



Revista

Cadernos de Finanças Públicas

03 | 2026



Previsão da Arrecadação de ICMS de São Paulo: Experimento com Econometria, *Deep Learning* e *Foundation Model*, sob Abordagens *Top-Down* e *Bottom-Up*.

Vinícius Inamura Yosimoto

Cássio Machiaveli Oishi

RESUMO

O Imposto sobre a Circulação de Mercadorias e Serviços (ICMS) é o principal componente da receita tributária do Estado de São Paulo e, com a promulgação da EC nº 132/2023, sua base integrará a transição para o Imposto sobre Bens e Serviços (IBS), em transição gradual até 2033. Assim, a previsão do ICMS torna-se fator crítico tanto para o planejamento quanto para a execução orçamentária, o que requer uma metodologia com bom desempenho preditivo, tanto em previsões anuais quanto mensais. Destarte, foram avaliadas sete metodologias de previsão, SARIMA, SARIMAX em diferenças, SARIMAX em nível, ECM, LSTM, TimesFM e SCINet, sob as abordagens *Top-Down* e *Bottom-Up*, em um experimento com validação em 12 janelas de previsão. O melhor desempenho preditivo foi alcançado pelos modelos econométricos SARIMAX e ECM, seguidos pela rede neural recorrente LSTM, sobretudo sob abordagem *Bottom-Up*, o que pode constituir valiosa lição para a futura previsão do IBS.

Palavras-Chave: Previsão; séries temporais; política tributária; aprendizado de máquina; econometria.

Classificação JEL: C53; H20; H71.

SUMÁRIO

1. INTRODUÇÃO	5
2. REVISÃO DE LITERATURA	6
3. METODOLOGIA	10
3.1. Base de Dados.....	11
3.2. Modelos de Previsão	14
3.3. Ajuste de Hiperparâmetros e Treinamento de Modelos de Deep Learning	15
3.4. Avaliação de Desempenho Preditivo.....	16
3.5. Experimento	18
4. RESULTADOS	21
4.1. Econometria	21
4.2. Deep Learning.....	29
4.3. Foundation Model.....	30
5. DISCUSSÃO	31
5.1. Ranking por Desempenho Ordinal.....	31
5.2. Ranking por Desempenho Cardinal	33
5.3. Conclusões	35
REFERÊNCIAS BIBLIOGRÁFICAS.....	38
APÊNDICE.....	43

FIGURAS

Figura 1 - Série histórica mensal do ICMS, por regime de arrecadação componente (R\$ bilhão)	6
Figura 2 - Séries temporais em números-índice (01/01/2004 = 100)	14
Figura 3 - Ranking anual, com base no EAPA médio.....	34
Figura 4 - Desempenho preditivo anual: Dispersão de médias e desvios padrão do EAPA.....	34
Figura 5 - Ranking mensal, com base no MAPE médio.....	35
Figura 6 - Desempenho preditivo mensal: Dispersão de médias e desvios padrão do MAPE.	35

TABELAS

Tabela 1 - Base de Dados Utilizada no Experimento.....	11
Tabela 2 – Exogeneidade	27
Tabela 3 - Variáveis exógenas da rede LSTM.....	29
Tabela 4 - Variáveis exógenas de cada regime de arrecadação.....	30
Tabela 5 - Ranking por desempenho ordinal anual.....	32
Tabela 6 - Ranking por desempenho ordinal mensal.	33
Tabela 7 - Síntese do desempenho das metodologias de previsão.....	43
Tabela 8 - Desempenho das metodologias por janela de previsão.....	43
Tabela 9 - Teste de Cointegração Johansen.....	46
Tabela 10 - Previsão da rede LSTM, sob abordagem Top-Down.	46
Tabela 11 - Previsão da rede LSTM, sob abordagem Bottom-Up, regime ICMS_RPA.....	47
Tabela 12 - Previsão da rede LSTM, sob abordagem Bottom-Up, regime ICMS_Importacao. ..	47
Tabela 13 - Previsão da rede LSTM, sob abordagem Bottom-Up, regime ICMS_ST	48
Tabela 14 - Previsão da rede SCINet, sob abordagem Top-Down.....	48
Tabela 15 - Previsão da rede SCINet, sob abordagem Bottom-Up, regime ICMS_RPA.	49
Tabela 16 - Previsão da rede SCINet, sob abordagem Bottom-Up, regime ICMS_Importacao. ..	49
Tabela 17 - Previsão da rede SCINet, sob abordagem Bottom-Up, regime ICMS_ST.	50

1. INTRODUÇÃO

A partir de dados da Secretaria da Fazenda e Planejamento do Estado de São Paulo¹, afere-se que a arrecadação do Imposto sobre Operações relativas à Circulação de Mercadorias e sobre Prestações de Serviços (ICMS) representou cerca de 82% da receita tributária paulista em 2025. Assim, atualmente, o ICMS figura como o principal tributo estadual. Em adição, consoante a Reforma Tributária promovida pela EC nº 132/2023, as bases tributárias do ICMS e do Imposto Sobre Serviços de Qualquer Natureza (ISS) irão compor o escopo de incidência do Imposto sobre Bens e Serviços (IBS), em transição gradual até 2033.

Dessa forma, boas práticas de previsão de arrecadação do ICMS tornam-se fundamentais para o planejamento e a gestão das contas públicas do Estado.

A arrecadação do ICMS é um fenômeno cuja natureza é tanto econômica quanto jurídica. Por um lado, o valor arrecadado depende da ocorrência das atividades econômicas que constituem o seu fato gerador, o que o torna sujeito às flutuações sistêmicas da economia. Por outro lado, inovações de legislação como benefícios fiscais e alterações de regras de cálculo e pagamento do imposto impactam diretamente o montante arrecadado. Como resultado (Figura 1), as séries históricas mensais do ICMS apresentam, além de tendência e sazonalidade, volatilidade e quebras estruturais, o que torna o trabalho de previsão desafiador. Os regimes de arrecadação antecipados na Figura 1 serão detalhados na seção 3.1.

Grosso modo, para fins orçamentários, o Estado deve apresentar uma previsão de receita tributária anual. Em adição, a gestão do caixa requer previsões mensais, para a elaboração da programação financeira e do fluxo de desembolsos ao longo do exercício. Portanto, este duplo mandato demanda uma metodologia que alcance excelência preditiva tanto em sua previsão anual quanto mensal. Historicamente, para tal propósito, tem-se utilizado metodologias de previsão econométrica, sobretudo modelos da família SARIMAX, cujo desempenho tem sido considerado satisfatório.

Todavia, a performance preditiva da metodologia atual carece de contraponto. Pois, até onde se sabe, a literatura não oferece comparação sistemática, validada em múltiplas janelas de previsão, que avalie o desempenho preditivo de metodologias econométricas, frente a modelos de *deep learning* e *foundation model*, em contexto de previsão de arrecadação tributária estadual.

1 Capítulo 3, Tabela 1.2a <https://portal.fazenda.sp.gov.br/acessoinformacao/Paginas/Relat%C3%B3rios-da-Receita-Tribut%C3%A1ria.aspx>. Acesso em: 24/06/2026.

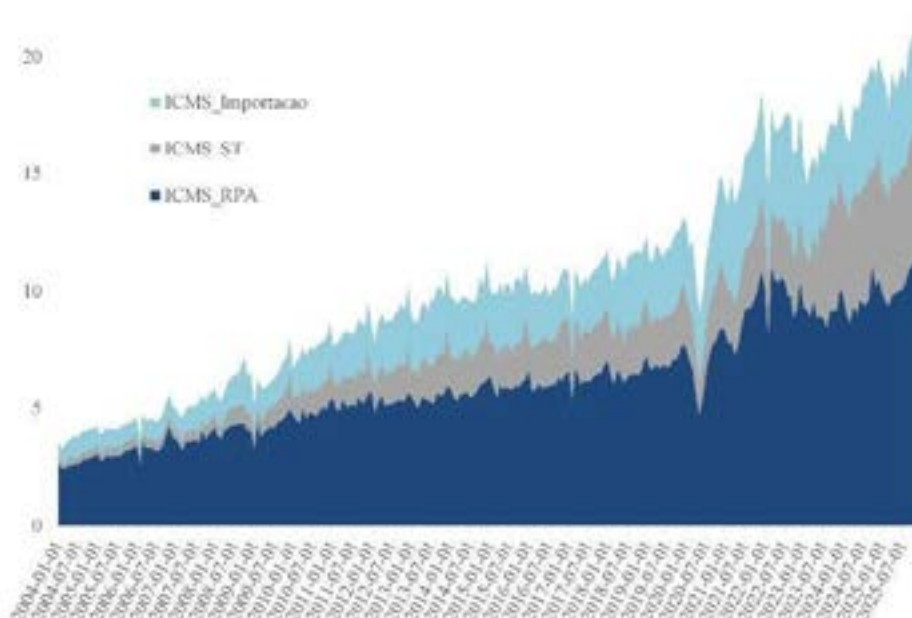
Na seção 2, revisou-se a literatura recente de previsão de ICMS, bem como as publicações sobre previsões de séries temporais, em geral. No segundo caso, privilegiaram-se revisões abrangentes, com a comparação de diversas metodologias de previsão.

Destarte, foram selecionadas 14 metodologias de previsão, constituídas pelos modelos SARIMA, SARIMAX em diferenças, SARIMAX em nível, ECM, rede neural recorrente LSTM, rede SCINet e TimesFM, sob as abordagens *Top-Down* e *Bottom-Up*.

Conforme detalhado na seção 3, as 14 metodologias foram avaliadas em experimento composto por 12 janelas de previsão, para mitigar eventual viés de seleção temporal. Em seguida, as performances preditivas resultantes (seção 4) foram ordenadas em duas modalidades distintas de *rankings* (seção 5).

Por conseguinte, concluiu-se que – de forma geral – as previsões mais precisas foram obtidas pelos modelos econométricos ECM e SARIMAX, seguidos da rede neural recorrente LSTM, sobretudo sob abordagem *Bottom-Up*, que acarreta a decomposição da receita do ICMS em regimes de arrecadação, cada qual com modelagem e previsão própria, com posterior consolidação.

Figura 1 - Série histórica mensal do ICMS, por regime de arrecadação componente (R\$ bilhão).



Fonte: SEFAZ-SP

2. REVISÃO DE LITERATURA

Em repositórios como os do *Institute of Electrical and Electronics Engineers* (IEEE) e *arxiv.org*, bem como em plataformas como o *Google Scholar*, buscou-se, de maneira específica, trabalhos sobre previsão de arrecadação do ICMS e, de modo mais amplo, artigos sobre metodologias de previsão de séries temporais, publicados a partir de 2015. Em primeiro lugar, objetivou-se identificar metodologias de previsão de séries temporais que já apresentavam desempenho promissor, quando aplicadas à arrecadação de ICMS. Em segundo lugar, buscou-se técnicas com precisão preditiva promissora, mesmo que aplicadas a séries temporais diversas.

As principais metodologias de previsão encontradas podem ser enquadradas em três grandes categorias, a econometria clássica – a qual pertencem o SARIMA e o SARIMAX –, o aprendizado profundo ou *deep learning* – que abarca a rede neural recorrente LSTM e a rede SCINet – e o *foundation model* que tem o TimesFM como exemplar.

Foram selecionadas cinco publicações sobre previsão de arrecadação de ICMS, tanto com técnicas estatísticas clássicas quanto com modelos de *deep learning*, que serão apresentados a seguir.

Bonato (2023) estudou duas abordagens para a previsão de arrecadação de ICMS para o Estado de São Paulo. Partiu-se da hipótese de que o processo gerador da série temporal do ICMS é não linear, por conta de suas quebras estruturais e, por isso, modelos não lineares apresentariam melhor desempenho na previsão de arrecadação do imposto. Sem mencionar referência específica, o autor afirmou que - atualmente - a Secretaria da Fazenda e Planejamento do Estado de São Paulo utiliza modelos econométricos *Seasonal Autoregressive Integrated Moving Average Exogenous* (SARIMAX) para realizar a previsão de receita de ICMS. Trata-se de modelo linear, em que as quebras estruturais são tratadas com variáveis *dummy*. Segundo o autor, isso resultaria em modelo sobreajustado, “pouco adequado para previsões de longo prazo”. Entretanto, tal hipótese não foi corroborada pelos resultados de seu estudo. Propôs-se - como contraponto - a previsão de ICMS com os modelos não lineares *Long Short-Term Memory* (LSTM) e *Markov-Switching*. Assim, foram realizadas diversas previsões, com versões univariada e de regressão múltipla para os dois grupos de modelos, lineares e não lineares. Como variáveis exógenas dos modelos de regressão múltipla foram utilizados IGP-M, IBC-BR e dólar. Para comparar o desempenho dos modelos, utilizou-se o *Model Confidence Set* (MCS) – conforme Hansen *et al.* (2011) – com o erro absoluto percentual médio (MAPE) como função perda. Por conseguinte, o ranqueamento resultante apontou o SARIMAX com regressão múltipla como o melhor modelo, com MAPE igual a 4,7%, seguido pelo SARIMAX univariado em segundo lugar (MAPE: 4,78%) e pelo *Markov-Switching* regressão múltipla (MAPE: 4,75%),

em terceiro. As redes neurais LSTM foram descartadas, por conta de seu desempenho irregular e dificuldade de parametrização.

Em um contexto de crise financeira do Estado do Rio de Janeiro, em que a precisão da previsão de arrecadação do ICMS se torna um fator ainda mais crítico para o planejamento das finanças públicas, Dornelas *et al.* (2022) propôs um experimento com redes neurais recorrentes *Long Short-Term Memory* (LSTM), para aprimorar a instrução da tomada de decisão das autoridades governamentais. O conjunto de dados compreendeu os valores mensais para as variáveis arrecadação de ICMS, PIB e Consumo de Energia Elétrica Comercial da Região Sudeste, de janeiro/2002 a dezembro/2019. Os dados de 2002 a 2016 integraram o conjunto de treinamento, os de 2017 e 2018 serviram o propósito de validação e os de 2019 restaram para o teste. A atual metodologia de previsão da Secretaria da Fazenda do Rio de Janeiro, que não é especificada no trabalho, apresentou erro relativo de 9% e 5,83% para os anos de 2017 e 2018. O modelo LSTM multivariado reduziu o erro relativo da previsão para 0,41% e 1,65%, respectivamente. Em trabalho anterior e correlato, Silva e Figueiredo (2020) utilizou uma rede LSTM univariada, alimentada apenas com valores derivados da arrecadação de ICMS do RJ, como médias móveis e valores defasados. Os dados de 2002 a 2016 compuseram o conjunto de treinamento, com validação em dados de 2017 e 2018 e teste com valores de 2019. Assim, obteve-se previsão para 2019 com erro relativo de 0,13%, em contraposição a erro relativo de 1,49%, resultante de previsão oficial da SEFAZ-RJ.

