
PREVISÃO DE TRIBUTOS ESTADUAIS COM USO DE MODELOS UNIVARIADOS COMBINADOS APLICADOS AO ESTADO DA PARAÍBA

State tax forecast with use of combined unisorted models applied to the state of Paraíba

Daniel Lins Batista Guerra

Economista. Doutor em Economia pela Universidade Federal da Paraíba – UFPB. Analista Ministerial do Ministério Público da Paraíba, Rua Rodrigues de Aquino, s/n, Centro, João Pessoa/PB, 58.013-030. danielconomia@hotmail.com

Cássio da Nóbrega Besarria

Economista. Doutor em Economia pela Universidade Federal de Pernambuco – UFPE. Professor adjunto da Universidade Federal da Paraíba, Cidade Universitária, João Pessoa/PB, 58.013-030. cassiodanobrega@yahoo.com.br

Resumo: O artigo objetiva apresentar modelos alternativos à previsão eficiente de tributos estaduais. Foram testados nove modelos univariados para a arrecadação de tributos do estado da Paraíba, com dados de 1997 a 2020. Os modelos foram alvo de combinações, dentre as quais as propostas em Figueiredo (2019), Hyndman e Athanasopoulos (2018) e Shaub e Ellis (2020). As combinações foram geradas com uso dos seguintes modelos: *Holt-Winters* com suavização exponencial por tratamento aditivo; SARIMA (*Seasonal Autoregressive Integrated Moving Average*); X-13-ARIMA-SEATS; o ETS (*Error Trend Seasonal*); NNAR (*Neural Network Autoregression*); TBATS (*Exponential Smoothing Method + Box-Cox Transformation + ARMA model for residuals + Trigonometric Seasonal*); modelo STLM (*Seasonal and Trend decomposition using Loess* THETAM (método *Theta de Assimakopoulos e Nikopoulos, 2000*); e, SNAIVE (*Seasonal Naive*). Os resultados das previsões foram comparados por meio dos indicadores de desempenho RMSE (*Root Mean Squared Error*); MAE (*Mean Absolute Error*); MPE (*Mean Percentual Error*); e, MAPE (*Mean Absolute Percentual Error*). Os achados apontam a necessidade de aperfeiçoamento na elaboração das previsões de tributos no estado da Paraíba. A incorporação do uso de modelos de previsão singulares ou combinados é capaz de oferecer ganhos às formulações de previsão, e ao planejamento e execução orçamentária estadual.

Palavras-chave: Séries temporais; Previsão; Política Fiscal.

Abstract: The article aims to present alternative models for efficient forecasting of state taxes. Nine univariate models were tested for tax collection in the state of Paraíba, with data from 1997 to 2020. The models were subject to combinations, including those proposed in Figueiredo (2019), Hyndman and Athanasopoulos (2018) and Shaub and Ellis (2020). The combinations were generated using the following models: Holt-Winters with exponential smoothing by additive treatment; SARIMA (Seasonal Autoregressive Integrated Moving Average); X-13-ARIMA-SEATS; o ETS (Error Trend Seasonal); NNAR (Neural Network Autoregression); TBATS (Exponential Smoothing Method + Box-Cox Transformation + ARMA model for residuals + Trigonometric Seasonal); model STLM (Seasonal and Trend decomposition using Loess THETAM (Theta method of Assimakopoulos and Nikopoulos, 2000); and, SNAIVE (Seasonal Naive). The forecast results were compared using the performance indicators RMSE (Root Mean Squared Error); MAE (Mean Absolute Error); MPE (Mean Percentual Error); and MAPE (Mean Absolute Percentual Error). The findings point to the need for improvement in the preparation of tax forecasts in the state of Paraíba. The incorporation of the use of singular forecast models or combined, can offer gains to forecast formulations and state budget planning and execution.

Keywords: time series; forecast; tax policy.

JEL Classification: C22; C53; H71

1 INTRODUÇÃO

A elaboração de previsões precisas da arrecadação de tributos pelo Estado é componente essencial para o bom planejamento orçamentário e a condução da política fiscal. Em razão disso, estudos recentes têm se dirigido ao teste de modelos de previsão de tributos no Brasil.

De um modo geral, as abordagens têm feito uso de modelos univariados, concentrando-se, principalmente, naqueles derivados da metodologia de Box-Jenkins, como são os casos dos estudos de Santos e Lima (2006), Azevedo, Silva e Gatsio (2017), Clemente e Clemente (2011), e Pessoa, Coronel e Lima (2013). Apenas mais recentemente é que estudos têm se voltado à combinação de um ou mais métodos de previsão, como discutido em Mendonça e Medrano (2016) e Moço (2017). Esses trabalhos relataram a vantagem dessa técnica em comparação aos modelos singulares (simples). Ocorre, porém, que ainda há um leque de possibilidades para a intersecção entre modelos a serem testados para a previsão de tributos.

Diante disso, o objetivo deste trabalho é apresentar modelos alternativos para a previsão eficiente de tributos estaduais. Para tanto, são testados diferentes modelos estatísticos voltados à previsão, incluindo-se os tradicionalmente abordados, e modelos combinados. As estimativas foram feitas para a arrecadação do ICMS e IPVA do estado da Paraíba.

Com dados de 1997 a 2020, foram utilizados nove modelos univariados: o modelo de *Holt-Winters* com suavização exponencial por tratamento aditivo; o SARIMA (*Seasonal Autoregressive Integrated Moving Average*) com sazonalidade; o *X-13-ARIMA-SEATS*; o ETS (*Error Trend Seasonal*); o NNAR (*Neural Network Autoregression*); o TBATS (*Exponential Smoothing Method + Box-Cox Transformation + ARMA model for residuals + Trigonometric Seasonal*); o modelo STLM (*Seasonal and Trend decomposition using Loess*); o THETAM (método *Theta de Assimakopoulos e Nikopoulos, 2000*); e, por fim, o SNAIVE (*Seasonal Naive*).

Esses modelos foram alvos de quatro combinações distintas, dentre as quais as propostas em Figueiredo (2019), Hyndman e Athanasopoulos (2018) e Shaub e Ellis (2020). A escolha dessas combinações ocorreu em razão de terem demonstrado superioridade frente aos modelos singulares em testes realizados com séries temporais distintas. Portanto, a expectativa é a de que o uso das combinações proporcione maior precisão nas previsões e se revele como opção para no planejamento orçamentário dos governantes.

Após esta introdução, segue uma breve revisão da literatura recente voltada à previsão de tributos no Brasil. Em seguida, é descrita a metodologia aplicada ao estudo. Os resultados são discutidos no capítulo 4. Por fim, as considerações finais são lançadas.

2 REVISÃO DA LITERATURA

A literatura voltada à previsão de tributos no País tem buscado encontrar a metodologia que gere melhor precisão. As técnicas de previsão majoritariamente propostas têm adotado os modelos derivados da metodologia *Box-Jenkins*. Mais recentemente, modelos combinados passaram a ser testados e vêm demonstrando maior capacidade de previsão quando comparados aos modelos tradicionais univariados.

Fazendo uso da metodologia *Box-Jenkins*, Santos e Lima (2006) estimaram modelos de séries temporais visando a identificar a especificação mais adequada à previsão do ICMS para o estado de Minas Gerais. A partir dos critérios de comparação AIC (*Akaike Information Criterion*) e SIC (*Schwarz Information Criterion*), encontraram como melhor especificação o modelo SARIMA $(12, 1, 12)(1, 1, 1)$, que apresentou um erro absoluto percentual médio de 4,56%, para 12 meses previstos. O mesmo objetivo foi proposto em Pessoa, Coronel e Lima (2013), que estimaram modelos

ARIMA E ARFIMA, para previsão da arrecadação do ICMS em Minas Gerais, e concluíram pela vantagem do último, que atingiu um MAPE de 6,03%.

O uso da estimação por meio da metodologia ARIMA também foi adotado em Azevedo, Silva e Gatsio (2007), para previsão de ICMS em seis estados da Federação. O estudo pôde constatar a maior eficiência de suas estimações em comparação às realizadas pelos próprios estados: em média, as projeções realizadas pelos estados foram 11,3% menores do que os valores arrecadados em 2012 e 11,6% menores em 2013; ao passo que as projeções realizadas pelos modelos de séries temporais foram em média 0,9% maiores em 2012 e 2,2% menores em 2013.

Em Duarte, Souza e Girão (2014), foi utilizado o modelo de suavização exponencial de *Holt-Winters* nos padrões aditivo e multiplicativo, para prever a arrecadação de ICMS na Paraíba, definindo como período de análise os anos de 1997 a 2013. O trabalho verificou que os métodos puderam estimar previsões próximas ao valor real para 2013, calculando erro absoluto percentual médio de 3,10% (12 meses).

O uso de modelos híbridos passou a surgir apenas mais recentemente na literatura de referência. Mendonça e Medrano (2016) combinaram previsões por meio de três modelos: o modelo fatorial dinâmico (MFD), o *SARIMA* e o *Holt-Winters* para tributos federais; e constataram que o uso de combinações fez com que as previsões alcançassem melhores resultados.

No trabalho de Moço (2017), foram estimados modelos de previsão para o ICMS do estado do Rio de Janeiro, utilizando-se de modelos univariados (*SARIMA* e *Holt-Winters*), multivariados – VAR (*Vector Autoregressive*) ou VEC (*Vector Error Correction*) – ou a combinação de ambos. A partir dos resultados, foi possível concluir que a combinação do melhor modelo univariado (*SARIMA*) com o melhor multivariado (VEC) proporciona um ganho na precisão das previsões. O modelo combinado vencedor atingiu um MAPE de 6,22%, bastante inferior ao alcançado na previsão feita pela Secretaria da Fazenda do Estado, que foi de 20,62%.

3 METODOLOGIA

Para atingir o propósito de estabelecer comparações entre alternativas de previsão de arrecadação de tributos estaduais, foram coletados dados mensais para o ICMS e IPVA¹ da Paraíba referentes ao período de 1997 a 2020. Inicialmente os dados foram separados em dois momentos: o período de treino, abrangendo o tempo de janeiro de 1997 a setembro de 2018; e o período de validação, que segue de outubro de 2018 a dezembro de 2019, perfazendo 15 meses.

A escolha do mês de setembro como fim do período de treino não se deu de forma aleatória ou *ad hoc*, mas porque é quando se iniciam os debates para aprovação da LOA (Lei Orçamentária Anual), cujo prazo final para encaminhamento pelo executivo à assembleia legislativa deve ocorrer em até quatro meses antes do fim do exercício anterior. Sabe-se que a LOA deve conter, necessariamente, a previsão de receitas para o exercício do ano fiscal seguinte. Portanto, é fundamental que o poder executivo conte com instrumentos apropriados e capacidade de elaboração de previsões, o mais próximas do que será concretizado possível.

