TEXTO PARA DISCUSSÃO

2611

PREVISÃO DE HOMICÍDIOS NO BRASIL: PROPOSTA DE VARIÁVEL ANTECEDENTE

Gabriel de Oliveira Accioly Lins
Daniel Ricardo de Castro Cerqueira
Danilo Coelho

ipea

2611
TEXTO PARA DISCUSSÃO
Brasília, novembro de 2020

PREVISÃO DE HOMICÍDIOS NO BRASIL: PROPOSTA DE VARIÁVEL ANTECEDENTE

Gabriel de Oliveira Accioly Lins[1]
Daniel Ricardo de Castro Cerqueira[2]
Danilo Coelho[3]

1. Pesquisador assistente na Diretoria de Estudos e Políticas do Estado, das Instituições e da Democracia (Diest) do Ipea. *E-mail*: <gablins@yahoo.com.br>.
2. Técnico de planejamento e pesquisa na Diest/Ipea. *E-mail*: <daniel.cerqueira@ipea.gov.br>.
3. Técnico de planejamento e pesquisa na Diest/Ipea. *E-mail*: <danilo.coelho@ipea.gov.br>.

Governo Federal

Ministério da Economia
Ministro Paulo Guedes

ipea **Instituto de Pesquisa Econômica Aplicada**

Fundação pública vinculada ao Ministério da Economia, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiros – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidente
Carlos von Doellinger

Diretor de Desenvolvimento Institucional
Manoel Rodrigues Junior

Diretora de Estudos e Políticas do Estado, das Instituições e da Democracia
Flávia de Holanda Schmidt

Diretor de Estudos e Políticas Macroeconômicas
José Ronaldo de Castro Souza Júnior

Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais
Nilo Luiz Saccaro Júnior

Diretor de Estudos e Políticas Setoriais de Inovação e Infraestrutura
André Tortato Rauen

Diretora de Estudos e Políticas Sociais
Lenita Maria Turchi

Diretor de Estudos e Relações Econômicas e Políticas Internacionais
Ivan Tiago Machado Oliveira

Assessor-chefe de Imprensa e Comunicação (substituto)
João Cláudio Garcia Rodrigues Lima

Ouvidoria: http://www.ipea.gov.br/ouvidoria
URL: http://www.ipea.gov.br

Texto para Discussão

Publicação seriada que divulga resultados de estudos e pesquisas em desenvolvimento pelo Ipea com o objetivo de fomentar o debate e oferecer subsídios à formulação e avaliação de políticas públicas.

© Instituto de Pesquisa Econômica Aplicada – **ipea** 2020

Texto para discussão / Instituto de Pesquisa Econômica Aplicada.- Brasília : Rio de Janeiro : Ipea , 1990-

ISSN 1415-4765

1.Brasil. 2.Aspectos Econômicos. 3.Aspectos Sociais. I. Instituto de Pesquisa Econômica Aplicada.

CDD 330.908

As publicações do Ipea estão disponíveis para *download* gratuito nos formatos PDF (todas) e EPUB (livros e periódicos). Acesse: http://www.ipea.gov.br/portal/publicacoes

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério da Economia.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

JEL: C53; K42; C22.