A SEFAZ/RS publicou texto para discussão (Dornelles *et al.*, 2022), com a aplicação de uma rede LSTM para a previsão de arrecadação, inspirada na metodologia apresentada em Silva e Figueiredo (2020). Foram utilizados os dados mensais de 2009 a 2017 para o treinamento do modelo e os dados 2018 a 2020 para teste. Por fim, realizou-se a previsão fora da amostra para o período de jan/2021 a ago/2021 e obteve-se erro relativo de -2,3%.

Em Fontenele (2017), utilizaram-se modelos econométricos do tipo SARIMA, selecionados pelos critérios de Akaike e Schwarz e estimados com dados mensais de jan/1998 a dez/2015, para prever a arrecadação de ICMS de 2016, para o Estado do Ceará. Os modelos apresentaram maior precisão preditiva nos primeiros 7 meses, em que os erros absolutos percentuais médios (MAPE) ficaram em torno de 1%. A precisão decaiu nos meses restantes, em que os MAPES superaram 6%. Assim, o erro relativo absoluto anual resultante gravitou em torno de 3,2%.

Em adição, foram encontrados estudos com modelos de *deep learning* e *foundation model*, aplicados à previsão de séries temporais, conforme detalhado abaixo.

O estudo de Masini *et al.* (2023) promoveu uma revisão não exaustiva da previsão de

séries temporais com modelos de *Machine Learning* (ML) e modelos estatísticos com alta dimensionalidade. Em específico, foi realizado exercício empírico de previsão da variância diária do Índice da Bolsa de Valores de São Paulo (IBOVESPA). Foram utilizados dados diários de 02/02/2000 a 21/05/2020, em um total de 4200 observações de variância diária do IBOVESPA, bem como de índices de bolsas de outros países como S&P500 (EUA), FTSE100 (Reino Unido), DAX (Alemanha), *Hang Seng* (Hong Kong) e *Nikkei* (Japão), elencados como variáveis regressoras candidatas. Em adição, para a estimação ou treinamento dos modelos, foi considerada uma janela deslizante de 1500 observações, para a previsão da variância um dia adiante. Como base de comparação (*benchmark*), utilizou-se um modelo de regressão *Heterogeneous Autoregressive* (HAR). Como modelos desafiantes, foram propostos um modelo de regressão penalizada HARX-LASSO, que é uma extensão do HAR com variáveis exógenas, um modelo de Regressão com Floresta Aleatória e redes neurais com diferentes números de camadas escondidas. Para comparar o resultado das previsões, recorreu-se às estatísticas QLIKE, segundo a qual nenhum modelo foi capaz de rejeitar a hipótese nula de equidade de capacidade preditiva em relação ao *benchmark*, e Erro Quadrático Médio (MSE), que indicou que as redes neurais apresentaram performance preditiva melhor que o *benchmark*, quando considerada toda a amostra, porém com poder preditivo menos expressivo durante períodos de maior volatilidade como a Crise do *Subprime* (2007-2008) e a pandemia de COVID-19 (2020). O modelo HARX-LASSO foi proclamado o vencedor entre os desafiantes, pois demonstrou maior precisão preditiva tanto para o conjunto de dados como um todo quanto para os períodos de maior volatilidade já mencionados. Ainda, concluiu-se que o modelo de Regressão com Floresta Aleatória não obteve boa performance preditiva em qualquer período.

Uma revisão de modelos *deep learning* para previsão de séries temporais foi apresentada em Han *et al.* (2021). Na indústria siderúrgica, o gerenciamento do fluxo de gases que carregam energia é um aspecto operacional crítico. Por isso, é importante prever o volume de gases gerado a cada instante. Assim, foram empregados os modelos de redes neurais recorrente (RNN), convolucional (CNN) e LSTM, classificados como discriminativos, bem como os modelos generativos *Generative Adversarial Networks* (GAN) e *Deep Belief Network* (DBN), além de um modelo híbrido que combinou *Autoencoder* (AE) com uma rede LSTM, para a previsão de geração de gases dos tipos *Blast Furnace Gas* (BFG) e *Coke Oven Gas* (COG), ao longo de 60 minutos. Para a geração do tipo BFG, os modelos com melhor performance preditiva foram LSTM (MAPE: 1,42%), DBN (MAPE: 1,87%) e híbrido (MAPE: 3,2%). Já para a geração COG, os modelos cujas previsões foram mais precisas foram LSTM (MAPE: 0,35%) e híbrido

(MAPE: 0,45%).

Li e Law (2024) apresentou revisão de modelos de *deep learning* para previsão de séries temporais desenvolvidos recentemente. Além de descrever os modelos teoricamente, os autores confrontaram os seus desempenhos preditivos, em um experimento empírico. Para tal, escolheram-se dois conjuntos de dados de temperaturas de transformadores elétricos, um com valores por hora e o outro por minuto, ambos se estendendo pelo período de 2 anos. Assim, foram comparados o Erro Quadrático Médio e o Erro Absoluto Médio de 9 modelos com design distintos. O modelo com a melhor performance preditiva, nos dois conjuntos de dados e segundo ambos os indicadores, foi o SCINet, que combina uma rede neural convolucional (CNN) e um *Encoder Decoder* do tipo Seq2Seq.

Em abordagem distinta daquelas apresentadas acima, Das *et al.* (2024) apresentou um modelo base, previamente treinado com conjunto de dados composto por milhões de séries temporais, tanto históricas e publicamente disponíveis (80%) quanto geradas sinteticamente (20%), balanceado por granularidade temporal. Dessa forma, tal abordagem se assemelha a dos *Large Language Models* (LLM), que são treinados em corpus textuais massivos e que - ao receber uma entrada textual (questão) que lhe fornece contexto - retornam uma saída textual como resposta à questão postulada. Assim, de forma análoga, o chamado *Foundation Model for Times-Series Forecasting* (TimesFM), pré-treinado com um corpus massivo de séries temporais, recebe como entrada uma série temporal histórica e - diante de tal contexto - gera a previsão para o horizonte estabelecido. Para avaliar a performance do TimesFM, os autores compararam as suas previsões com a de modelos estatísticos e de aprendizado profundo, relativas a 3 bases de dados propositalmente deixadas de fora do seu corpus de treinamento. Concluiu-se que o TimesFM obteve maior precisão de previsão em duas das bases de dados e ficou dentro do intervalo de confiança dos três modelos mais precisos na terceira base de dados.

Logo, a partir da revisão literária supra, SARIMA, SARIMAX, LSTM, SCINet e TimesFM foram selecionados para integrarem um experimento de avaliação, no contexto da previsão de ICMS do Estado de São Paulo, em que seus desempenhos preditivos serão colocados à prova, em um arcabouço isonômico, composto por múltiplos períodos de treinamento e teste, para mitigar eventual viés de seleção temporal, e avaliados segundo as mesmas métricas de precisão preditiva. Em adição, o *Error Correction Model* (ECM) foi incorporado à avaliação, pelos motivos apresentados na seção 3.2, apesar de não ter figurado na revisão da literatura.

3. METODOLOGIA

Conforme já enunciado, o objetivo deste trabalho é avaliar metodologias de previsão de séries temporais, quando aplicadas à série de arrecadação de ICMS do estado de São Paulo. Assim, a precisão preditiva das metodologias em análise será avaliada por meio de janelas deslizantes crescentes, conforme detalhado abaixo. Grosso modo, a ideia é repetir o procedimento metodológico de preparação de dados, modelagem ou ajuste dos hiperparâmetros, estimação ou treinamento dos modelos e previsão 12 meses adiante, com conjuntos de treinamento distintos e crescentes. Assim, a avaliação não fica restrita a uma única seleção de conjuntos de treinamento e teste, mitigando-se eventual viés de seleção.

A seguir, encontram-se a descrição da base de dados, breve exposição teórica das metodologias avaliadas, bem como os critérios de avaliação das previsões resultantes.

3.1. Base de Dados

A base de dados selecionada possui periodicidade mensal e estende-se de 01/01/2004 a 01/12/2025, o que perfaz 264 períodos. Um resumo sobre as variáveis que a compõem encontra-se na Tabela 1 abaixo.

Tabela 1 - Base de Dados Utilizada no Experimento

Variável	Descrição	Observações	Fonte
ICMS	Arrecadação de ICMS ² do Estado de São Paulo.	Medido em R\$ correntes.	Sefaz/SP
ICMS_RPA	Arrecadação de ICMS ² do Estado de São Paulo, sob Regime Periódico de Apuração (RPA), somada àquela do Simples Nacional e de outros regimes menos representativos.	Medido em R\$ correntes.	Sefaz/SP
ICMS_ST	Arrecadação de ICMS ² do Estado de São Paulo, sob regime de Substituição Tributária (ST).	Medido em R\$ correntes.	Sefaz/SP
ICMS_Importacao	Arrecadação de ICMS do Estado de São Paulo decorrente de Importações de bens do exterior.	Medido em R\$ correntes.	Sefaz/SP
PIBI	Índice mensal de Produto Interno Bruto do Brasil, construído a partir do PIB trimestral (IBGE, cód. 6612), do PIB mensal (BCB, cód. 4380) e do IPCA (BCB, cód. 433).	Número-índice, em que jan/2004 = 100.	Elaboração própria, a partir de dados do IBGE e do BCB
IPCA	Índice Nacional de Preços ao Consumidor Amplo (BCB, cód. 433).	Número-índice, em que jan/2004 = 100.	IBGE

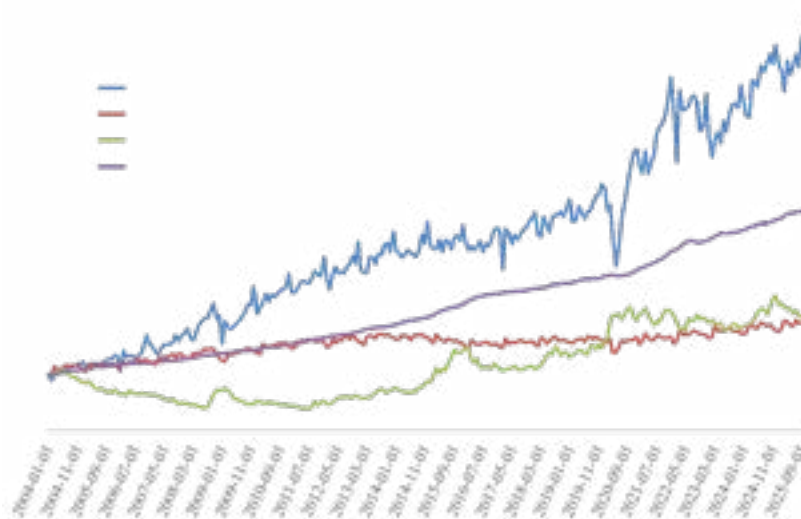
Câmbio	Taxa de câmbio - Livre - Dólar americano (venda) - Fim de período (cód. 3696)	Unidade monetária corrente (R\$)/US\$	Sisbacen PTAX800
DU	Número mensal de dias úteis	SGS12_NDIASUTEIS-PAS12	IPEADATA
<i>Dummies</i> diversas	Variáveis binárias, que só assumem os valores 0 ou 1.		

Fonte: Elaboração própria

O ICMS incide – grosso modo – sobre circulação de mercadorias, prestações interestaduais e intermunicipais de transporte e serviços de comunicação (Lei 6.374/89). Assim, o seu escopo de incidência é bastante abrangente, o que suscita a existência de regimes diversos de recolhimento. Trata-se de um imposto sobre valor agregado, então – de maneira geral – o seu recolhimento decorre da diferença entre o valor do imposto incidente sobre saídas (débitos) e entradas (créditos), sob o Regime Periódico de Apuração (RPA). Assim, tal regime congrega a maior parte da arrecadação do ICMS, representando 40,7% do total anual em 2025. A maioria dos contribuintes sujeitos ao RPA efetua o recolhimento do imposto no mês subsequente ao mês de referência da apuração de débitos e créditos. Logo, a arrecadação dos RPA de um dado mês tende a ser impactada por indicadores de atividade econômica do mês anterior. Em adição, agrega-se à variável ICMS_RPA o recolhimento de contribuintes enquadrados no Simples Nacional (3,2% do total), por conta da similaridade de prazos de recolhimento. Ainda, regimes de recolhimento menos expressivos, cujo valor recolhido representou 9% do total em 2025, também foram consubstanciados com o RPA. O segundo regime de recolhimento mais significativo é a Substituição Tributária (ST), com cerca de 26% da arrecadação de 2025. Em certos setores, há grande diferença de concentração econômica entre os diversos elos da cadeia de valor. Assim, setores mais concentrados como indústrias intensivas em capital tendem a apresentar um número menor de contribuintes, cuja fiscalização demanda menos recursos públicos. Assim, atribui-se a contribuintes de elos mais concentrados a responsabilidade de recolhimento do imposto incidente sobre elos mais pulverizados, cuja fiscalização seria mais onerosa ao erário. Quando isso ocorre, estabelece-se o regime de ST. Por exemplo, quando uma indústria se enquadra na ST, ela deve recolher antecipadamente o ICMS que incidirá sobre elos mais difusos como o varejo. Por conseguinte, quando levamos em consideração o prazo de recolhimento, a arrecadação a título de ST tende a ocorrer contemporaneamente à ocorrência da saída correspondente. De fato, constata-se que a variável ICMS_ST correlaciona-se com indicadores de atividade econômica contemporâneos. Por fim, há o regime relativo às importações, que representou 21,1%

da arrecadação de 2025. Tal regime estabelece que o recolhimento deve ser efetuado quando ocorrer o desembaraço aduaneiro. Assim, a variável ICMS_Importacao também estabelece relação contemporânea com indicadores de atividade econômica. Além disso, especificamente, o recolhimento referente a importações é impactado pela taxa de câmbio e pelo número de dias úteis do mês. A variável PIBI é o índice de atividade econômica, construído a partir da série de PIB trimestral do IBGE, mensalizada por meio da proporção que o valor de cada mês representa em seu respectivo trimestre, consoante PIB mensal nominal calculado pelo BCB, atualizado pelo IPCA até a data mais recente disponível. Por fim, a série mensal resultante foi transformada em número-índice, iniciado com o valor de jan/2004 igual a 100. As variáveis *dummies* foram utilizadas nos modelos econométricos para modelar ocorrências excepcionais ou mudanças de regimes. São variáveis binárias que assumem valores 0 ou 1. Em caso de choques econômicos, concentrados em períodos específicos, como foi a pandemia de COVID-19, utilizou-se *dummies* do tipo *Additive Outlier* (AO), que assume valor igual 1 durante o período da ocorrência e 0 no restante dos meses. Para tratar efeitos calendários pontuais, em que uma data de recolhimento significativo do ICMS é postergada para o mês subsequente, utilizou-se a diferença de *dummy* pontual (AO). Por exemplo, em certos anos, feriados de datas variáveis como o carnaval podem acarretar o deslocamento do recolhimento de um mês para o mês seguinte. Dessa forma, a diferença da *dummy* assumiria o valor 1 no mês que sofreu a redução de arrecadação e -1 no mês que observou elevação de arrecadação. Assim, o seu coeficiente seria negativo e teria ordem de grandeza que indica a magnitude da transferência de arrecadação de um mês para o outro. Por fim, foram utilizadas *dummies* do tipo *Level Shift* (LS), para capturar mudanças de regime como alterações da legislação tributária com impacto sobre a arrecadação do ICMS. Assim, a *dummy* LS assumiria valor igual a 0 antes da alteração de regime e 1 dali em diante. Outros tipos de *dummy* serão acompanhadas de descrição *ad hoc*, quando utilizadas.