Espera-se que em condições “normais” o uso de combinações de modelos possa conferir maior eficácia na previsão de séries tributárias. Mas essa vantagem se mantém em situações de choques exógenos intensos, como o ocasionado pela insurgência da pandemia da Covid-19? Para responder a essa questão, o período de análise de previsão foi modificado: o período de treino passou a abranger o tempo de janeiro de 1997 a setembro de 2019; e o período de validação, de outubro de 2019 a dezembro de 2020. Por último, ainda foram feitas previsões para fora da amostra, compreendendo todos os meses do ano de 2021.

1 A opção por esses dois tributos se deu em razão de que ambos representam cerca de 97% da arrecadação própria do estado – excluindo-se as transferências e participação em tributos federais. Os dados são da Secretaria de Estado da Fazenda da Paraíba.

Foram feitas previsões a partir dos dados de treino, que foram relacionados aos dados reais correspondentes ao período de treino, com o fim de avaliar quais métodos apresentaram melhores ajustes aos dados. Inicialmente foram feitas estimações por meio de nove modelos distintos individualizados, aqui denominados de modelos simples. Os métodos utilizados foram: o modelo de *Holt-Winters* com suavização exponencial por tratamento aditivo; o SARIMA (*Seasonal Autoregressive Integrated Moving Average*) com sazonalidade; o *X-13-ARIMA-SEATS*; o ETS (*Error Trend Seasonal*); o NNAR (*Neural Network Autoregression*); o TBATS (*Exponential Smoothing Method + Box-Cox Transformation + ARMA model for residuals + Trigonometric Seasonal*); o modelo STLM (*Seasonal and Trend decomposition using Loess*); o THETAM (método *Theta de Assimakopoulos e Nikopoulos, 2000*); e, por fim, o SNAIVE (*Seasonal Naive*).

Em seguida, foram realizadas previsões por meio de modelos híbridos, ou seja, fazendo a junção de mais de um método de estimação. Foram gerados quatro modelos híbridos alternativos, aqui chamados de modelos combinados, sendo três com uso da média das estimativas e o último por meio de combinação com pesos diferenciados visando à minimização de erros.

Não obstante, em que pesem outros esquemas de ponderação visando à otimização possam ser derivados para combinar previsões, estudos têm demonstrado que o uso da média simples de previsões geralmente funciona bem em relação a combinações mais complexas (Armstrong, 2001; Clemen, 1989).

O primeiro modelo híbrido (C1) foi formado pela combinação dos modelos *Holt-Winters* (HW) e SARIMA:

$$C1 = a_1HW + a_2SARIMA \quad (1)$$

Em que C1 representa a combinação 1 e $a_1 = a_2 = 0,50$.

O segundo modelo híbrido (C2) foi formado a partir da combinação de cinco métodos de estimação, conferindo pesos iguais a cada um, seguindo o que foi apresentado em Figueiredo (2019):

$$C2 = \beta_1ETS + \beta_2SARIMA + \beta_3STL-ETS + \beta_4NNAR + \beta_5TBATS \quad (2)$$

em que $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0,166$, são os pesos de cada método na combinação.

O terceiro modelo combinado (C3) acresce ao C2 o método Thetan, mantendo pesos iguais entre eles:

$$C3 = \Phi_1ETS + \Phi_2SARIMA + \Phi_3STL-ETS + \Phi_4NNAR + \Phi_5TBATS + \Phi_6THETAM \quad (3)$$

em que $\Phi_1 = \Phi_2 = \Phi_3 = \Phi_4 = \Phi_5 = \Phi_6 = 0,143$, são os pesos de cada método na combinação.

Por último, a partir de Hyndman e Athanasopoulos (2018) e Shaub e Ellis (2020), foi gerado um último modelo combinado (C4) com pesos diferenciados a partir da minimização de erros da amostra, no que se refere ao erro absoluto médio. Nesse caso, os pesos conferidos foram diferentes para as séries ICMS e IPVA:

$$C4(a) = \gamma_1 ETS + \gamma_2 SARIMA + \gamma_3 STL-ETS + \gamma_4 NNAR + \gamma_5 TBATS + \gamma_6 THETAM \quad (4.a)$$

$$C4(b) = \tau_1 ETS + \tau_2 SARIMA + \tau_3 STL-ETS + \tau_4 NNAR + \tau_5 TBATS + \tau_6 THETAM \quad (4.b)$$

em que C4(a) é o modelo combinado para o ICMS com pesos diferenciados: $\gamma_1 = 0,18$; $\gamma_2 = 0,16$; $\gamma_3 = 0,19$; $\gamma_4 = 0,16$; $\gamma_5 = 0,09$; $\gamma_6 = 0,06$; $\gamma_7 = 0,14$. E C4(b) é o modelo combinado para o IPVA com pesos também diferenciados: $\tau_1 = 0,18$; $\tau_2 = 0,16$; $\tau_3 = 0,19$; $\tau_4 = 0,16$; $\tau_5 = 0,09$; $\tau_6 = 0,06$; $\tau_7 = 0,14$.

Após a estimação das previsões de todos os modelos propostos, seus resultados foram comparados por meio de medidas de desempenho RMSE (*Root Mean Squared Error*); MAE (*Mean Absolute Error*); MPE (*Mean Percentual Error*); e MAPE (*Mean Absolute Percentual Error*) calculados por:

$$RMSE = \sqrt{N^{-1} \sum_{t=1}^N e_t^2} \quad (5)$$

$$MAE = 1/n \sum_{t=1}^N |e_t| \quad (6)$$

$$MPE = [N^{-1} \sum_{t=1}^N e_t / y_t] \times 100 \quad (7)$$

$$MAPE = [1/n \sum_{t=1}^N |e_t| / |y_t|] \times 100 \quad (8)$$

Em que e_t é o erro de previsão calculado a partir da diferença entre o valor real Y_t e a previsão estimada F_t : $e_t = y_t - f_t$. Essas medidas têm sido empregadas com frequência em trabalhos que abordam previsão, dentre os quais Fan *et al.* (2020) e Mendonça e Medrano (2016).

As previsões geradas pelos métodos de combinação foram confrontadas com os valores realizados nos exercícios de 2019, 2020 e 2021 a fim de calcular os seus erros. Estes também foram calculados para a previsão oficial contida nas Leis Orçamentárias Anuais de 2019, 2020 e 2021 (Paraíba, 2019, 2020 e 2021). A ideia subjacente a esse exercício é a de verificar se há superioridade dos modelos combinados em relação à previsão oficial e, também, em relação aos modelos simples.

Segundo explica Wooldridge (2014), para realizar previsões é necessário utilizar critérios de escolha fora da amostra, como aqui adotado, já que a previsão é fundamentalmente um problema fora da amostra. Assim, a comparação fora da amostra faz uso da primeira parte (treino) para estimar os parâmetros do modelo, e da parte restante (validação) para avaliar a capacidade de previsão. Para avaliação, ele sugere o uso de RMSE e o MAE, sendo melhor o modelo que apresentar o valor menor dessas medidas para previsões fora da amostra. Assim, a partir da comparação das medidas de erros de previsão, é que serão indicados os melhores modelos para a previsão de tributos estaduais.

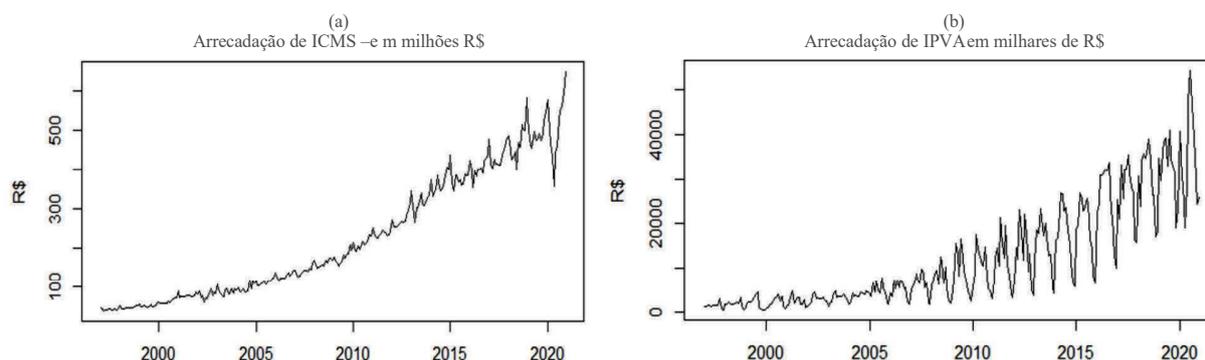
4 RESULTADOS E DISCUSSÕES

4.1 Evolução da série e teste de raiz unitária

A arrecadação de tributos da Paraíba vem crescendo ao longo do período da amostra. O ICMS saltou de um valor nominal de R\$ 491 milhões para cerca de R\$ 6,1 bilhões, entre 1997 e o final de 2020. Enquanto o IPVA passou de R\$ 16,371 milhões para cerca de R\$ 423,800 milhões, ambos em termos nominais.

A Figura 1 apresenta os gráficos com a evolução desses tributos, onde se percebe, nitidamente, a forte tendência de alta para o ICMS e o IPVA, principalmente após 2005, e a presença do componente de sazonalidade, sendo este mais visível para o ICMS. Outro movimento que se destaca é o da expressiva queda de arrecadação do ICMS nos primeiros meses de 2020, que resulta na forma de um vale profundo, em consequência dos efeitos adversos oriundos da pandemia da Covid-19, mas que passa por rápida recuperação em forma de “v”. Para o ICMS é possível notar um comportamento semelhante: há queda abrupta entre março e maio de 2020 – comportamento diferente do padrão da série – e franca recuperação em seguida. Também é possível notar um aumento na variância das séries, o que, em conjunto com as demais características, pode sinalizar a ausência de estacionariedade. Essa questão é particularmente importante para a modelagem de Box-Jenkins com uso de modelos autoregressivos de médias móveis (ARMA).

Figura 1 – Gráficos para as arrecadações de ICMS e IPVA na Paraíba de 1997/2020



Fonte: Secretaria de Estado da Fazenda da Paraíba. Elaboração própria.

Para checar a estacionariedade das séries, foram feitos os testes de raiz unitária KPSS apresentados em Kwiatkowski *et al.* (1992) e o teste Dickey-Fuller-GLS (ERS) sugerido por Elliott *et al.*, cujos resultados seguem no Quadro 1. Os testes foram realizados para dois períodos: o primeiro abrangendo a série até 2019; e o segundo com a série completa.