DOI: http://dx.doi.org/10.38116/td2611

SUMÁRIO

SINOPSE

ABSTRACT

1 INTRODUÇÃO 7

2 DADOS, METODOLOGIA E ESPECIFICAÇÃO DO MODELO 8

3 PRINCIPAIS RESULTADOS 18

4 *CHECK* DE ROBUSTEZ 26

5 EXERCÍCIO DE PREVISÃO PSEUDO FORA DA AMOSTRA 27

6 CONCLUSÕES 30

REFERÊNCIAS 31

APÊNDICE A 36

SINOPSE

Neste estudo, investigamos a capacidade de variáveis antecedentes, entre elas internações por agressão, na previsão do número de homicídios no Brasil. O objetivo principal desta pesquisa é suprimir a lacuna referente à defasagem de informações na divulgação sobre homicídios no país, permitindo assim análises conjunturais atualizadas. Para tanto, por intermédio do esquema *rolling window* e da abordagem *model confidence set* (MCS), investigamos se modelos de variáveis antecedentes apresentam desempenho preditivo superior ao conjunto de modelos univariados. Ao aplicar a abordagem MCS, considerando diferentes estatísticas de avaliação, funções de perda e janelas de estimação, encontramos fortes evidências da capacidade das variáveis antecedentes utilizadas fornecerem conteúdo informacional adicional na previsão da dinâmica criminal brasileira, com modelos de variáveis antecedentes sistematicamente superando modelos univariados. Na média, os melhores modelos de variáveis antecedentes apresentam melhorias relativas ao *benchmark random walk*, de 60% em termos de raiz do erro quadrado médio (RMSE), erro absoluto médio (MAE) e desvio absoluto médio da média (MAD).

Palavras-chave: previsão; previsão de homicídios; *model confidence set*; séries temporais; violência.

ABSTRACT

In this article we investigate the capacity of antecedent variables, among them admissions for aggression, in the prediction of the number of homicides in Brazil. The main objective is to eliminate a gap regarding the large time lag in the dissemination of information about these deaths in the country, allowing updated conjuncture analyzes. For this, through a "rolling window" scheme and "Model Confidence Set" approach, we investigate whether multivariate models with leading variables show forecast performance superior to a set of univariate models. In applying the MCS approach, considering different evaluation statistics, loss functions and estimation windows, we find strong evidence of the ability of the leading variables used to provide additional information content in the prediction of the Brazilian criminal dynamics, with models of leading variables systematically surpassing univariate models, especially in extended periods of forecasting. In general, improvements related to the benchmark Random walk model, in terms of RMSE, MAE and MAD, are of the order of 60%.

Keywords: forecasting; homicide forecasting; model confidence set; time series; violence.

1 INTRODUÇÃO

No Brasil, o número de homicídios apresenta tendência de crescimento ao longo dos últimos vinte anos, com taxa média de crescimento na ordem de 8,0% ao ano (a.a.). À luz deste fenômeno, diversas análises investigaram as causas e consequências da criminalidade (Cerqueira, 2014; Cerqueira e Soares, 2016). A ausência de homogeneidade metodológica entre as Unidades da Federação (UFs) ao produzir estatísticas criminais, no entanto, impede a utilização agregada das estatísticas estaduais e obriga interessados no tema a utilizar informações do Sistema de Informações sobre Mortalidade (SIM) do Ministério da Saúde (MS).

Criado com o intuito de produzir informações regulares sobre a mortalidade no Brasil, para fins de análise epidemiológica, o SIM se tornou referência em diversas áreas da gestão pública e pesquisa acadêmica, bem como no campo da criminologia (Cerqueira, 2012; Cerqueira *et al.*, 2018; 2019; Murray, Cerqueira e Kahn, 2013). Apesar da ampla utilização, as informações disponibilizadas pelo SIM são divulgadas com defasagem. Com isso, existe uma lacuna aos interessados em acompanhar a conjuntura da criminalidade violenta. Na ausência de informações atualizadas, a previsão do número de homicídios no Brasil aparece como alternativa aos interessados em elaborar, implementar e avaliar a possível Política Nacional de Segurança Pública. Similar ao observado pelos *policymakers* dos bancos centrais na condução da política monetária.

Este trabalho explora a lacuna existente na disponibilização de informações atualizadas e tem como objetivo mitigar a escassez de dados sobre a criminalidade violenta, demonstrando destarte os ganhos preditivos ao utilizar o conjunto de variáveis antecedentes, entre elas, internações por agressão em modelos preditivos de homicídios. Assim sendo, construiu-se um modelo previsor multivariado de desempenho preditivo superior ao observado em modelos univariados, contribuindo assim com a produção de evidências preliminares acerca da dinâmica criminal brasileira e permitindo aos *policymakers* da segurança pública ajustarem o impacto de políticas nacionais de segurança, aumentado o bem-estar social, além de estabelecer um modelo previsor *benchmark* aos interessados em previsão de dinâmica criminal.