Figura 2 - Séries temporais em números-índice (01/01/2004 = 100)



Fonte: Elaboração própria, com dados da SEFAZ/SP.

Para ilustrar a variação mensal das principais séries temporais da Tabela 1, de 2004 a 2025, construiu-se números-índice para as variáveis ICMS (R\$ correntes), PIB (R\$ dez/2025), Câmbio (R\$ correntes) e IPCA, em que o primeiro valor é igual a 100 e os demais refletem as variações subsequentes, conforme Figura 2.

3.2. Modelos de Previsão

Consoante a revisão da literatura apresentada na seção 2, destacaram-se modelos econométricos, de aprendizado de máquina e fundacional, por seu excelso desempenho preditivo.

Da econometria clássica, Fontenele (2017) obteve bons resultados, ao prever a arrecadação de ICMS do Estado do Ceará, com modelo da classe *Seasonal Autoregressive Integrated Moving Average* (SARIMA), que – consoante fundamentado em Box e Jenkins (1970) – contempla apenas componentes auto-regressivos (AR) e de médias móveis (MA), inclusive sazonais.

Já em Bonato (2023), com boa precisão em previsões de arrecadação de ICMS do Estado de São Paulo, destacou-se o modelo *Seasonal Autoregressive Integrated Moving Average with exogenous regressors* (SARIMAX), que - além dos componentes do modelo SARIMA - incorpora variáveis exógenas.

Embora não conste da revisão de literatura, o *Error Correction Model* (ECM), proposto por Engle e Granger (1987), constitui uma alternativa natural para o presente experimento de

previsão, pois admite a modelagem expressa da relação de longo prazo entre as séries temporais cointegradas, por meio do termo de correção de erro (ECT).

Salienta-se que apenas a performance preditiva em dados que não foram utilizados em sua estimação será considerada para atestar o valor de um modelo de previsão. Com efeito, relega-se a *status* secundário a avaliação das propriedades estatísticas dos estimadores dos parâmetros dos modelos. Não obstante, consoante Hyndman e Athanasopoulos (2021), um bom método de previsão produzirá resíduos de inovação não correlacionados. Pois, se houver autocorrelação dos resíduos, isso implica que ainda há informação a ser extraída dos dados e utilizada na previsão. Ademais, a média de tais resíduos deve ser zero, caso contrário, produz-se previsão viesada. Ao falhar em satisfazer tais condições, um método de previsão indica que pode ser aperfeiçoado. Dessa forma, tais propriedades dos resíduos serão consideradas condições necessárias, no âmbito da seleção dos modelos.

Dentre os modelos de *deep learning*, sobressaíram-se a rede neural recorrente *Long Short-Term Memory* (LSTM) e a rede SCINet (Liu *et al.*, 2022). A rede LSTM (Hochreiter e Schmidhuber, 1997) obteve resultados notáveis na previsão de arrecadação de ICMS dos Estados do RJ (Dornelas *et al.*, 2022) e RS (Dornelles *et al.*, 2022), bem como na previsão de geração de gases siderúrgicos (Han *et al.*, 2021). Ela incorpora a persistência de informações, por meio de um laço, o que torna propícia a sua utilização em estruturas de dados que apresentam algum encadeamento sequencial, como as séries temporais. Já a rede SCINet foi desenhada para capturar características de séries temporais como tendência e sazonalidade. Baseada em uma estrutura hierárquica em árvore com operações de subamostragem, convolução e interação, destacou-se em Li e Law (2024), ao apresentar o menor erro absoluto médio entre múltiplos modelos comparados.

Por fim, o TimesFM (Das *et al.*, 2024) é um *foundation model*, pré-treinado em corpus de séries temporais constituído por 300 bilhões de pontos, construído para realizar previsões *zero-shot*, ou seja, sem necessidade de qualquer preparação de dados de entrada, ajuste de hiperparâmetros ou mesmo treinamento. O modelo apenas recebe uma série temporal como entrada, que lhe provê contexto, e retorna a previsão para o horizonte temporal definido. Em Das *et al.* (2024), apresenta-se o resultado de performance comparativa, em que as previsões do TimesFM figuram entre os menores erros absolutos médios obtidos.

3.3. Ajuste de Hiperparâmetros e Treinamento de Modelos de Deep Learning

Previamente ao seu treinamento, é necessário estabelecer-se os valores dos hiperparâmetros dos modelos de *deep learning* como as redes LSTM e SCINet, que definem a sua arquitetura. Para tal, foi utilizado o algoritmo de *Otimização Bayesiana* (Jones *et al.*, 1998), cuja busca é – grosso modo – orientada por um modelo probabilístico, com fito de equilibrar a descoberta do desempenho de combinações de valores de hiperparâmetros em vizinhanças menos conhecidas e o refinamento da busca em vizinhanças em que o modelo probabilístico estima bom desempenho.

Por alto, uma vez estabelecidos os valores dos hiperparâmetros, o treinamento de modelos de aprendizado profundo como as redes LSTM e SCINet se dão de forma iterativa, em sucessivas épocas, em que os parâmetros dos modelos são ajustados, de forma a minimizar a função-perda. Com o fito de obter maior robustez e uma convergência mais célere, utilizou-se o algoritmo *Adaptive Moment Estimation* - Adam (Kingma e Ba, 2015), que estabelece taxas de aprendizado adaptativas e eficientes para cada parâmetro e época de treinamento.

Além disso, para mitigar a ocorrência do sobreajuste utilizou-se a parada antecipada, que interrompe o treinamento, quando o decréscimo marginal da função-perda se mostra consistentemente insignificante. Dessa forma, reduz-se o risco de o modelo tornar-se sobreajustado ao conjunto de treinamento, o que visa melhorar sua performance preditiva fora da amostra.

3.4. Avaliação de Desempenho Preditivo

Reitera-se que o principal critério de avaliação de metodologias de previsão é o seu desempenho preditivo fora da amostra de dados utilizados em sua estimação ou treinamento. A seguir, apresentam-se as métricas utilizadas para se aferir a performance da previsão, o arcabouço da janela deslizante crescente, cujo propósito é mitigar o viés de seleção de períodos específicos de treinamento ou estimação e teste ou previsão, bem como a metodologia para ordenar o desempenho dos modelos.

No que tange à gestão fazendária, duas competências se apoiam intensivamente na previsão de arrecadação. Para fins orçamentários, a previsão de receita tributária anual constitui elemento fulcral de planejamento, tanto material quanto legalmente. Logo, a precisão da previsão anual torna-se um aspecto crítico. Dessa forma, adotou-se o Erro Absoluto Percentual Agregado (EAPA) de 12 meses (Equação 2), para se aferir a performance da previsão anual. Além disso, a gestão de caixa é diretamente impactada pela precisão da previsão da arrecadação mensal. Deste modo, também foi aferido o erro absoluto percentual médio (Equação 1), cujo

acrônimo em inglês é MAPE (*Mean Absolute Percentual Error*), que reflete a precisão mensal da previsão.

$$MAPE = \frac{1}{H} \sum_H \left(\frac{Y_t - \hat{Y}_t}{Y_t} \right)$$

(Equação 1)

$$EAPA = \frac{\frac{1}{H} \sum_H Y_t - \frac{1}{H} \sum_H \hat{Y}_t}{\frac{1}{H} \sum_H Y_t}$$

(Equação 2)

H: número de períodos do horizonte de previsão (H=12)

Y_t : valor da série no instante t

$\hat{Y}_t$: valor previsto pelo modelo para o instante t

Ao avaliar o desempenho de modelos de previsão, tanto econométricos quanto de aprendizado de máquina, há um conjunto de dados de treinamento e um de teste. Assim, grosso modo, utiliza-se o conjunto de treinamento para estimar ou treinar os modelos e – em seguida – vale-se do conjunto de teste para realizar a previsão e avaliar a sua precisão. Contudo, o resultado da avaliação condiciona-se à escolha dos conjuntos de treinamento e teste, o que pode acarretar viés de seleção.

Diante de uma estrutura de dados de séries temporais, em que os dados possuem ordem cronológica ou dependência temporal, não é possível segmentar aleatoriamente os dados em conjuntos de treinamento e teste, sem risco de vazamento temporal. Pois, em uma seleção aleatória de conjuntos de treinamento e teste, há o risco de violação da ordem cronológica dos dados. Assim, corre-se o risco de se utilizar dados do futuro para se prever o passado.

Por conseguinte, recorre-se à validação por meio de janelas crescentes e deslizantes. Nessa abordagem, é definida uma janela inicial de treinamento do período t_1 ao período t_j . Daí, treina-se o modelo com os dados dessa janela, realiza-se a previsão de H pontos à frente, para o período de t_{j+1} a t_{j+H} , e calcula-se o erro. A seguir, a janela de estimação cresce em um período, estendendo-se do período t_1 ao período t_{j+1} , e reitera-se a execução de treinamento. Já a janela de previsão desliza um período adiante para o intervalo de t_{j+2} a t_{j+2+H} e – assim – sucessivamente. As iterações prosseguem até que o último período do horizonte de previsão alcance t_T , em que T é o número total de períodos da série temporal. Por fim, calcula-se a média dos erros

resultantes de cada iteração. Dessa forma, mitiga-se o viés de seleção temporal e obtém-se uma avaliação mais robusta do desempenho preditivo dos modelos.

Além das sobreditas métricas de precisão preditiva, para ordenar o desempenho das metodologias de previsão, considerou-se a posição relativa da previsão de cada metodologia, em relação ao *ranking* de seus desempenhos do menor ao maior erro preditivo, em cada passo de previsão, ao longo de todas as janelas descritas em 3.5.

Para cada modelo, são obtidas 12 previsões de 12 passos cada. Assim, por exemplo, no caso da previsão mensal, por meio de uma função perda, que afere o erro de previsão, a performance de cada metodologia foi ordenada, em cada passo de cada janela.

Ademais, para se mitigar pelo menos parcialmente o efeito de se ter relativamente poucas janelas de previsão, de modo a melhorar a qualidade da inferência, realizou-se a reamostragem por bootstrap estacionário (Politis e Romano, 1994).

Assim, com base nos *rankings* obtidos, o *Model Confidence Set* (MCS), proposto em Hansen *et al.* (2011), foi utilizado para circunscrever o conjunto de modelos cujo desempenho preditivo é estatisticamente indistinguível, a um dado nível de significância. Em seguida, realizou-se o teste de Friedman (1937), para avaliar a homogeneidade da performance de tal conjunto, e utilizou-se o *rank* médio de cada modelo, para ordenar as previsões remanescentes, quando a heterogeneidade foi constatada. A comparação pós-hoc de Nemenyi (1963) foi calculada, porém com baixo poder de discernir as performances preditivas dos modelos.

3.5. Experimento

A construção do experimento buscou assegurar a comparabilidade das metodologias de previsão, ao estabelecer parâmetros uniformes de janela de treinamento, horizonte de previsão, número de passos de deslizamento e número total de deslizamentos. Além disso, logrou-se obter o maior número de deslizamentos, diante das restrições computacionais, de modo que a realização do experimento fosse factível. As metodologias de previsão que utilizam os modelos de *deep learning* LSTM e SCINet provaram-se onerosas, com tempo de execução de até 4 horas, para ensaios com 12 janelas de treinamento e teste. Com efeito, limitou-se o experimento de avaliação de performance preditiva a 12 janelas deslizantes crescentes.

Por conseguinte, a janela de treinamento inicial se estendeu de 01/01/2004 até 01/01/2024 e a primeira previsão 12 meses adiante se deu para o período de 01/02/2024 a 01/01/2025. Em seguida, a janela de treinamento estendeu-se um período adiante, alongando-se até 01/02/2024,

e a previsão 12 meses à frente deslizou adiante, tendo início em 01/03/2024. E, assim, sucessivamente, até que o horizonte de previsão alcançasse 01/12/2025, de modo que o último período de teste coincidissem com os 12 meses de 2025, o último ano calendário para o qual os dados de arrecadação estavam integralmente disponíveis. Dessa forma, o experimento constituiu-se de 12 janelas crescentes de estimação ou treinamento e 12 janelas deslizantes de previsão e teste, com deslizamentos de 1 mês e horizontes de previsão de 12 meses.

Por construção, o modelo econométrico SARIMA e o *foundation model* TimesFM não admitem variáveis exógenas, conforme descrito na seção 3.2 e, para a rede SCINet, elas não foram utilizadas, por conta de restrições computacionais. Já os modelos econométricos SARI-MAX e ECM, bem como a rede neural recorrente LSTM, comportaram a utilização de variáveis exógenas. Assim, nesses casos, foram admitidas como potenciais variáveis exógenas aquelas descritas na seção 3.1.

A variável objetivo, em última instância, é a arrecadação total de ICMS do Estado de São Paulo. Assim, a sua previsão pode ser obtida de forma direta, utilizando-a expressamente como variável objetivo dos modelos, o que constitui a abordagem *Top-Down* ou agregada. Alternativamente, é possível decompor a arrecadação de ICMS nos regimes de recolhimento ICMS_RPA, ICMS_ST e ICMS_Importacao, já descritos na seção 3.1, o que constitui a abordagem *Bottom-Up*. Ao estabelecer um modelo distinto para cada regime, a abordagem permite capturar o padrão de cada componente específico de arrecadação, bem como evidenciar sua relação com conjuntos distintos de variáveis exógenas. Dessa forma, a previsão resultante dos três regimes componentes seria agregada, para se obter a previsão de ICMS. As duas estratégias de previsão foram incorporadas ao experimento.