Quadro 1 - Testes ETS e KPSS para verificar presença de raiz unitária

Período 1997/2019								
Teste	ICMS				IPVA			
	ERS	KPSS	ERS	KPSS	ERS	KPSS	ERS	KPSS
Ordem de Int.	I(0)	I(0)	I(1)	I(1)	I(0)	I(0)	I(1)	I(1)
Estatística de teste	-0,71	1,06	-5,74	0,01	-5,53	0,62	-9,07	0,00
Período 1997/2020								
Teste	ERS	KPSS	ERS	KPSS	ERS	KPSS	ERS	KPSS
Ordem de Int.	I(0)	I(0)	I(1)	I(1)	I(0)	I(0)	I(1)	I(1)
Estatística de teste	-1,08	1,08	-8,91	0,02	-5,21	0,70	-8,91	0,02

Fonte: Elaboração própria a partir dos dados da pesquisa com uso do R-Studio.

Notas: Valor crítico para 5% do ETS é -2,89 e para o KPSS é 0,14.

Conforme se observa nos resultados dos testes, as características das séries para os dois períodos são iguais. A série do ICMS é integrada de ordem 1, ou seja, possui raiz unitária, que é dissipada após a diferenciação de primeira ordem. Os resultados para a série IPVA foram conflitantes, pois enquanto o método ERS indica a ausência de raiz unitária, o teste KPSS faz o contrário. Todavia,

a observação da Função de Autocorrelação² das séries aponta para um processo não-estacionário, de modo que a série deverá ser integrada para a estimação do modelo ARMA.

4.2 Comparativo das medidas de desempenho dos modelos de previsão

A fim de identificar o melhor método de estimação para a previsão das séries de arrecadação do estado da Paraíba, foram utilizadas quatro medidas de desempenho relacionadas aos erros de previsão de todos os modelos estimados, cujos valores encontram-se discriminados na Tabela 1 para o caso do IPVA de 2019.

A partir dos valores calculados para as medidas de desempenho de cada uma das estimações, é possível perceber que os modelos combinados apresentaram previsões, de um modo geral, com melhor ajuste aos valores reais, em comparação aos modelos simples. Destaque deve ser dado ao modelo “Combinação 3”, que quantificou uma estatística MPE que importa em um erro médio de previsão de 0,05%. Por outro lado, o modelo simples SARIMA obteve bons resultados, com métricas de precisão para as medidas RMSE, MAE e MAPE, que superam em qualidade inclusive as calculadas para os modelos combinados, assim como o modelo *Holt-Winters* com suavização exponencial sazonal.

Tabela 1 – Medidas de desempenho para os erros de projeção dos modelos para o IPVA da PB – 2019

Modelos/Med.	RMSE	MAE	MPE	MAPE	(M1)	(M2)	(M3)	(M4)	(M5)	(M6)	(M7)	(M8)	(M9)
Modelos Simples	(R\$)	(R\$)	(%)	(%)	Pesos								
Holt-Winters (M1)	2632,00	2072,00	-4,51	7,52	1								
Sarima (M2) (0,1,1)(0,1,1)	2428,07	2044,49	-1,91	7,31		1							
X-13ARIMA-SEATS (M3)	4317,64	2805,90	-21,96	43,16			1						
ETS (M4)	3394,47	2567,61	-1,21	8,68				1					
STL-ETS (M5)	6804,57	6139,34	1,30	21,34					1				
NNAR (M6)	5090,16	3731,88	-11,28	15,84						1			
TBATS (M7)	5531,49	4801,94	-7,38	15,79							1		
SNAIVE(M8)	3903,8	3262,23	9,88	11,28								1	
TETHAM (M9)	5715,42	4352,81	13,49	15,97									1
Combinações													
C1	2505,47	2057,75	-3,21	7,41	0,5	0,5							
C2	2986,69	2420,18	-1,76	8,45		0,166		0,166	0,166	0,166	0,166	0,166	
C3	2679,14	2194,94	0,05	7,85		0,143		0,143	0,143	0,143	0,143	0,143	0,143
C4	2572,37	2154,43	-1,93	7,93		0,11		0,1	0,13	0,41	0,1	0,08	0,06

Fonte: Elaboração própria a partir das estimativas pelo pacote R-Studio.

Assim, é possível dizer que a utilização de modelos combinados para previsão do IPVA pode ser uma boa estratégia, pois tende a reduzir os erros de previsão quando comparada a alternativas de estimação via modelos simples. Contudo, a alternativa pela escolha de um modelo mais parcimonioso deve ser dirigida, para o caso do IPVA, para o *SARIMA*, que mostrou ter uma eficiência próxima ou superior aos combinados.

A Tabela 2 expõe as medidas de desempenho para as previsões realizadas sobre o ICMS. Os resultados, nesse caso, apontam de forma mais explícita para a superioridade dos métodos de combinação sobre o uso de modelos simples. Dentre os modelos híbridos aquele que obteve maior precisão foi a “Combinação 4”, que vence todos os demais para todos os parâmetros, com exceção

2 Apêndice A.

da medida RMSE do modelo *TETHAM*. Dentre os modelos simples, aqueles que foram mais eficientes foram o *ETS*, o *TETHAM* e o *SARIMA*, com destaque para o primeiro.

Tabela 2 – Medidas de desempenho para os erros de projeção dos modelos para o ICMS da PB – 2019

Modelos/Med.	RMSE	MAE	MPE	MAPE	(M1)	(M2)	(M3)	(M4)	(M5)	(M6)	(M7)	(M8)	(M9)
Modelos Simples	(R\$)	(R\$)	(%)	(%)	Pesos								
Holt-Winters (M1)	28,70	20,60	-2,10	4,09	1								
Sarima (M2) (0,1,1)(0,1,1)	26,93	19,36	-0,03	3,80		1							
X-13ARIMA-SEATS (M3)	42,35	39,30	-6,36	7,89			1						
ETS (M4)	24,92	17,38	-0,50	3,40				1					
STL-ETS (M5)	41,37	37,74	-5,70	7,55					1				
NNAR (M6)	43,84	34,24	-0,79	6,75						1			
TBATS (M7)	28,6	22,50	-1,96	4,49							1		
SNAIVE(M8)	52,65	47,91	8,36	9,44								1	
TETHAM (M9)	23,41	18,82	-1,31	3,72									1
Combinações													
C1	26,95	18,41	-1,06	3,63	0,5	0,5							
C2	26,07	17,61	-0,10	3,43		0,166		0,166	0,166	0,166	0,166	0,166	
C3	25,53	16,65	0,49	3,22		0,143		0,143	0,143	0,143	0,143	0,143	0,143
C4	24,47	15,6	0,01	3,03		0,16		0,18	0,19	0,16	0,09	0,06	0,14

Fonte: Elaboração própria a partir das estimativas pelo pacote R-Studio.

Em suma, as medidas de desempenho para os diferentes modelos mostraram que, para situações de “normalidade” como foi o ano de 2019, a realização de combinações pode resultar em maior eficiência na geração de previsões e reduzir os seus erros. Mas essa vantagem se mantém quando a série passa por um choque agudo como o que houve em 2020 em razão da pandemia da Covid-19?

Tabela 3 – Medidas de desempenho para os erros de projeção dos modelos para o IPVA da PB – 2020

Modelos/Med.	RMSE	MAE	MPE	MAPE	(M1)	(M2)	(M3)	(M4)	(M5)	(M6)	(M7)	(M8)	(M9)
Modelos Simples	(R\$)	(R\$)	(%)	(%)	Pesos								
Holt-Winters (M1)	8712,12	6462,30	-7,30	21,90	1								
Sarima (M2) (0,1,1)(0,1,1)	8768,41	6733,01	-5,48	22,50		1							
X-13ARIMA-SEATS (M3)	4510,95	2966,16	-22,86	44,01			1						
ETS (M4)	9160,39	6846,69	-6,47	24,06				1					
STL-ETS (M5)	13481,50	9252,52	-17,13	36,09					1				
NNAR (M6)	8698,95	6681,65	-4,75	20,98						1			
TBATS (M7)	10358,67	7525,63	-10,20	27,56							1		
SNAIVE(M8)	9917,77	8682,82	5,77	28,04								1	
TETHAM (M9)	11063,98	9464,67	5,53	32,17									1
Combinações													
C1	8735,18	6597,66	-6,39	22,20	0,50	0,50							
C2	9316,58	7127,59	-6,38	24,89		0,17		0,17	0,17	0,17	0,17	0,17	
C3	9100,26	7248,63	-3,31	24,30		0,14		0,14	0,14	0,14	0,14	0,14	0,14
C4	8960,65	7007,05	-4,66	23,54		0,16		0,15	0,08	0,15	0,19	0,14	0,12

Fonte: Elaboração própria a partir das estimativas pelo pacote R-Studio.

A Tabela 3 apresenta as medidas de desempenho para previsões do IPVA considerando o período de outubro de 2019 a dezembro de 2020 e ajuda a responder a questão. O primeiro ponto que se destaca é o de que todas as medidas de erro para todos os modelos se mostraram consideravelmente maiores em comparação à previsão realizada no período anterior. Esse comportamento era esperado, pois seria completamente improvável que qualquer modelo econométrico pudesse ane-ter um choque exógeno envolvendo uma pandemia de tamanhas proporções, com repercussões profundas nas economias de todo o mundo. Outro ponto que se realça é o de que há uma ligeira queda na franca superioridade dos modelos combinados para o conjunto dos indicadores de desempenho frente aos simplificados.

Considerando esses indicadores, sugere-se que as opções pelo uso do *SARIMA*, *Holt-Winters* ou *NNAR* podem conferir resultados tão bons ou melhores que os modelos combinados. Isso não invalida a avaliação de que o uso de combinações pode ser uma importante alternativa na confecção de previsões de séries tributárias, mesmo porque, para ambos os períodos, a combinação apresentou, no geral, superioridade no desempenho.

A análise dos indicadores de desempenho para as previsões do ICMS, cujos valores constam na Tabela 4, não difere muito da que foi feita para o IPVA. É possível perceber que alguns modelos simples podem gerar previsões com eficiência próxima aos modelos combinados, com destaque para o *X-13-ARIMA-SEATS*, *ETS*, *TBATS* e, novamente, o *Holt-Winters*.

Importante ainda registrar que o modelo *X-13-ARIMA-SEATS* apresentou, com folga, o pior desempenho na previsão para o período até 2019. O fato de que no período seguinte, que sofreu notáveis choques externos profundos e inesperados na série, o modelo tenha conseguido apresentar bom ajuste na previsão, pode ser um indício de que houve “premiação” ao erro. Ou seja, os erros de previsão podem ter sido em parte compensados ou “corrigidos” pelo comportamento exótico decorrente do choque.

