

TEXTO PARA DISCUSSÃO

001/2023

Previsão e
determinantes do
equilíbrio fiscal
para o Estado
de Goiás

Governo do Estado de Goiás

Ronaldo Ramos Caiado

Secretaria-Geral de Governo

Adriano da Rocha Lima

Diretoria-Executiva

Erik Alencar de Figueiredo

Assessoria-Executiva

Alex Felipe Rodrigues Lima

Superintendência de Inteligência

Evando Natal Fernandes de Oliveira

Superintendência de Estudos e Projeções Macroeconômica

Rennan Kertlly de Medeiros

Gerência de Estudos Ambientais e Agronegócio

Érica Basílio Tavares Ramos

Superintendência de Estudos Sociais e Ambientais

Evelyn de Castro Cruvinel

Gerência de Indicadores Conjunturais e Estruturais

Rafael dos Reis Costa

Equipe técnica

Regis Augusto

Fernanda dos Anjos Felix

Alex Felipe Rodrigues Lima

Capa e Diagramação

Ricceli Alencar Cardoso

Revisão

Kimberly Magalhães Moreira

Ana Luíza de Souza Pereira Carvalho

Todos os direitos deste trabalho reservados ao Instituto Mauro Borges de Estatísticas e Estudos Socioeconômicos (IMB).

Praça Dr. Pedro Ludovico Teixeira (Praça Cívica), Setor Central (Antiga Chefatura de Polícia), Goiânia – GO.

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto Mauro Borges de Estatísticas e Estudos Socioeconômicos.

E-mail: imb@goias.gov.br

As publicações do IMB estão disponíveis para download gratuito nos formatos PDF.

Acesse: <https://www.imb.go.gov.br>

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte.

Reproduções para fins comerciais são proibidas.

ELY, R. A.; FELIZ, F. A.; LIMA, A. F. R.

Previsão e determinantes do equilíbrio fiscal para o Estado de Goiás. Goiânia-GO: Instituto Mauro Borges de Estatísticas e Estudos Socioeconômicos – IMB, 2023.

Índices para catálogo sistemático:

1. Equilíbrio Fiscal.
2. Goiás.

RESUMO

Este projeto de pesquisa teve como objetivo utilizar técnicas de análise de séries de tempo e aprendizagem estatística para construir uma metodologia de projeção e análise do equilíbrio fiscal do Estado de Goiás. Para isso, foram identificados os principais previsores das contas públicas relativas as receitas e despesas do Estado, estimados os diversos modelos de previsão e comparados os seus rendimentos em termos de acurácia da previsão no curto e no longo prazo, e por fim, realizadas as projeções das receitas tributárias, despesas correntes e da capacidade de pagamento do Estado. Com os resultados desta pesquisa, foram identificados os melhores modelos de previsão para as contas públicas no curto e no longo prazo, bem como as principais variáveis previsoras, possibilitando uma melhoria da projeção das contas públicas e com isso assessorando a tomada de decisão da Secretaria de Estado da Economia.

INTRODUÇÃO

O aumento dos gastos discricionários dos governos com benefícios previdenciários, gasto com pessoal e incentivos fiscais tem contribuído para a deterioração fiscal dos Municípios, Estados e da União. Quando estes instrumentos de despesas são utilizados sem se pautar no princípio do equilíbrio fiscal intergeracional, isto compromete cada vez mais o espaço no orçamento para a manutenção de gastos públicos essenciais para o bem-estar da população e da qualidade da infraestrutura necessária para o crescimento econômico de longo prazo.

A situação atual aponta para a necessidade de uma reforma fiscal ampla e robusta contemplando todos os entes da Federação. O foco deverá ser na estabilização da dívida pública, redução do déficit orçamentário, elevação dos níveis de investimento público, previdenciária, pessoal e por fim, a reforma tributária.

Em relação aos Estados, o quadro fiscal é tão grave quanto o da União. Em Goiás, o comprometimento do gasto com pessoal como proporção da receita corrente líquida (RCL) apresentou uma razão na ordem de 50,7% em agosto de 2018.¹ Ainda, levando em consideração o perfil demográfico do Estado de Goiás, a razão de dependência vem apresentando um crescimento gradual, porém, consistente.² Deste modo, o gasto com aposentadorias vem consumindo a maior parte do gasto com inativos. Nos Estados da região Centro-Oeste, a média é de 70,7% sendo a maior média entre as regiões brasileiras.

Além dos gastos com pessoal, os incentivos fiscais também são utilizados como importante instrumento de política fiscal e industrial. O objetivo central é minimizar as desigualdades regionais, por meio de atração de empresas, bem como o fomento de cadeias produtivas. A lógica desse instrumento é compensar a falta de competitividade, infraestrutura local e deficiências na qualidade do capital humano por meio de subsídios tributários.

¹ O limite legal de gasto com pessoal no poder Estadual é de 60% da RCL, de acordo com a Lei de Responsabilidade Fiscal (LRF).

² Conforme as projeções Populacionais do IBGE – 2010-2060, a razão de dependência que reflete a divisão entre a população com 65 anos ou mais pela população de 15 até 64 anos (PEA), em 2019, foi de 11.70, passando para 12.09, 12.50, 12.93 nos anos subsequentes.

Entretanto, os incentivos fiscais recebem críticas por dois principais motivos. O primeiro está relacionado ao baixo nível de transparência e regras claras para definição dos beneficiários. Isso impõe um custo elevado para o Estado, por meio da perda de competitividade, com geração de atividades improdutivas, redução de emprego e renda, e com isso redução significativa da arrecadação tributária. O segundo postula que tal instrumento de política fiscal abre espaço para que grupos de interesse pressionem as autoridades econômicas e políticas locais com o intuito de obter privilégios, restringindo a concorrência e o tratamento mais equânime da cunha tributária. Assim, quando isso ocorre, observa-se uma transferência de renda dos mais pobres e menos organizados para estes grupos de interesse, fenômeno caracterizado como *rent-seeking*³ na literatura econômica.

Diante desse quadro fiscal, este projeto de pesquisa pretende utilizar técnicas de análise de séries de tempo e aprendizagem estatística para realizar previsões das variáveis mais relevantes das receitas tributárias e das despesas e encargos, e então agregar as previsões utilizando o algoritmo de compatibilização de Hyndman et al. (2011) de forma a obter as estimativas da Capacidade de Pagamento (CAPAG) para o Estado de Goiás, encontrando também as principais variáveis determinantes do resultado fiscal para o Estado de Goiás. Os modelos de previsões elaborados neste projeto de pesquisa podem servir de apoio para a construção de orçamentos, permitindo a estimativa dos resultados futuros e auxiliando gestores na tomada de decisão e no planejamento da arrecadação tributária. Com os resultados desta pesquisa, espera-se construir melhores modelos de previsão para o resultado fiscal, incluindo a arrecadação de ICMS e as despesas com pessoal, bem como identificar as principais variáveis determinantes do mesmo no curto, médio e longo prazo.

Este relatório está estruturado da seguinte forma. Além desta introdução, na seção dois são apresentados os objetivos; na seção três é elaborada uma justificativa da importância do tema; na seção quatro é feita uma breve revisão de literatura sobre previsão de equilíbrio fiscal; na seção cinco são descritos os métodos científicos a serem utilizados; na seção seis são detalhadas as variáveis e as bases de dados utilizadas; na seção sete são listadas as ações realizadas e os resultados obtidos até o momento na evolução do projeto; e, por fim, na seção oito são contempladas as próximas etapas da realização do trabalho.

³ Podemos definir rent seeking como a busca de renda econômica, por meio da manipulação do ambiente econômico, social ou político no qual as atividades econômicas são desenvolvidas. Para mais detalhes, ver Krueger (1974) e Lisboa e Latif (2013).

OBJETIVOS

Este projeto de pesquisa tem como objetivo geral construir modelos de previsão para o resultado fiscal do Estado de Goiás utilizando técnicas de séries de tempo e aprendizagem estatística.

Entre os objetivos específicos destacam-se a análise, identificação e classificação das principais variáveis preditivas da arrecadação de tributos e gastos com pessoal. Adicionalmente, propõe-se estimar como os resultados das previsões variam para diferentes horizontes de tempo, incluindo curto, médio e longo prazo. Por fim, pretende-se elaborar modelos de previsão que possam ser utilizados em outros cenários para subsidiar a tomada de decisão e o planejamento tributário.

As principais questões a serem respondidas pelo estudo a ser realizado são:

- (i) Quais são as principais variáveis que determinam a previsão das receitas do Estado de Goiás, em especial a arrecadação do ICMS?
- (ii) Quais são as principais variáveis que determinam a previsão das despesas do Estado de Goiás?
- (iii) Quais são os melhores modelos de previsão a serem utilizados para projeção do resultado fiscal do Estado de Goiás?
- (iv) Há diferenças significativas entre as variáveis preditivas para o curto, médio e longo prazo?

JUSTIFICATIVA

Este projeto justifica-se tendo em vista o grave quadro fiscal em que os Municípios, Estados e a União se encontram, em especial o caso do Estado de Goiás. O aumento dos gastos discricionários compromete a manutenção dos gastos públicos essenciais e afeta negativamente a qualidade da infraestrutura necessária para garantir o desenvolvimento de longo prazo. Nesse sentido, há a necessidade de gerar dados confiáveis e robustos para o assessoramento no processo de tomada de decisão dos gestores nas decisões orçamentárias.

Como ressalta Mendonça e Medrano (2016), um bom modelo de previsão de arrecadação de tributos é de importância crítica para o governo. A elaboração de modelos de previsão para o resultado fiscal, bem como identificação das principais variáveis preditivas devem servir de apoio para a construção de orçamentos, viabilizando a redução do nível de incerteza. No curto prazo, a previsão gerada por estes modelos auxilia na gestão do fluxo de receita e atua como indicador estratégico de mudanças conjunturais. No longo prazo, estas estimativas fornecem subsídios para a adoção de políticas econômicas que viabilizem mudanças estruturais.

No Brasil, a necessidade de previsão da arrecadação tributária federal, estadual e municipal é ainda mais relevante, uma vez que a Lei de Responsabilidade Fiscal⁴, em seu artigo 12, estabelece a obrigatoriedade da previsão de receitas tributárias, através de métodos e premissas adequadas.

Assim, o presente estudo, ao responder as questões e objetivos delineados na seção anterior, procura contribuir para um ajuste mais adequado das contas públicas, assessorando a tomada de decisão da Secretaria de Estado da Economia de Goiás.

⁴ Lei Complementar 101 de 4/5/2000.

REVISÃO DE LITERATURA

A utilização de modelos econométricos para previsão de arrecadação e resultado fiscal se alinha a estudos como os de Asimakopoulos e Paredes (2013), Favero e Marcellino (2005), Imrohorgly et al. (2011) e Botric e Vizek (2012). Mais especificamente para o caso do Brasil, Mendonça e Medrano (2016) combinaram modelos de previsão para a receita tributária da União, enquanto que trabalhos como os de Pessoa et al. (2013), Clemente e Clemente (2012), Scheffer et al. (2014), Chain (2015), Santos e Lima (2006), Mendonça et al. (2020) e Gadelha et al. (2020) focaram na previsão de receitas tributárias e de ICMS dos Estados.

A utilização de previsões desagregadas baseadas nos componentes das contas públicas e a posterior agregação através de métodos de reconciliamento hierárquico como o descrito em Hyndman et al. (2011) é uma novidade em relação aos trabalhos realizados para o Brasil. Embora artigos como Asimakopoulos e Paredes (2013) tenham utilizado um princípio semelhante para previsão de dados fiscais, os autores investigam apenas dados dos Estados Unidos e utilizam uma metodologia de agregação conhecida como “bottom-up”, que é um caso especial do procedimento hierárquico de Hyndman et al. (2011), sendo este último capaz de produzir menores erros de previsão.

Lütkepohl (2010) indica que uma previsão que utiliza componentes desagregados antes de realizar a agregação dos componentes pode levar a uma redução considerável dos erros de previsão quadráticos médios devido as informações mais ricas contidas nestes componentes.

O uso desta abordagem é mais comum no contexto da previsão do PIB. Em particular, Marcellino et al. (2003) mostram que é melhor prever o PIB da zona do euro por meio da agregação de previsões de países individuais (abordagem desagregada), enquanto Perevalov e Maier (2010) indicam que a previsão da atividade econômica dos Estados Unidos por meio dos componentes de despesas pode gerar previsões melhores.

Para o caso brasileiro, existem uma série de artigos que aplicaram metodologias de séries de tempo univariadas a fim de prever arrecadação tributária, como é o caso de Melo (2001) e Siqueira (2002). Estes autores utilizaram métodos como modelos ARIMA e suavização de Holt-Winters. Outros métodos alternativos para previsão de arrecadação tributária possibilitam a inclusão de variáveis explicativas, como é o caso dos modelos vetoriais autorregressivos (VAR), dos modelos vetoriais com correção de erros (VEC), e de outros modelos estruturais e de equações simultâneas. Alternativamente, Mendonça, Sachsida e Medrano (2013) utilizaram um modelo linear dinâmico bayesiano com parâmetros variáveis visando modelar séries individuais que correspondiam a cerca 80% da carga tributária bruta brasileira.

Uma constante nos trabalhos que realizam projeções de arrecadação tributária é a utilização de combinações de previsão. Mendonça e Medrano (2016) combinaram previsões obtidas a partir de três modelos, o modelo fatorial dinâmico (MFD), o modelo sazonal autorregressivo integrado com média móvel (SARIMA) e o modelo de suavização de Holt-Winters, adotando para isso critérios de combinação baseados na performance das previsões, em regressões simples e em médias e medianas. Já Dias e Ichikava (2011) também utilizaram combinações de previsões para analisar a relação entre risco de crédito e spread cobrado em títulos da dívida privada.