A elaboração de previsões de variáveis criminais não é inédita e serve a diferentes objetivos, tais como: desdobramento tático de recursos policiais, alocação de recursos e planejamento estratégico (Gorr, Olligschlaeger e Thompson, 2003). Por ser um campo pouco explorado na literatura nacional, os estudos existentes a esse respeito utilizam

técnicas de previsão e métodos de avaliação modestos (Souza, *et al.*, 2007; Provenza *et al.*, 2015; Provenza, Costa e Silva, 2015). Por sua vez, a literatura internacional utiliza crescente sofisticação técnica nas pesquisas (Gorr e Harries, 2003; Gorr, Olligschlaeger e Thompson, 2003; Klepinger e Weis, 1985; Pepper, 2008; Shoesmith, 2013). No entanto, a maioria destes trabalhos utiliza abordagem univariada. Sendo assim, esta pesquisa inova ao introduzir inédita variável antecedente capaz de superar a capacidade previsora de modelos univariados. Além disso, estabelecemos *benchmark* de melhor *performance* em diferentes horizontes de previsão.

O desempenho preditivo de modelos com a variável antecedente internação por agressão foi comparado a um conjunto de modelos univariados. A precisão fora da amostra foi avaliada, em diferentes horizontes de previsão, por meio de estatísticas de precisão e da abordagem *model confidence set* (MCS) proposta em Hansen, Asger e James (2011), utilizando distintas estatísticas de avaliação, função de perda e largura de janelas de estimação.

As evidências encontradas sugerem que a inclusão de variáveis antecedentes gera modelos de previsão com melhor desempenho fora da amostra daqueles do conjunto de modelos univariados. Especificamente, os modelos de variáveis antecedentes apresentam o melhor desempenho preditivo em todos os horizontes de previsão considerados. Entretanto, não existe evidência de modelo multivariado superior aos demais em todos os horizontes de previsão. Na média, os melhores modelos de variáveis antecedentes apresentam melhorias relativas ao modelo de *benchmark random walk* de 60%, em termos de raiz do erro quadrado médio (RMSE), erro absoluto médio (MAE) e desvio absoluto médio da média (MAD).

Este estudo está organizado da seguinte forma. Além desta introdução, a seção 2 apresenta os dados utilizados no trabalho, as evidências iniciais e a metodologia. Na seção 3, são apresentados os principais resultados. Elaboramos exercício de robustez na seção 4; por fim, a seção 5 conclui.

2 DADOS, METODOLOGIA E ESPECIFICAÇÃO DO MODELO

Nossa base de dados compreende duas séries temporais mensais, agregadas nacionalmente, abrangendo o período entre 1999m01 e 2017m12, isto é, o número de óbitos causados por agressão e intervenção legal no Brasil, doravante chamado de homicídios,

e o número de internações por agressão (por exemplo, agressões por arma de fogo, lesões corporais e ferimentos perfurantes).[1]

As informações sobre homicídios são oriundas do Sistema de Informações sobre Mortalidade (SIM), sistema gerido pela Secretaria de Vigilância à Saúde (SVS) do Ministério da Saúde (MS), em conjunto com as secretarias estaduais e municipais de saúde. Estas secretarias coletam as declarações de óbitos dos cartórios e conferem, no SIM, as informações nelas contidas. Criado com o objetivo de produzir informações regulares sobre a mortalidade no país, o SIM possui abrangência universal.

Uma das informações primordiais do SIM é a causa básica de óbito, isto é, doença ou lesão que iniciou a cadeia de acontecimentos patológicos que conduziram diretamente à morte, ou as circunstâncias do acidente ou violência que produziram a lesão fatal, a qual é codificada a partir do declarado pelo médico atestante, segundo regras estabelecidas pela Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde (CID-10), publicada pela Organização Mundial de Saúde (OMS).