Em síntese, comparando os indicadores de previsão dos diferentes modelos estimados para ambos os períodos e séries, verifica-se que, embora modelos singulares possam ser capazes de oferecer previsões tributárias com razoável robustez, o uso dos modelos combinados aqui propostos ajuda a imprimir, no geral, maior precisão independentemente da conjuntura.

Tabela 4 – Medidas de desempenho para os erros de projeção dos modelos para o ICMS da PB – 2020

Modelos/Med.	RMSE	MAE	MPE	MAPE	(M1)	(M2)	(M3)	(M4)	(M5)	(M6)	(M7)	(M8)	(M9)
Modelos Simples	(R\$)	(R\$)	(%)	(%)	Pesos								
Holt-Winters (M1)	56,49	45,10	-0,14	9,32	1								
Sarima (M2) (0,1,1)(0,1,1)	62,51	48,31	-1,39	10,07		1							
X-13ARIMA-SEATS (M3)	10,92	6,94	-0,21	3,33			1						
ETS (M4)	57,06	43,99	-0,34	9,02				1					
STL-ETS (M5)	62,73	39,42	-8,05	9,15					1				
NNAR (M6)	63,61	48,93	-3,22	10,35						1			
TBATS (M7)	56,60	42,78	-3,07	9,15							1		
SNAIVE(M8)	58,90	48,35	0,65	9,85								1	
TETHAM (M9)	58,92	46,27	1,25	9,28									1
Combinações													
C1	59,25	45,97	-0,77	9,54	0,5	0,5							
C2	57,28	42,75	-2,57	9,11		0,166		0,166	0,166	0,166	0,166	0,166	
C3	58,01	44,44	-1,21	9,26		0,143		0,143	0,143	0,143	0,143	0,143	0,143
C4	56,99	43,76	-1,56	9,16		0,18		0,18	0,14	0,15	0,2	0,09	0,06

Fonte: Elaboração própria a partir das estimativas pelo pacote R-Studio.

4.3 Comparativo entre previsões oficiais e as geradas pelos modelos combinados

Foram comparadas as previsões geradas pelos modelos combinados aqui propostos com as apresentadas pelo executivo estadual para o exercício de 2019, com o fim de conferir se, de fato, as combinações são mais eficazes do que a metodologia de previsão oficial.

A Tabela 5 expõe os valores previstos para o IPVA e os erros nominal e percentual para cada uma das técnicas empregadas. O valor arrecadado no ano de 2019 com IPVA na Paraíba foi de R\$ 393.118.220,00 (trezentos e noventa e três milhões, cento e dezoito mil, duzentos e vinte reais), enquanto a previsão oficial calculou arrecadação inferior em R\$ 43.115.047,65 (quarenta e três milhões, cento e quinze mil, quarenta e sete reais e sessenta e cinco centavos). Assim, a receita projetada com a fonte de IPVA para o ano de 2019 na Paraíba foi subestimada em quase 11%.

Tabela 5 – Erros de previsão para o IPVA para o exercício de 2019

Modelos	Valores R\$	Erros de previsão R\$	Erros de previsão %	Erros de previsão % (Absoluto)	Rank
REALIZADO	393.118.220,65	-	-	-	-
C3	390.728.870,00	- 2.389.350,65	-0,61%	0,61%	1
SARIMA	395.953.740,00	2.835.519,35	0,72%	0,72%	2
NNAR	399.330.020,00	6.211.799,35	1,58%	1,58%	3
C1	400.461.810,00	7.343.589,35	1,87%	1,87%	4
C2	401.413.200,00	8.294.979,35	2,11%	2,11%	5
X13-ARIMA-SEATS	402.892.500,00	9.774.279,35	2,49%	2,49%	6
ETS	404.787.140,00	11.668.919,35	2,97%	2,97%	7
HW	404.969.850,00	11.851.629,35	3,01%	3,01%	8
C4	405.058.850,00	11.940.629,35	3,04%	3,04%	9
STL-ETS	421.845.440,00	28.727.219,35	7,31%	7,31%	10
TBATS	430.344.550,00	37.226.329,35	9,47%	9,47%	11
SNAIVE	353.564.880,00	- 39.553.340,65	-10,06%	10,06%	12
Previsão Oficial (LOA)	350.003.173,00	- 43.115.047,65	-10,97%	10,97%	13
THETAM	348.590.960,00	- 44.527.260,65	-11,33%	11,33%	14

Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Por outro lado, os valores calculados a partir dos modelos de combinação tiveram resultados francamente superiores. O uso das combinações permitiu encontrar previsões bastante aproximadas dos valores realizados, cometendo erros de previsão situados entre -0,61% (C3) e 3,04% (C4). A eficácia das combinações fica ainda mais patenteada quando se observa que dentre as 14 diferentes previsões, três dos modelos combinados ocupam as cinco previsões mais aproximadas do valor realizado, sendo a combinação 3 a previsão mais aproximada frente a todas as demais.

Dentre os modelos simples, os que se destacaram foram o *SARIMA*, com erro de previsão de apenas 0,72% e o *NNAR* com erro de 1,58%, números que possibilitaram aos referidos modelos ocupar a segunda e terceira melhores previsões para o tributo.

A previsão oficial foi superada em termos de desempenho por 12 das 13 distintas opções de previsão testadas, incorrendo em menor erro apenas que o modelo *Tetham*, que reportou erro de 11,33%, que é um patamar substantivamente elevado.

A Tabela 6 apresenta o mesmo exercício, mas aplicado à arrecadação do ICMS. A constatação é similar à anterior tanto os modelos combinados como a maior parte dos singulares conseguiram captar melhor a trajetória da série, cometendo erros expressivamente inferiores ao da previsão oficial.

A previsão oficial subestimou a arrecadação em R\$ 318.888.258,89 (trezentos e dezoito milhões, oitocentos e oitenta e oito mil, duzentos e cinquenta e oito reais e oitenta e nove centavos), ou seja, projetou um valor 5,40% abaixo do realizado. Assim, embora a previsão oficial para o ICMS tenha sido mais bem-sucedida do que a feita para o IPVA, foi superada em termos de desempenho por 10 dos 13 modelos alternativos testados. Isso sugere que há nítida deficiência na metodologia empregada pelo executivo estadual em seu planejamento tributário e reforça a necessidade de apontar alternativas mais eficientes para a gestão pública.

Por outro lado, os modelos combinados atingiram erros que variaram de um mínimo de 1,04% ao máximo de 2,34%, ganhando, novamente, a Combinação 3, que obteve o menor erro dentre todas as previsões. Com isso, três das quatro melhores previsões foram geradas por meio de modelos híbridos, o que indica que as combinações se mostram como opções vantajosas para o planejamento tributário estatal.

Tabela 6 – Erros de previsão para o ICMS para o exercício de 2019

Modelos	Valores RS	Erros de previsão R\$	Erros de previsão %	Erros de previsão % (Absoluto)	Rank
REALIZADO	5.904.371.745,89	-	-	-	-
C3	5.965.942.900,00	61.571.154,11	1,04%	1,04%	1
SARIMA	5.965.959.700,00	61.587.954,11	1,04%	1,04%	2
C4	5.975.114.900,00	70.743.154,11	1,20%	1,20%	3
C2	5.995.285.400,00	90.913.654,11	1,54%	1,54%	4
NNAR	6.023.657.800,00	119.286.054,11	2,02%	2,02%	5
ETS	6.026.363.600,00	121.991.854,11	2,07%	2,07%	6
C1	6.042.610.000,00	138.238.254,11	2,34%	2,34%	7
THETAM	6.042.985.400,00	138.613.654,11	2,35%	2,35%	8
HW	6.119.259.900,00	214.888.154,11	3,64%	3,64%	9
TBATS	6.132.713.700,00	228.341.954,11	3,87%	3,87%	10
Previsão Oficial (LOA)	5.585.483.487,00	- 318.888.258,89	-5,40%	5,40%	11
SNAIVE	5.457.661.200,00	- 446.710.545,89	-7,57%	7,57%	12
STL-ETS	6.382.005.100,00	477.633.354,11	8,09%	8,09%	13
X13-ARIMA-SEATS	6.417.311.700,00	512.939.954,11	8,69%	8,69%	14

Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Dentre os modelos simples, novamente o *SARIMA* e o *NNAR* foram os que geraram melhores previsões, cujos erros foram de 1,04% e 2,02%, em sequência. Esses resultados colocaram os modelos na segunda e quinta melhores classificações dentre todos os testados, respectivamente. Importante destacar que o modelo *SARIMA* ficou posicionado como gerador da segunda melhor previsão para os dois tributos no período, com erros percentuais relativamente baixos, portanto, pode ser indicado como a melhor opção dentre os modelos simples testados.

A partir da análise dos resultados, pode-se afirmar que houve falhas expressivas na metodologia empregada pelo executivo estadual para estimar as receitas tributárias. Ao mesmo tempo, é possível conferir maior precisão nas previsões com uso de métodos testados, com destaque para modelos combinados e os singulares *SARIMA* e *NNAR*.

As previsões até aqui apresentadas foram feitas, porém, para um período de relativa normalidade e um exercício. O ano de 2020, sabidamente, foi de muita instabilidade com reflexos sobre a atividade econômica e receitas tributárias, em função do choque inesperado resultante da pandemia da Covid-19. Será então que os resultados se repetem quando aplicados a 2020?

A Tabela 7 exhibe os valores previstos para o IPVA em 2020. O valor arrecadado para o ano com IPVA na Paraíba foi de R\$ 423.803.854,39 (quatrocentos e vinte e três milhões, oitocentos e três mil, oitocentos e cinquenta e quatro reais e trinta e nove centavos).

A previsão oficial subestimou a arrecadação em espantosos R\$ 269.732.060,39 (duzentos e sessenta e nove milhões, setecentos e trinta e dois mil, sessenta reais e trinta e nove centavos). Ou seja, foi cometido um erro superior aos 63% ao projetar as receitas com o IPVA. A falha em projetar as receitas do IPVA é flagrante e inexplicável, pois o valor calculado chega a ser cerca de 60% inferior ao realizado no ano anterior.

Ressalta-se que o IPVA tem nítida tendência ascendente e na altura da elaboração da LOA, não havia nenhum sinal de pandemia ou grave crise econômica que pudesse indicar uma depressão tão aguda nas receitas.

Em contraposição, os modelos combinados conseguiram realizar boas previsões para o IPVA com erros absolutos máximos de 2,38% e mínimo de apenas 0,49%, gerando duas das quatro previsões mais eficientes dentre as testadas.