Entretanto, os trabalhos de previsão de arrecadação tributária e equilíbrio fiscal aplicados para o Brasil não costumam utilizar metodologias de agregação de previsões individuais. Nesse sentido, a utilização do modelo hierárquico de Hyndman et al. (2011) para a estimação do equilíbrio fiscal do Estado de Goiás é uma contribuição perante a literatura, podendo ter o potencial de melhorar as estimativas, conforme sugerido por autores como Lütkepohl (2010), Asimakopoulos e Paredes (2013), Marcellino et al. (2003) e Perevalov e Maier (2010).

METODOLOGIA

A principal medida de equilíbrio fiscal de um Estado é a sua Capacidade de Pagamento (CAPAG), que é determinada através de três indicadores econômico-financeiros: i) Endividamento; ii) Poupança Corrente; e iii) Liquidez. A classificação da saúde financeira do Estado e o cálculo de cada um destes indicadores é detalhado na Nota Técnica nº 31788/2020 do Ministério da Economia. Pode-se descrever estes indicadores através das seguintes fórmulas:

$$\text{Endividamento} = \frac{\text{Dívida Consolidada Bruta}}{\text{Receita Corrente Líquida}} \quad (1)$$

$$\text{Poupança Corrente} = \frac{\text{Despesas Correntes}}{\text{Receitas Correntes Ajustadas}} \quad (2)$$

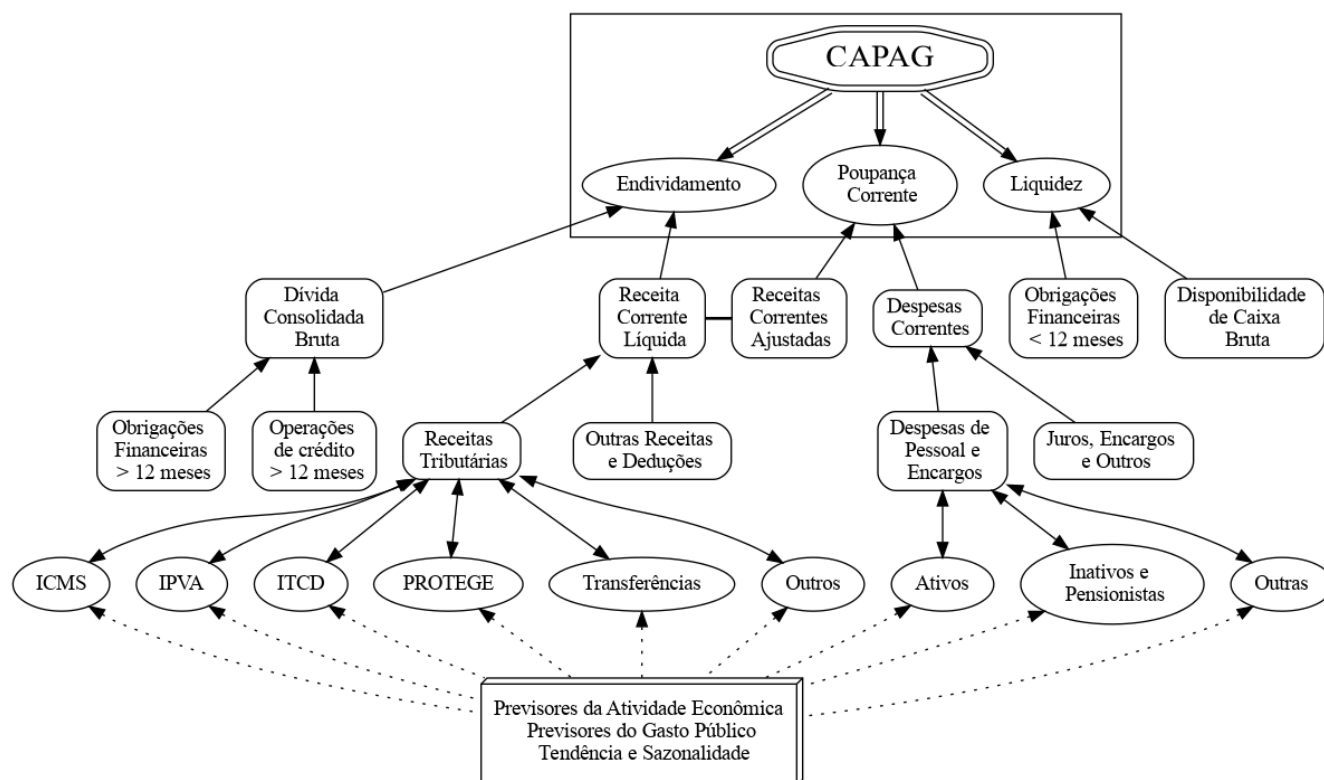
$$\text{Liquidez} = \frac{\text{Obrigações Financeiras}}{\text{Disponibilidade de Caixa Bruta}} \quad (3)$$

Entretanto, os componentes destes indicadores são calculados agregando-se outras contas relativas as receitas e despesas do Estado, de modo que as séries temporais que compõem o equilíbrio fiscal estão organizadas em uma estrutura hierárquica. Para se realizar a previsão da capacidade de pagamento de um Estado, o ideal seria começar com estimativas para as contas mais desagregadas, uma vez que a capacidade preditiva tende a ser maior nos níveis mais desagregados. Entretanto, a soma das previsões dos componentes desagregados não necessariamente corresponde ao todo, podendo inclusive amplificar a existência de viés nas projeções desagregadas. Isto leva a necessidade de reconciliar as previsões em toda a hierarquia (ou seja, garantir que as previsões somam de forma adequada entre os níveis).

Wickramasuriya et al. (2019) propuseram um novo método estatístico para a previsão de séries temporais hierárquicas que fornece previsões pontuais que são reconciliadas entre os níveis da hierarquia, incorporando as informações de uma matriz de covariância completa de erros de previsão na obtenção de um conjunto de previsões coerentes que minimiza o erro quadrático médio. Além disto, este método também reconcilia as estimativas dos intervalos de previsão entre os níveis da hierarquia e é flexível para que outros ajustes ad hoc possam ser incorporados.

Dessa forma, são realizadas previsões iniciais para as variáveis mais relevantes das receitas tributárias e das despesas de pessoal e encargos através do uso de diversos modelos univariados e multivariados de séries de tempo e de aprendizagem estatística, e então são selecionados os melhores modelos com base no erro quadrático médio e agregadas as previsões utilizando o algoritmo de compatibilização de Wickramasuriya et al. (2019) de forma a obter as estimativas de todos os componentes necessários para o cálculo da Capacidade de Pagamento do Estado (CAPAG). A Figura 1 apresenta uma esquematização da agregação das contas do Estado nos vários níveis.

Figura 1 - Estrutura hierárquica para previsão do equilíbrio fiscal



Fonte: elaboração do autor.

Inicialmente são coletados dados relativos aos previsores da atividade econômica e do gasto público do Estado, e então são construídos modelos de previsão tanto para as receitas tributárias e as despesas de pessoal agregadas, como para cada um de seus componentes (ex: ICMS, IPVA, Despesas de Ativos, Inativos, etc.). Para isso são utilizando os previsores como variáveis explicativas nos modelos, bem como adicionados componentes autorregressivos, de tendência e de sazonalidade. Uma vez estimados estes modelos, são calculados os erros de previsão e então selecionados os modelos que minimizam o erro quadrático médio, e por fim, as previsões são compatibilizadas nos níveis hierárquicos através do algoritmo de minimização de traço (MinT) de Wickramasuriya et al. (2019). Isto será feito sucessivamente para os outros níveis hierárquicos até chegarmos na estimativa da CAPAG.

Para a seleção dos modelos a serem estimados em cada uma das etapas é elaborada uma estratégia de validação que servirá como referência para testar a qualidade das previsões. É importante lembrar que a construção da etapa de validação é essencial para o sucesso dos modelos de previsão. Para isso, inicialmente separamos a nossa amostra em uma parte utilizada para estimar e parametrizar o modelo, conhecida como conjunto treino, e uma parte para testar a qualidade das previsões, conhecida como conjunto teste. Dentro do conjunto treino é ainda realizada a etapa de validação cruzada, que

normalmente inclui a construção de diversas janelas de estimação e teste com o objetivo de encontrar os melhores parâmetros e modelos a serem utilizados. Existem diversos meios de realizar esta etapa de validação cruzada, sendo que a maneira mais adequada depende da característica dos dados disponíveis. Como os dados utilizados serão séries de tempo, a estratégia mais recomendada é a construção de janelas móveis, definindo assim um número mínimo de observações iniciais e um horizonte de previsão utilizado para descolar as janelas e obter diversos conjuntos de validação (HYNDMAN e ATHANASOPOULOS, 2018). Será utilizado o horizonte de um período a frente nestes conjuntos de validação, com a janela inicial compreendendo os valores de janeiro de 2010 a dezembro de 2018. Já a separação entre conjunto treino e conjunto teste se dará em fevereiro de 2020, de modo que o conjunto teste será composto por um período de grandes choques nas variáveis de interesse, para podermos avaliar o comportamento das previsões em um cenário de estresse.

Uma vez geradas as previsões fora da amostra de um período a frente dos modelos nos conjuntos de validação, os resultados são comparados aos dados reais para obtermos uma estimativa de acurácia de curto prazo (um mês a frente) dos modelos. Da mesma forma, todos os modelos também são estimados no conjunto treino completo, de modo a gerar previsões fora da amostra para o conjunto teste, que contém cerca de 24 meses (2 anos). A comparação dos resultados com o conjunto teste nos dá uma medida da acurácia de longo prazo. Entretanto, para a escolha do modelo a ser utilizado, deve-se considerar o clássico trade-off entre viés e variância. Ou seja, há dois problemas que podem surgir:

I. Pode-se utilizar um modelo complexo que é capaz de reduzir consideravelmente o erro de previsão no conjunto treino, mas ao mesmo tempo não é tão generalizável a ponto de apresentar um bom resultado no conjunto teste. Se este for o caso, o modelo apresentará viés baixo e variância alta. Assim, o modelo terá um erro de previsão baixo no conjunto treino, mas quando avaliamos o modelo em um novo conjunto de dados (conjunto teste), o erro sobe consideravelmente. Este é o caso típico de overfitting, ou seja, um modelo com variância alta;

II. Pode-se, por outro lado, optar por um modelo simples que é bem generalizável, mas que não reduz consideravelmente o erro de previsão no conjunto treino. Um modelo deste tipo apresenta viés alto mas variância baixa. Assim, permite-se um erro de previsão maior no conjunto treino de modo a obter um erro de previsão semelhante (com pouca variância) no conjunto teste. Esse é o típico caso de underfitting, ou sejam um modelo com viés alto.

Do ponto de vista matemático, o erro de previsão de um modelo é dado por:

$$E \left[\left(y - \hat{f}(x) \right)^2 \right] = E \left[\hat{f}(x) - f(x) \right]^2 + E \left[\hat{f}(x)^2 \right] - E \left[\hat{f}(x) \right]^2 + \sigma^2, \quad (4)$$

onde $f(x)$ é a função de previsão a ser estimada, y a variável aleatória original a ser prevista, E é o operador de expectativas, e σ^2 é a variância do componente aleatório e independente da variável y . Considerando que a expressão $E[f(x)-f(x)]$ corresponde ao viés do modelo de previsão, e a expressão $E[f(x)^2]-E[f(x)]^2$ corresponde a variância do mesmo, a ideia é que ao selecionar um modelo com menor erro de previsão, deve-se tentar diminuir o máximo possível o viés e a variância. Entretanto, pode-se ver nas expressões que há um *trade-off* claro entre os dois termos.

Assim, para obter o modelo de previsão mais adequado, devemos escolher uma função $f(x)$, que apresenta uma boa relação entre viés e variância, representando um ponto intermediário entre um modelo complexo e bem generalizável. Para isso, serão estimados diversos modelos de séries de tempo e de aprendizagem estatística, incluindo diferentes conjuntos de previsores, de modo a obter os menores erros de previsão.

Para avaliar a qualidade das previsões, é necessária a utilização de critérios para o ranqueamento dos modelos elaborados. A escolha de uma medida para avaliar a precisão das previsões em séries temporais, apesar de controversa, é de grande importância prática. Devido às suas propriedades estatísticas, as medidas populares nem sempre garantem resultados facilmente interpretáveis quando aplicadas na prática (HYNDMAN e KOEHLER, 2006). Entre as medidas que podem ser utilizadas neste projeto estão o Erro Quadrático Médio (MSE), o Erro Absoluto Médio (MAE), o Erro Absoluto Médio Ponderado (WMAE), o Erro Médio Percentual Absoluto (MAPE), o Erro Absoluto Relativo Médio (MRAE), entre outros.