Em adição, detalha-se adiante a configuração do ajuste dos hiperparâmetros das redes LSTM e SCINet. Em ambas, utilizou-se o algoritmo de *Otimização Bayesiana*, apresentado na seção 3.3. Inicializou-se o processo Gaussiano com 5 combinações de hiperparâmetros e limitou-se a otimização a 25 pontos adicionais. Utilizou-se também a parada antecipada, com paciência de 5 iterações.

No caso da rede LSTM, com base em (Dornelas *et al.*, 2022), restringiu-se o espaço de hiperparâmetros aos intervalos enunciados a seguir. Limitou-se a busca do número de neurônios da camada escondida entre 32 e 700, da taxa de dropout entre 5% e 75% e do tamanho da sequência entre 3 e 24. Além disso, utilizou-se a função de ativação *ReLU* e o otimizador Adam com taxa de aprendizado inicial de 0,001 (seção 3.3).

A estrutura da rede SCINet impõe uma restrição intrínseca a dois hiperparâmetros. O ta-

manho da sequência de entrada ou *look-back window* deve ser divisível por 2^L , em que L é o número de níveis da árvore binária de *SCI-Blocks*. Assim, estabeleceu-se intervalo para L entre 1 e 3. Logo, a *look-back window* deve ser divisível por 8 (2^3). Dessa forma, adotou-se o tamanho de sequência de entrada igual a 24. Para se encontrar os valores dos demais hiperparâmetros que minimizam a função-perda, utilizou-se o algoritmo de *Otimização Bayesiana*. O espaço de busca destes hiperparâmetros foi delimitado conforme os intervalos a seguir, tomando-se como referência (Liu *et al.*, 2022). Dentro do bloco SCI, o tamanho do filtro de convolução e da camada escondida foi limitado entre 3 e 12 e a taxa de dropout entre 20% e 72%. Restringiu-se a taxa de aprendizado inicial entre 0,00005 e 0,002 e o tamanho do *batch* entre 2 e 10. Em teoria, é possível empilhar mais de uma rede SCINet. Porém, a implementação disponibilizada pelos autores (Liu *et al.*, 2022) permite apenas a pilha com 1 rede SCINet. Assim, não foi possível realizar ensaios com tamanho de pilha maior que 1.

O modelo TimesFM não admite ajuste de hiperparâmetros ou treinamento, pois trata-se de modelo base pré-treinado.

Os modelos econométricos SARIMA, SARIMAX e ECM foram implementados com componentes da biblioteca *Statsmodels*. A seleção de variáveis exógenas foi realizada consoante a relevância econômica e a significância estatística, quando válida, garantida a relação de cointegração. A arrecadação do ICMS pressupõe a prática de atividades econômicas enquadradas no fato gerador do tributo, o que se reflete no Produto Interno Bruto (PIB). Além disso, a variação de preços dos bens e serviços tributados também impactam o montante de imposto recolhido. No caso da importação, a taxa de câmbio afeta diretamente o valor de recolhimento. Dessa forma, a inclusão dessas três variáveis macroeconômicas, PIB, IPCA e cotação do dólar, foi aventada durante a modelagem econométrica. Em adição, componentes auto-regressivos (AR) e de média móvel (MA), bem como suas instâncias sazonais (SAR e SMA) foram testadas, consoante a função de autocorrelação (ACF) e a função de autocorrelação parcial dos resíduos (PACF). Como resultado, os modelos devem apresentar resíduos que se comportam como ruído branco. Todavia, o teste último do valor de um modelo preditivo é a precisão de sua previsão para dados fora do conjunto utilizado na estimação de seus parâmetros.

A rede neural recorrente LSTM foi implementada por componentes provenientes da biblioteca *tensorflow.keras* (*Sequential*, *Input*, *LSTM*, *Dropout* e *Dense*). O ajuste dos hiperparâmetros utilizou as classes Bayesian Optimization da biblioteca *bayes_opt*. Além disso, a classe *Early Stopping* da biblioteca *tensorflow.keras* foi utilizada no treinamento do modelo, para incorporar a parada antecipada.

O modelo fundacional TimesFM é implementado pela biblioteca *timesfm*, cujas classes carregam o modelo pré-treinado de um repositório do *Hugging Face*, recebem uma série histórica como entrada e realizam a previsão.

Para a rede SCINet, utilizou-se a implementação dos próprios autores (Liu *et al.*, 2022), disponível em <https://github.com/cure-lab/SCINet> (acesso em: 17/12/2025). Em adição, recorreu-se à biblioteca *torch*, para a preparação dos tensores de entrada e para incorporar o otimizador Adam.

Por fim, ressalta-se que foi estabelecida como axioma do experimento a simulação de uma situação real de previsão. Dessa forma, nas previsões com metodologias que utilizaram variáveis exógenas, foram considerados os seus valores futuros esperados à época, conforme registrados pelas expectativas de mercado do Relatório Focus do Banco Central do Brasil (BCB).

4. RESULTADOS

Conforme descrito na seção 3, 14 metodologias de previsão foram avaliadas, compostas por 7 modelos, SARIMA, SARIMAX em diferenças, SARIMAX em nível, ECM, rede LSTM, rede SCINet e TimesFM, sob as abordagens *Top-Down* e *Bottom-Up*, ao longo de 12 janelas de treinamento e teste. Os resultados apurados nas 12 janelas foram registrados na Tabela 8 e sintetizados na média e desvio padrão observados para EAPA e MAPE, conforme apresentado na Tabela 7, ambas situadas no Apêndice (seção 7).

A seguir, analisam-se as características das séries temporais que compõem o estudo e discorre-se sobre as especificações dos modelos utilizados, bem como sobre a performance preditiva observada para cada um deles.

4.1. Econometria

Considerando-se a revisão literária (seção 2) e as séries temporais que compõem a base de dados utilizada (seção 3.1), foram selecionados 3 modelos econométricos, SARIMA, SARIMAX e ECM.

Como será apresentado adiante, todas as séries de arrecadação de ICMS são integradas de grau 1. Assim, suas primeiras diferenças são estacionárias, o que torna viável a estimação de modelos SARIMA, cujos estimadores dos parâmetros são consistentes e assintoticamente não-viesados e eficientes (Box e Jenkins, 1970). Outrossim, para combinações de séries tempo-

rais que apresentam relação de cointegração, torna-se justificável utilizar o SARIMAX, tanto em primeiras diferenças quanto em nível. Por um lado, em primeiras diferenças, as variáveis tornam-se estacionárias e, por conseguinte, os estimadores dos parâmetros beneficiam-se de consistência e eficiência assintótica, porém a relação de longo prazo entre as variáveis deixa de ser expressamente modelada. Por outro lado, em nível, a não estacionariedade acarreta a superconsistência, ou seja, os estimadores dos coeficientes de cointegração convergem para o valor verdadeiro a uma taxa mais rápida que o normal, não obstante, apresentam viés, ou seja, deixam de ser eficientes e a inferência com as estatísticas *t* e *F* deixa de ser válida (Engle e Granger, 1987). Se houvesse interesse no valor dos parâmetros, o SARIMAX em nível seria descartado. No entanto, como o presente expediente é a avaliação do desempenho preditivo, o SARIMAX em nível irá compor o experimento e será avaliado por sua precisão preditiva fora da amostra de estimação. Por fim, por modelar explicitamente a relação de longo prazo entre as variáveis cointegradas, bem como os desvios de curto prazo, o ECM desponta como uma opção natural, cujos estimadores da relação de cointegração são superconsistentes, assintoticamente não viesados e eficientes (Engle e Granger, 1987), uma vez que seja constatada a exogeneidade fraca.

Consoante Hyndman e Athanasopoulos (2021), um bom método de previsão produzirá resíduos de inovação não correlacionados. Se houver autocorrelação dos resíduos, isso implica que ainda há informação a ser extraída dos dados e utilizada na previsão. Ainda, a média de tais resíduos deve ser zero, caso contrário, produz-se previsão viesada. Ao falhar em satisfazer essas duas condições, um método de previsão indica que pode ser aperfeiçoado. Dessa forma, tais propriedades dos resíduos foram consideradas condições necessárias, no âmbito da seleção dos modelos.

Já caracterizadas na seção 3.1, as séries temporais admitidas como variáveis dependentes são a arrecadação de ICMS do Estado de São Paulo, bem como a arrecadação segmentada pelos regimes componentes RPA (ICMS_RPA), Importação (ICMS_Importacao) e ST (ICMS_ST). Em adição, as séries temporais PIBI, IPCA e Câmbio foram incorporadas como variáveis exógenas.

Em primeiro lugar, avaliou-se a ordem de integração das sobreditas séries temporais, por meio dos testes de hipótese ADF (Said e Dickey, 1984), PP (Phillips e Perron, 1988) e KPSS (Kwiatkowski *et al.*, 1992). Assim, uma série temporal foi considerada estacionária, quando as hipóteses nulas de ADF e PP são rejeitadas e a do KPSS deixa de ser rejeitada, a um nível de significância de 5%.

Concluiu-se que as séries temporais da base de dados caracterizada na seção 3.1 são in-

tegradas de grau 1, $I(1)$, com a exceção de IPCA, para a qual os testes apresentaram resultados contraditórios e, por conseguinte, inconclusivos.

Em modelos com variáveis exógenas, é fundamental assegurar a cointegração entre as variáveis endógenas e exógenas, para não recair em regressão espúria. Assim, avaliaram-se as combinações de variáveis dependentes e independentes utilizadas adiante, por meio do teste de Johansen (1988). O resultado pode ser visto na Tabela 9 (Apêndice).

Em adição, avaliou-se a não linearidade das séries temporais, o que foi feito por intermédio do teste BDS (Brock *et al.*, 1987), aplicado aos resíduos de modelos ARIMA, preliminarmente ajustados às séries temporais, com base em critérios de informação. Em adição, é examinou-se a sua volatilidade ou heterocedasticidade, pois – além de impactar as propriedades estatísticas dos estimadores – também é uma característica que pode ser mitigada pela transformação logarítmica. A avaliação da heterocedasticidade foi realizada em dois níveis, o primeiro constituído pelo Teste ARCH-LM (Engle, 1982) – que detecta heterocedasticidade condicional autorregressiva – e o segundo pelo teste de White (1980) – que captura heterocedasticidade geral –, acionado apenas quando o primeiro nível não rejeita a hipótese nula de homocedasticidade, para todas as defasagens.

Destarte, constatou-se indícios de que as séries temporais apresentam não linearidades, pois rejeitou-se a hipótese nula de que os resíduos resultantes se comportam como ruído branco. Ainda, a transformação logarítmica não afetou o resultado do teste BDS. No que tange à heterocedasticidade, a hipótese nula de homocedasticidade foi rejeitada ao nível de significância de 5% em todos os casos. Entretanto, a transformação logarítmica mostrou-se efetiva, ao mitigar a heterocedasticidade, seja com a elevação do p-valor observado ou mesmo com a não rejeição da hipótese de homocedasticidade.

Dessa forma, a seguir, com base na análise supra das séries temporais, os modelos ARIMA, SARIMAX e ECM serão especificados, para o experimento de previsão da arrecadação de ICMS do Estado de São Paulo.

A aplicação da metodologia de Box e Jenkins (1970) pressupõe que as séries temporais sejam estacionárias. Por meio dos testes descritos acima, constatou-se que todas as séries de arrecadação de ICMS (ICMS, ICMS_RPA, ICMS_Importacao e ICMS_ST) são integradas de grau 1. Assim, tomou-se a primeira diferença de cada uma delas, para se estimar os modelos SARIMA. Além disso, com o fito de atenuar os efeitos da volatilidade das séries temporais, recorreu-se à transformação logarítmica.

Iniciou-se pela abordagem *Top-Down*, em que foi estimado o modelo SARIMA para a

série de arrecadação total de ICMS especificado na Equação 3 (TD_SARIMA).

$$(1 - \Phi_1 L^{12})(1 - L)\ln(ICMS_t) = c + (1 + \theta_1 L)(1 + \Theta_1 L^{12})\varepsilon_t + (1 - L)Dummies + \delta_5 D_{202004}$$

$$Dummies = \delta_1 D_{200602} + \delta_2 D_{200901} + \delta_3 D_{201703} + \delta_4 D_{202202}$$

(Equação 3)

As *dummies* que modelam eventos pontuais apresentam a sintaxe $D_{AAAAAMM}$, em que AAAA representa o ano e MM o mês para os quais ela assume o valor 1, sendo igual a 0 nos demais períodos.

Ao longo das 12 janelas crescentes (Tabela 8) descritas na seção 3.5, o modelo SARIMA obteve EAPA médio igual a 3,13% e MAPE médio igual a 3,63% (Tabela 7).

Em seguida, avaliou-se a performance do modelo SARIMA, sob abordagem *Bottom-Up*, com a segmentação da previsão em três modelos, um para cada regime de arrecadação, conforme especificações contidas nas Equações 4, 5 e 6 (BU_SARIMA).

$$(1 - \Phi_1 L)(1 - \Phi_1 L^{12})(1 - L)\ln(ICMS_{RPA}_t) = c + (1 + \theta_1 L)(1 + \Theta_1 L^{12})\varepsilon_t + (1 - L)D_{RPA}$$

$$D_{RPA} = \delta_1 D_{200602} + \delta_2 D_{200901} + \delta_3 D_{201703} + \delta_4 D_{202202}$$

(Equação 4)

$$(1 - L)\ln(ICMS_{Importacao}_t) = c + (1 + \theta_1 L)(1 + \Theta_1 L^{12})\varepsilon_t + \delta(1 - L)D_{200403}$$

(Equação 5)

$$(1 - \Phi_2 L^2)(1 - \Phi_1 L^{12})(1 - L)\ln(ICMS_{ST}_t) = c + (1 + \theta_1 L)(1 + \Theta_1 L^{12})\varepsilon_t + \delta(1 - L)D_{200901}$$

(Equação 6)

Como resultado, observou-se deterioração da performance preditiva com a elevação do EAPA médio para 4,11% e do MAPE para 4,71%.

Sob abordagem *Top-Down*, em primeiras diferenças, a especificação do modelo SARI-MAX resultante foi a da Equação 7 (TD_SARIMAX_d). Como as séries temporais – endógena e exógenas – apresentam volatilidade, valeu-se de suas transformações logarítmicas, um recurso que será reiteradamente utilizado no decorrer do estudo, porém sem menções adicionais, para efeito de concisão.

$$(1 - \Phi_1 L^{12})(1 - L)\ln(ICMS_t) = c + L(1 - L)[\beta_1 \ln(PIBI_t) + \beta_2 \ln(IPCA_t)] + (1 + \theta_1 L)(1 + \Theta_1 L^{12})\varepsilon_t$$

(Equação 7)

Após a execução das 12 janelas deslizantes crescentes de previsão, observou-se EAPA₂₄

médio igual a 1,95%, com desvio padrão de 1,63% (Tabela 7), o que constitui significativo ganho de precisão da previsão anual em relação ao modelo SARIMA. Já o MAPE da previsão mensal apresentou média de 3,48% e desvio padrão de 0,99%.