Tabela 7 – Erros de previsão para o IPVA para o exercício de 2020

Modelos	Valores R\$	Erros de previsão R\$	Erros de previsão %	Erros de previsão % (Absoluto)	Rank
REALIZADO	423.803.854,39	-	-	-	-
SARIMA	422.951.880,00	- 851.974,39	-0,20%	0,20%	1
C1	425.887.480,00	2.083.625,61	0,49%	0,49%	2
C4	419.950.810,00	- 3.853.044,39	-0,91%	0,91%	3
HW	428.823.110,00	5.019.255,61	1,18%	1,18%	4
ETS	432.420.530,00	8.616.675,61	2,03%	2,03%	5
C2	432.554.940,00	8.751.085,61	2,06%	2,06%	6
C3	413.696.490,00	- 10.107.364,39	-2,38%	2,38%	7
NNAR	410.302.080,00	- 13.501.774,39	-3,19%	3,19%	8
TBATS	448.641.800,00	24.837.945,61	5,86%	5,86%	9
THETAM	387.269.230,00	- 36.534.624,39	-8,62%	8,62%	10
SNAIVE	381.162.330,00	- 42.641.524,39	-10,06%	10,06%	11
STL-ETS	508.424.990,00	84.621.135,61	19,97%	19,97%	12
X13-ARIMA-SEATS	322.917.940,00	- 100.885.914,39	-23,80%	23,80%	13
Previsão Oficial (LOA)	154.071.794,00	- 269.732.060,39	-63,65%	63,65%	14

Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Para os modelos unitários, destacam-se positivamente o *SARIMA* e o *HW*, com erros de -0,20% e 1,18%, que lhes conferem a primeira e quarta melhores classificações. Por outro lado, os modelos *STL-ETS* e *X13-ARIMA-SEATS* entregaram resultados muito ruins, com erros de projeção de 19,97% e -23,80%, em sequência.

A Tabela 8 mostra que o executivo estadual não subestimou apenas a arrecadação para o IPVA, mas também para o ICMS, cujo erro atingiu quase elevados 35%. Nesse caso, também foram projetados, estranhamente, valores inferiores ao ano anterior em cerca de 40%. Ocorre que a arrecadação do ICMS sofreu mais intensamente com a pandemia do que a de IPVA, tendo, de fato, reduzido seu valor em comparação ao ano anterior. Contudo, novamente, a previsão oficial cometeu um erro muito acima dos padrões, atingindo um percentual de quase 35%.

Já os modelos combinados conseguiram apresentar resultados amplamente melhor ajustados do que a previsão oficial, com erros percentuais situando-se na faixa entre 11,05% a 12,13%. Percebe-se que esses erros foram bem maiores que os encontrados na previsão do ano anterior, o que se justifica exatamente pelo forte impacto causado na economia pela pandemia. Apesar disso, o uso de previsões por modelos combinados se mostrou explicitamente mais acertado do que o método que vem sendo utilizado pelo executivo estadual, haja vista as margens de erro aqui demonstradas.

Contudo, nesse caso, os modelos simples preencheram as primeiras quatro melhores classificações de desempenho. As quatro previsões mais ajustadas foram geradas pelos modelos *SNAIVE*,

THETAN, *HW* e *ETS*, que obtiveram erros de 8,48%, 8,92%, 10,47% e 10,70%, respectivamente. Nota-se que esses modelos não foram bem-posicionados para a previsão do tributo para o ano pré-pandêmico de 2019, o que pode sugerir uma premiação ao erro. Ademais, é imperioso ressaltar que todas as previsões para o ICMS de 2020 pioraram o desempenho, o que é natural dado o forte efeito do choque gerado pela pandemia.

Tabela 8 – Erros de previsão para o ICMS para o exercício de 2020

Modelos	Valores R\$	Erros de previsão R\$	Erros de previsão %	Erros de previsão % (Absoluto)	Rank
REALIZADO	5.459.667.586,40	-	-	-	-
SNAIVE	5.922.459.800,00	462.792.213,60	8,48%	8,48%	1
THETAM	5.946.825.200,00	487.157.613,60	8,92%	8,92%	2
HW	6.031.143.600,00	571.476.013,60	10,47%	10,47%	3
ETS	6.044.017.700,00	584.350.113,60	10,70%	10,70%	4
C1	6.062.924.700,00	603.257.113,60	11,05%	11,05%	5
C3	6.091.845.400,00	632.177.813,60	11,58%	11,58%	6
SARIMA	6.094.705.600,00	635.038.013,60	11,63%	11,63%	7
C4	6.122.180.600,00	662.513.013,60	12,13%	12,13%	8
C2	6.182.195.500,00	722.527.913,60	13,23%	13,23%	9
NNAR	6.226.827.200,00	767.159.613,60	14,05%	14,05%	10
TBATS	6.239.147.500,00	779.479.913,60	14,28%	14,28%	11
X13-ARIMA-SEATS	6.554.838.200,00	1.095.170.613,60	20,06%	20,06%	12
STL-ETS	6.581.780.000,00	1.122.112.413,60	20,55%	20,55%	13
Previsão Oficial (LOA)	3.576.162.468,00	- 1.883.505.118,40	-34,50%	34,50%	14

Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Confrontando-se as capacidades de previsão demonstradas para as receitas tributárias dos anos de 2019 e 2020, nota-se relevante alteração na ordem de classificação e poder preditivo dos modelos, sendo essa mudança mais intensa para o ICMS. Esse comportamento é justificável, como já delineado, em face dos efeitos adversos sobre a atividade econômica exercida pela pandemia da Covid-19. Esses efeitos foram menos impactantes sobre o IPVA, cuja arrecadação se manteve crescente, o que pode ser explicado pela natureza de sua base de incidência tributária.

Tais resultados lançam dúvidas sobre a melhor estratégia de previsão a ser adotada pelos executivos estaduais, ao mesmo tempo em que ratificam a necessidade de se buscar aperfeiçoar o método utilizado, ao menos para o caso em específico. Mas, em razão das discrepâncias quanto às eficiências preditivas dos modelos, impõe-se que se extrapolem os testes para um período à frente.

Segue-se assim a repetição da experiência, considerando o período de 2021³. Naquele ano, admite-se que os efeitos pandêmicos sobre a atividade econômica ainda foram bastante relevantes, particularmente em seu primeiro semestre, o que exige dos modelos de previsão forte “capacidade de aprendizado”. A Tabela 9 exhibe os valores previstos para o IPVA em 2021. O valor arrecadado para o ano com IPVA na Paraíba foi de R\$ 468.721.000,00 (quatrocentos e sessenta e oito milhões, setecentos e vinte e um mil reais).

A previsão oficial subestimou a arrecadação em espantosos R\$ 269.732.060,39 (duzentos e sessenta e nove milhões, setecentos e trinta e dois mil, sessenta reais e trinta e nove centavos). A falha desmesurada em projetar as receitas do IPVA foi repetida em 2021, cujo valor projetado, inclusive, é mais do que 63% inferior ao realizado no ano anterior.

3 Nesse caso, a previsão é feita fora da amostra e para o período de 12 meses.

Os modelos combinados proporcionaram, mais uma vez, previsões muito mais eficientes do que a oficial, com destaque para C2 e C1, que registraram menores erros de 0,33% e 2,09%, ficando classificadas na segunda e quinta melhores previsões.

Para os modelos simples têm-se como destaques o *ETS*, o *X13-ARIMA-SEATS* e o *SARIMA*, que ficaram posicionados na primeira, terceira e quarta colocação dentre as melhores previsões com erros de -0,24%, 1,51% e -1,93%.

Tabela 9 – Erros de previsão para o IPVA para o exercício de 2021

Modelos	Valores R\$	Erros de previsão R\$	Erros de previsão %	Erros de previsão % (Absoluto)	Rank
REALIZADO	468.721.000,00	-	-	-	-
ETS	467.584.060,00	- 1.136.940,00	-0,24%	0,24%	1
C2	467.183.750,00	- 1.537.250,00	-0,33%	0,33%	2
X13-ARIMA-SEATS	475.776.770,00	7.055.770,00	1,51%	1,51%	3
SARIMA	459.687.710,00	- 9.033.290,00	-1,93%	1,93%	4
C1	458.945.070,00	- 9.775.930,00	-2,09%	2,09%	5
HW	458.202.420,00	- 10.518.580,00	-2,24%	2,24%	6
C3	481.514.140,00	12.793.140,00	2,73%	2,73%	7
TBATS	495.813.300,00	27.092.300,00	5,78%	5,78%	8
C4	427.605.420,00	- 41.115.580,00	-8,77%	8,77%	9
SNAIVE	423.803.850,00	- 44.917.150,00	-9,58%	9,58%	10
NNAR	376.278.280,00	- 92.442.720,00	-19,72%	19,72%	11
STL-ETS	579.935.240,00	111.214.240,00	23,73%	23,73%	12
THETAM	702.396.020,00	233.675.020,00	49,85%	49,85%	13
Previsão Oficial (LOA)	171.299.211,00	- 297.421.789,00	-63,45%	63,45%	14

Elaboração própria a partir das estimações com uso do R-Studio.

Por fim, a Tabela 10 mostra que o executivo estadual também subestimou fortemente a arrecadação para o ICMS de 2021, cujo erro alcançou a cifra de quase R\$ 4 bilhões ou o equivalente a um erro superior a 52%. Tal falha de projeção deixa mais uma vez patenteado que há sérios problemas na metodologia adotada pelo executivo do estado da Paraíba para previsão de tributos, principalmente naquela que foi utilizada para os anos de 2020 e 2021. Com isso, todos os modelos de previsão estimados superaram com larga folga a eficiência da previsão oficial.

Ressalta-se que, à exceção do modelo *SNAIVE*, todas as previsões testadas melhoraram o desempenho em comparação às realizadas para o ano anterior, o que sugere que conseguiram recuperar parte das perdas de eficácia decorrentes da quebra da série ocorrida em 2020.