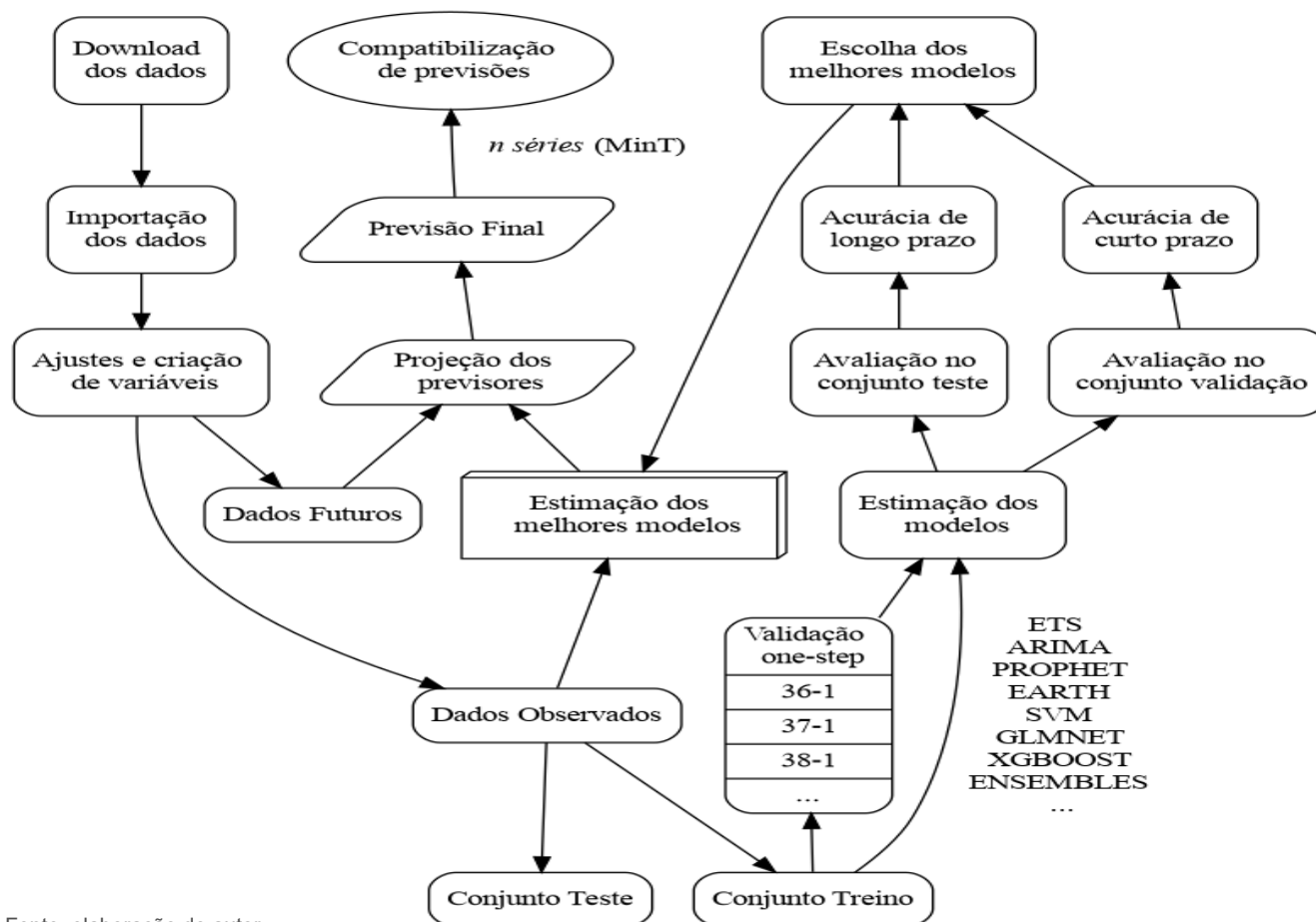
Houve diferentes tentativas na literatura para identificar as propriedades mais importantes pelas quais uma medida de erro deve ser julgada. Em particular, Fildes (1992) justifica as propriedades de interpretabilidade e sensibilidade a valores extremos. Alguns autores, como Zellner (1986) argumentam que o critério pelo qual avaliamos as previsões deve corresponder ao critério pelo qual otimizamos as previsões. Em outras palavras, se otimizarmos estimativas usando alguma função de perda fornecida, devemos usar a mesma função de perda para avaliação empírica, a fim de descobrir qual modelo é melhor.

Tendo isto em vista, a estratégia utilizada neste projeto de pesquisa será utilizar diferentes medidas e assim checar a congruência entre elas, dando maior peso para medidas mais utilizadas na literatura, em especial o Erro Quadrático Médio.

Uma vez estimados e selecionados os melhores modelos para cada um dos componentes das receitas e despesas do Estado, o modelo de compatibilização hierárquica proposto por Wickramasuriya *et al.* (2019) é construído, de modo a gerar previsões consistente ao longo de toda a hierarquia descrita na Figura 1.

Na Figura 2 pode-se observar uma representação esquemática de todas as etapas realizadas para a estimação dos modelos.

Figura 2 - Etapas para a estimação dos modelos de previsão



Fonte: elaboração do autor.

Inicialmente o download e a coleta de dados são realizados de forma automatizada, sendo os mesmos importados e então realizados ajustes e criação de variáveis auxiliares. Então os dados observados são separados em dois conjuntos (treino e teste), e a etapa de validação é conduzida no conjunto treino. Após a estimação de diversos modelos, são calculados erros de previsão de curto prazo, obtidos através dos conjuntos de validação, e erros de previsão de longo prazo, obtidos através da comparação com o conjunto teste. Por fim, selecionam-se os melhores modelos e então realiza-se a projeção futura de cada um dos componentes das receitas e despesas do Estado. Após a estimação de todos os componentes, é feita a etapa de conciliação hierárquica.

É importante ressaltar que a estimação final dos modelos para horizontes de longo prazo envolve a utilização de previsores cujos dados ainda não estão disponíveis. Em geral, há duas possibilidades para lidar com esse problema, a construção de cenários ou a estimação de modelos de previsão. Adotamos a segunda estratégia, sendo que os previsores serão estimados através de um modelo linear simples com tendência e sazonalidade, de forma a replicar a evolução dos dados observada no passado.

Para buscar os modelos que minimizam os erros de previsão propõe-se a utilização de diversas metodologias baseadas em séries de tempo e aprendizagem estatística. Os modelos estimados se dividem em cinco categorias:

I) Modelos Benchmark: são os modelos mais simples a serem estimados, servindo apenas para comparação, de modo que os outros modelos de previsão devem possuir erro quadrático médio menor que estes modelos para podermos considerar que há uma melhoria adequada nas previsões ao utilizar estas metodologias mais sofisticadas. São utilizados dois modelos como benchmark, o modelo de passeio aleatório (NAIVE) e o modelo de passeio aleatório sazonal (SNAIVE);

II) Modelos de Séries de Tempo Univariados: são os modelos de séries de tempo que utilizam apenas os valores passados da variável dependente para modelar os componentes de autorregressividade, sazonalidade e tendência. Neste grupo são estimados quatro modelos distintos, um modelo de suavização exponencial (ETS), um modelo autorregressivo com médias móveis e componentes sazonais (ARIMA), um modelo ARIMA com termos de Fourier para modelar componentes sazonais, e um modelo aditivo com tendências não-lineares e componentes sazonais e autorregressivos (PROPHET);

III) Modelos de Séries de Tempo Multivariados: são os modelos de séries de tempo com a adição de previsores (descritos na seção seguinte). Em geral, esses modelos utilizam os previsores para explicar o resíduo após a estimação dos componentes autorregressivos, sazonais e de tendência. Neste grupo são estimados dois modelos distintos, um modelo ARIMA com erros explicados por um modelo de gradient boosting (XGBOOST), e um modelo PROPHET com erros explicados por um modelo de gradient boosting (XGBOOST);

IV) Modelos multivariados de aprendizagem estatística: são modelos que envolvem o uso de previsores e algoritmos de aprendizagem estatística. Quatro modelos distintos serão estimados nesse grupo, um modelo de regressão linear com penalização (ELASTIC NET), um modelo de gradient boosting (XGBOOST), um modelo de árvores de decisão aleatórias (RANDOM FOREST) e um modelo de regressão com vetores de suporte (KERNLAB).

V) Ensembles: por fim, é construído um modelo que combina as previsões geradas pelos melhores modelos estimados através de uma regressão linear com penalização do tipo LASSO, uma vez que a combinação de previsões de modelos tende a gerar melhores previsões, especialmente quando os modelos são estimados por métodos distintos e possuem baixa correlação.

Entre os modelos benchmark, o modelo de passeio aleatório apenas considera que a previsão da série tempo é o valor imediatamente anterior, enquanto que o modelo sazonal de passeio aleatório considera que a previsão é o valor do mesmo mês do ano anterior. Já o modelo ETS é um modelo de suavização exponencial com um espaço de estado subjacente que consiste em um

componente de nível, um componente de tendência, um componente sazonal e um termo de erro (HYNDMAN et al., 2008).

Entre os outros modelos de previsão a serem estimados, o modelo ARIMA inclui componentes dinâmicos autorregressivos e de médias móveis, podendo também incluir termos sazonais e variáveis explicativas. Em geral, um modelo ARIMA(p,d,q) pode ser representado como:

$$y_t = c + \beta_0 x_t + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + \rho_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \dots + \theta_q \varepsilon_{t-q} \quad (5)$$

onde c é um termo constante; β_0 é um vetor de coeficientes das variáveis explicativas incluídas no modelo; ρ_i são os coeficientes da parcela autorregressiva do modelo que inclui as defasagens da variável dependente y_t ; θ_i são os coeficientes da parcela de média móvel do modelo que inclui as variáveis estocásticas independentemente distribuídas que descrevem o comportamento dos resíduos (ε_t); e, por fim, d se refere ao número de diferenças da variável dependente realizadas antes da estimação da regressão. Este modelo também pode incluir componentes sazonais caso esta propriedade esteja presente nas séries de tempo.

Já o modelo PROPHET é um procedimento para prever dados de séries de tempo com base em um modelo aditivo em que tendências não lineares são ajustadas com sazonalidade anual, semanal e diária, além de efeitos de feriados (TAYLOR e LETHAM, 2018). Este modelo é robusto para dados ausentes, valores extremos e mudanças na tendência.

Outros modelos de séries de tempo que também foram considerados incluem um modelo dinâmico linear aditivo, que é um tipo de modelo de regressão linear em que os parâmetros são tratados como variáveis no tempo, em vez de estáticos (PETRIS et al., 2009); e um modelo bayesiano estrutural, que consiste em três componentes: i) um filtro de Kalman, utilizado para decompor a série de tempo em componentes de tendência, sazonalidade e outros previsores; ii) um método de *spike-and-slab*, em que os previsores mais relevantes são selecionados; e iii) uma média de modelos bayesianos com base em diferentes combinações de previsores (SCOTT e VARIAN, 2014).

Entre os modelos multivariados de aprendizagem estatística, o modelo Elastic Net ajusta modelos lineares generalizados e semelhantes por meio de máxima verossimilhança penalizada através de estratégias de regularização como LASSO e RIDGE (ZOU e HASTIE, 2005). Este modelo possibilita a inclusão de um grande número de previsores, selecionando aqueles com maior poder preditivo através do processo de penalização e regularização. Já os modelos de árvores (florestas) de decisão aleatória são métodos de aprendizado que operam construindo um grande número de árvores de decisão através de subamostras dos dados originais e do cálculo de média das árvores individuais (HASTIE et

al., 2009). Por fim, os modelos de *gradient boosting* utilizam *ensembles* de modelos de previsão fracos permitindo a otimização de uma função de perda. Os modelos utilizados nos *ensembles* normalmente são árvores de decisão, de modo que este tipo de modelo costuma superar os modelos de floresta aleatória (CHEN e GUESTRIN, 2016).

É importante ressaltar que para a estimação da maioria desses modelos é necessária a definição de uma série de parâmetros. A parametrização ótima será escolhida avaliando a qualidade preditiva dos modelos no conjunto de validação, usando para isso a estratégia de construção de janelas móveis mencionada acima. A janela móvel utilizada nos resultados deste relatório é de um período a frente. Já para a seleção dos parâmetros dos modelos ARIMA, será utilizado o algoritmo de seleção automática construído por Hyndman e Khandakar (2008).

Assim, a construção dos modelos de previsão do equilíbrio fiscal para o Estado de Goiás será dividida em quatro etapas. Inicialmente será elaborada uma estratégia de validação que servirá como referência para testar a qualidade das previsões. Em uma segunda etapa, serão estimados diversos modelos de previsão para os componentes da receita tributária e das despesas de pessoal de modo a comparar os erros quadráticos médios no conjunto de validação e selecionar os melhores modelos e parâmetros. Na terceira etapa, serão realizadas as previsões futuras destes componentes, utilizando para isso os modelos selecionados na segunda etapa e as projeções do conjunto de previsores da atividade econômica e do gasto público obtidas através de modelos lineares com tendência e sazonalidade. Por fim, na quarta etapa, será realizada a compatibilização das previsões nos níveis hierárquicos superiores, conforme Wickramasuriya *et al.* (2019).

DADOS

Considerando a estrutura hierárquica dos modelos de previsão apresentadas na Figura 1, para estimação das previsões da Capacidade de Pagamento do Estado de Goiás, inicialmente foram obtidas as rubricas contábeis das contas públicas do Estado e então selecionados e obtidos os previsores da atividade econômica e do gasto público.

De acordo com a disponibilidade dos dados, foram utilizadas observações com frequência mensal a partir de janeiro de 2010. Inicialmente foram obtidas as contas relativas as receitas tributárias do Estado e as despesas de pessoal e encargos.

Para construir os modelos de previsão relativos às receitas tributárias, foram coletados os dados de arrecadação mensal de Imposto sobre Circulação de Mercadorias e Serviços (ICMS), Imposto sobre a propriedade de veículos automotores (IPVA), Imposto sobre a Transmissão Causa Mortis e Doação (ITCD), do Fundo PROTEGE, das Transferências Governamentais e outras receitas e tributos, enquanto para as despesas de pessoal e encargos foram obtidos dados das despesas com ativos, inativos e outras despesas decorrentes de contratos de terceirização. Já para as estimações da Capacidade de Pagamento do Estado foram utilizados os dados computados pela Secretaria de Estado da Economia de Goiás, relativos aos três indicadores que compõe a Capacidade de Pagamento, conforme Resolução 501/2017 do Ministério da Fazenda. Para cada uma destas séries de tempo são estimados os modelos de previsão elencados na seção anterior, e então realizada a compatibilização hierárquica conforme a Figura 1.

Para os modelos que envolvem a utilização de previsores, foram levantados os seguintes dados da atividade econômica: i) variação mensal do IPCA para Goiânia, ii) número de emplacamentos de veículos por tipo, iii) índice de atividade econômica regional (IBCR-GO), iv) índice de volume de vendas no varejo (IVV-GO), v) índice de produção industrial (PI-GO), vi) empregos formais gerados no Estado de Goiás, e vii) faturamento relativo as notas fiscais eletrônicas emitidas no Estado de Goiás por setor macro. Ao todo, estes previsores correspondem a 28 séries de tempo, sendo que nos modelos multivariados, além destes 28 previsores, também são utilizados mais 19 previsores relativos as características temporais das séries analisadas (*dummies* sazonais para o mês, o trimestre, feriados, etc.), totalizando assim 47 previsores.

