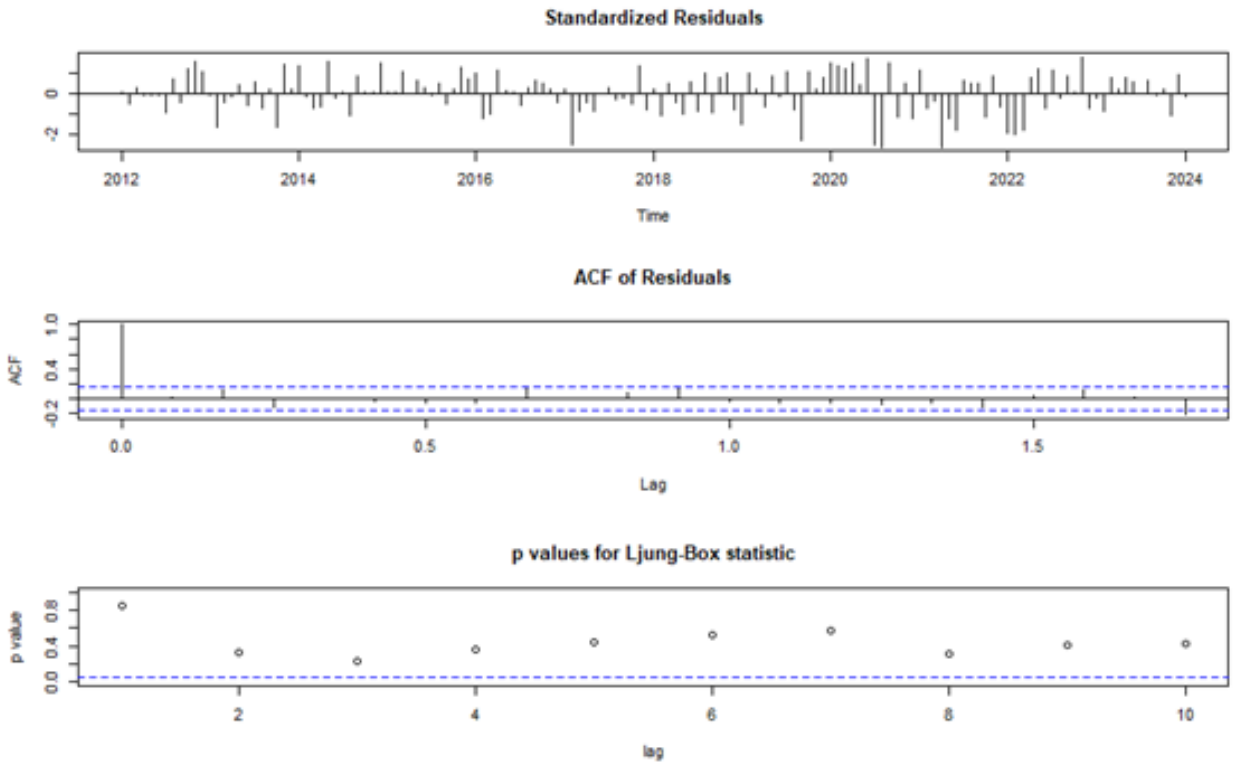
Fonte: Elaborado pelo autor através do Software RStudio.

A estimação do modelo revelou os seguintes coeficientes significativos:

- AR(1) = 0,8071
- SAR(12) = 0,3327
- SAR(24) = 0,4604

Com AIC = -108,39, o modelo demonstrou boa qualidade de ajuste. O diagnóstico dos resíduos, apresentado no gráfico 5, mostra comportamento aproximadamente aleatório, sem padrões discerníveis. O teste de Ljung-Box (p = 0,3294) não detectou autocorrelação serial, o que sugere que os resíduos são ruído branco. O teste de Shapiro-Wilk (p > 0,05) indicou normalidade dos resíduos, e o teste ARCH-LM confirmou ausência de heterocedasticidade significativa.