Tabela 10 – Erros de previsão para o ICMS para o exercício de 2021

Modelos	Valores R\$	Erros de previsão R\$	Erros de previsão %	Erros de previsão % (Absoluto)	Rank
REALIZADO	7.495.185.000,00	-	-	-	-
X13-ARIMA-SEATS	7.457.493.600,00	- 37.691.400,00	-0,50%	0,50%	1
SARIMA	7.442.864.800,00	- 52.320.200,00	-0,70%	0,70%	2
C1	7.401.939.600,00	- 93.245.400,00	-1,24%	1,24%	3
HW	7.361.014.600,00	- 134.170.400,00	-1,79%	1,79%	4
TBATS	7.326.797.700,00	- 168.387.300,00	-2,25%	2,25%	5
THETAM	7.310.375.600,00	-184.809.400,00	-2,47%	2,47%	6
ETS	7.273.254.300,00	- 221.930.700,00	-2,96%	2,96%	7
STL-ETS	7.258.608.100,00	- 236.576.900,00	-3,16%	3,16%	8
C3	7.011.906.000,00	- 483.279.000,00	-6,45%	6,45%	9
C2	6.984.514.900,00	- 510.670.100,00	-6,81%	6,81%	10
C4	6.932.024.500,00	- 563.160.500,00	-7,51%	7,51%	11
NNAR	6.497.484.100,00	- 997.700.900,00	-13,31%	13,31%	12
SNAIVE	6.108.081.800,00	- 1.387.103.200,00	-18,51%	18,51%	13
Previsão Oficial (LOA)	3.572.045.842,00	- 3.923.139.158,00	-52,34%	52,34%	14

Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Os resultados trazidos para o ano de 2021 não permitem afiançar de forma peremptória a melhor técnica de previsão a ser utilizada para os tributos estaduais em discussão. Não obstante, não houve qualquer coincidência entre os melhores modelos durante os três anos abrangidos pelo estudo. Por outro lado, confirma-se uma vez mais a exigência pela busca do aprimoramento dos métodos de previsão de tributos adotados pelos gestores estatais, em especial para o estado da Paraíba.

Considerando os três períodos, depreende-se que o uso de modelos combinados demonstrou boa eficácia, especialmente para o ano pré-pandêmico de 2019. Nesse período, o modelo C3 venceu todos os demais para ambos os tributos, enquanto outras combinações também atingiram boas classificações. Para os exercícios seguintes, embora sempre um ou mais modelos híbridos tenham ficado entre os cinco que apresentaram menores erros de previsão, há disparidade entre os vencedores, sendo que todos foram modelos singulares.

Diante da ausência de unicidade para as melhores previsões entre os períodos e tributos, foram calculadas as médias dos erros absolutos de previsão de todos os modelos testados para cada um dos exercícios. Os resultados constam na Tabela 11 e indicam que, considerando a média para os três exercícios, o modelo *SARIMA* foi capaz de gerar as melhores previsões tanto para o IPVA como para o ICMS, com erros médios de 0,95% e 4,46%, respectivamente. No caso dos modelos híbridos, aquele que obteve maior êxito foi o modelo *CI*, que pontuou como a segunda e terceira melhor previsão média para o IPVA e o ICMS, respectivamente, com erros médios de 1,48% e 4,88%. O erro médio das previsões oficiais alcançou patamares muito acima de qualquer uma das opções de modelos testados, sinalizando nitidamente falha no processo de construção do orçamento pelo executivo estadual, o que pode vir a comprometer seriamente o planejamento e execução orçamentária e gerar, ao fim, possíveis reflexos danosos sobre a prestação do serviço público à população.

Por fim, considerando que no período abrangido pelas previsões houve uma quebra intensa decorrente da pandemia da Covid-19 a afetar o desempenho da economia e as arrecadações tributárias, as previsões geradas pelos modelos testados conseguiram, de um modo geral, captar de forma relativamente bem aproximada o comportamento das séries, com destaque para o IPVA. Nesse caso, três das cinco melhores previsões foram obtidas por modelos híbridos, C2 e C3 ocupando a terceira e quinta melhores posições com erros médios de apenas 1,50% e 1,91%, o que confirma a ideia de que a combinação de modelos pode ser capaz de superar os modelos singulares.

Tabela 11 – Erro absoluto percentual médio das previsões para o ICMS e o IPVA do estado da Paraíba para os anos de 2019, 2020 e 2021

Rank	IPVA		ICMS	
	Modelo	Média	Modelo	Média
1	SARIMA	0,95%	SARIMA	4,46%
2	C1	1,48%	THETAM	4,58%
3	C2	1,50%	C1	4,88%
4	ETS	1,75%	ETS	5,24%
5	C3	1,91%	HW	5,30%
6	HW	2,15%	C3	6,36%
7	C4	4,24%	TBATS	6,80%
8	TBATS	7,04%	C4	6,95%
9	NNAR	8,16%	C2	7,20%
10	X13-ARIMA-SEATS	9,27%	X13-ARIMA-SEATS	9,75%
11	SNAIVE	9,90%	NNAR	9,79%
12	STL-ETS	17,00%	STL-ETS	10,60%
13	THETAM	23,27%	SNAIVE	11,52%
14	Previsão Oficial (LOA)	46,02%	Previsão Oficial (LOA)	30,75%

Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

5 CONSIDERAÇÕES FINAIS

Este trabalho teve como escopo a aplicação de modelos de previsões para estimar a arrecadação de ICMS e IPVA na Paraíba, com uso de dados de 1997 a 2020. Foram utilizados 09 modelos univariados distintos – chamados de modelos simples – e quatro modelos híbridos, formados por diferentes combinações dos modelos simples. A precisão das previsões foi comparada, a fim de se identificar o melhor método a ser utilizado, por meio de quatro medidas de desempenho: RMSE (*Root Mean Squared Error*); MAE (*Mean Absolute Error*); MPE (*Mean Percentual Error*); e MAPE (*Mean Absolute Percentual Error*). Ainda foram feitas comparações entre as previsões geradas pelos modelos simples e híbridos para os exercícios de 2019, 2020 e 2021 e os valores realizados, e calculados os erros cometidos.

Os resultados obtidos apontam que o método de previsão adotado pelo executivo estadual da Paraíba requer revisão, em face de estar produzindo estimativas para a arrecadação tributária consideravelmente distantes dos valores realizados. Tal constatação realça a importância de se identificar modelos de previsão que possam oferecer robusta capacidade preditiva a fim de que sejam utilizados pelas gestões estatais.

Ficou demonstrado, ainda, que os modelos combinados foram capazes de gerar previsões em valores bastante aproximados dos dados reais, especialmente para o período de normalidade (anterior ao choque pandêmico), conferindo precisão muito superior às estimativas oficiais, o que comunga com os achados de Mendonça e Medrano (2016). Contudo, o poder preditivo dos modelos híbridos testados não foi uniforme e perdeu força a partir de 2020, o que pode decorrer da quebra das séries, resultantes dos efeitos da Covid-19 sobre a atividade econômica.

Ademais, considerando o ajuste médio das previsões para os três exercícios, ficou demonstrado que o modelo *SARIMA* foi capaz de gerar as melhores previsões, tanto para o IPVA como para o ICMS, confirmando a utilidade do uso do modelo para aferição de tributos estaduais, conforme anteriormente assinalado nos estudos de Santos e Lima (2006) e Azevedo, Silva e Gatsio (2017). Outros modelos singulares também obtiveram bons ajustes, com destaque para o *ETS* e *Holt-Winters*, o que sinaliza que suas utilizações também podem ser proficientes.

Como consequência, é possível dizer que há necessidade de aperfeiçoamento na elaboração das previsões de tributos no estado da Paraíba. Para isso, tanto o uso de modelos de previsão combinados quanto o de modelos singulares podem ser indicados como ferramentas importantes a serem incorporadas nas formulações de previsão orçamentária dos governos estaduais para fins de planejamento, elaboração de leis orçamentárias ou como suporte a decisões de políticas fiscais.

Os resultados abrem outras perspectivas de estudo. A extrapolação da pesquisa para outros períodos, com inclusão e comparação de previsões aplicadas a outros estados federativos, é um encaminhamento que pode permitir avaliar, com maior segurança, a opção do modelo de previsão a ser prioritariamente adotado. Além disso, outras combinações podem ser testadas a fim de verificar se oferecem ganhos na robustez das previsões. A incorporação de modelos multivariados e de aprendizagem de máquina (*machine learning*) pode ser um caminho a ser percorrido como continuação a esse estudo. Por fim, a verificação dos efeitos de ciclos econômicos sobre o desempenho dos modelos de previsões é um direcionamento para pesquisa futura.

REFERÊNCIAS

- ARMSTRONG, J. S. **Principles of Forecasting**. Norwell, MA: Kluwer Academic, 2001.
- ASSIMAKOPOULOS, V.; NIKOPOULOS, K. The theta model: a decomposition approach to forecasting. **International Journal of Forecasting**, v. 16, p. 521-530, 2000.
- AZEVEDO, R. R.; SILVA, J. M.; GATSIOS, R. C. Análise crítica dos modelos de previsão de série temporal com base no ICMS estadual. **Revista de gestão, Finanças e Contabilidade**. UNEB, v. 7, n. 1, p. 164-184, 2017.
- BOX, G. E. P; JENKINS, G. M. **Time series analysis: forecasting and control**. Holden-Day, San Francisco. 1970.
- BRATU, M. New methods for comparing the forecasts accuracy. **American Journal of Applied Mathematics and Statistics**, v. 1, n. 1-5, 2013.
- CAMARGOS, A. A. B. **Modelos de Previsão da arrecadação tributária do estado de São Paulo: ICMS, IPVA, ITCMD, e taxas**. Dissertação (Mestrado) - Faculdade IBMEC - Programa de Mestrado Profissional em Economia, São Paulo, 2008.
- CLEMEN, R. T. Combining forecasts: A review and annotated bibliography. **International Journal of Forecasting**, 5, p. 59-583, 1989.
- 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. **Economia & Tecnologia**, v. 27, n. 07, 2011.
- CONTRERAS, C. S. **Previsão de arrecadação do ICMS através de redes neurais no Brasil**. 2005. Dissertação (Mestrado) - Programa de Pós-Graduação em Estatística, Universidade Federal de Pernambuco, Recife, 2005.
- DUARTE, F. C. de L.; DE SOUZA, M. F.; GIRÃO L. F. de A. Previsão de arrecadação do ICMS no estado da Paraíba. **Revista UNIABEU**, v.7, n. 17, 2014.