As séries de tempo contemplam frequência mensal a partir de janeiro de 2010 e até março de 2022, totalizando 147 observações. As únicas exceções são os dados de emplacamentos de veículos por tipo e os dados de faturamentos das notas fiscais eletrônicas por setor, que estão disponíveis apenas a partir de janeiro de 2015 e janeiro de 2018, respectivamente. Para utilização nos modelos de previsão,

os eventuais valores faltantes das séries de dados serão interpolados com base em modelos lineares com *dummies* sazonais e tendências.

Os dados coletados utilizaram como fonte a Secretaria de Estado da Economia de Goiás, o portal transparência do Governo de Goiás, o Sistema de Séries Temporais do Banco Central do Brasil e o Cadastro Geral de Empregados e Desempregados e o Departamento de Trânsito de Goiás. Para deflacionamento dos dados será utilizado o Índice de Preços ao Consumidor Amplo (IPCA). Na estimação dos modelos de previsão, alterações em algumas das variáveis serão realizadas para contemplar problemas relacionados a dados faltantes, sazonalidade, não estacionariedade e estabilização da variância, porém as alterações serão sempre revertidas para o cálculo final dos valores previstos. A Tabela 1 resume as variáveis utilizadas como previsores nos modelos, incluindo a frequência e a fonte dos dados.

Tabela 1 - Previsores utilizados nos modelos e fonte dos dados

Variável	Frequência	Fonte dos dados
IPCA Goiânia	Janeiro/2010 a Março/2022	Sistema de Séries Temporais do Banco Central do Brasil
Número de emplacamentos de veículos por tipo	Janeiro/2015 a Março/2022	Departamento de Trânsito do Estado de Goiás
Índice de atividade econômica regional (IBCR-GO)	Janeiro/2010 a Fevereiro/2022	Sistema de Séries Temporais do Banco Central do Brasil
Índice de volume de venda no varejo	Janeiro/2010 a Fevereiro/2022	Sistema de Séries Temporais do Banco Central do Brasil
Índice de produção industrial	Janeiro/2010 a Março/2022	Sistema de Séries Temporais do Banco Central do Brasil
Número de empregos formais gerados	Janeiro/2010 a Janeiro/2022	Cadastro Geral de Empregados e Desempregados
Faturamento das notas fiscais eletrônicas por setor macro	Janeiro/2018 a Março/2022	Secretaria de Economia do Estado de Goiás

Nota: Todas as variáveis são relativas ao Estado de Goiás.

Com estes dados, inicialmente serão estimados os modelos de previsão da primeira estrutura hierárquica descrita na parte inferior da Figura 1, sendo posteriormente feita a compatibilização das previsões para obtermos os grupos contábeis mais agregados. Para esta compatibilização também serão utilizados os dados relativos ao histórico destas contas agregadas.

Na primeira etapa, serão realizadas as previsões para as receitas tributárias. Pode-se observar na Tabela 2 a evolução das principais contas da receita tributária do Estado de Goiás durante o primeiro trimestre dos anos de 2020, 2021 e 2022.

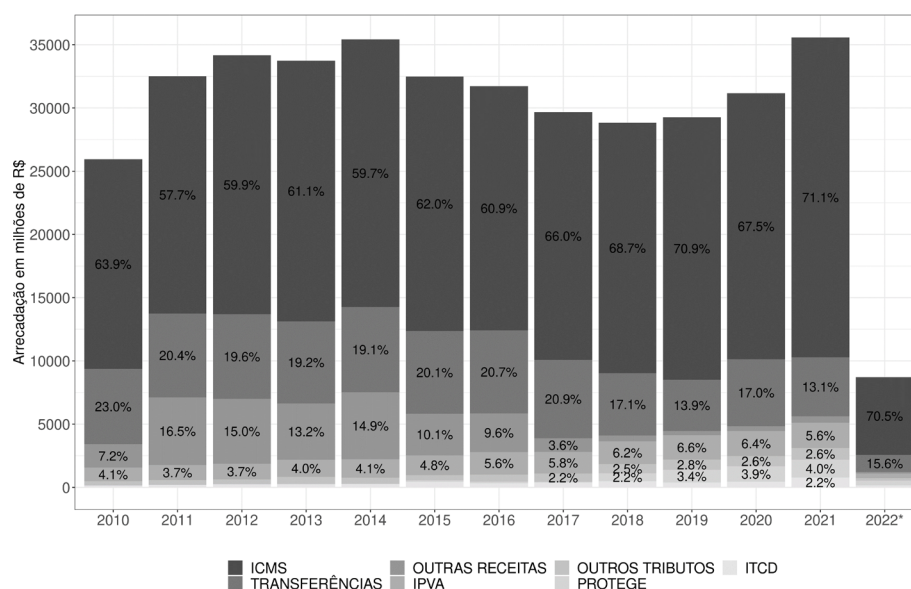
Tabela 2 - Variação trimestral das receitas tributárias no Estado de Goiás

Tipo	Jan-Mar 2020	Jan-Mar 2021	Jan-Mar 2022	Var. (%) 2021/2020	Var. (%) 2022/2021
ICMS	5.110,49	5.964,27	6.144,04	16,71	3,01
IPVA	331,97	345,63	368,22	4,12	6,53
ITCD	84,23	232,97	163,40	176,60	-29,86
PROTEGE	100,15	95,99	109,54	-4,15	14,11
Transferências	200,93	222,55	217,98	10,76	-2,05
Outras Receitas	256,76	298,64	352,76	16,31	18,12
Outros Tributos	1.062,32	1.360,60	1.363,12	28,08	0,19
Total	7.146,84	8.520,66	8.719,06	19,22	2,33

Nota: Os valores estão em R\$ milhões e foram deflacionados pelo IPCA com data base em março de 2022.

Fonte: Secretaria de Estado da Economia de Goiás.

Até o mês de março do ano de 2022, as receitas tributárias totais de 2022, deflacionadas pelo IPCA, cresceram 2,33% em relação ao mesmo período do ano anterior, totalizando R\$ 8,72 bilhões, sendo que a arrecadação de ICMS teve um percentual mais alto de crescimento de 3,01%, totalizando R\$ 6,14 bilhões. Estes valores são significativamente maiores do que o mesmo período no ano de 2020. O primeiro trimestre de 2021 mostrou um aumento na arrecadação total de 19,22% em relação ao primeiro trimestre de 2020. Isto demonstra a forte recuperação da arrecadação tributária dos Estados nos anos seguintes à pandemia de COVID-19. Na Figura 2 podemos ver a composição dessas receitas tributárias, em termos reais, ao longo do período da amostra. A arrecadação do ICMS é o principal componente, representando atualmente mais de 70% do total da receita tributária.

Figura 3 - Composição das receitas tributárias do Estado de Goiás

Nota: Os valores do ano de 2022 vão até o mês de março e todos os valores estão deflacionados pelo IPCA na data base de março de 2021.

Fonte: Secretaria de Estado da Economia de Goiás.

Utilizando os dados de arrecadação de ICMS por grupo macro, disponíveis até o mês de fevereiro de 2022, pode-se observar na Tabela 3 as variações setoriais na arrecadação do período de janeiro a fevereiro de 2022 em relação ao mesmo período nos anos de 2021 e 2020, também deflacionadas pelo IPCA. Praticamente todos os setores tiveram um aumento de arrecadação de ICMS maior que 10% nos dois primeiros meses do ano de 2021, com exceção dos setores de comunicação, energia elétrica e combustível. Já no ano de 2022, apenas os setores de combustível e produção agropecuária mostraram crescimento acima de 10% na arrecadação.

Tabela 3 - Arrecadação de ICMS no Estado de Goiás por grupo macro

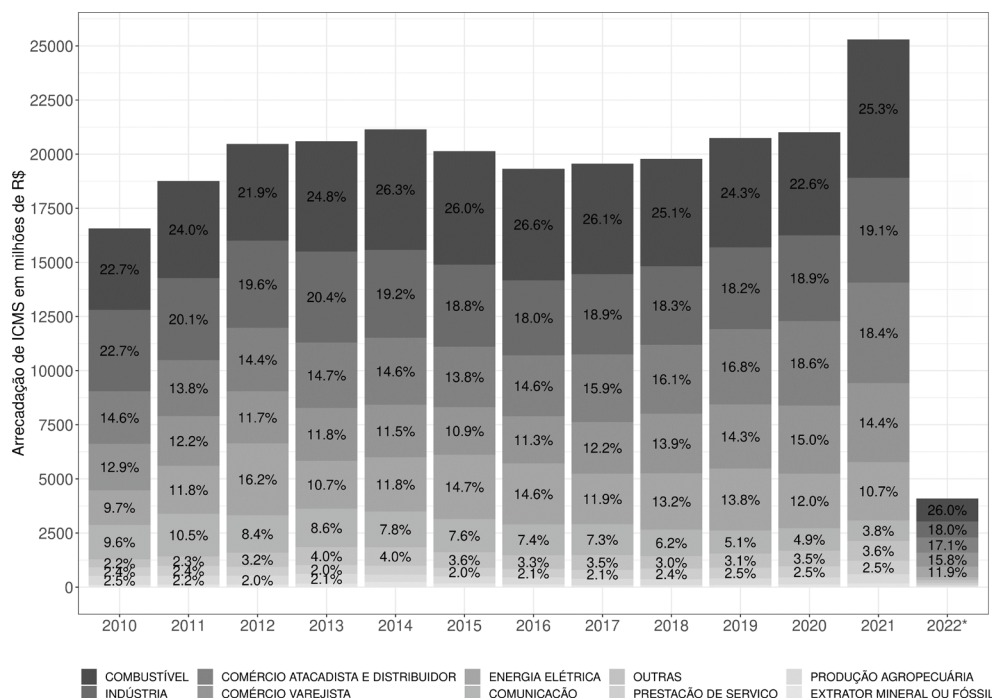
Grupo Macro	Jan-Fev 2020	Jan-Fev 2021	Jan-Fev 2022	Var. (%) 2021/2020	Var. (%) 2022/2021
Combustível	829,89	850,32	1.064,05	2,46	25,14
Comércio Atacadista E Distribuidor	602,64	749,78	700,64	24,42	-6,55
Comércio Varejista	578,35	662,73	646,98	14,59	-2,38
Comunicação	176,82	171,04	141,08	-3,27	-17,52
Energia Elétrica	430,12	443,46	485,93	3,10	9,58
Extrator Mineral Ou Fóssil	16,24	23,14	24,63	42,47	6,41
Indústria	629,96	811,85	739,61	28,87	-8,90
Outras	102,57	143,12	118,85	39,54	-16,96
Prestação de Serviços	77,88	103,20	96,69	32,52	-6,30
Produção Agropecuária	55,93	66,53	80,59	18,94	21,14

Nota: Os valores estão em R\$ milhões e foram deflacionados pelo IPCA com data base em março de 2022.
Fonte: Secretaria de Estado da Economia de Goiás.

Na Figura 3 pode-se ver a participação destes setores na arrecadação de ICMS ao longo do período da amostra, bem como a evolução da arrecadação de ICMS em termos reais. A maior parte da arrecadação de ICMS advém dos setores de combustíveis, industrial, de comércio varejista e atacadista e de energia elétrica.

Assim, os valores reportados nas tabelas e figuras anteriores nos indicam que houve um aumento significativo das receitas tributárias do Estado de Goiás no ano de 2021 em termos reais, especialmente relacionado a um aumento de arrecadação no ICMS. A tendência de crescimento diminuiu em 2022, porém a arrecadação tributária total ainda apresenta crescimento de 2,33% no primeiro trimestre. Os dados das receitas tributárias por tipo e os previsores da atividade econômica mencionados nesta seção serão utilizados para a estimação dos modelos de previsão e, posteriormente, será realizada a compatibilização da receita tributária total através do método de conciliação hierárquica.

Figura 4 - Arrecadação de ICMS no Estado de Goiás por setor



Nota: Os valores do ano de 2022 vão até o mês de fevereiro e todos os valores estão deflacionados pelo IPCA na data base de março de 2022.
Fonte: Secretaria de Estado da Economia de Goiás.

Na Tabela 4 pode-se observar a evolução das despesas de pessoal e encargos do Estado de Goiás durante os anos de 2020 e 2021 em termos reais. Ao deflacionar os dados pelo IPCA, as despesas do Estado mostram uma diminuição consistente em quase todas as contas no ano de 2021, com exceção dos contratos de terceirização.