Como já ressaltado anteriormente, o ICMS incide sobre um espectro amplo de atividades econômicas, em múltiplos elos da cadeia de valor. Ademais, em princípio, trata-se de um imposto sobre valor agregado (IVA). Dessa forma, a variação da arrecadação do ICMS guarda relação estreita com a variação do Produto Interno Bruto (PIB) do Estado de São Paulo, que – por sua vez – tem como *proxy* a variação do PIB do Brasil. Por conseguinte, utiliza-se o número índice do PIB (PIBI) como variável explicativa dos três regimes, embora sob defasagens temporais distintas, consoante elaborado na seção 3.1. Além disso, o ICMS incidente sobre as importações é impactado direta e indiretamente pela taxa de câmbio. Por um lado, o aumento da taxa de câmbio acarreta um aumento contemporâneo do valor de ICMS arrecadado. Por outro lado, a mesma elevação também pode reduzir o quantum de importação em períodos posteriores. Por isso, o modelo econométrico do regime ICMS_Importacao apresenta a variável câmbio em dois componentes distintos, um contemporâneo e um com defasagem de 4 períodos. O número de dias úteis do mês (DU) também figura como *driver* do modelo pertinente ao regime ICMS_Importacao.

Assim, prosseguindo para a abordagem *Bottom-Up*, os modelos para cada regime de arrecadação foram especificados conforme as Equações 8, 9 e 10 (BU_SARIMAX_d).

$$(1 - \Phi_1 L^{12})(1 - L)\ln(ICMS_{RPA}_t) = c + L(1 - L)\beta_1 \ln(PIBI_t) + (1 + \Theta_1 L)(1 + \Theta_1 L^{12})\varepsilon_t + (1 - L)D_{RPA} + \delta_5 D_{202004} \\ D_{RPA} = \delta_1 D_{200602} + \delta_2 D_{200901} + \delta_3 D_{201703} + \delta_4 D_{202202}$$

(Equação 8)

$$(1 - \Phi_3 L^{12})(1 - L)\ln(ICMS_{Importacao}_t) = c + (1 - L)[\beta_1 \ln(PIBI_t) + \beta_2 \ln(Câmbio_t) + \beta_3 \ln(DU_t) + \beta_4 L^4 \ln(Câmbio_t)] + (1 + \Theta_1 L)\varepsilon_t$$

(Equação 9)

$$(1 - \Phi_1 L)(1 - \Phi_1 L^{12})(1 - L)\ln(ICMS_{ST}_t) = c + (1 - L)[\beta_1 \ln(PIBI_t) + \delta D_{200901}] + (1 + \Theta_1 L)(1 + \Theta_1 L^{12})\varepsilon_t$$

(Equação 10)

Por conseguinte, em primeiras diferenças, a abordagem *Bottom-Up* resultou em piora do desempenho preditivo (Tabela 7), com elevação do EAPA médio para 2,81% e sutil redução do MAPE médio para 3,42%, ao longo das 12 janelas de estimação e previsão.

As séries temporais ICMS, PIBI e IPCA são cointegradas (Tabela 9). Dessa forma, nos termos expostos anteriormente, viabiliza-se a especificação de modelo SARIMAX em nível, consoante a Equação 11 (TD_SARIMAX), apesar da não-estacionariedade das séries tempo-

rais.

$$(1 - \phi_1 L)(1 - \phi_1 L^{12}) \ln(ICMS_t) = c + \beta_1 \ln(PIBI_t) + \beta_2 \ln(IPCA_t) + (1 + \theta_1 L)(1 + \theta_1 L^{12}) \varepsilon_t$$

(Equação 11)

Com efeito, sob abordagem *Top-Down*, o SARIMAX em nível apresentou previsões mais precisas (Tabela 8) que aquelas em diferenças, com EAPA médio igual a 1,74% e MAPE médio igual a 2,6%.

Ademais, modelou-se o SARIMAX em nível, sob abordagem *Bottom-Up*, consoante as Equações 12, 13 e 14 (BU_SARIMAX). Ressalta-se que a variável *dummy* D_{ST} reflete uma alteração de prazos de recolhimento do regime de Substituição Tributária, ocorrido em 2016.

$$(1 - \phi_1 L)(1 - \phi_1 L^{12}) \ln(ICMS_{RPA_t}) = c + \beta_1 L \ln(PIBI_t) + (1 + \theta_1 L) \varepsilon_t + (1 - L) D_{RPA} + \delta_4 D_{201703} + \delta_5 D_{202004}$$

$$D_{RPA} = \delta_1 D_{200602} + \delta_2 D_{200901} + \delta_3 D_{202202}$$

(Equação 12)

$$(1 - \phi_1 L)(1 - \phi_1 L^{12}) \ln(ICMS_{Importacao_t}) = c + \beta_1 \ln(PIBI_t) + [\beta_2 + \beta_4 L^4] \ln(Câmbio_t) + \beta_3 \ln(DU_t) + (1 + \theta_1 L) \varepsilon_t + \delta_1 D_{200805} + \delta_2 D_{202011}$$

(Equação 13)

$$(1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3)(1 - \phi_1 L^{12}) \ln(ICMS_{ST_t}) = c + \beta_1 \ln(PIBI_t) + \delta_1 D_{200901} + \delta_2 D_{ST} + \delta_3 D_{12} + \varepsilon_t$$

(Equação 14)

Consequentemente, o desempenho preditivo dos modelos supra representou aprimoramento em relação à abordagem *Top-Down*, com EAPA médio igual a 1,4% e MAPE médio igual a 2,03%, conforme disposto na Tabela 8 (Apêndice). Salienta-se que, além de mais precisa, a previsão do modelo SARIMAX sob abordagem *Bottom-Up* também foi mais estável, com desvios-padrão do EAPA e do MAPE iguais a 0,6% e 0,46%, respectivamente.

Uma vez constatada a cointegração entre as séries temporais, sabe-se que o estimador do vetor de cointegração é superconsistente (Engle e Granger, 1987), o que é atestado pelo teste de Johansen (1988), conforme Tabela 9 (Apêndice). Ademais, a exogeneidade fraca (Engle *et al.*, 1983) assegura a validade da inferência das estatísticas t e F sobre os parâmetros de longo e curto prazo, bem como a eficiência assintótica de seus estimadores. Considerando-se as séries temporais componentes da relação de cointegração, uma vez estimado um *Vector Error Correction Model* (VECM) completo, caso o coeficiente de uma variável em teste não seja estatisticamente significativo, conclui-se que ela é fracamente exógena. No entanto, mesmo diante de significância estatística, se tal coeficiente for pequeno o suficiente, os efeitos da endogeneidade sobre eficiência dos estimadores tornam-se negligenciáveis (Stock, 1987; Phillips e Hansen,

1990; Phillips, 1991). Em termos econômicos, é razoável esperar que variações de indicadores como PIBI, IPCA e Câmbio impactem a arrecadação de ICMS de forma relevante, porém o contrário não é verdadeiro. As variações da arrecadação do ICMS de SP dificilmente seriam uma força explicativa saliente das variações do PIB, da cotação do dólar ou do índice de preços nacionais, o que é corroborado pelo valor diminuto de seus coeficientes (Tabela 2). Portanto, mesmo afastando a hipótese de exogeneidade fraca para as variáveis econômicas como PIBI, IPCA e Câmbio, na maioria dos modelos, como pode ser visto na Tabela 2, a ínfima magnitude de seus coeficientes (α) sugere impacto desprezível da endogeneidade sobre a eficiência dos estimadores, o que podemos chamar de *quasi-exogeneidade fraca*, que justifica a utilização do ECM.

Tabela 2 – Exogeneidade

Modelo	Variável	Rank r	α	p-valor α
ICMS	PIBI	3	0.0000	0.00%
ICMS	IPCA_ind	3	0.0000	0.01%
ICMS	Câmbio	3	0.0000	0.24%
ICMS_RPA	log(PIBI)	1	0.0133	0.00%
ICMS_RPA	log(IPCA_ind)	1	0.0007	0.29%
ICMS_Importacao	PIBI	3	0.0000	0.08%
ICMS_Importacao	Câmbio	3	0.0000	0.00%
ICMS_Importacao	IPCA_ind	3	0.0000	2.49%
ICMS_ST	PIBI	3	0.0000	3.08%
ICMS_ST	IPCA_ind	3	0.0000	0.01%
ICMS_ST	Câmbio	3	0.0000	11.47%

Fonte: Elaboração própria

Iniciando-se pela abordagem *Top-Down*, modelou-se o termo de correção de erro (ECT) e o ECM com as especificações expressas nas Equações 15 e 16 (TD_ECM), respectivamente.

$$ECT_t = ICMS_t - \beta_0 - \beta_1 PIBI_t - \beta_2 IPCA_t - \beta_3 Câmbio_t - \beta_4 SLP_{200909} \cdot t - \beta_5 SLP_{202008} \cdot t$$

(Equação 15)

$$(1 - \phi_1 L - \phi_2 L^2 - \phi_{12} L^{12})(1 - L)ICMS_t = c + \alpha ECT_{t-1} + L(1 - L) [\theta_1 PIBI_t + \theta_2 SLP_{202008} \cdot t] + \varepsilon_t$$

(Equação 16)

Salienta-se que as variáveis SLP_{200909} e SLP_{202008} são *dummies* de inclinação, que assumem o valor 1 a partir da data que as subscrevem e 0 antes disso, em função do número de períodos transcorridos desde a data subscrita (t). Tais variáveis foram utilizadas para modelar quebras

estruturais que sucederam a Crise do *Subprime* e a Pandemia de Covid-19, respectivamente. Em adição, ressalta-se que o componente autorregressivo $\Phi_2 L^2$ foi mantido a despeito de insignificância estatística, por neutralizar a autocorrelação dos resíduos.

Efetivamente, obteve-se previsões para as 12 janelas deslizantes crescentes, cuja média do EAPA foi igual a 3,05% e a do MAPE igual 3,87% (Tabela 8).

A seguir, sob a abordagem *Bottom-Up*, foram estabelecidos os modelos ECM para os regimes RPA, Importação e ST, consoante as Equações 18, 20 e 22 (BU_ECM), respectivamente. Observa-se que, para o ICMS_ST, houve uma quebra estrutural em torno de junho/2023, decorrente de alterações de regras de recolhimento relativas a operações com combustíveis, que acarretou elevação da arrecadação à título de Substituição Tributária. Assim, para modelar tal quebra estrutural, utilizou-se a variável *dummy* LS_{202306} , que assume valor igual a 1 a partir de junho/2023 e zero antes disso.

$$ECT_t^{RPA} = ICMS_{RPA}_t - \beta_0 - \beta_1 PIBI_t - \beta_2 IPCA_t$$

(Equação 17)

$$(1 - \phi_1 L - \phi_{12} L^{12})(1 - L)ICMS_{RPA}_t = c + \alpha ECT_{t-1}^{RPA} + L(1 - L) \theta_1 PIBI_t + \varepsilon_t$$

(Equação 18)

$$ECT_t^{Importacao} = ICMS_{Importacao}_t - \beta_0 - \beta_1 PIBI_t - \beta_2 IPCA_t - \beta_3 Câmbio_t - \beta_4 SLP_{201104}$$

(Equação 19)

$$(1 - \phi_1 L - \phi_2 L^2 - \phi_{12} L^{12})(1 - L)ICMS_{Importacao}_t = c + \alpha ECT_{t-1}^{Importacao} + (1 - L)[L^{12} \theta_1 PIBI_t + L \theta_1 Câmbio_t] + \varepsilon_t$$

(Equação 20)

$$ECT_t^{ST} = ICMS_{ST}_t - \beta_0 - \beta_1 PIBI_t - \beta_2 IPCA_t - \beta_3 LS_{202306}$$

(Equação 21)

$$(1 - \phi_1 L - \phi_3 L^3 - \phi_{12} L^{12})(1 - L)ICMS_{ST}_t = c + \alpha ECT_{t-1}^{ST} + L(1 - L) \theta_1 PIBI_t + \delta_1 D_{12} + \varepsilon_t$$

(Equação 22)

Ao final das 12 janelas crescentes e deslizantes, restou claro que o desempenho da previsão com ECM sob abordagem *Bottom-Up* (Tabela 8) superou a precisão anual do modelo SARIMAX, com EAPA médio igual a 1,2%, porém sua precisão mensal apresentou degradação em relação ao SARIMAX (2,03%), com MAPE médio igual a 3,27%. Tal resultado é consistente com a intuição econométrica, dado que o ECM modela expressamente a relação de longo prazo entre as séries temporais, por meio do ECT, e o SARIMAX apresenta recursos especializados para capturar padrões de curto prazo como a sazonalidade.

4.2. Deep Learning

Modelos de Deep Learning têm demonstrado bom desempenho em previsões de séries temporais financeiras (Masini *et al.*, 2023), em contexto de quantidade massiva de dados, diários ou até horários. Inclusive em previsões de arrecadação de ICMS, tais modelos têm demonstrado performance promissora (Dornelas *et al.*, 2022). Contudo, a literatura ainda carece de avaliações validadas em múltiplas janelas de previsão, no âmbito de dados mensais, cuja escala tende a ser menor. Os ensaios a seguir têm o propósito de aferir a performance preditiva das redes SCINet e LSTM, diante de base de dados com apenas 264 períodos.

A rede SCINet possui 7 hiperparâmetros, porém apenas 5 são objeto de ajuste para cada janela de treinamento, por meio do algoritmo de *Otimização Bayesiana*, consoante exposto na seção 3.2. Os valores resultantes dos hiperparâmetros são apresentados nas Tabela 14, Tabela 15, Tabela 16 e Tabela 17 (Apêndice).

Sob abordagem *Top-Down*, a rede SCINet apresentou EAPA com média igual a 2,69%, que figura como o segundo melhor desempenho anual entre modelos univariados, e MAPE com média igual a 4,61%, conforme exibido na Tabela 8 (seção 7).

Já sob abordagem *Bottom-Up*, a rede SCINet teve a sua performance preditiva degradada, com EAPA médio igual a 5,48% e MAPE médio igual a 7,28%.

A rede neural recorrente LSTM possui três hiperparâmetros, o que pode ser considerado um número enxuto, quando comparado aos sete da SCINet. Porém, foi dimensionada com um número hiperbólico de neurônios na camada escondida (*units*), cuja busca – por meio da *Otimização Bayesiana* – pode atingir o limite de 700, conforme detalhado na seção 3.5, o que acarretou elevado custo de processamento.