- ELLIOTT, G.; ROTHENBERG, T.; STOCK, J. Efficient test for an autoregressive unit root. **Econometrica**, v. 64, n. 4, p. 813-838, 1996.
- FAN, D., SUN, H.; YAO, J. et al. Well production forecasting based on ARIMA-LTSM model considering manual operations. **Energy**, n. 220, 2021.
- FERREIRA, P. G. C. **Análise de séries temporais em R: curso introdutório**. 1. ed. São Paulo: GEN. Publicado pelo selo Editora Atlas Ltda., FGV IBRE, 2020.
- FERRO, W. A.; LIMA, J. D.; TRENTIN, M. G. Combinações de Métodos Quantitativos na Previsão de Demanda de Vendas de Eletrodomésticos. **GEPROS: Gestão da Produção, Operações e Sistemas**, v. 14, n. 5, p. 67-88, 2019.
- FIGUEIREDO, A. M. R. **Séries Temporais com R: Modelo híbrido de Hyndman e Athanasopoulos**. Campo Grande-MS, Brasil: RStudio/Rpubs, 2019. Disponível em: http://rpubs.com/amrofi/combination_forecast. Acesso em: 29 jun. 2021.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. 2nd ed. Otexts, 2018. Disponível em: <https://otexts.org/fpp2/>. Acesso em: 29 jun. 2021.
- MENDONÇA, M. J.; MEDRANO, L. A. Um modelo de combinação de previsões para arrecadação de receita tributária no Brasil. **Instituto de Pesquisa Econômica Aplicada. Texto para Discussão n. 2186**, mar. 2016.
- MENDONÇA, M. J.; SACHIDA, A. Um modelo econométrico para previsão de Receita Tributária no Brasil. **Economia Aplicada**, v. 17, n. 2, p. 295-329, 2013.
- KWIATKOWSKI, D.; PHILLIPS, P. C.; SCHIMIDT, P.; SHIN, Y. Testing the null hypothesis of stationarity against the alternative of unit root: how sure are we that economic time series have a unit root? **Journal of Econometrics**, v. 54, n. 1, p. 159-178, 1992.
- MANCUSO, A. C. B.; WERNER, L. Review of combining forecasts approaches. **Independent Journal of Management & Production**, v. 4, n. 1, jun. 2013.
- MOÇO, J. R. C. **Estimação de um modelo de previsão de arrecadação tributária do estado do Rio de Janeiro**. 2016. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Instituto de Economia, Programa de Pós-Graduação em Economia da Indústria e da Tecnologia, 2017.
- PARAÍBA. Lei nº 11.295, de 15 de janeiro de 2019. **Estima a receita e fixa e despesa do estado para o exercício financeiro de 2019, e dá outras providências**. João Pessoa: Assembleia Legislativa, 2019.
- PARAÍBA. Lei nº 11.627, de 14 de janeiro de 2020. **Estima a receita e fixa e despesa do estado para o exercício financeiro de 2020, e dá outras providências**. João Pessoa: Assembleia Legislativa, 2020.
- PARAÍBA. Lei nº 11.831, de 07 de janeiro de 2021. **Estima a receita e fixa e despesa do estado para o exercício financeiro de 2021, e dá outras providências**. João Pessoa: Assembleia Legislativa, 2021.
- PESSOA, F. 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. 9, p. 47-64, 2013.

SANTOS, C. M.; LIMA, J. E. Análise de previsões de arrecadação do ICMS no estado de Minas Gerais. **Revista de Economia e Administração/IBMEC São Paulo**, v. 5, n. 4, 2006.

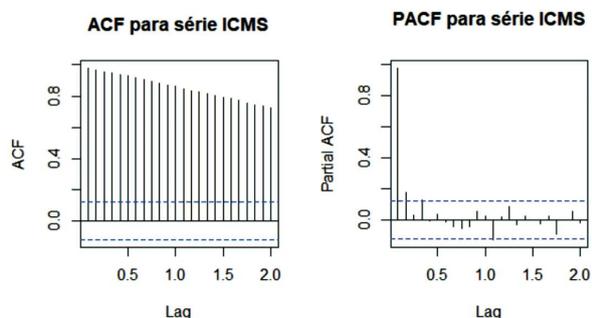
SCHEFFER, D.; SOUZA, A. M.; ZANINI, R. R. Utilização de modelos ARIMA para previsão da arrecadação de ICMS do Estado do Rio Grande do Sul. *In*: SIMPÓSIO DE PESQUISA OPERACIONAL E LOGÍSTICA DA MARINHA – SPOLM, 17., **Anais [...]** São Paulo: 2014.

SHAUB, D.; ELLIS, P. **forecastHybrid: Convenient Functions for Ensemble Time Series Forecasts. R package version 5.0.19, 2020**. Disponível em: <https://cran.r-project.org/package=forecastHybrid>. Acesso em: 29 mar. 2021.

WOOLDRIDGE, J. M. **Introdução à econometria: uma abordagem moderna**. São Paulo: Cengage Learning, 2014.

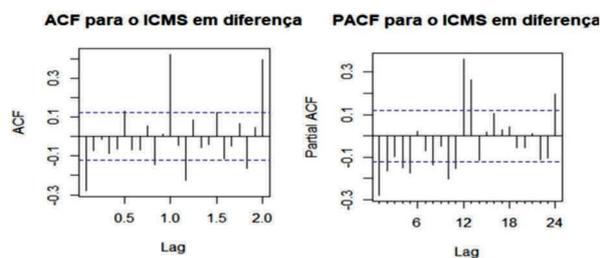
APÊNDICE A – FUNÇÕES DE AUTOCORRELAÇÃO – 1997/2018

Figura A.1. – Funções de Autocorrelação e de Autocorrelação parcial para o ICMS da série original (1997-09.2018)



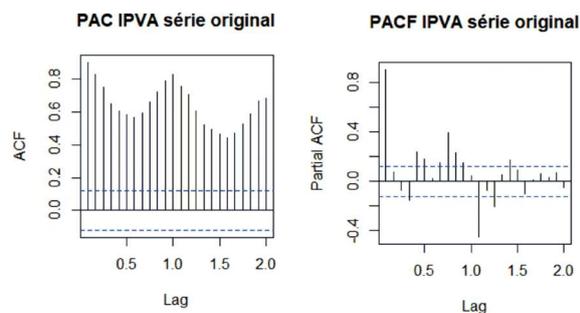
Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Figura A.2. – Funções de Autocorrelação e de Autocorrelação parcial para o ICMS da série em primeira diferença (1997-09.2018)



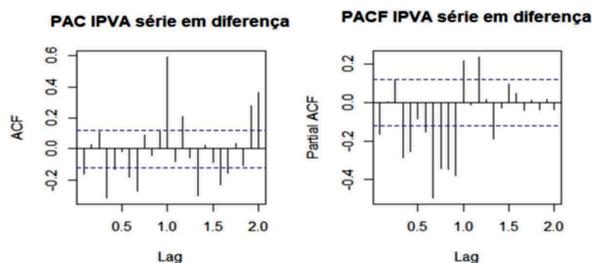
Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Figura A.3. – Funções de Autocorrelação e de Autocorrelação parcial para o IPVA da série original (1997-09.2018)



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Figura A.4. – Funções de Autocorrelação e de Autocorrelação parcial para o IPVA da série em 1 diferença (1997-09.2018)



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

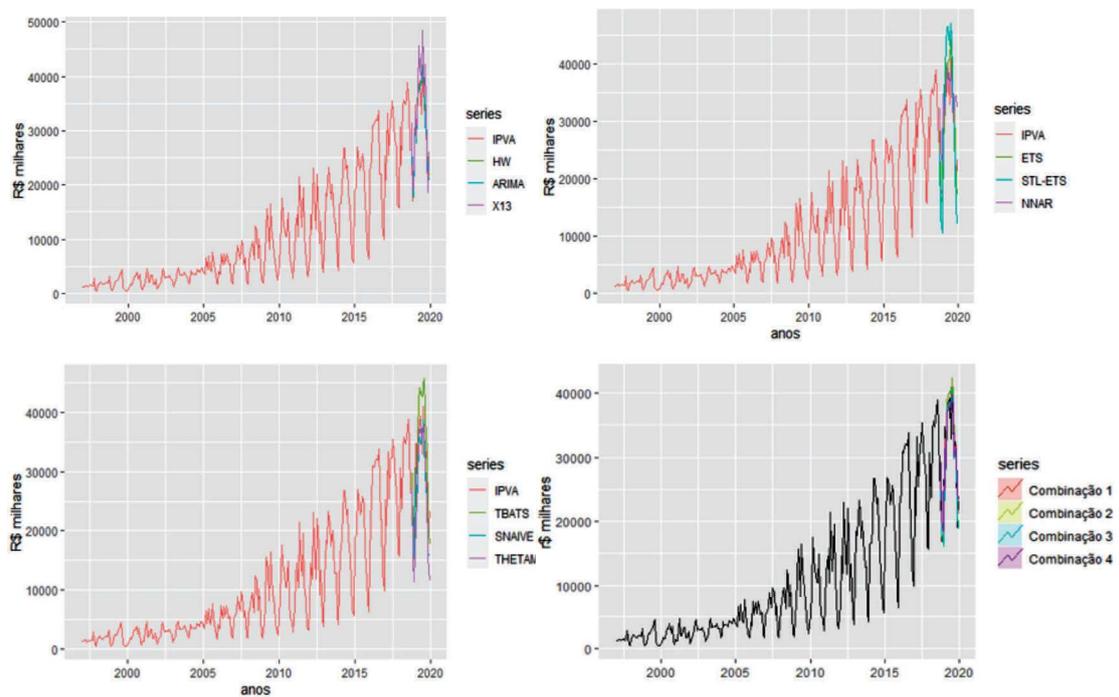
APÊNDICE B – GRÁFICOS DAS PREVISÕES - 10.2018 A 12.2019

Figura B.1. – Gráficos das previsões para o ICMS do estado da Paraíba por modelos simples e combinados para os meses de 10.2018 a 12.2019



Elaboração própria a partir das estimações com uso do R-Studio.