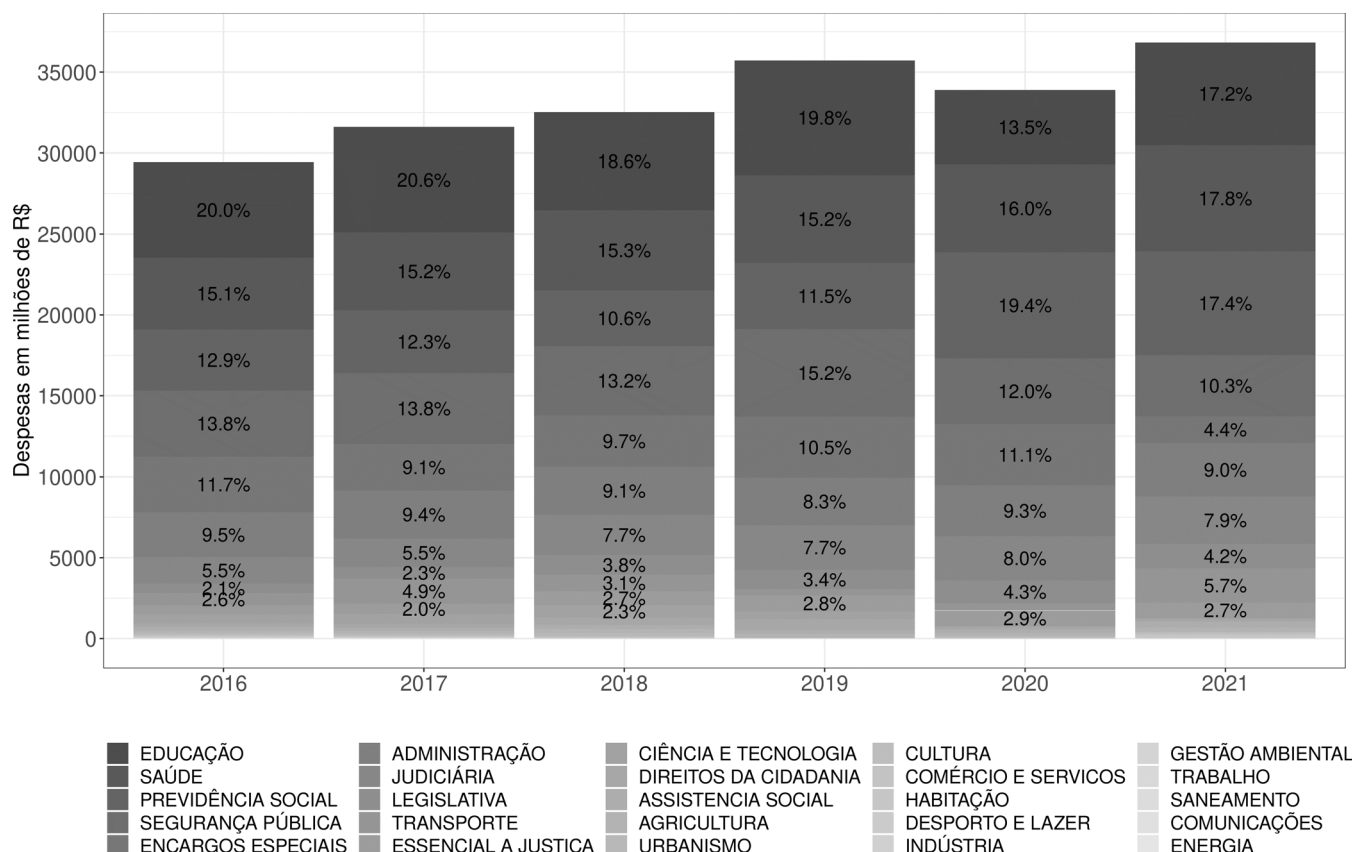
Tabela 4 - Despesas de Pessoal e Encargos no Estado de Goiás

Tipo	2020	2021	Diferença	Var. (%) 2021/2020
Pessoal Ativo	9.536,00	8.942,60	-593,39	-6,22
Outras Despesas Variáveis	8.046,56	7.660,62	-385,94	-4,80
Obrigações Patronais	1.489,44	1.281,98	-207,45	-13,93
Inativos e Pensionistas	6.268,24	5.956,09	-312,15	-4,98
Aposentadorias, Reserva e Reformas	5.156,61	4.932,79	-223,82	-4,34
Pensões	1.110,06	1.023,31	-86,75	-7,81
Contratos de Terceirização	38,21	60,01	21,80	57,06
Despesa Bruta com Pessoal	15.842,44	14.958,71	-883,74	-5,58

Nota: Os valores estão em R\$ milhões e foram deflacionados pelo IPCA com data base em março de 2022.
Fonte: Secretaria de Estado da Economia de Goiás.

A Figura 5 mostra a evolução da execução de despesas por função no Estado de Goiás, com os dados deflacionados pelo IPCA. Pode-se verificar que os maiores gastos estão nos setores de educação, saúde, previdência social e segurança pública.

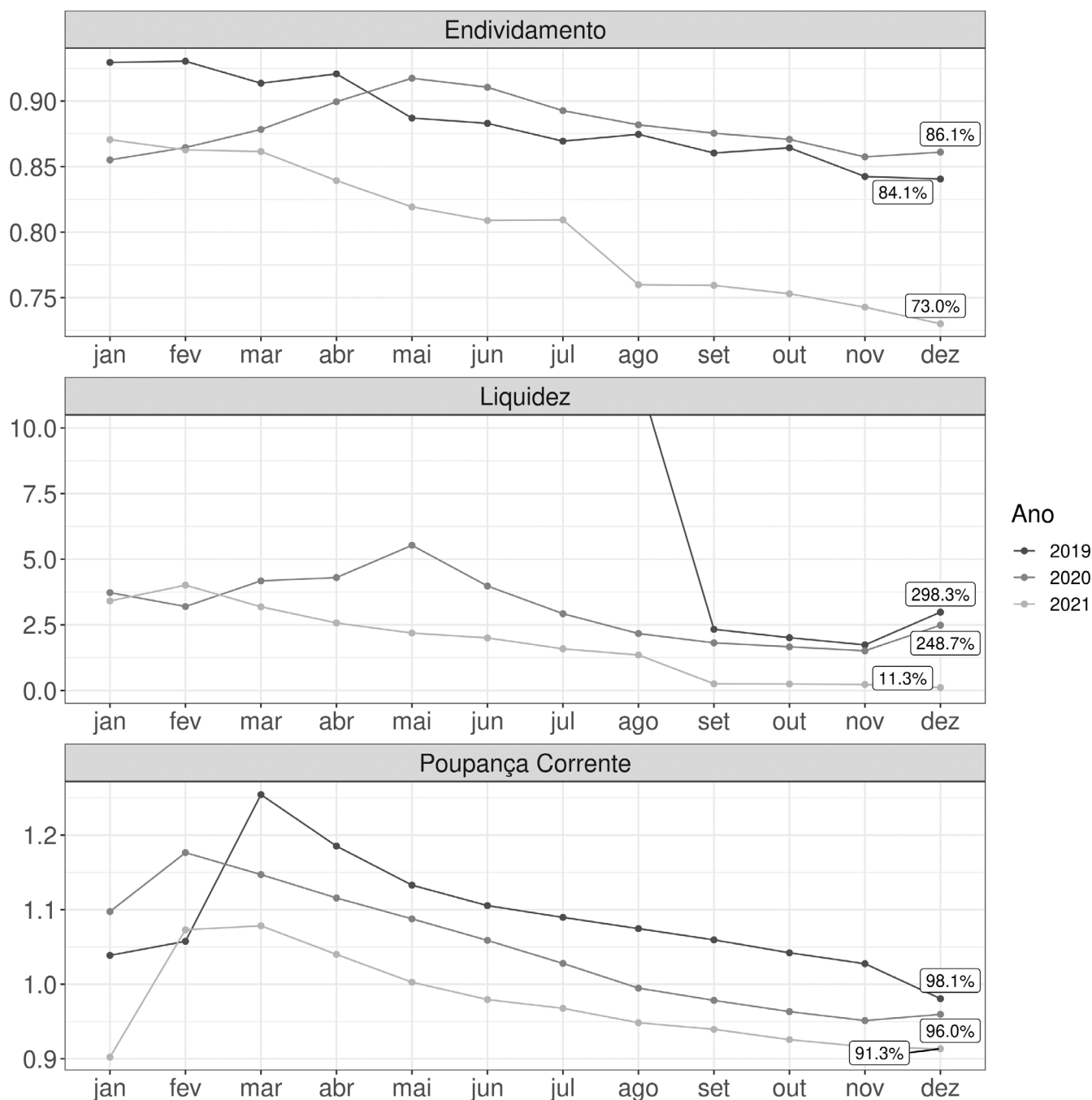
Figura 5 - Execução das despesas por função no Estado de Goiás



Nota: Todos os valores estão deflacionados pelo IPCA na data base de março de 2022.
 Fonte: Secretaria de Estado da Economia de Goiás.

Por fim, a Figura 5 mostra a evolução dos três principais indicadores que compõem a capacidade de pagamento do Estado de Goiás, calculados conforme as fórmulas 1, 2 e 3 e a resolução 501/2017 do Ministério da Fazenda.

Figura 6 - Evolução dos indicadores da capacidade de pagamento do Estado de Goiás



Nota: O cálculo dos indicadores foi elaborado de acordo com a Resolução 501/2017 do Ministério da Fazenda.

Fonte: Secretaria de Estado da Economia de Goiás.

Pode-se observar que houve uma melhora recente em todos os três indicadores ao longo do ano de 2021, o que conferiu ao Estado de Goiás a classificação B em relação a capacidade de pagamento. Todas as variáveis apresentadas nessa seção serão utilizadas para as estimativas dos modelos de previsão propostos na seção de 5.

RESULTADOS

Nesta seção são apresentados os resultados das previsões das receitas tributárias do Estado de Goiás e as projeções para a capacidade de pagamento. Para isso, são realizadas as estimações dos modelos descritos na seção 5 utilizando-se um horizonte final de previsão de 21 períodos, de modo a obter as projeções para o fechamento dos anos de 2022 e 2023.

A Tabela 5 mostra a acurácia dos modelos no curto prazo, estimada a partir da média da raiz do erro quadrático médio (RMSE) da previsão de um período a frente nos conjuntos de validação para cada modelo estimado e cada tipo de receita tributária. Para o cálculo destes dados foi utilizada uma amostra que contém os dados até fevereiro de 2020, excluindo o período pós-pandemia. Os modelos com o menor RMSE foram destacados em negrito.

Tabela 5 - Erros de previsão de curto prazo das Receitas Tributárias (pré-COVID)

Modelo	ICMS	IPVA	ITCD	Outras Receitas	Outros Tributos	Protege	Transfe- rências	Total
ARIMA	51.93	18.59	13.49	102.82	2.14	6.89	56.77	96.24
ARIMA_FOURIER	48.20	23.35	14.04	114.12	2.76	7.22	56.35	146.31
ARIMA_XGBOOST	48.45	17.32	13.72	75.06	2.00	6.96	55.62	84.89
ENSEMBLE (LASSO STACK)	33.21	5.68	10.91	6.31	1.60	6.79	44.73	46.81
ETS	46.24	13.65	14.79	32.89	8.35	6.89	56.24	126.06
GLMNET	54.94	23.57	14.87	79.37	4.09	20.67	107.26	143.44
MARS	62.87	26.67	13.2	74.45	2.79	14.47	80.40	92.20
NAIVE	64.13	23.19	15.63	12.77	2.80	6.79	60.51	77.12
PROPHET	48.15	19.33	11.82	107.50	4.50	23.39	75.14	143.23
PROPHET_XGBOOST	52.98	23.49	11.97	69.61	2.94	16.34	78.90	126.33
RANDOM_FOREST	144.67	30.00	10.89	103.40	9.23	21.08	66.25	119.63
SNAIVE	88.80	<u>11.57</u>	16.26	23.64	7.10	32.99	92.34	112.30
SVM	61.24	26.48	10.86	33.69	4.37	22.48	69.64	102.39
XGBOOST	63.5	21.49	9.29	11.96	3.68	12.62	53.69	86.56

Nota: Os valores estão em R\$ milhões por mês e os erros de previsão foram calculados como a raiz do erro quadrático médio (RMSE) utilizando um período a frente de previsão fora da amostra.

Fonte: elaboração do autor.

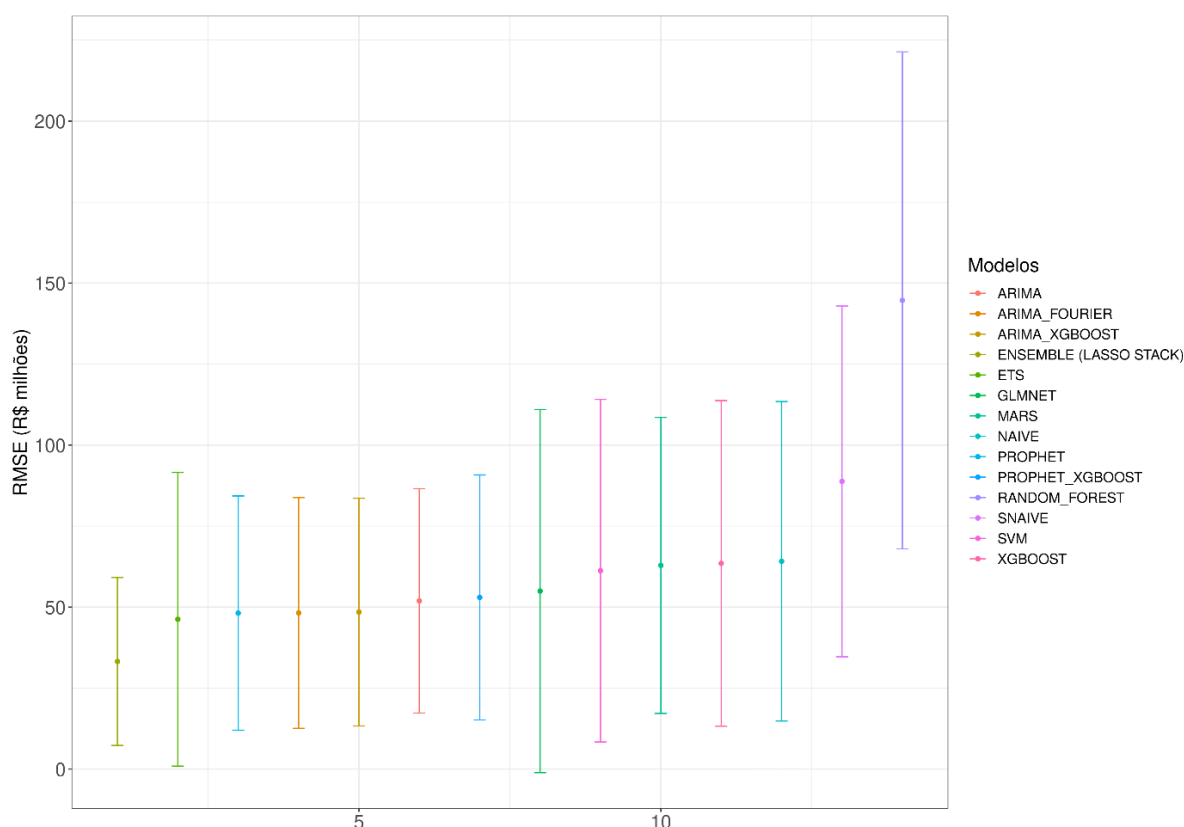
Pode-se observar que na maioria dos casos, o modelo gerado através das combinações das previsões (ENSEMBLE) mostrou-se o modelo com menor RMSE. No caso do ICMS, que compreende mais de 70% das receitas tributárias, o erro médio de estimação deste modelo ficou em cerca de R\$ 33,21 milhões, considerando a frequência mensal. Isso representa uma diminuição significativa de

cerca de 48% do erro de previsão em relação ao modelo benchmark, onde utilizamos apenas o valor passado como previsão, bem como uma diminuição de cerca de 36,05% do erro de previsão em relação aos modelos ARIMA, que é a metodologia mais utilizada para previsões de séries de tempo.