Para a abordagem *Top-Down*, considerou-se uma rede LSTM, em que as entradas são as variáveis apresentadas na Tabela 3.

Tabela 3 - Variáveis exógenas da rede LSTM

Variável	Descrição
PIBI(-1)	Número-índice do PIB, defasado em 1 período.
ICMS(-1)	Arrecadação do ICMS de SP, defasada em 1 período.
MA(12)	Média móvel da arrecadação de ICMS dos últimos 12 meses
DU	Dias úteis do mês atual
Câmbio	Taxa de câmbio contemporânea

Fonte: Elaboração própria

Por conseguinte, obteve-se desempenho preditivo pouco preciso ao longo das 12 janelas (Tabela 8), cujo EAPA médio foi igual a 4,77% e MAPE médio igual a 5,49%, conforme apresentado na Tabela 10 (Apêndice), que contém também os valores dos hiperparâmetros.

Já sob a abordagem *Bottom-Up*, foram consideradas para cada regime de arrecadação as variáveis exógenas elencadas na Tabela 4.

Tabela 4 - Variáveis exógenas de cada regime de arrecadação

Regime	Variáveis exógenas
ICMS_RPA	PIBI(-1), ICMS_RPA(-1)
ICMS_Importação	PIBI, Câmbio, DU, ICMS_Importação(-1)
ICMS_ST	PIBI, ICMS_ST(-1)

Fonte: Elaboração própria

Executadas as 12 janelas de previsão, a performance preditiva da rede LSTM sob abordagem *Bottom-Up* demonstrou ganho de precisão em relação a *Top-Down*, com EAPA médio igual a 1,69% e MAPE médio igual a 3,12% (Tabela 7). Destaca-se também o reduzido desvio padrão apresentado pelas métricas EAPA (0,89%) e MAPE (0,46%), que demonstra estabilidade de performance preditiva. Os valores dos hiperparâmetros constam das Tabelas 11, 12 e 13 do Apêndice (seção 7).

4.3. Foundation Model

Sistemas de inteligência artificial gigantescos, treinados com uma quantidade massiva de dados, a exemplo dos grandes modelos de linguagem (LLMs) para processamento de linguagem natural, são chamados de Foundation Models. Recentemente, o mesmo arcabouço tem sido aplicado para grandes bases de séries temporais, com o fito de gerar previsões sem qualquer outra entrada que não seja a própria série temporal (*zero-shot*).

O modelo TimesFM (Das *et al.*, 2024) é pré-treinado em corpus de séries temporais constituído por 300 bilhões de pontos, construído para realizar previsões, sem necessidade de qualquer preparação de dados de entrada, ajuste de hiperparâmetros ou mesmo treinamento. O modelo apenas recebe uma série temporal como entrada, que lhe provê contexto, e retorna a previsão para um horizonte temporal definido. Ele possui 500 milhões de parâmetros e admite uma entrada de dados ou contexto de até 2048 pontos. Dessa forma, por um lado, o tempo de

processamento é drasticamente reduzido, pois trata-se de modelo que não requer ajuste de hiperparâmetros ou treinamento. Por outro lado, tal modelo reclama maior espaço de memória, cerca de 4 Gb, o que é consistente como o dilema espaço-tempo da computação.

Tendo em conta que o TimesFM exige esforço analítico mínimo e pouco processamento, sem ter qualquer treinamento na série temporal de entrada, o desempenho de sua previsão *Top-Down* foi surpreendente, com EAPA médio igual a 2,66%, que foi o melhor desempenho anual de modelo univariado, e MAPE médio igual a 3,72% (Tabela 8).

Já sob abordagem *Bottom-Up*, houve uma piora significativa do desempenho preditivo do modelo TimesFM, com EAPA médio igual a 4,25% e MAPE médio igual a 4,58% (Tabela 8).

5. DISCUSSÃO

A comparação do desempenho preditivo das 14 metodologias de previsão, cujos resultados foram apresentados na seção 4, se deu de duas formas. A primeira se valeu do *Model Confidence Set* (Hansen *et al.*, 2011), em conjunto com o teste de Friedman (1937), para ordenar a performance das previsões, consoante a posição relativa de seus erros frente aos erros das demais previsões, o que foi chamado de *ranking* por desempenho ordinal. Já a segunda utilizou a média das próprias métricas de precisão, apresentadas na seção 3.4, para ranquear o desempenho preditivo das metodologias, a qual se deu o nome de *ranking* por desempenho cardinal.

5.1. *Ranking por Desempenho Ordinal*

Considerando-se as 12 janelas de previsão, em cada um dos 12 passos previstos, o desempenho mensal das 14 metodologias enumeradas na seção 4 foi ranqueado, a cada passo de previsão. Ademais, considerou-se o somatório dos 12 passos de cada janela de previsão, para se ranquear o desempenho da previsão anual das mesmas metodologias.

No que tange à previsão anual, foi obtido o *ranking* da Tabela 5, que ordenou o desempenho preditivo consoante a média dos *ranks* obtidos nas 12 janelas. Ainda, é apresentado o p-valor do teste EPA (*Equal Predictive Ability*) do *Model Confidence Set* (Hansen *et al.*, 2011). Assim, sob um nível de significância de 25%, modelos cujo p-valor é superior a tal patamar estariam dentro do conjunto remanescente.

Para tal conjunto, realizou-se o teste de Friedman (1937). Com p-valor igual a 0,0001, rejeitou-se a hipótese nula de equivalência de performance preditiva. Em seguida, utilizou-se a

sua posição média, considerando-se as 12 janelas de previsão, para ordenar o desempenho das previsões anuais. A comparação pós-hoc de Nemenyi (1963) não foi informativa, pois permitiu concluir apenas que a performance das metodologias que compõem as 10 primeiras posições do *ranking* anual foi superior à performance das 4 últimas posições.

Tabela 5 - Ranking por desempenho ordinal anual

Experimento	Rank médio	p-valor MCS
BU_ECM	3,583	1
TD_SARIMAX	4,5	0,72
BU_SARIMAX	4,667	0,97
TD_SARIMAX_d	4,917	0,475
BU_LSTM	5,75	0,72
TD_SCINet	6,667	0,38
TD_TimesFM	6,667	0,38
BU_SARIMAX_d	7,25	0,32
TD_ECM	7,667	0,21
TD_SARIMA	8,75	0,04
TD_LSTM	9,75	0,035
BU_SCINet	10,25	0
BU_TimesFM	11	0
BU_SARIMA	13,583	0

Fonte: Elaboração própria

Desta feita, aferiu-se que a previsão do *Error Correction Model*, sob abordagem *Bottom-Up* (BU_ECM), atingiu a melhor performance de previsão anual, ocupando a primeira posição do *ranking*, seguido pelo SARIMAX, sob abordagens *Top-Down* e *Bottom-Up* (TD_SARIMAX e BU_SARIMAX), na segunda e terceira posição, respectivamente.

Em adição, avaliou-se o desempenho da previsão mensal, por meio da posição relativa de cada uma das 14 metodologias, em cada um dos 12 passos, ao longo das 12 janelas de previsão. Assim, obteve-se o *rank* médio da Tabela 6. Na mesma tabela, apresentou-se em quantos horizontes de previsão cada modelo integrou o grupo remanescente do MCS (Hansen *et al.*, 2011).

Por conseguinte, o SARIMAX sob abordagem Bottom Up (BU_SARIMAX) despontou com a melhor performance preditiva mensal, seguido do TD_SARIMAX e da rede LSTM sob abordagem Bottom Up (BU_LSTM), em segundo e terceiro lugar, respectivamente.

Tabela 6 - Ranking por desempenho ordinal mensal

Metodologia	Rank médio	Nº Horizontes no MCS ($\alpha=0,25$)
BU_SARIMAX	4,62	12
TD_SARIMAX	5,35	12
BU_LSTM	5,81	12
BU_ECM	6,55	12
TD_SARIMAX_d	6,67	12
BU_SARIMAX_d	6,8	12
TD_TimesFM	6,85	11
TD_SARIMA	7,03	12
TD_ECM	7,19	12
TD_SCINet	7,9	8
BU_TimesFM	8,76	5
TD_LSTM	9,23	4
BU_SCINet	9,99	3
BU_SARIMA	12,25	1

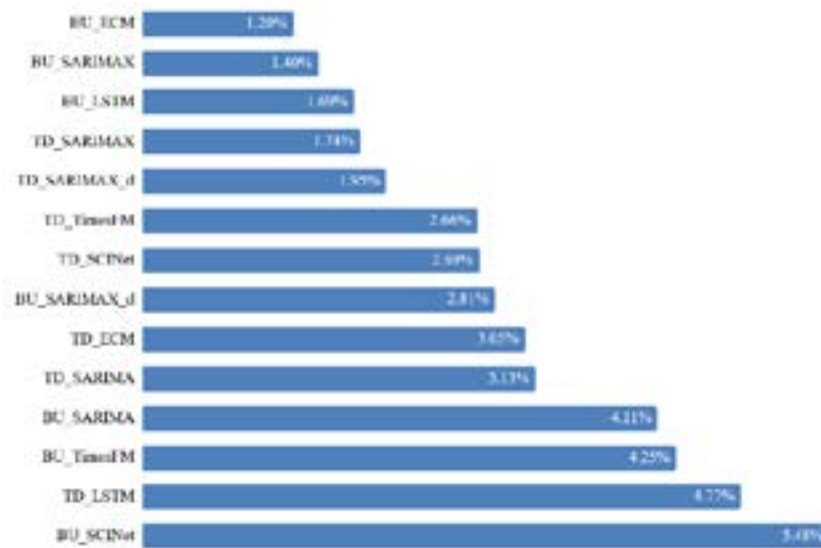
Fonte: Elaboração própria

5.2. Ranking por Desempenho Cardinal

De maneira complementar, utilizou-se as métricas de precisão preditiva anual (EAPA) e mensal (MAPE), apresentadas na seção 3.4, para comparar o desempenho preditivo das 14 metodologias.

Novamente, como é disposto na Figura 3, o BU_ECM figurou na primeira posição do *ranking* de previsão anual, com EAPA médio de 1,2%, seguido do BU_SARIMAX (1,4%) e do BU_LSTM (1,69%). O ECM modela expressamente a relação de longo prazo das séries temporais cointegradas. Assim, o seu desempenho preditivo anual, frente aos demais, confirma a intuição econométrica.

Figura 3 - Ranking anual, com base no EAPA médio.



Fonte: Elaboração própria

No entanto, ao olharmos também para o desvio padrão do EAPA anual, constata-se que o desempenho do BU_SARIMAX foi mais estável, com desvio padrão igual a 0,6%, quando comparado com o BU_LSTM e o BU_ECM, cujos EAPAs apresentaram desvios padrão iguais a 0,89% e 1,1%, respectivamente (Figura 4).

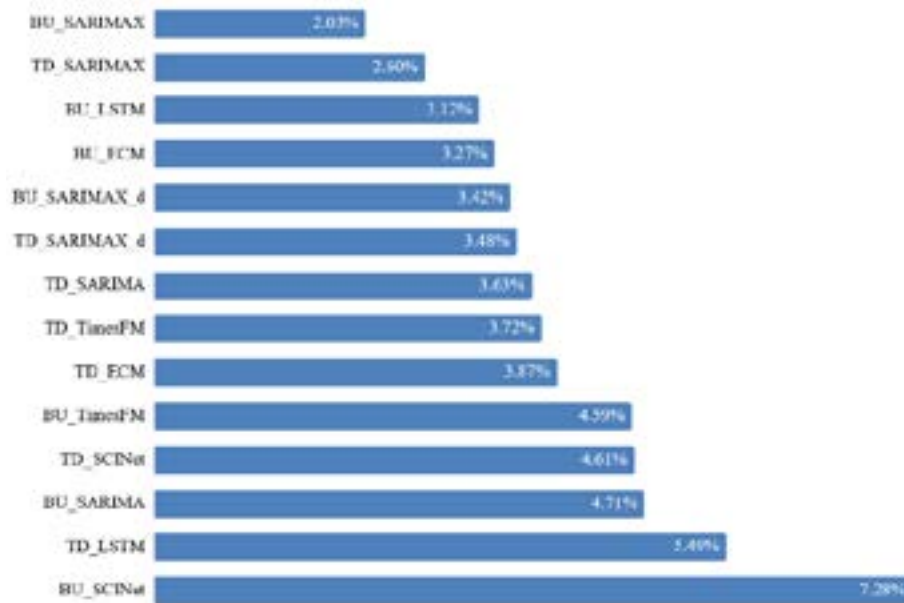
Figura 4 - Desempenho preditivo anual: Dispersão de médias e desvios-padrão do EAPA.



Fonte: Elaboração própria

Corroborando o *ranking* por desempenho ordinal da seção 5.1, a previsão do BU_SARIMAX apresentou a melhor performance mensal, aferida pelo MAPE, com média igual 2,03%, seguida pelas previsões do TD_SARIMAX e do BU_LSTM, cujas médias do MAPE foram iguais a 2,6% e 3,12%, respectivamente (Figura 5).

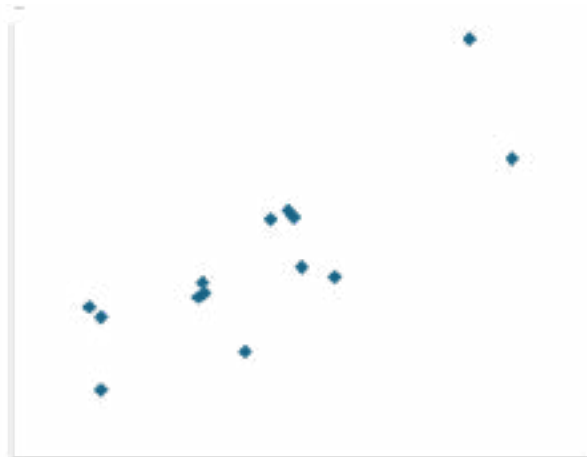
Figura 5 - Ranking mensal, com base no MAPE médio.



Fonte: Elaboração própria

Em termos de estabilidade da previsão mensal, o BU_SARIMAX apresentou MAPE com desvio padrão igual a 0,46%. Assim, ficou bem próximo do BU_LSTM (0,46%) e do BU_ECM (0,40%), conforme Figura 6.

Figura 6 - Desempenho preditivo mensal: Dispersão de médias e desvios padrão do MAPE.