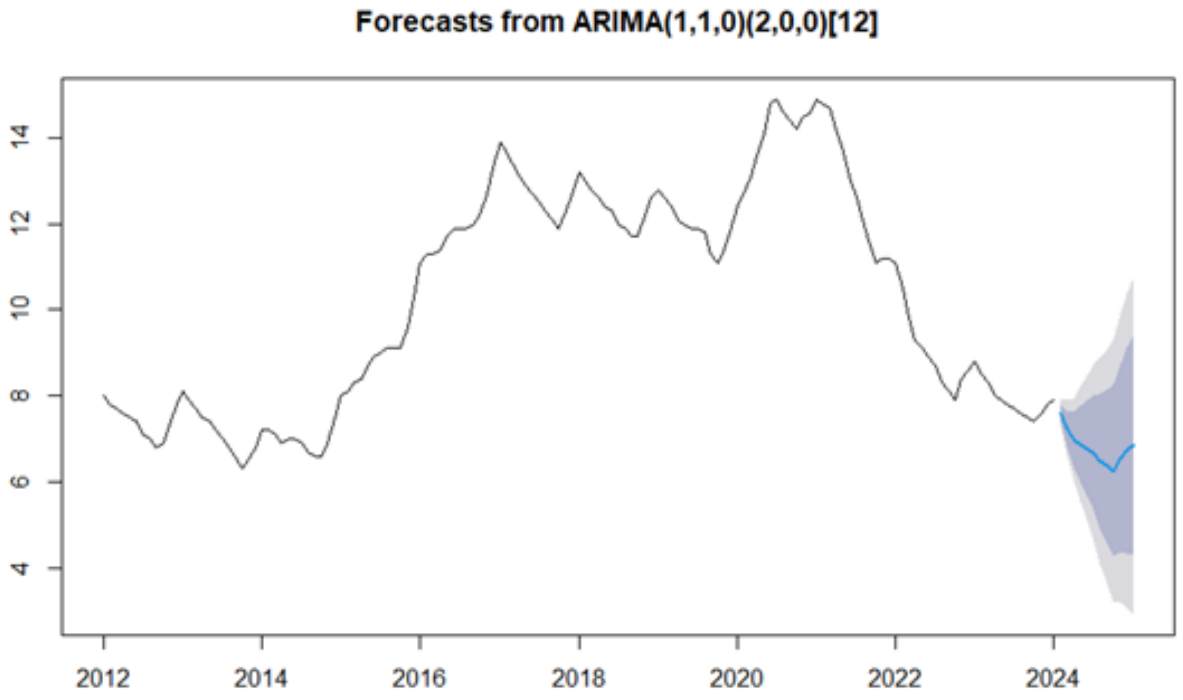
Gráfico 5 – Resíduos do Modelo SARIMA Ajustado



Fonte: Elaborado pelo autor através do Software RStudio.

A previsão mensal da taxa de desemprego para o período de fevereiro de 2024 a janeiro de 2025 é apresentada no gráfico 6. Nota-se uma tendência de queda gradual ao longo do ano, compatível com a recuperação econômica esperada e com a sazonalidade típica do mercado de trabalho brasileiro, que tende a aquecer no segundo semestre.

Gráfico 6 – Previsão da Taxa de Desemprego (fev/2024 a jan/2025)



Fonte: Elaborado pelo autor através do Softaware RStudio.

As estimativas variam de 7,59% em fevereiro para 6,85% em janeiro de 2025, com intervalos de confiança relativamente estreitos nos primeiros meses e progressivamente mais amplos à medida que o horizonte de previsão se estende. A Tabela 1 resume os valores pontuais e seus respectivos intervalos de confiança.