O modelo benchmark (NAIVE) foi superado em todos os casos por outros modelos com melhor performance, mostrando assim o aumento na capacidade preditiva dos modelos utilizados em relação a este benchmark. Entretanto, em algumas séries, como é o caso dos valores relacionados ao fundo Protege, o modelo benchmark se mostrou difícil de ser superado, sendo que os modelos utilizados não foram capazes de trazer previsões significativamente melhores. Muitas vezes isso pode ser melhorado através da desagregação em mais níveis de cada uma destas rubricas. Por outro lado, é importante mencionar que, conforme a Figura 2, a participação deste componente nas receitas tributárias totais é relativamente pequena.

Para o caso do ICMS, o principal componente das receitas tributárias, a Figura 6 mostra não só a média dos erros de previsão para todos os modelos como também a variância do mesmo nos conjuntos de validação.

Figura 7 - Erros de previsão dos modelos para a arrecadação de ICMS (pré-COVID)



Nota: Os valores estão em R\$ milhões por mês e os erros de previsão foram calculados como a raiz do erro quadrático médio (RMSE) utilizando um período a frente de previsão fora da amostra.

Fonte: elaboração do autor.

Conforme a Figura 6, a combinação de previsões não só apresenta um erro quadrático médio menor nos conjuntos de validação, como também apresenta uma das menores variâncias. A acurácia dos modelos no curto prazo também foi estimada utilizando-se a amostra completa dos dados, incluindo o período pós-pandemia. A Tabela 6 reporta estes resultados.

Tabela 6 - Erros de previsão de curto prazo das Receitas Tributárias (amostra completa)

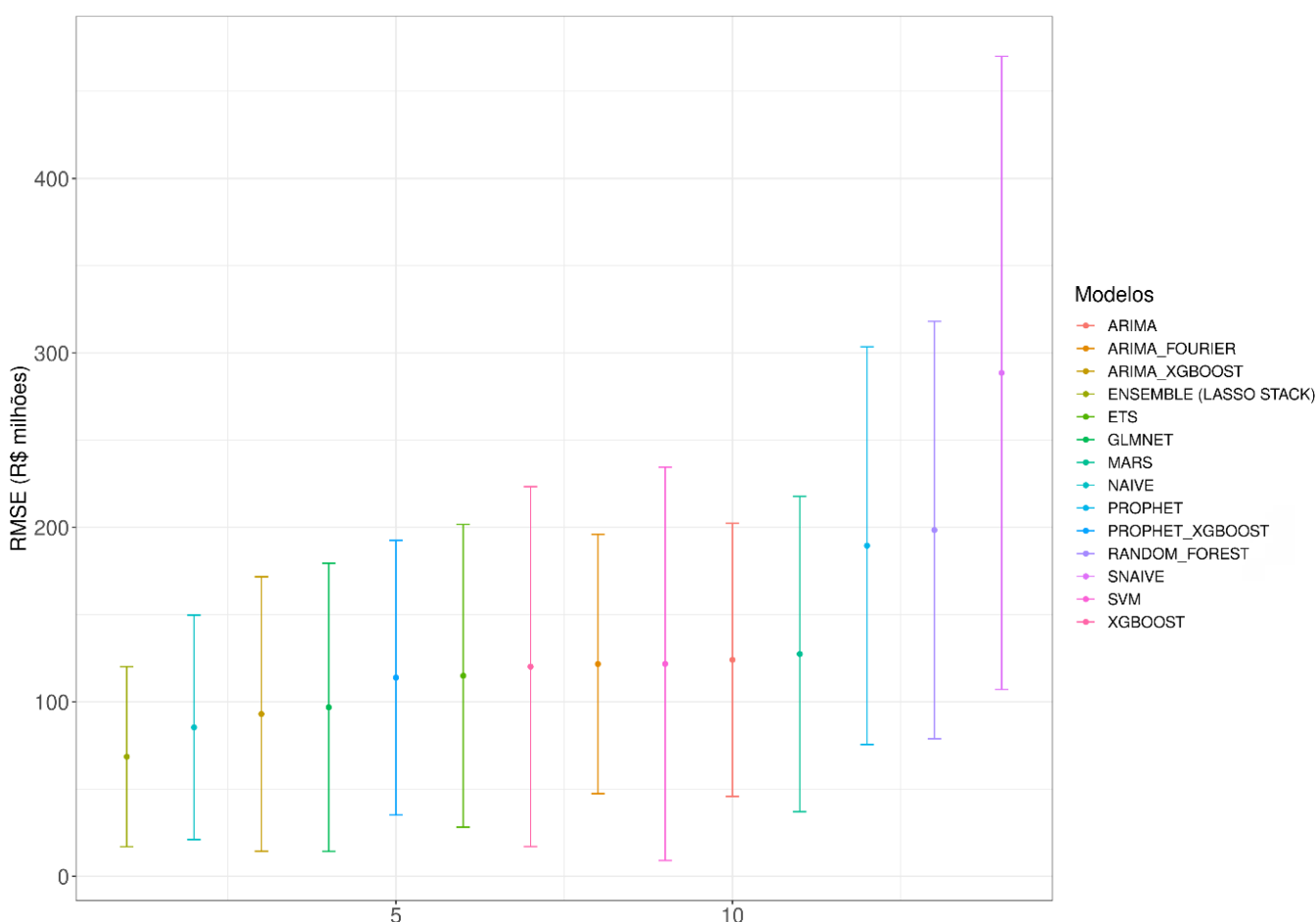
Modelo	ICMS	IPVA	ITCD	Outras Receitas	Outros Tributos	Protege	Transfe- rências	Total
ARIMA	124.08	32.44	10.51	10.71	4.07	6.38	89.67	219.42
ARIMA_FOURIER	121.67	33.45	10.31	50.45	4.36	6.62	87.72	234.76
ARIMA_XGBOOST	93.03	32.15	8.94	10.03	3.94	6.34	88.92	167.75
ENSEMBLE (LASSO STACK)	68.53	24.51	7.57	6.44	2.94	3.55	78.43	122.13
ETS	114.93	44.67	13.08	25.76	7.68	6.57	84.65	373.50
GLMNET	96.84	46.84	10.95	83.57	5.34	22.45	112.12	210.20
MARS	127.41	49.3	15.82	52.87	6.57	11.47	87.75	251.83
NAIVE	85.34	48.22	12.25	12.31	3.27	6.57	83.10	151.76
PROPHET	189.50	48.22	12.82	84.22	4.39	8.12	85.19	394.99
PROPHET_XGBOOST	113.84	50.89	10.82	14.37	3.98	7.05	88.75	168.89
RANDOM_FOREST	198.46	50.96	9.33	88.53	8.01	18.85	94.81	245.45
SNAIVE	288.54	31.28	18.56	11.76	8.63	21.67	138.82	407.10
SVM	121.77	50.72	13.53	22.06	6.71	21.62	85.54	205.42
XGBOOST	120.16	46.04	9.77	14.09	4.97	9.07	88.02	208.91

Nota: Os valores estão em R\$ milhões por mês e os erros de previsão foram calculados como a raiz do erro quadrático médio (RMSE) utilizando um período a frente de previsão fora da amostra.

Fonte: elaboração do autor.

Em todos os casos, o modelo gerado através das combinações das previsões (ENSEMBLE) também se mostrou o modelo com menor RMSE. Entretanto, todos os modelos apresentaram um RMSE médio maior, uma vez que o período amostral completo compreende uma maior variância nas séries das receitas tributárias. Outro ponto importante é que os modelos multivariados tiveram melhor rendimento do que os modelos univariados, uma vez que o período pós-pandemia compreende grandes choques que são incorporados nos modelos através dos previsores.

Para a arrecadação de ICMS, na Figura 7, podemos ver que, embora os modelos tiveram uma média maior do erro quadrático nos conjuntos de validação, a variância destes erros permaneceu razoavelmente controlada.

Figura 8 - Erros de previsão dos modelos para o ICMS (amostra completa)

Nota: Os valores estão em R\$ milhões por mês e os erros de previsão foram calculados como a raiz do erro quadrático médio (RMSE) utilizando um período a frente de previsão fora da amostra.

Fonte: elaboração do autor.

Os erros de previsão reportados até aqui são referentes a previsões realizadas para apenas um mês a frente. De modo a obter uma estimativa dos erros de previsão para um horizonte mais longo, em um cenário com grandes choques, os modelos foram estimados para a amostra dos dados de janeiro de 2010 a fevereiro de 2020, e comparados com o período pós-pandemia de março de 2020 a março de 2022. Este seria um dos piores cenários possíveis de comparação, porém pode-se considerar estes erros como um limite superior para as estimativas. A Tabela 7 reporta esses resultados. É importante mencionar que como estas estimativas foram obtidas apenas utilizando-se um conjunto de comparação, em um período com grandes choques nas variáveis de interesse, não é recomendado comparar a acurácia de modelos distintos com base nestes resultados. O objetivo é apenas ter uma estimativa do erro quadrático médio dos modelos em um cenário com grandes quebras estruturais.

Tabela 7 - Erros de previsão de longo prazo das Receitas Tributárias (amostra completa)

Modelo	ICMS	IPVA	ITCD	Outras Receitas	Outros Tributos	Protege	Transfe- rências	Total
ARIMA	309.01	65.81	28.76	15.25	6.51	34.97	124.36	490.34
ARIMA_FOURIER	320.29	64.70	29.54	60.09	6.69	34.93	128.39	494.49
ARIMA_XGBOOST	306.95	65.17	28.39	86.53	6.33	34.55	125.06	492.03
ENSEMBLE (LASSO STACK)	336.34	59.76	26.38	18.76	9.18	16.08	135.20	535.77
ETS	325.21	68.58	30.55	90.33	15.21	34.23	135.33	699.11
GLMNET	334.21	64.16	22.48	81.53	5.64	13.19	123.56	509.98
MARS	428.55	71.77	27.70	63.07	6.10	30.56	239.10	577.29
NAIVE	478.61	93.53	32.03	21.32	10.69	34.23	109.59	646.93
PROPHET	356.49	67.98	23.93	105.16	5.49	13.75	189.59	600.07
PROPHET_XGBOOST	332.04	64.29	23.49	30.16	5.74	15.59	164.50	505.97
RANDOM_FOREST	466.49	69.17	27.52	26.77	13.56	33.69	138.32	672.20
SNAIVE	432.58	61.84	31.03	18.69	12.06	33.51	162.12	632.89
SVM	336.74	61.98	27.95	21.87	6.08	15.78	205.03	619.21
XGBOOST	437.94	70.76	28.78	21.45	10.99	27.31	154.97	675.13

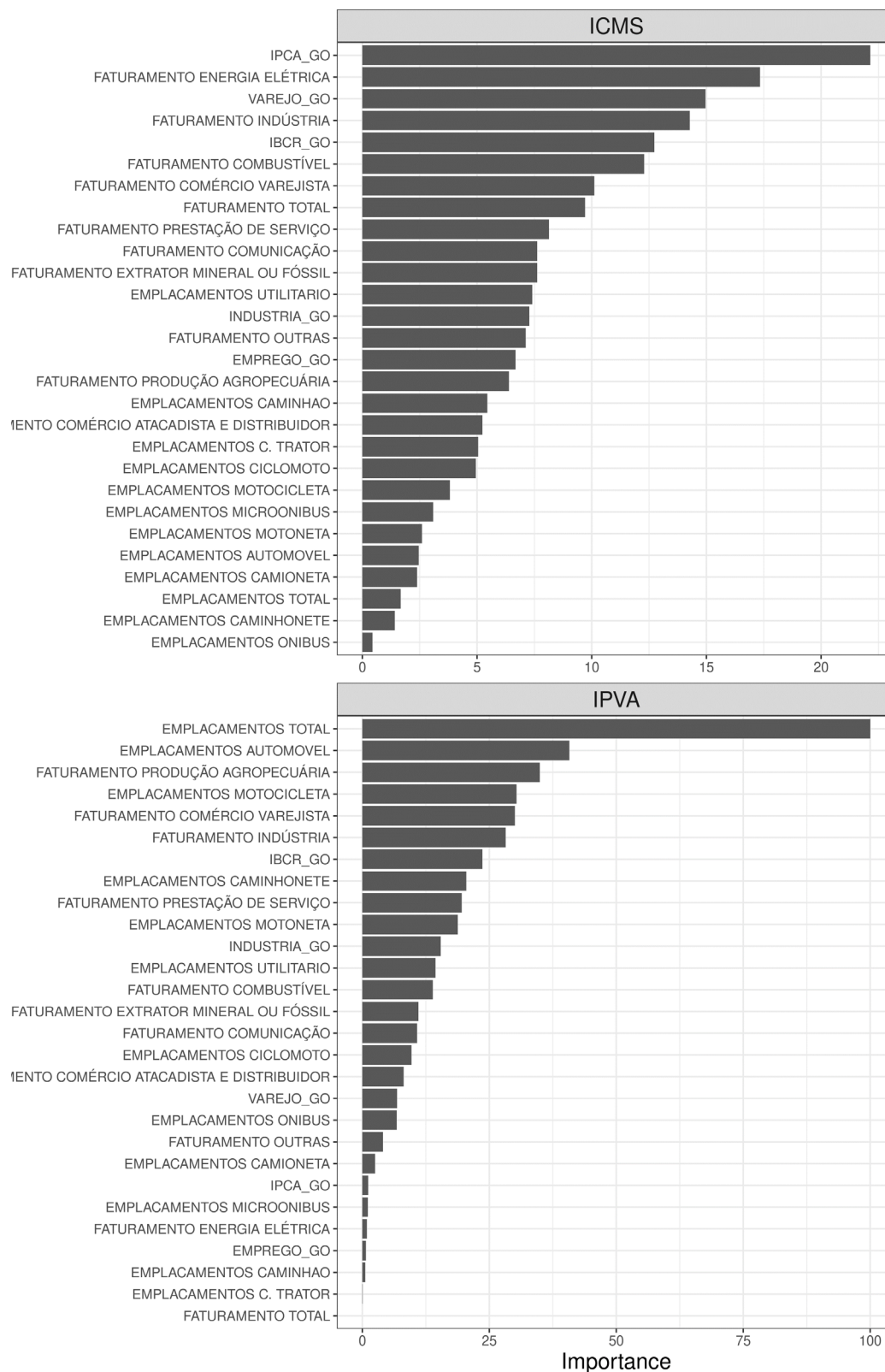
Nota: Os valores estão em R\$ milhões por mês e os erros de previsão foram calculados como a raiz do erro quadrático médio (RMSE) utilizando 24 períodos a frente de previsão fora da amostra (2 anos).