Fonte: Elaboração própria

5.3. Conclusões

Diante dos *rankings* por desempenho ordinal (seção 5.1) e cardinal (seção 5.2), tanto em termos de previsão mensal quanto anual, observou-se que os modelos que incorporam variáveis exógenas (ECM, SARIMAX, SARIMAX_d e LSTM) apresentaram performance preditiva mais

precisa, ocupando os primeiros lugares. Já os modelos univariados (SARIMA, SCINet e Times-FM) ocuparam posições de menor destaque. Salienta-se que as previsões foram realizadas apenas com as informações disponíveis na época do último período de estimação ou treinamento dos modelos, para simular uma situação real de previsão, conforme descrito na seção 3.5. Desse modo, foram utilizados os valores esperados das variáveis exógenas, conforme registrados pelas expectativas coletadas pelo BCB (Relatório FOCUS), o que confere condição de isonomia competitiva entre os modelos univariados e aqueles que admitiram variáveis exógenas.

Ademais, constatou-se – de maneira geral – que a abordagem *Bottom-Up* (BU) proporcionou melhora da performance preditiva dos modelos que incorporaram variáveis exógenas, frente à abordagem *Top-Down* (TD). Já para os modelos univariados, a estratégia BU acarretou degradação da precisão preditiva, quando comparada à TD.

Com efeito, no que diz respeito à previsão anual, os modelos econométricos ECM e SARIMAX se destacaram, figurando nas primeiras posições tanto no *ranking* por performance ordinal quanto cardinal, sobretudo sob abordagem BU. Dentre os modelos de aprendizado de máquina, a rede LSTM sob abordagem BU foi a que apresentou maior precisão preditiva. Apesar de não figurar nas primeiras posições, o TimesFM sob abordagem TD se destacou por apresentar desempenho mediano, mesmo exigindo esforço analítico e custo de processamento negligenciáveis.

Já no que tange à previsão mensal, as metodologias que dominaram as primeiras colocações, consistentemente, tanto no *ranking* de desempenho ordinal quanto cardinal, foram BU_SARIMAX, TD_SARIMAX, BU_LSTM e BU_ECM, nesta ordem.

Por fim, as metodologias que apresentaram desempenho geral mais auspicioso foram constituídas dos modelos econométricos SARIMAX e ECM e da rede neural recorrente LSTM, sobretudo sob abordagem BU. A decomposição da arrecadação de ICMS em ICMS_RPA, ICMS_Importacao e ICMS_ST, cada regime com seu modelo específico, com variáveis exógenas e defasagens próprias, parece ter conferido maior precisão às previsões, tanto em horizonte anual quanto mensal. Portanto, no âmbito das competências fazendárias discutidas na introdução (seção 1), tais metodologias figuram como as mais aptas a instruir tanto o planejamento orçamentário anual quanto a programação de desembolsos ao longo do exercício.

Ademais, diante da reforma tributária promulgada pela EC nº 132/2023, em que as bases tributárias do ICMS e do Imposto sobre Serviços de Qualquer Natureza (ISS) serão consubstanciadas e ampliadas no escopo do Imposto sobre Bens e Serviços (IBS), em transição gradual até 2033, a abordagem *Bottom-Up* tende a se apresentar como alternativa natural para a previsão da futura arrecadação do IBS. Pois, vislumbra-se que, diante da inexistência de série histórica do

IBS, a previsão segmentada do ICMS e do ISS, posteriormente consolidada, pode constituir uma primeira estimativa de arrecadação do novo imposto.

REFERÊNCIAS BIBLIOGRÁFICAS

BONATO, A. S. F. **Utilização de Modelos Não Lineares para Previsão do ICMS Paulista**. 2023. 170 p. Dissertação (Dissertação (Mestrado em Economia)) — Fundação Getúlio Vargas/ Escola de Economia de São Paulo, São Paulo, 2023. Disponível em: <https://repositorio.fgv.br/items/15a860b4-0458-4c68-b5aa-5f9b0789a5e4>. Acesso em: 24/06/2026.

BOX, George E. P.; JENKINS, Gwilym M. **Time Series Analysis: Forecasting and Control**. San Francisco: Holden-Day, 1970.

BROCK, W.; DECHERT, W.; SCHEINKMAN, J. **A test for independence based on the correlation dimension**. SSRN Working Paper, 1987. Disponível em: <https://doi.org/10.1080/07474939608800353>. Acesso em: 24/06/2026.

DAS, A. et al. **A decoder-only foundation model for time-series forecasting**. In: PROCEEDINGS OF THE 41st INTERNATIONAL CONFERENCE ON MACHINE LEARNING. [S. l.]: PMLR, 2024. p. 10148-10167. Disponível em: <https://proceedings.mlr.press/v235/das24c.html>. Acesso em: 24 jun. 2026.

DICKEY, David A.; FULLER, Wayne A. **Distribution of the estimators for autoregressive time series with a unit root**. Journal of the American Statistical Association, [s. l.], v. 74, n. 366, p. 427-431, 1979. DOI: <https://doi.org/10.1080/01621459.1979.10482531>. Acesso em: 24/06/2026.

DORNELAS, A. F.; CAMPOS, L. D.; FIGUEIREDO, K. **Modelos para previsão tributária utilizando redes neurais lstm**. In: Anais do XIX Encontro Nacional de Inteligência Artificial e Computacional. SBC, 2022. p. 95–105. Disponível em: <https://doi.org/10.5753/eniac.2022.227595>. Acesso em: 24/06/2026.

DORNELLES, G. Z.; SCHWARTZER, F. R.; BRAATZ, J. **Redes Neurais Aplicadas na Previsão de Receita de ICMS no Rio Grande do Sul**. Textos para Discussão TE/RS Nº 19. Rio Grande do Sul, Brasil, 2022. Disponível em: <https://tesouro.fazenda.rs.gov.br/upload/arquivos/202508/29094521-1643376127-artigo-modelo-redes-neurais-texto-de-discussao.pdf>. Aces-

so em: 30/06/2026.

ENGLE, Robert F. **Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation**. *Econometrica*, v. 50, n. 4, p. 987–1007, 1982. DOI: 10.2307/1912773. Disponível em: <https://doi.org/10.2307/1912773>. Acesso em: 24/06/2026.

ENGLE, Robert F.; GRANGER, Clive W. J. **Co-integration and Error Correction: Representation, Estimation, and Testing**. *Econometrica*, [s. l.], v. 55, n. 2, p. 251-276, mar. 1987. Disponível em: <https://doi.org/10.2307/1913236>. Acesso em: 28 maio 2026.

ENGLE, R. F.; HENDRY, D. F.; RICHARD, J.-F. **Exogeneity**. *Econometrica*, [S. l.], v. 51, n. 2, p. 277–304, mar. 1983. DOI: 10.2307/1911990. Disponível em: <https://doi.org/10.2307/1911990>. Acesso em: 10 jun. 2026.

FONTENELE, S. V. P. **Arrecadação Tributária: Modelos de Previsão para o Estado do Ceará**. 2017. Dissertação (Dissertação (Mestrado Profissional em Economia do Setor Público – MESP)) — Programa de Pós-Graduação em Economia (CAEN), Universidade Federal do Ceará, Fortaleza, 2017.

FRIEDMAN, M. **The use of ranks to avoid the assumption of normality implicit in the analysis of variance**. *Journal of the American Statistical Association*, v. 32, n. 200, p. 675–701, 1937. DOI: <https://doi.org/10.1080/01621459.1937.10503522>. Acesso em: 24/06/2026.

HAMILTON, James D. **Time Series Analysis**. Princeton: Princeton University Press, 1994.

HAN, Z. et al. **A review of deep learning models for time series prediction**. *IEEE Sensors Journal*, v. 21, n. 6, p. 7833–7848, 2021. Disponível em: <https://doi.org/10.1109/JSEN.2019.2923982>. Acesso em: 24/06/2026.

HANSEN, P. R.; LUNDE, A.; NASON, J. M. **The model confidence set**. *Econometrica*, v. 79, n. 2, p. 453–497, 2011. DOI: <https://doi.org/10.3982/ECTA5771>.

HOCHREITER, S.; SCHMIDHUBER, J. **Long short-term memory**. *Neural Computation*, v. 39

9, n. 8, p. 1735–1780, 1997.

HYNDMAN, R. J.; KOEHLER, A. B. **Another look at measures of forecast accuracy**. *International Journal of Forecasting*, v. 22, n. 4, p. 679–688, 2006. DOI: <https://doi.org/10.1016/j.ijforecast.2006.03.001>. Acesso em: 24/06/2026.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. 3rd. ed. Melbourne, Australia: OTexts, 2021. Disponível em: [OTexts.com/fpp3](https://otexts.com/fpp3). Acesso em: 23 out. 2025.

JOHANSEN, Søren. **Statistical analysis of cointegration vectors**. *Journal of Economic Dynamics and Control*, [s. l.], v. 12, n. 2-3, p. 231-254, 1988. DOI: [https://doi.org/10.1016/0165-1889\(88\)90041-3](https://doi.org/10.1016/0165-1889(88)90041-3). Acesso em: 24/06/2026.

JOHANSEN, Søren. **Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models**. *Econometrica*, [s. l.], v. 59, n. 6, p. 1551-1580, 1991. DOI: <https://doi.org/10.2307/2938278>. Acesso em: 24/06/2026.

JOHANSEN, S. **Likelihood-based inference in cointegrated vector autoregressive models**. Oxford: Oxford University Press, 1995.

JONES, D. R.; PERTTUNEN, C. D.; STUCKMAN, B. E. **Efficient global optimization of expensive black-box functions**. *Journal of Global Optimization*, Kluwer Academic Publishers, v. 13, n. 4, p. 455–492, 1998.

KINGMA, D. P.; BA, J. **Adam: a method for stochastic optimization**. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS (ICLR), 3., 2015, San Diego. Proceedings [...]. San Diego: ICLR, 2015. p. 1–15. Disponível em: <https://arxiv.org/abs/1412.6980>. Acesso em: 30 jun. 2026.

KWIATKOWSKI, Denis; PHILLIPS, Peter C. B.; SCHMIDT, Peter; SHIN, Yongcheol. **Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?** *Journal of Econometrics*, [s. l.], v. 54, n. 1-3, p. 159-178, 1992. DOI: [https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y). Acesso em: 24/06/2026.

LI, W.; LAW, K. L. E. **Deep learning models for time series forecasting: A review**. IEEE Access, Piscataway, v. 12, p. 92306–92327, 2024. Disponível em: <https://doi.org/10.1109/ACCESS.2024.3422528>. Acesso em: 24/06/2026.

LIU, M. et al. **SCINet: Time Series Modeling and Forecasting with Sample Convolution and Interaction**. 2022. ArXiv. Disponível em: <http://arxiv.org/abs/2106.09305v3>. Acesso em: 8 set. 2025.

MASINI, R. P.; MEDEIROS, M. C.; MENDES, E. F. **Machine learning advances for time series forecasting**. Journal of Economic Surveys, v. 37, n. 1, p. 76–111, fev. 2023.

NEMENYI, P. **Distribution-free multiple comparisons**. 1963. Tese (Doutorado) — Princeton University, Princeton, 1963.

PHILLIPS, Peter C. B.; PERRON, Pierre. **Testing for a unit root in time series regression**. Biometrika, [s. l.], v. 75, n. 2, p. 335–346, 1988. DOI: <https://doi.org/10.1093/biomet/75.2.335>. Acesso em: 24/06/2026.

PHILLIPS, P. C. B.; HANSEN, B. E. **Statistical inference in instrumental variables regression with I(1) processes**. Review of Economic Studies, v. 57, n. 1, p. 99–125, 1990.

PHILLIPS, P. C. B. **Optimal inference in cointegrated systems**. Econometrica, v. 59, n. 2, p. 283–306, 1991.

POLITIS, D. N.; ROMANO, J. P. **The stationary bootstrap**. Journal of the American Statistical Association, v. 89, n. 428, p. 1303–1313, 1994. DOI: <https://doi.org/10.1080/01621459.1994.10476870>. Acesso em: 24/06/2026.

SAID, Said E.; DICKEY, David A. **Testing for unit roots in autoregressive-moving average models of unknown order**. Biometrika, [s. l.], v. 71, n. 3, p. 599–607, 1984. DOI: <https://doi.org/10.1093/biomet/71.3.599>. Acesso em: 24/06/2026.

SILVA, P.; FIGUEIREDO, K. **Aprendizado profundo aplicado na previsão de receita tributária utilizando variáveis endógenas**. In: Anais do XVII Encontro Nacional de Inteligência Artificial e Computacional. Porto Alegre, RS, Brasil: SBC, 2020. p. 414–425. ISSN 2763-9061. Disponível em: <https://sol.sbc.org.br/index.php/eniac/article/view/12147>. Acesso em: 24/06/2026.

STOCK, J. H. **Asymptotic properties of least squares estimators of cointegrating vectors**. *Econometrica*, v. 55, n. 5, p. 1035–1056, 1987.

WHITE, H. **A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity**. *Econometrica*, v. 48, n. 4, p. 817–838, 1980. DOI: 10.2307/1912934. Disponível em: <https://doi.org/10.2307/1912934>. Acesso em: 24/06/2026.