Tabela 1 - Análise da Previsão Taxa de Desemprego no Brasil

Mês/Ano	Previsão Pontual	Intervalo 80% (inferior)	Intervalo 80% (superior)	Intervalo 95% (inferior)	Interv 95% (super
Fev 2024	7.588131	7.387860	7.788401	7.281843	7.8944
Mar 2024	7.251109	6.737480	7.664739	6.618518	7.8837
Abr 2024	6.962912	6.319854	7.605971	5.979439	7.9463
Mai 2024	6.871306	5.993387	7.749225	5.528645	8.2139
Jun 2024	6.773192	5.661088	7.885295	5.072376	8.4740

⚠ Esta tabela possui muitas colunas e foi cortada para impressão. Para visualizá-la completa, acesse o artigo original em: https://revistatopicos.com.br/artigos/previsao-da-taxa-de-desemprego-no-brasil-com-modelo-estatistico-sarima-univariado-e-componente-sazonal-aplicacao-aos-dados-da-pnad-continua-2012-2023?noblockage

Fonte: Elaboração própria com auxílio do Software RStudio.

Para aferir a qualidade das previsões, os valores mensais foram agregados em médias trimestrais e comparados aos dados reais do IBGE para os trimestres de 2024. A Tabela 2 apresenta essa comparação, revelando erros modestos, variando de -0,65 p.p. no primeiro trimestre a +0,64 p.p. no último. Essas variações, embora dentro dos intervalos de confiança previstos, indicam que o modelo teve menor precisão em momentos de mudança mais abrupta, como no pico de contratações temporárias do fim do ano.

Tabela 2 - Comparação entre Previsão do Modelo e Desemprego Efetivo

Trimestre 2024	Desemprego Real	Previsão Pontual
Mar	7,9%	7,2%
Jun	6,9%	6,7%
Set	6,4%	6,3%
Dez	6,2%	6,8%

Fonte: Elaboração própria com resultados obtidos no Software RStudio.

Tais discrepâncias reforçam um ponto já consolidado na literatura: modelos univariados como o SARIMA são eficazes para capturar

padrões cíclicos e sazonais, mas são limitados na resposta a choques exógenos não contemplados pela estrutura da série (Marasca et al., 2017; Cavalcante & Silva, 2011). Ainda assim, o MAPE de 4,1% e o RMSE de 0,27 são considerados indicadores muito bons para previsões macroeconômicas trimestrais.

Do ponto de vista teórico, a capacidade do SARIMA de antecipar a tendência geral e os ciclos sazonais reforça sua aplicabilidade em contextos emergentes, onde a qualidade dos dados e a estabilidade institucional são desafiadoras. Do ponto de vista prático, os resultados sugerem que o modelo pode ser incorporado a sistemas de monitoramento do mercado de trabalho, fornecendo subsídios técnicos para a tomada de decisão por parte de governos e empresas.

Entre as limitações, destaca-se a ausência de variáveis explicativas, o que impede a sensibilidade do modelo a políticas fiscais, monetárias ou alterações estruturais no mercado. Recomenda-se, portanto, a ampliação do modelo para uma estrutura ARIMAX ou híbrida, incorporando machine learning de forma supervisionada, para obter maior responsividade sem sacrificar interpretabilidade.

Em suma, a análise dos gráficos e estatísticas confirma que o modelo SARIMA aplicado capturou adequadamente os principais padrões da série de desemprego brasileira e apresentou desempenho preditivo robusto, dentro de suas limitações estruturais.

5. CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo avaliar a capacidade do modelo SARIMA em prever a taxa de desemprego no Brasil, utilizando a série

da PNAD Contínua no período de 2012 a 2023, com projeções mensais para 2024. Para isso, a série trimestral foi interpolada por spline cúbico, possibilitando a aplicação de uma modelagem sazonal mais sensível às variações intraanuais do mercado de trabalho. O modelo estimado, SARIMA (1,1,0)(2,0,0)[12], demonstrou bom desempenho estatístico, com resíduos bem-comportados e métricas de erro dentro de padrões aceitáveis (MAPE = 4,1%; RMSE = 0,27), além de captar corretamente a tendência geral de queda da taxa de desemprego no ano previsto.