O número de homicídios agrega falecidos cuja causa básica está classificada nas seguintes CIDs-10: X85 até Y09 e Y35 até Y36, ou seja, o número de óbitos causados por agressão e intervenção legal. Falecimentos por complicações decorrentes da agressão ou intervenção legal são classificados como homicídios, se a causa básica informada da morte for alguma das CIDs anteriores. Por exemplo, agressão por arma de fogo pode levar a internação prolongada e posterior morte por septicemia. Neste caso, a causa básica da morte foi perfuração por arma de fogo (CID-10 X93), no entanto, a causa terminal ou secundária foi septicemia (CID-10 A41). Assim, tempos após a agressão o falecimento é classificado como homicídio na data da morte pela causa secundária septicemia, pois a causa básica da morte informa perfuração por arma de fogo.

Dado o formato de construção da base SIM, existe a possibilidade de os homicídios decorrentes de internação de longa duração distorcerem o número de homicídios no mês. Por construção, a base do SIM informa a data de falecimento, mas não indica o tempo de internação do paciente, não sendo possível identificar, na base SIM, mortes cuja causa básica seja agressão após internação de longa duração. No exemplo anterior, a morte pela causa básica perfuração de arma de fogo ocorreu meses após a internação. A morte é registrada meses após a ocorrência, viesando o nível de violência capturado

1. Ver na tabela A.1 do apêndice a fonte e descrição sumária das variáveis utilizadas e na tabela A.2 as estatísticas descritivas.

pela variável homicídios. No entanto, tal possibilidade parece remota, pois, entre 2006 e 2017, dos homicídios com local de óbito determinado, 72% ocorreram fora do ambiente hospitalar, inexistindo assim possibilidade de internação de longa duração. Além disso, segundo os dados de internação, o tempo médio de internação daqueles mortos com causa base agressão ou intervenção legal é de 4,93 dias, sugerindo escassa ocorrência de mortes por homicídios após internações de longa duração e, portanto, residual influência na dinâmica do resultado final.

Por sua vez, a variável internação por agressão advém do Sistema de informações Hospitalares (SIH), cujas informações são originárias das autorizações de internação hospitalar, destinadas ao pagamento de internações de hospitais públicos e privados conveniados ao Sistema Único de Saúde (SUS). Ou seja, o SIH armazena dados sobre as internações hospitalares no âmbito do SUS, informadas mensalmente por todos os estabelecimentos de saúde públicos, conveniados e contratados que realizam internações.

A variável internação por agressão foi construída por meio de microdados do SIH. Ocorre a agregação mensal de novas internações cuja causa primária ou secundária seja codificada pelos códigos CID-10 X85 até Y09, Y35, Y36 e W50. Ou seja, agressões, intervenções legais e operações de guerra e exposição a forças mecânicas animadas (isto é, golpe, pancada e pontapé).

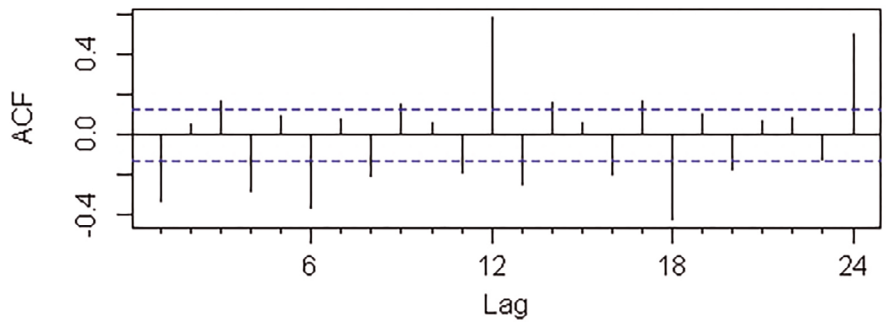
Como usualmente acontece em análises de série temporal, é necessário verificar a existência de raiz unitária, evitando a possibilidade de regressão espúria. Com este objetivo, foram realizados os testes Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) e Kwiatkowski-Phillips-Schmidt-Shin (KPSS).[2] Os resultados indicam que apenas internações é série I (0).