Fonte: elaboração do autor.

Para os modelos de previsão multivariados, pode-se determinar quais são os principais previsores que contribuem para reduzir o erro quadrático. A Figura 8 calcula um indicador de importância para o modelo GLMNET que varia de 0 a 100. São reportados os dois principais componentes das receitas tributárias, o ICMS e o IPVA. Neste indicador, cada um dos previsores é removido do modelo e então o modelo é reestimado para se obter um novo erro quadrático de previsão, de modo a obter a diferença entre o RMSE com e sem a inclusão do previsor. Aqueles previsores que apresentaram um maior aumento do RMSE após a exclusão são os primeiros a serem ranqueados no indicador.

Pode-se observar que os previsores associados ao índice de inflação e o faturamento de notas fiscais para diversos setores são os mais relevantes para o ICMS, enquanto que os previsores relativos a emplacamento de veículos têm maior importância para a arrecadação do IPVA. Este resultado está dentro do esperado, mostrando que o conjunto dos previsores utilizados está sendo efetivo para contribuir com o poder preditivo do modelo.

Figura 9 - Importância dos previsores no modelo GLMNET para ICMS e IPVA



Nota: Os valores estão em R\$ milhões por mês e os erros de previsão foram calculados como a raiz do erro quadrático médio (RMSE) utilizando um período a frente de previsão fora da amostra.
 Fonte: elaboração do autor.

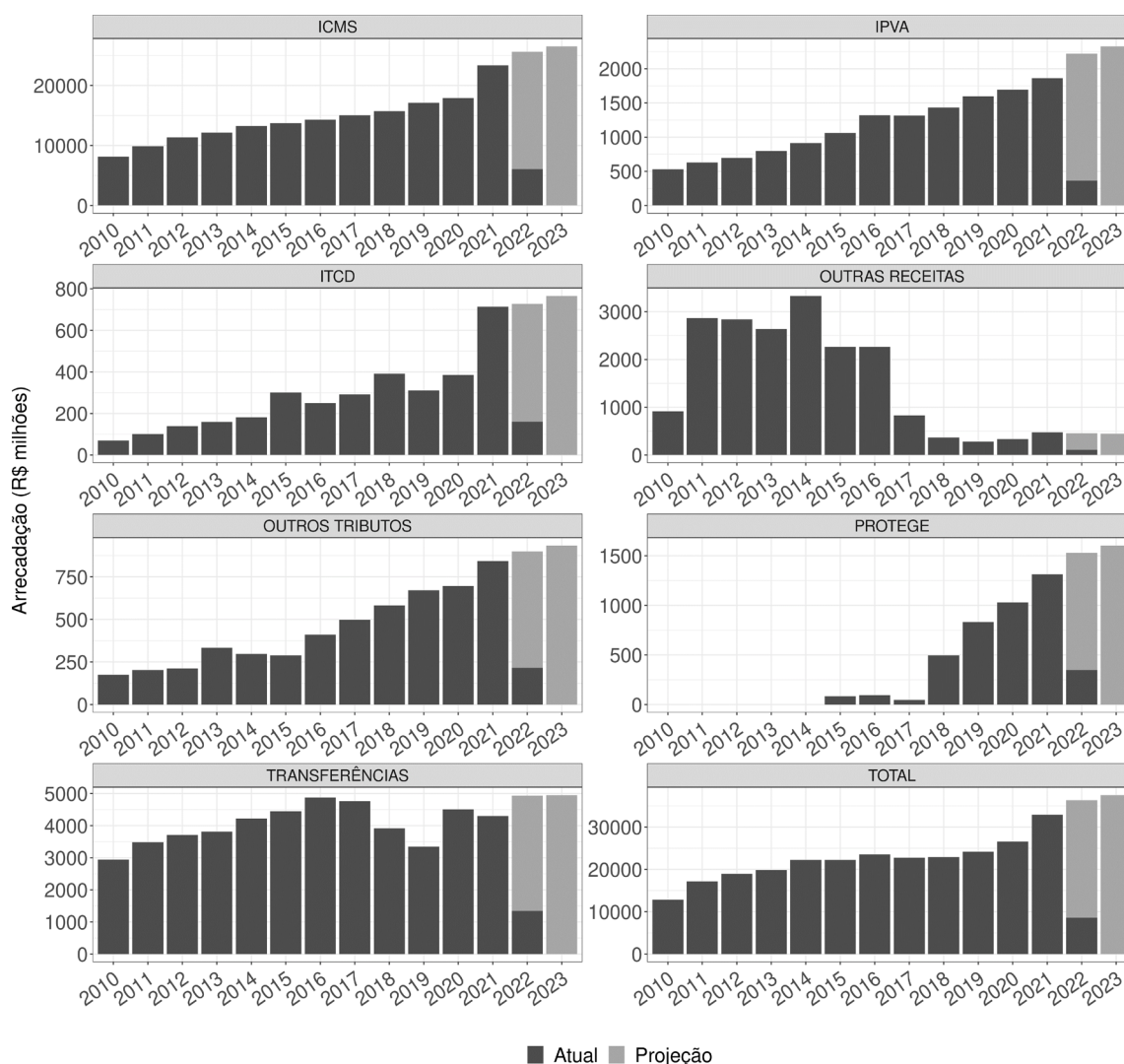
Uma vez analisada a acurácia de curto e longo prazo dos modelos de previsão e as principais variáveis determinantes, pode-se fazer a seleção dos melhores modelos de previsão e estimá-los na amostra completa dos dados para obter as projeções finais. Para isso, foi selecionado o melhor modelo para cada tipo de receita tributária com base no RMSE de curto prazo considerando-se a amostra sem o período da pandemia. Após a escolha dos melhores modelos, utiliza-se a amostra completa para reestimá-los e gerar previsões futuras com horizonte de 21 meses (a fim de completar os anos de 2022 e 2023). Por fim, os resultados das previsões são compatibilizados dentro da estrutura hierárquica através do algoritmo desenvolvido por Wickramasuriya et al. (2019). Os resultados das previsões são reportados na Tabela 8 de forma agregada para os anos de 2022 e 2023. Os valores das receitas tributárias na estimação dos modelos foram utilizados em termos nominais, pois o deflacionamento pode distorcer os cálculos dos erros de previsão dependendo da data base do índice utilizado. A evolução dos valores projetados também pode ser observada na Figura 9.

Tabela 8 - Projeções das Receitas Tributárias para os anos de 2022 e 2023

Ano	ICMS	IPVA	ITCD	Outras Receitas	Outros Tributos	Protege	Transfe- rências	Total
2010	8170.08	526.87	69.35	915.89	174.3	0	2941.41	12797.89
2011	9875.18	630.03	101.6	2866.51	201.87	0	3486.28	17161.48
2012	11369.28	695.51	138.64	2840.94	212.05	0	3708.09	18964.51
2013	12137.75	796.87	159.33	2638.88	332.7	0	3817.89	19883.42
2014	13252.85	914.36	181.37	3329.03	297.29	0	4224.00	22198.91
2015	13745.22	1063.05	301.09	2263.48	288.74	83.44	4451.55	22196.57
2016	14334.66	1319.63	249.22	2265.89	411.52	93.46	4873.77	23548.15
2017	15022.54	1316.3	291.82	829.32	497.53	48.03	4769.29	22774.83
2018	15754.65	1431.21	391.17	366.12	582.69	495.40	3914.86	22936.10
2019	17125.88	1594.77	311.17	276.2	670.89	833.74	3350.84	24163.50
2020	17921.74	1693.69	385.62	330.91	696.24	1029.69	4506.37	26564.27
2021	23376.44	1862.17	713.67	473.93	842.93	1311.74	4298.59	32879.49
2022*	25619.22	2218.32	726.82	456.63	899.37	1529.88	4937.75	36388.00
2023*	26529.62	2326.86	765.37	444.17	934.02	1603.56	4952.61	37556.21

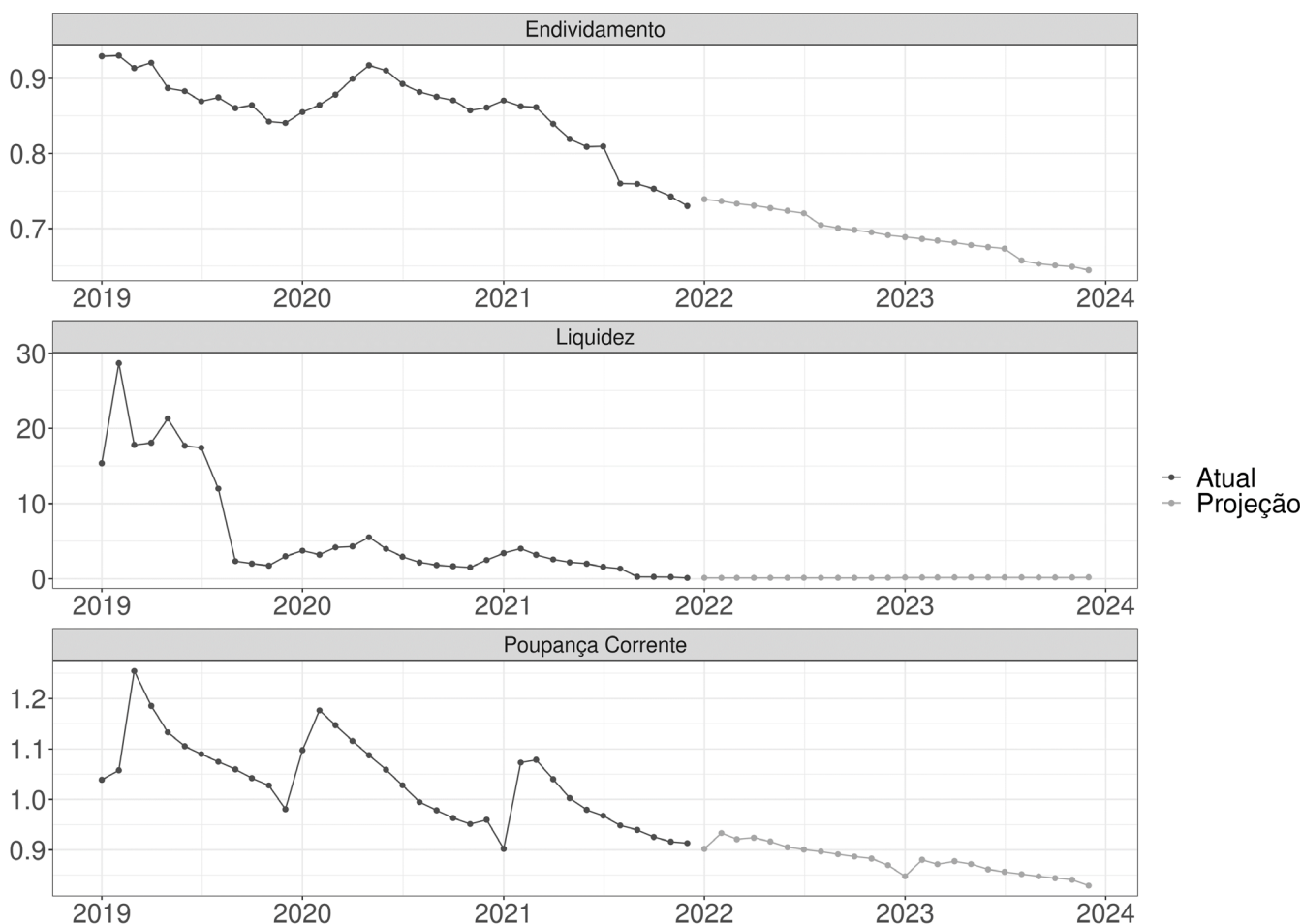
Nota: Os valores são nominais e estão em R\$ milhões por ano, sendo que o horizonte final corresponde a 21 períodos a frente e o modelo escolhido foi o que teve o melhor desempenho na amostra pré-covid, sendo realizada a conciliação hierárquica após a estimação das previsões. Fonte: elaboração do autor.

Figura 10 - Projeções das Receitas Tributárias para os anos de 2022 e 2023



Nota: Os valores são nominais e estão em R\$ milhões por ano, sendo que o horizonte final corresponde a 21 períodos a frente e o modelo escolhido foi o que teve o melhor desempenho na amostra pré-covid, sendo realizada a conciliação hierárquica após a estimação das previsões. Fonte: elaboração do autor.