APÊNDICE

Tabela 7 - Síntese do desempenho das metodologias de previsão

Metodologia	EAPA		MAPE	
	Média	Desvio Padrão	Média	Desvio Padrão
TD SARIMA	3,13%	1,37%	3,63%	0,98%
BU SARIMA	4,11%	1,87%	4,71%	1,42%
TD_SARIMAX_d	1,95%	1,63%	3,48%	0,99%
BU_SARIMAX_d	2,81%	1,24%	3,42%	0,96%
TD SARIMAX	1,74%	1,77%	2,60%	1,20%
BU SARIMAX	1,40%	0,60%	2,03%	0,46%
TD ECM	3,05%	2,10%	3,87%	1,49%
BU ECM	1,20%	1,10%	3,27%	0,40%
TD SCINet	2,69%	2,18%	4,61%	1,45%
BU SCINet	5,48%	3,23%	7,28%	2,35%
TD LSTM	4,77%	3,35%	5,49%	2,57%
BU LSTM	1,69%	0,89%	3,12%	0,46%
TD TimesFM	2,66%	2,28%	3,72%	1,66%
BU TimesFM	4,25%	1,54%	4,58%	1,33%

Fonte: Elaboração própria

Tabela 8 - Desempenho das metodologias por janela de previsão

Janela	TD_SARIMA		BU_SARIMA		TD_SARIMAX_d	
	EAPA	MAPE	EAPA	MAPE	EAPA	MAPE
01/02/2024	2,36%	3,11%	1,04%	2,71%	5,07%	5,42%
01/03/2024	2,40%	3,21%	2,64%	3,76%	3,51%	4,57%
01/04/2024	5,00%	5,34%	2,66%	4,54%	4,56%	5,06%
01/05/2024	0,16%	1,90%	6,06%	6,21%	2,67%	3,69%
01/06/2024	3,58%	3,69%	7,82%	7,87%	0,58%	2,75%
01/07/2024	2,36%	2,98%	3,56%	3,95%	0,89%	2,73%
01/08/2024	4,52%	4,57%	4,54%	4,72%	0,61%	2,51%
01/09/2024	3,73%	3,99%	4,62%	4,96%	0,42%	2,75%
01/10/2024	3,59%	4,00%	4,76%	5,17%	0,55%	2,79%
01/11/2024	4,85%	4,89%	5,41%	5,47%	1,82%	3,32%
01/12/2024	2,86%	3,20%	2,00%	2,90%	1,34%	3,20%
01/01/2025	2,19%	2,65%	4,18%	4,24%	1,32%	2,98%

	BU_SARIMAX_d		TD_SARIMAX		BU_SARIMAX	
Janela	EAPA	MAPE	EAPA	MAPE	EAPA	MAPE
01/02/2024	2,00%	2,25%	4,77%	4,78%	0,38%	1,59%
01/03/2024	2,54%	2,77%	3,68%	3,89%	0,25%	1,60%
01/04/2024	3,29%	3,46%	4,74%	4,72%	1,51%	2,01%
01/05/2024	3,57%	3,72%	2,62%	2,72%	1,70%	2,02%
01/06/2024	4,38%	4,43%	1,17%	1,85%	2,32%	2,38%
01/07/2024	1,51%	2,14%	1,38%	1,88%	0,89%	1,24%
01/08/2024	1,41%	2,48%	0,18%	1,37%	1,46%	1,69%
01/09/2024	1,80%	3,38%	0,49%	1,78%	1,75%	2,14%
01/10/2024	1,87%	3,60%	0,13%	1,67%	1,78%	2,20%
01/11/2024	5,31%	5,45%	1,05%	1,98%	1,56%	2,13%
01/12/2024	2,31%	3,23%	0,63%	1,95%	1,65%	2,34%
01/01/2025	3,73%	4,18%	0,00%	2,64%	1,60%	3,00%

	TD_ECM		BU_ECM		TD_SCINet	
Janela	EAPA	MAPE	EAPA	MAPE	EAPA	MAPE
01/02/2024	3,60%	4,77%	0,04%	3,70%	2,72%	5,77%
01/03/2024	5,78%	6,21%	1,89%	3,26%	0,13%	3,94%
01/04/2024	6,95%	6,88%	3,35%	3,67%	5,69%	6,34%
01/05/2024	4,59%	4,68%	2,01%	3,06%	1,99%	3,62%
01/06/2024	2,96%	3,51%	1,69%	3,37%	0,93%	4,71%
01/07/2024	4,52%	4,47%	1,14%	2,97%	3,19%	3,84%
01/08/2024	1,88%	2,39%	0,04%	2,60%	1,42%	4,16%
01/09/2024	2,53%	2,70%	0,53%	2,98%	5,25%	5,38%
01/10/2024	1,72%	2,38%	0,05%	2,80%	2,17%	3,02%
01/11/2024	0,17%	2,11%	2,59%	3,62%	6,86%	7,69%
01/12/2024	1,38%	2,77%	0,39%	3,30%	0,26%	4,07%
01/01/2025	0,47%	3,56%	0,73%	3,89%	1,61%	2,74%

Janela	BU_SCINet		TD_LSTM		BU_LSTM	
	EAPA	MAPE	EAPA	MAPE	EAPA	MAPE
01/02/2024	1,71%	3,32%	8,56%	8,39%	1,60%	3,04%
01/03/2024	4,91%	6,18%	9,95%	9,87%	0,90%	2,31%
01/04/2024	5,77%	6,29%	8,87%	8,80%	0,91%	3,23%
01/05/2024	10,32%	10,59%	8,06%	7,99%	0,76%	2,44%
01/06/2024	6,82%	7,03%	4,52%	5,08%	2,90%	3,53%
01/07/2024	4,60%	6,03%	4,96%	5,29%	2,46%	3,40%
01/08/2024	3,86%	6,55%	3,85%	4,36%	1,53%	2,80%
01/09/2024	2,00%	4,81%	2,70%	3,26%	0,07%	2,70%
01/10/2024	1,68%	8,07%	1,33%	3,13%	2,63%	3,41%
01/11/2024	3,90%	7,45%	0,31%	2,81%	2,58%	3,76%
01/12/2024	9,60%	9,49%	0,90%	3,05%	2,17%	3,38%
01/01/2025	10,55%	11,56%	3,21%	3,86%	1,74%	3,46%

Janela	TD_TimesFM		BU_TimesFM	
	EAPA	MAPE	EAPA	MAPE
01/02/2024	4,75%	5,27%	5,05%	5,57%
01/03/2024	5,25%	5,85%	5,93%	6,39%
01/04/2024	7,52%	7,49%	6,79%	6,75%
01/05/2024	3,25%	3,44%	4,53%	4,54%
01/06/2024	1,65%	2,72%	5,71%	5,64%
01/07/2024	3,46%	3,76%	5,29%	5,24%
01/08/2024	0,31%	1,82%	2,87%	3,34%
01/09/2024	0,35%	2,30%	2,38%	3,01%
01/10/2024	0,28%	2,67%	2,05%	2,79%
01/11/2024	0,78%	2,60%	3,55%	3,72%
01/12/2024	2,16%	3,34%	2,86%	3,61%
01/01/2025	2,19%	3,39%	4,02%	4,42%

Fonte: Elaboração própria

Tabela 9 - Teste de Cointegração Johansen

Endógena	Exógenas	Lags (BIC)	Rank (traço)	Rank (λ_{max})
ICMS	PIBI, IPCA	2	1	1
ICMS	PIBI, Câmbio, IPCA	1	3	3
ICMS_RPA	PIBI	1	2	1
ICMS_RPA	PIBI, IPCA	1	1	1
ICMS_Importação	PIBI, Câmbio	2	3	2
ICMS_Importação	PIBI, Câmbio, IPCA	1	3	3
ICMS_ST	PIBI, IPCA, LS202306	2	1	2
ICMS_ST	PIBI, Dst	1	1	1

Fonte: Elaboração própria

Tabela 10 - Previsão da rede LSTM, sob abordagem Top-Down

Janela	units	dropouts	seq_length	EAPA	MAPE
01/02/2024	585	0,05	6	8,56%	8,39%
01/03/2024	556	0,45	4	9,95%	9,87%
01/04/2024	409	0,05	7	8,87%	8,80%
01/05/2024	587	0,24	6	8,06%	7,99%
01/06/2024	594	0,05	7	4,52%	5,08%
01/07/2024	434	0,57	9	4,96%	5,29%
01/08/2024	499	0,05	24	3,85%	4,36%
01/09/2024	587	0,45	6	2,70%	3,26%
01/10/2024	142	0,05	19	1,33%	3,13%
01/11/2024	506	0,05	20	0,31%	2,81%
01/12/2024	504	0,06	23	0,90%	3,05%
01/01/2025	504	0,06	23	3,21%	3,86%

Fonte: Elaboração própria

Tabela 11 - Previsão da rede LSTM, sob abordagem Bottom-Up, regime ICMS_RPA

Janela	units	dropouts	seq_length	MAPE	EAPA
2024-02-01	588	17,54%	6	3,95%	0,62%
2024-03-01	512	65,91%	10	4,43%	0,79%
2024-04-01	431	15,92%	6	4,42%	0,03%
2024-05-01	692	5,00%	17	4,18%	0,23%
2024-06-01	588	19,86%	6	4,54%	0,90%
2024-07-01	431	15,92%	6	4,82%	1,69%
2024-08-01	594	75,00%	23	3,44%	2,27%
2024-09-01	601	5,00%	16	3,46%	0,98%
2024-10-01	504	6,44%	23	4,32%	0,57%
2024-11-01	430	5,00%	9	4,38%	1,00%
2024-12-01	433	5,00%	6	4,50%	0,49%
2025-01-01	432	5,00%	10	4,82%	0,65%

Fonte: Elaboração própria

Tabela 12 - Previsão da rede LSTM, sob abordagem Bottom-Up, regime ICMS_Importacao

Janela	units	dropouts	seq_length	MAPE	EAPA
2024-02-01	607	19,29%	4	7,14%	2,59%
2024-03-01	182	5,00%	3	6,13%	3,46%
2024-04-01	569	25,15%	3	6,48%	2,70%
2024-05-01	619	5,71%	3	7,25%	2,88%
2024-06-01	592	5,00%	3	7,11%	0,68%
2024-07-01	589	63,89%	4	5,92%	0,24%
2024-08-01	504	6,44%	23	5,91%	1,52%
2024-09-01	465	47,68%	3	6,87%	1,57%
2024-10-01	428	5,00%	8	7,06%	2,44%
2024-11-01	431	15,92%	6	7,78%	6,25%
2024-12-01	419	5,00%	3	7,95%	6,41%
2025-01-01	156	27,51%	3	9,11%	6,96%

Fonte: Elaboração própria

Tabela 13 - Previsão da rede LSTM, sob abordagem Bottom-Up, regime ICMS_ST

Janela	units	dropouts	seq_length	MAPE	EAPA
2024-02-01	518	75.00%	3	8.06%	2.75%
2024-03-01	544	75.00%	3	8.12%	2.12%
2024-04-01	552	75.00%	3	8.79%	1.17%
2024-05-01	218	5.00%	3	8.98%	4.91%
2024-06-01	374	5.00%	3	11.16%	8.87%
2024-07-01	610	5.00%	3	9.42%	6.35%
2024-08-01	431	15.92%	6	12.77%	11.86%
2024-09-01	178	55.07%	3	7.51%	0.39%
2024-10-01	365	5.00%	3	9.77%	7.07%
2024-11-01	317	5.00%	3	5.43%	2.74%
2024-12-01	460	5.00%	3	5.27%	2.08%
2025-01-01	466	5.00%	3	4.25%	2.37%

Fonte: Elaboração própria

Tabela 14 - Previsão da rede SCINet, sob abordagem Top-Down

Janela	kernel_size	dropout	batch_size	learning_rate	num_levels	EAPA	MAPE
01/02/2024	10	32,30%	5	0,51%	1	2,72%	5,77%
01/03/2024	6	32,40%	7	0,92%	1	0,13%	3,94%
01/04/2024	8	61,30%	5	0,76%	2	5,69%	6,34%
01/05/2024	10	32,30%	5	0,51%	1	1,99%	3,62%
01/06/2024	11	39,50%	4	0,62%	2	0,93%	4,71%
01/07/2024	9	63,20%	5	0,75%	2	3,19%	3,84%
01/08/2024	10	72,00%	7	1,00%	3	1,42%	4,16%
01/09/2024	10	32,30%	5	0,51%	1	5,25%	5,38%
01/10/2024	12	20,00%	6	0,67%	1	2,17%	3,02%
01/11/2024	8	20,00%	5	0,77%	1	6,86%	7,69%
01/12/2024	9	72,00%	5	1,00%	3	0,26%	4,07%
01/01/2025	8	62,70%	5	0,91%	2	1,61%	2,74%

Fonte: Elaboração própria

Tabela 15 - Previsão da rede SCINet, sob abordagem Bottom-Up, regime ICMS_RPA

Janela	kernel_size	dropout	batch_size	learning_rate	num_levels	MAPE	EAPA
2024-02-01	7	54,22%	3	0,53%	2	16,31%	16,25%
2024-03-01	6	40,00%	5	0,96%	2	10,26%	9,78%
2024-04-01	5	56,04%	4	0,85%	2	6,06%	4,85%
2024-05-01	4	60,00%	5	1,00%	2	5,78%	3,94%
2024-06-01	7	60,00%	5	1,00%	3	7,63%	5,03%
2024-07-01	4	60,00%	5	1,00%	3	4,10%	1,44%
2024-08-01	6	41,27%	4	0,54%	2	3,99%	1,68%
2024-09-01	5	43,47%	4	1,00%	2	5,32%	0,12%
2024-10-01	6	44,72%	4	0,60%	2	5,79%	2,23%
2024-11-01	6	40,00%	3	0,19%	2	7,07%	5,73%
2024-12-01	7	40,00%	4	1,00%	2	7,73%	6,61%
2025-01-01	5	54,64%	4	0,32%	2	4,92%	1,77%

Fonte: Elaboração própria

Tabela 16 - Previsão da rede SCINet, sob abordagem Bottom-Up, regime ICMS_Importacao

Janela	kernel_size	dropout	batch_size	learning_rate	num_levels	MAPE	EAPA
2024-02-01	6	44,58%	4	0,56%	2	9,33%	6,87%
2024-03-01	7	40,80%	3	0,80%	2	31,45%	32,47%
2024-04-01	7	40,00%	3	1,00%	3	30,31%	30,89%
2024-05-01	6	41,27%	4	0,54%	2	23,55%	24,15%
2024-06-01	7	40,80%	3	0,80%	2	13,73%	14,12%
2024-07-01	7	40,00%	3	1,00%	3	11,11%	10,27%
2024-08-01	7	40,80%	3	0,80%	2	14,44%	14,71%
2024-09-01	5	42,45%	3	0,32%	2	13,31%	13,82%
2024-10-01	7	40,00%	3	1,00%	3	8,04%	3,38%
2024-11-01	6	41,27%	4	0,54%	2	9,17%	6,29%
2024-12-01	6	56,91%	5	0,66%	3	7,93%	3,78%
2025-01-01	7	40,00%	3	1,00%	3	8,70%	5,99%

Fonte: Elaboração própria

Tabela 17 - Previsão da rede SCINet, sob abordagem Bottom-Up, regime ICMS_ST

Janela	kernel_size	dropout	batch_size	learning_rate	num_levels	MAPE	EAPA
2024-02-01	6	44.72%	4	0.60%	2	19.64%	20.27%
2024-03-01	6	44.72%	4	0.60%	2	14.10%	11.08%
2024-04-01	6	56.91%	5	0.66%	3	23.74%	5.81%
2024-05-01	6	60.00%	5	1.00%	3	27.82%	27.10%
2024-06-01	6	58.52%	5	0.62%	2	25.67%	24.59%
2024-07-01	7	40.00%	3	1.00%	3	10.64%	6.10%
2024-08-01	6	59.11%	4	0.67%	3	23.66%	5.79%
2024-09-01	7	43.56%	5	0.94%	2	13.17%	4.25%
2024-10-01	7	40.00%	3	1.00%	2	27.55%	0.94%
2024-11-01	6	49.74%	3	0.90%	3	33.11%	31.79%
2024-12-01	6	44.72%	4	0.60%	2	29.21%	26.67%
2025-01-01	6	59.11%	4	0.67%	3	40.95%	42.04%

Fonte: Elaboração própria