Entre os diferenciais da pesquisa, destaca-se a validação das previsões com dados reais efetivamente observados em 2024, abordagem ainda pouco explorada na literatura nacional. Essa comparação revelou consistência entre as estimativas e os valores observados, com erros pontuais explicados, sobretudo, pela ausência de variáveis exógenas. Tal resultado evidencia que, mesmo com estrutura univariada e sem considerar choques conjunturais diretos, o modelo SARIMA é capaz de oferecer previsões de curto prazo com nível razoável de precisão e aplicabilidade prática.

Do ponto de vista metodológico, a interpolação mensal, embora introduza valores não observados diretamente, mostrou-se eficaz para adaptar a série às exigências da modelagem sazonal. Reconhece-se, no entanto, que esse procedimento pode suavizar variações abruptas e reduzir a capacidade do modelo em capturar eventos pontuais, especialmente em contextos de alta volatilidade ou instabilidade institucional. Ainda assim, trata-se de uma escolha tecnicamente defensável e compatível com práticas consolidadas na literatura.

Do ponto de vista aplicado, os resultados obtidos demonstram o potencial do modelo SARIMA como ferramenta de monitoramento e análise conjuntural para instituições públicas, centros de pesquisa e agentes do mercado. A previsibilidade da taxa de desemprego, mesmo com dados univariados, representa um ganho importante para a formulação de políticas públicas mais responsivas às dinâmicas sazonais e estruturais do mercado de trabalho brasileiro.

Como direções futuras de pesquisa, recomenda-se o aprimoramento do modelo com a incorporação de variáveis explicativas via ARIMAX ou modelos estruturais, bem como o uso de abordagens híbridas com técnicas de aprendizado de máquina supervisionado. A análise desagregada por regiões, faixas etárias ou níveis educacionais também pode ampliar a capacidade diagnóstica da modelagem e gerar subsídios mais direcionados às políticas de emprego.

Por fim, espera-se que a presente pesquisa contribua para um melhor entendimento sobre os movimentos da taxa de desemprego no Brasil. A previsão baseada em modelos estatísticos contribui com a compreensão das variáveis que influenciam o desemprego no país. Essa compreensão é fundamental para o desenvolvimento de políticas públicas que incentivem a empregabilidade e a formalização das ocupações. Assim, este trabalho reforça a importância de se investir em políticas que priorizem a educação, a qualificação profissional, a inovação tecnológica e o empreendedorismo como estratégias de enfrentamento do desemprego.

REFERÊNCIAS BIBLIOGRÁFICAS

ANTUNES, R. **O privilégio da servidão: o novo proletariado de serviços na era digital**. São Paulo: Boitempo, 2010.

ARAÚJO, E. B.; PEREIRA, R. G.; SILVA, R. F. **Desigualdades no mercado de trabalho brasileiro: impactos da pandemia e perspectivas de recuperação**. Brasília: IPEA, 2022.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Time Series Analysis: Forecasting and Control**. 5. ed. Hoboken: Wiley, 2016.

CARLEIAL, L. A. **O mercado de trabalho e a persistência do desemprego estrutural no Brasil.** Revista de Economia Contemporânea, Rio de Janeiro, v. 19, n. 2, p. 179–202, 2015.

CARNEIRO, F. G.; GALVÃO, A. F. **Long-run determinants of the Brazilian unemployment rate.** Revista Brasileira de Economia, Rio de Janeiro, v. 65, n. 2, p. 157–176, 2011.

CAVALCANTE, L. R.; SILVA, R. S. **A crise econômica de 2008 e os efeitos sobre o mercado de trabalho brasileiro**. Revista de Economia Contemporânea, Rio de Janeiro, v. 15, n. 3, p. 481–506, 2011.

DIAMOND, P.; MORTENSEN, D. T.; PISSARIDES, C. A. **The Beveridge Curve**. Brookings Papers on Economic Activity, Washington, v. 1, p. 23–74, 1994.

ENDERS, W. **Applied Econometric Time Series**. 4. ed. Hoboken: Wiley, 2015.