Podemos observar que praticamente todas as receitas tributárias tem uma projeção de aumento para 2022 e 2023 em termos nominais. Cabe mencionar que a previsão para a arrecadação de ICMS em 2022 alcançou o patamar de cerca de R\$ 25,6 bilhões, sendo que a previsão para o fechamento da receita tributária no ano de 2022 alcança a cifra de cerca de R\$ 36,4 bilhões, incluindo os valores já reportados até o mês março e as previsões do modelo para os meses de abril a dezembro. Esses resultados representam o primeiro nível de agregação proposto na Figura 1. Após estas estimativas são elaboradas as previsões para os demais níveis de agregação seguindo a estrutura das Figuras 1 e 2, de forma a obter as projeções para os indicadores da Capacidade de Pagamento do Estado. Estas projeções são reportadas na Figura 10.

Figura 11 - Projeções da Capacidade de Pagamento para os anos de 2022 e 2023

Nota: Os valores dos indicadores são calculados conforme a resolução 501/2017 do Ministério da Fazenda, sendo que o horizonte final das previsões corresponde a 21 períodos a frente.
 Fonte: elaboração do autor.

Pode-se observar uma projeção de melhoria nos três indicadores que compõem a Capacidade de Pagamento do Estado de Goiás. Entretanto, vale ressaltar que essas projeções apenas compreendem as características associadas as tendências e sazonalidades prévias existentes nos dados, uma vez que para projeções de longo prazo, as principais variáveis preditoras ainda não estão disponíveis, e são então substituídas por previsões univariadas para o cômputo das projeções finais. Possíveis mudanças no redirecionamento das políticas públicas podem afetar significativamente as projeções de longo prazo.

DESCRIÇÃO DA EVOLUÇÃO DO PROJETO

O projeto começou a sua vigência a partir do dia 23 de abril de 2021. No dia 12 de maio de 2021 foi realizada uma reunião inicial com o Instituto Mauro Borges para apresentação e alinhamento dos projetos de pesquisa. Durante esta reunião foram definidos os papéis de cada integrante no auxílio aos projetos de pesquisa em andamento, bem como o coordenador do projeto “Previsão e determinantes do equilíbrio fiscal para o Estado de Goiás” apresentou o mesmo para a equipe da Secretaria de Estado da Economia de Goiás.

No mês de maio foi elaborado um edital de seleção de bolsista de Pós-Graduação para auxiliar na realização do projeto, bem como realizados os trâmites internos relativos a aprovação desta chamada. Assim, no dia 31 de maio de 2021 foi publicada a Chamada Pública n. 4 de 2021 do Instituto Mauro Borges, com o intuito de contratação de Bolsista de Pós-Graduação no Projeto “Previsão e determinantes do equilíbrio fiscal para o Estado de Goiás”. O resultado final da chamada acabou sendo divulgado no dia 23 de julho de 2021, sendo selecionada a discente de Pós-Graduação Fernanda dos Anjos Felix, que começou as suas atividades a partir do mês de agosto de 2021.

Durante os meses de maio e junho de 2021 também foram realizadas as etapas de prospecção da literatura sobre previsão de resultados fiscais, com especial atenção aos métodos de previsão adotados na literatura. Com isso, foi construída uma base bibliográfica para referências futuras nas próximas etapas do projeto e elaborada uma nova estrutura hierárquica para a estimação dos modelos de previsão do equilíbrio fiscal do Estado de Goiás. Estas etapas estão contempladas nas seções 4 e 5 deste relatório.

Durante o mês de julho foi feito um esforço em conjunto com a Secretaria de Estado da Economia para a coleta dos dados a serem utilizados nos modelos de previsão. Em 20 de julho de 2021 foi realizada uma nova reunião com a equipe do Instituto Mauro Borges em que foi apresentada a evolução do projeto até o momento e foram listadas as variáveis que ainda faltavam ser obtidas para a realização do projeto. Com o esforço conjunto da equipe, todas as variáveis necessárias foram obtidas durante as semanas seguintes. A descrição detalhada das variáveis e das fontes dos dados está na seção 6 deste relatório.

Também nesta reunião do dia 20 de julho de 2021, foi revelado um interesse mais imediato na projeção da arrecadação do ICMS por parte da Secretaria de Estado da Economia. Sendo assim, foi definido que o foco das semanas seguintes seria na construção de modelos de previsão preliminares para a arrecadação do ICMS.

No mês de agosto foram entregues os resultados iniciais das estimativas dos modelos para a arrecadação do ICMS à Secretaria da Economia de Goiás, sendo realizada uma reunião no dia 31 de agosto de 2021. Nesta ocasião, o coordenador do projeto fez uma apresentação sobre os avanços do projeto de

pesquisa e os resultados preliminares de previsão do ICMS, conforme demanda da Secretaria da Economia em reunião anterior. Também foram demandadas previsões preliminares para o fechamento do ICMS de 2021 para o final do dia, demanda que foi prontamente atendida pelo coordenador do projeto.

No mês de setembro o coordenador do projeto trabalhou nas previsões do restante das receitas tributárias e na elaboração dos gráficos e tabelas a serem reportados no terceiro relatório. No dia 18 de outubro de 2021 foi feita uma nova reunião com a Secretaria da Economia, com o objetivo de discutir as principais variáveis previsoras das receitas tributárias. Na ocasião foi exposto a intenção da Secretaria da Economia em obter previsões de curto prazo da receita tributária total para os próximos meses de 2021. Estas previsões foram realizadas e adicionadas no relatório semestral do projeto. Os resultados deste relatório foram apresentados ao público no dia 25 de novembro, sob ocasião da Apresentação Técnica dos Projetos de Pesquisa da Parceria - IBM/Fapeg, contando com a contribuição de profissionais da Secretaria da Economia. A apresentação foi gravada e disponibilizada ao público através do link <https://www.facebook.com/imbgoias/videos/360310192531090/>.

Durantes os meses de outubro a dezembro de 2021, foram construídas as previsões para os outros componentes do equilíbrio fiscal do Estado e produzido um relatório interativo através do Power BI. Nos dias 26 de janeiro de 2022 e 28 de março de 2022 foram feitas duas reuniões de alinhamento e apresentação de resultados preliminares com o IMB e a Secretaria da Economia, também com o intuito de obter os últimos dados necessários para a operacionalização final do projeto e, conseqüentemente, foi realizado a entrega do Relatório 3 (na forma de painel interativo). Com isso, durante os meses de fevereiro e março foram feitos os ajustes finais ao projeto de pesquisa.

Os resultados finais do projeto contemplam um relatório de pesquisa, um framework de previsão e projeção do equilíbrio fiscal do Estado implementado através do software R, e um dashboard interativo para visualização dos resultados construídos utilizando-se o Power BI. Todos estes arquivos foram disponibilizados ao Instituto Mauro Borges na ocasião da entrega do relatório final do projeto. Cada etapa do projeto e as entregas dos relatórios foram documentadas no SEI, através do processo número 202118037002655.

Os resultados finais foram apresentados em duas ocasiões, a primeira, no dia 18 de abril de 2022, para um público mais amplo, e a segunda, no dia 26 de abril de 2022 para o corpo técnico, com o objetivo de transmissão da tecnologia desenvolvida e capacitação técnica, conforme previsto no Plano de Trabalho. A capacitação com os técnicos do IMB ocorreu no dia 06 de junho de 2022.

REFERÊNCIAS

- ASIMAKOPOULOS, S.; PAREDES, J.; WARMEDINGER, T. Forecasting fiscal time series using mixed frequency data. ECB Working Paper, n. 1550, 2013.
- BOTRIĆ, V.; VIZEK, M. Forecasting fiscal revenues in a transition country: the case of Croatia. Zagreb International Review of Economics & Business, v. 15, n. 1, p. 23-36, 2012.
- CLEMENTE, A.; CLEMENTE, L. T. Aplicação da metodologia Box-Jenkins para previsão do ICMS do Estado do Paraná de agosto de 2011 a julho de 2012. Revista Economia & Tecnologia, v. 7, n. 4, 2011.
- CHAIN, C. P. et al. Contribuição da modelagem de valores atípicos na previsão da arrecadação do ICMS do Estado de Minas Gerais. Exacta, v. 13, n. 2, p. 239-249, 2015.
- CHEN, Tianqi; GUESTRIN, Carlos. Xgboost: A scalable tree boosting system. CoRR abs/1603.02754 (2016). arXiv:1603.02754, 2016.
- DIAS, V. P.; ICHIKAVA, E. Uma análise empírica da relação entre spread e risco. Revista do BNDES, Rio de Janeiro, n. 36, p. 177-213, 2011.
- GADELHA, S. R. de B.; LIMA, A. F. R.; POLLI, D. A.. Uso da Metodologia de Combinação de Previsões para Projeções da Arrecadação de Receitas Brutas Primárias de Tributos Federais. Cadernos de Finanças Públicas, v. 1, n. 1, 2020.
- FAVERO, C. A.; MARCELLINO, M. Modelling and forecasting fiscal variables for the Euro area. Oxford Bulletin of Economics and Statistics, v. 67, p. 755-783, 2005.
- FILDES, R. The evaluation of extrapolative forecasting methods. International Journal of Forecasting, v. 8, n. 1, p. 81-98, 1992.
- HASTIE, Trevor et al. The elements of statistical learning: data mining, inference, and prediction. New York: springer, 2009.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. Forecasting: principles and practice. OTexts, 2018.
- HYNDMAN, R. J.; KHANDAKAR, Y. Automatic Time Series Forecasting: The forecast Package for R. Journal of Statistical Software, v. 27, n. 3, p. 1–22, 2008.

HYNDMAN, R.; KOEHLER, A. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679-688, 2006.

HYNDMAN, R. et al. *Forecasting with exponential smoothing: the state space approach*. Springer Science & Business Media, 2008.

IMROHOROGLU, S. et al. Productivity and fiscal policy in Japan: short-term forecasts from the standard growth model. *Monetary and Economic Studies*, v. 29, p. 73-106, 2011.

KRUEGER, A. O. The Political Economy of the Rent-Seeking Society. *American Economic Review*, v. 64, n. 3, p. 291–303, 1974.

LISBOA, M. B.; LATIF, Z. A. Democracy and growth in Brazil. In: *Democracy Consensus workshop*, Rio de Janeiro, 2013.

LJUNG, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. *Biometrika*, v. 65, n. 2, p. 297-303, 1978.

LÜTKEPOHL, Helmut. Forecasting aggregated time series variables: A survey. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, v. 2010, n. 2, p. 1-26, 2011.

MARCELLINO, Massimiliano; STOCK, James H.; WATSON, Mark W. Macroeconomic forecasting in the euro area: Country specific versus area-wide information. *European Economic Review*, v. 47, n. 1, p. 1-18, 2003.

MELO, Bruno. Modelo de previsão para arrecadação tributária. SECRETARIA DO TESOURO NACIONAL. *Finanças Públicas: VI Prêmio Tesouro Nacional*, 2001.

MENDONÇA, M. J. C.; MEDRANO, L. A. Um modelo de combinação de previsões para arrecadação de receita tributária no Brasil. *Texto para Discussão IPEA*, n. 2186, 2016.

MENDONÇA, M. J. et al. Aplicação de modelos dinâmicos bayesianos para projeção de arrecadação tributária. *Texto para Discussão IPEA*, n. 2590, 2020.

MENDONÇA, M. J. C.; SACHSIDA, A.; MEDRANO, L. A. T. Um modelo econométrico para previsão de impostos no Brasil. *Economia Aplicada*, v. 17, n. 2, p. 295-329, 2013.

PEREVALOV, Nikita; MAIER, Philipp. On the advantages of disaggregated data: Insights from forecasting the US economy in a data-rich environment. *Bank of Canada*, 2010.

PESSOA, F. de M. C.; CORONEL, D. A.; LIMA, J. E.. Previsão de arrecadação de ICMS para o estado de Minas Gerais: uma comparação entre modelos Arima e Arfima. *Revista Brasileira de Gestão e Desenvolvimento Regional*, v. 9, n. 2, 2013.

PETRIS, G.; PETRONE, S.; CAMPAGNOLI, P. Dynamic linear models. In: *Dynamic Linear Models with R*. Springer, New York, NY, 2009. p. 31-84.

SANTOS, C. M. dos; LIMA, J. E. Análise de previsões da arrecadação do ICMS no Estado de Minas Gerais. *Revista de Economia e Administração*, v. 5, n. 4, p. 413- 423, 2006.

SCHEFFER, D.; SOUZA, A. M.; ZANINI, R. R. Utilização de modelos ARIMA para previsão da arrecadação de ICMS do Estado do Rio Grande do Sul. *Anais do XVII Simpósio de Pesquisa Operacional e Logística da Marinha – SPOLM*, São Paulo, 2014.

SCOTT, Steven L.; VARIAN, Hal R. Predicting the present with Bayesian structural time series. *International Journal of Mathematical Modelling and Numerical Optimisation*, v. 5, n. 1-2, p. 4-23, 2014.

SIQUEIRA, Marcelo Lettieri. Melhorando a previsão de arrecadação tributária federal através da utilização de modelos de séries temporais. SECRETARIA DO TESOURO NACIONAL. *Finanças Públicas: VII Prêmio Tesouro Nacional*, 2002.

TAYLOR, Sean J.; LETHAM, Benjamin. Forecasting at scale. *The American Statistician*, v. 72, n. 1, p. 37-45, 2018.

WICKRAMASURIYA, Shanika L.; ATHANASOPOULOS, George; HYNDMAN, Rob J. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, v. 114, n. 526, p. 804-819, 2019.

ZELLNER, A. A tale of forecasting 1001 series: The Bayesian knight strikes again. *International Journal of Forecasting*, v. 2, p. 491-494, 1986.

ZOU, Hui; HASTIE, Trevor. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, v. 67, n. 2, p. 301-320, 2005.