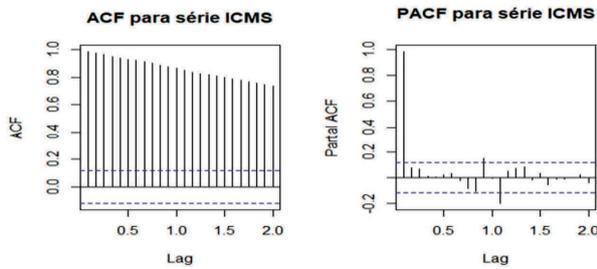
Figura B.2. – Gráficos das previsões o IPVA do estado da Paraíba por modelos simples e combinados para os meses de 10.2018 a 12.2019



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

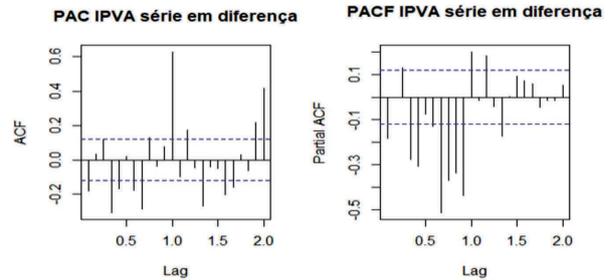
APÊNDICE C - FUNÇÕES DE AUTOCORRELAÇÃO – 1997/2019

Figura C.1. – Funções de Autocorrelação e de Autocorrelação parcial para o ICMS da série original - (1997-09.2019)



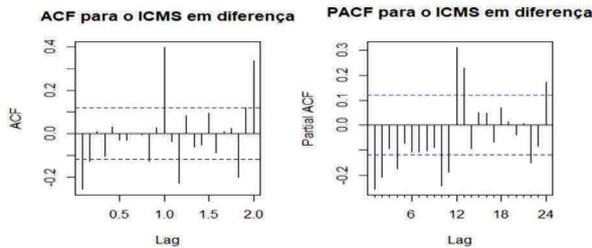
Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Figura C.4. – Funções de Autocorrelação e de Autocorrelação parcial para o IPVA da série em 1 diferença - (1997-09.2019)



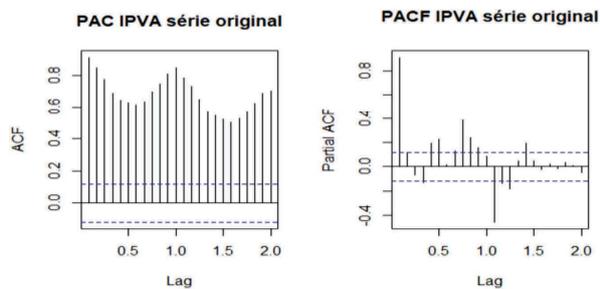
Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Figura C.2. – Funções de Autocorrelação e de Autocorrelação parcial para o ICMS da série em 1 diferença - (1997-09.2019)



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Figura C.3. – Funções de Autocorrelação e de Autocorrelação parcial para o IPVA da série original - (1997-09.2019)



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

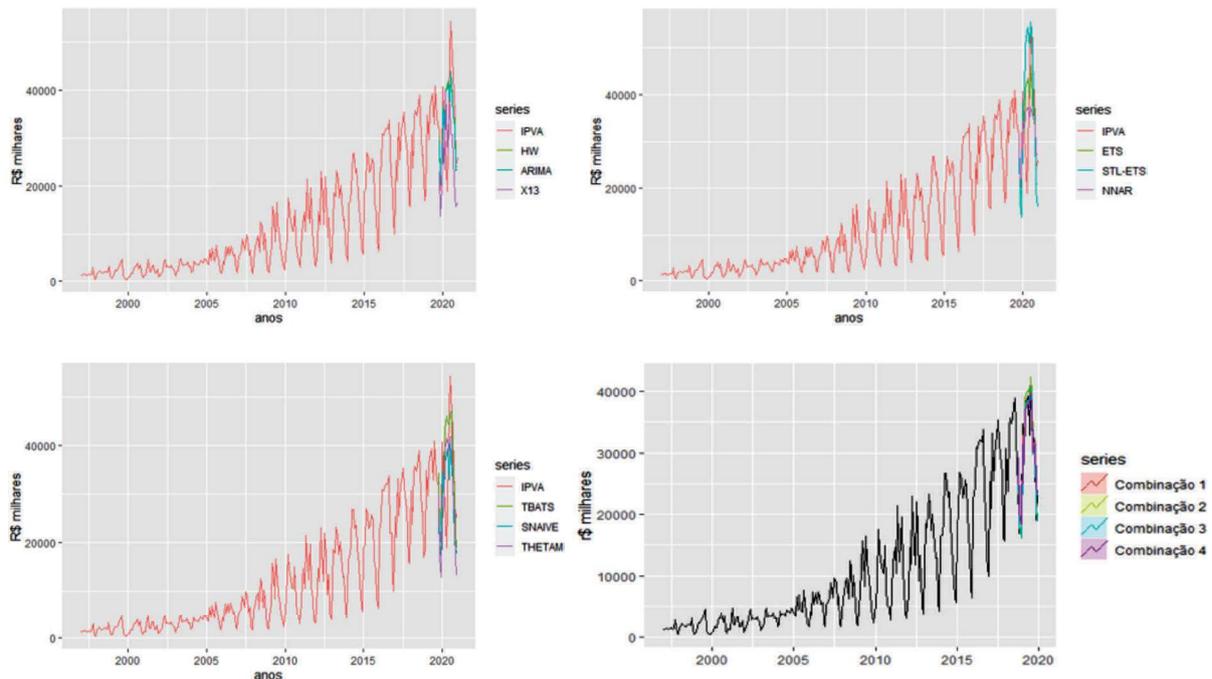
APÊNDICE D – GRÁFICOS DAS PREVISÕES - 10.2019 A 12.2020

Figura D.1. – Gráficos das previsões para o ICMS do estado da Paraíba por modelos simples e combinados - 10.2019 a 12.2020



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

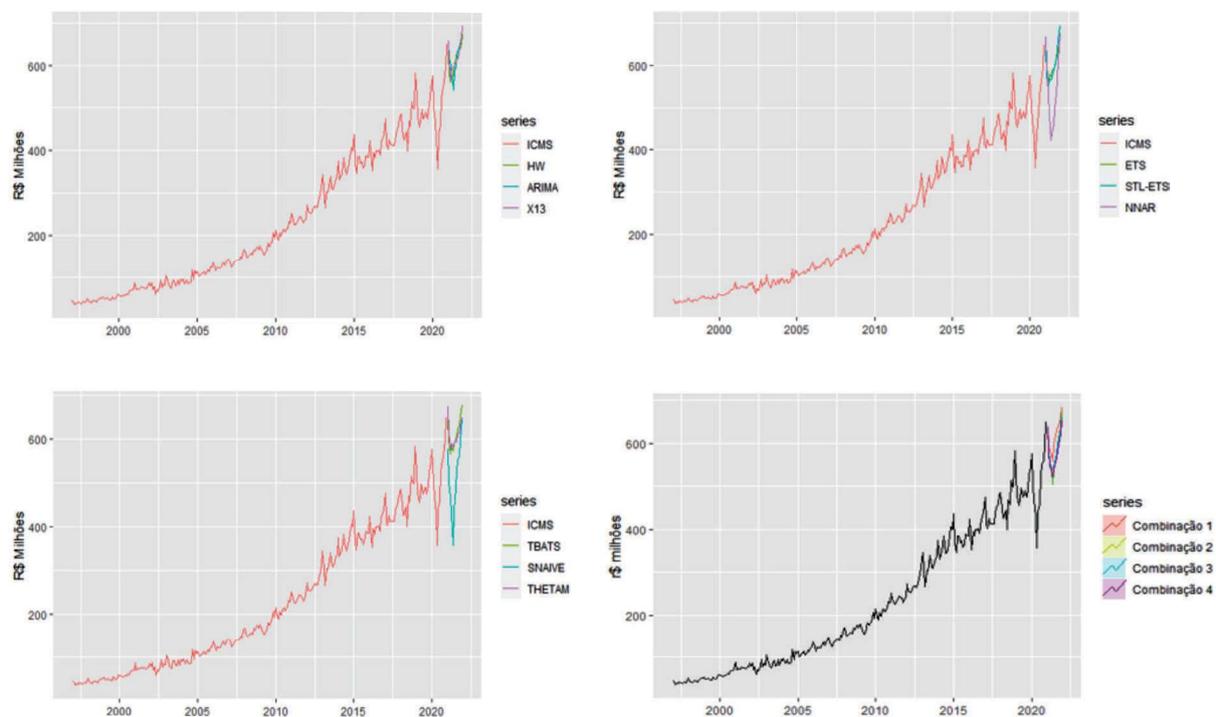
Figura D.2. – Gráficos das previsões para o IPVA do estado da Paraíba por modelos simples e combinados - 10.2019 a 12.2020



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

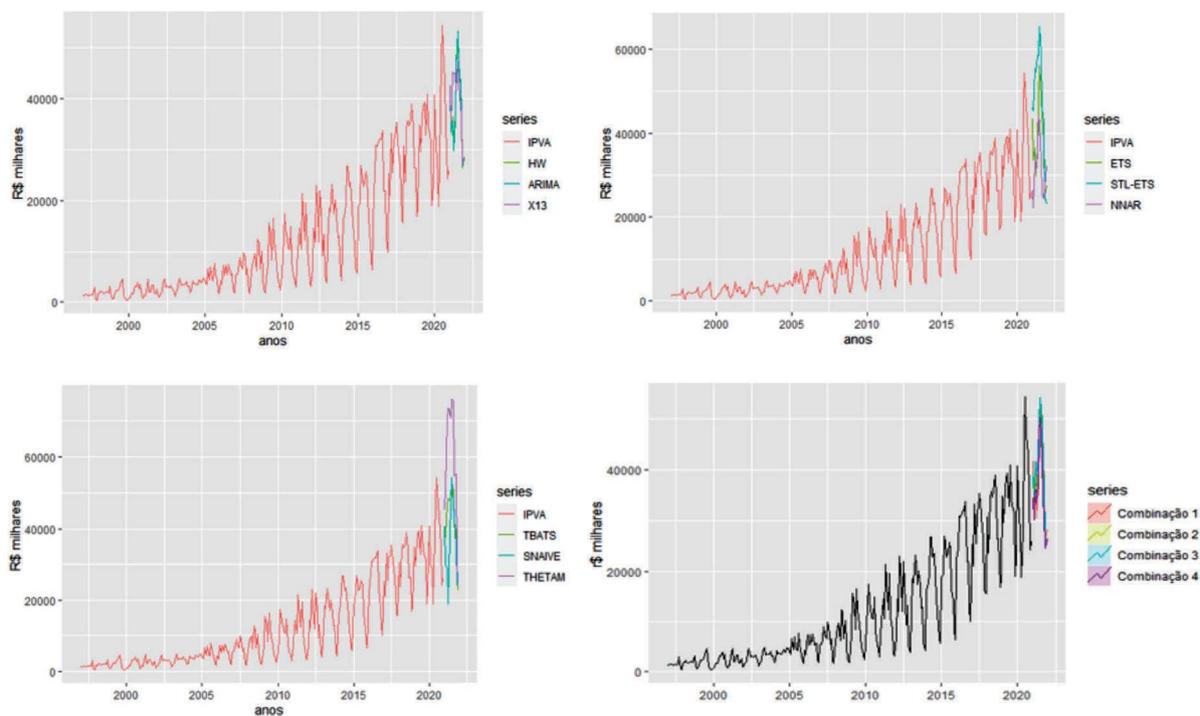
APÊNDICE E – GRÁFICOS DAS PREVISÕES - 10.2020 A 12.2021

Figura E.1 – Gráficos das previsões para o ICMS do estado da Paraíba por modelos simples e combinados – 10.2020 a 12.2021



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.

Figura E.2. – Gráficos das previsões para o IPVA do estado da Paraíba por modelos simples e combinados – 10.2020 a 12.2021



Fonte: Elaboração própria a partir das estimações com uso do R-Studio.