FRIEDMAN, M. **The Role of Monetary Policy**. American Economic Review, v. 58, n. 1, p. 1–17, 1968.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: Principles and Practice**. 3. ed. Melbourne: OTexts, 2021. Disponível em: https://otexts.com/fpp3/. Acesso em: 10 mai. 2025.

IBGE – Instituto Brasileiro de Geografia e Estatística. **PNAD Contínua – Taxa de Desocupação**. Rio de Janeiro, 2023. Disponível em: https://www.ibge.gov.br. Acesso em: 5 mai. 2025.

IBGE – Instituto Brasileiro de Geografia e Estatística. **Taxa de desocupação no Brasil atinge menor nível desde 2014**. Notícia IBGE, 2025. Disponível em: https://www.ibge.gov.br. Acesso em: 20 mai. 2025.

IPEA – Instituto de Pesquisa Econômica Aplicada. **Mercado de Trabalho: conjuntura e análise**. Brasília: IPEA, 2023. (Boletim Mercado de Trabalho, n. 71).

JANSEN, M.; UEKULL, E. VON. **Productivity, technology and structural unemployment**. International Labour Review, v. 149, n. 3, p. 217–232, 2010.

KEYNES, J. M. **A teoria geral do emprego, do juro e da moeda.** São Paulo: Nova Cultural, 1985. (Original publicado em 1936).

MAKRIDAKIS, S.; SPILIOTIS, E.; ASSIMAKOPOULOS, V. **Statistical and machine learning forecasting methods: Concerns and ways forward**. PLoS ONE, v. 13, n. 3, e0194889, 2018.

MARASCA, R. B.; LIMA, R. C.; FAVERO, V. A. **Previsão do desemprego no Brasil: uma aplicação dos modelos ARIMA e SARIMA.** Revista da FAE, Curitiba, v. 20, n. 2, p. 187–202, 2017.

MAZON, D. L. **Previsão da taxa de desemprego no Brasil utilizando modelos SARIMA.** Trabalho de Conclusão de Curso – Universidade Estadual do Centro-Oeste, Guarapuava, 2024.

OIT – Organização Internacional do Trabalho. **Perspectivas sociais e do emprego no mundo – Tendências 2022**. Genebra: OIT, 2022.

PAUL, K. I.; MOSER, K. **Unemployment impairs mental health: Meta-analyses.** Journal of Vocational Behavior, v. 74, n. 3, p. 264–282, 2009.

SILVA, F. T.; COURSEUIL, C. H. F.; COSTA, A. C. **Trabalho informal e desigualdade: evidências recentes no Brasil.** Texto para Discussão n. 2605. Brasília: IPEA, 2020.

SOUZA, R. M. **Econometria básica com aplicações usando o software R**. São Paulo: Atlas, 2016.

WINKELMANN, L**. Unemployment, social capital, and subjective well-being**. Journal of Happiness Studies, v. 10, n. 4, p. 421–430, 2009.

ZHANG, G.; PATUWO, B. E.; HU, M. Y. **Forecasting with artificial neural networks: The state of the art**. International Journal of Forecasting, v. 14, p. 35–62, 1998.

---

[1] Mestrando do Programa de Pós-Graduação em Economia da Universidade Estadual de Ponta Grossa (PPGECO/UEPG). E-mail: acesse o artigo original para visualizar o e-mail

[2] Contratação via PJ: Forma de contratação em que a empresa contrata como Pessoa Jurídica, sem vínculo CLT, evitando encargos

trabalhistas.

[3] Modelo de trabalho mediado por plataformas digitais, como Uber ou iFood, em que o trabalhador atua como autônomo, é pago por tarefa realizada e não tem vínculo empregatício nem garantias da CLT.

[4] Quando um modelo aprende excessivamente os dados de treinamento, incluindo padrões irrelevantes ou ruídos, o que prejudica seu desempenho em novos dados.

[5] Spline cúbico: Função polinomial por partes composta por segmentos de grau 3 que interpola pontos dados, garantindo continuidade até a segunda derivada (suavidade C^2). Usado em modelagem de curvas, análise numérica e gráficos computacionais.