A primeira evidência, apresentada na figura 1, sugere autocorrelação de primeira ordem na série histórica de homicídios brasileiros, evidência alinhada aos achados de Santos (2009). Por estas evidências, incluímos uma defasagem da variável dependente no conjunto de variáveis previsoras. Além disso, por causa do pequeno comportamento sazonal, incluímos *dummies* mensais como possíveis previsoras de homicídios, seguindo Medeiros, Vasconcelos e Freitas (2015).[3]

2. Ver resultados na tabela A.2 (apêndice).

3. Aplicar filtro de ajuste sazonal está sujeito à alteração na dinâmica da variável (Canova e Ghysels, 1994; Davidson e MacKinnon, 2004) e redução do poder de testes de raiz unitária (Harries e Sollis, 2003). Portanto, opta-se em capturar o comportamento sazonal por meio das variáveis utilizadas.

FIGURA 1
Função de autocorrelação: homicídios



Elaboração dos autores.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

A fim de estabelecer *benchmark* aos modelos de variáveis antecedentes, na ausência de consenso na literatura acerca do melhor método univariado capaz de prever homicídios e, por causa da flexibilidade do método de avaliação empregado, elaborou-se um conjunto de modelos univariados. Os métodos aplicados foram os seguintes.

1) *Random walk*: conhecido como passeio aleatório, trata-se do modelo mais simples no conjunto dos modelos estimados. Neste caso, o número de homicídios no mês é igual ao número do mês anterior, $\widehat{y}_{T+h|t} = y_{T+h-1|t}$. Este modelo não apresenta desempenho preditivo adequado na presença de elevada aleatoriedade ou forte tendência e sazonalidade. No entanto, será previsor adequado de séries com frequentes mudanças de padrão. Este modelo será utilizado como *benchmark*.

2) *Snaive*: neste caso, o número de homicídios é igual ao número de doze meses antes $y_{T+h|t} = y_{T+h-12|t}$. Método utilizado por departamentos policiais na previsão do nível de crime, considerando variações sazonais (Gorr, Olligschlaeger e Thompson, 2003).

3) ETS: modelos de suavização exponencial, aplica o algoritmo proposto em Hyndman *et al*. (2002) ao modelar componente sazonal, tendência e termo de erro da série a partir de classificação proposta em Hyndman *et al*. (2008). Seguindo a abordagem de Ord, Koehler e Snyder (1997), deriva-se modelo espaço-estado com erros de origem única, equivalente ao modelo de suavização exponencial selecionado de acordo com o Critério de Informação de Akaike (AIC). Parâmetros de suavização e vetor de estados iniciais são estimados pela maximização de verossimilhança. É possível obter diversos métodos, tais como: suavização exponencial simples, modelo linear de Holt e diferentes formas do método de Holt-Winter.

4) Arima: aplicação de algoritmo proposto por Hyndman e Khandakar (2008). Trata-se de regressão com erros Arima (p, d, q), estimada por máximo verossimilhança.

O algoritmo seleciona a ordem do termo autorregressivo *p* e média móvel *q* de acordo com critério de informação AICc.[4]

5) STL: neste caso, aplica-se a decomposição do tipo *seasonal and trend decomposition using loess* (STL), proposto por Cleveland, Cleveland e Terpenning (1990). Isto é, decompomos a série analisada em componente sazonal, tendência e termo de erro e realizamos separadamente ajuste e previsão destes componentes. Ajustam-se séries sazonalmente (tendência e termo de erro) pelo modelo ETS (STLe) e Arima (STLa). Em seguida, é feita a previsão da série ajustada sazonalmente, então os resultados são ressazonalidados. Como o padrão sazonal repete-se ao longo da série analisada, adicionamos à série ajustada sazonalmente estimativa do componente sazonal do ano anterior, ou seja, a previsão do componente sazonal é realizada pelo método *random walk*.

6) *Bootstrap aggregating* (*bagging*): método proposto por Bergmeir, Hyndman e Benitez (2016). Ele aplica *bootstrap* aos resíduos da série, simulando valores futuros desta. Após transformação do tipo *box-cox*, a série é decomposta em tendência, sazonalidade e resíduo, por decomposição do tipo STL. Então, a previsão de cada componente é realizada pelo algoritmo de Hyndman e Khandakar (2008). Neste caso, o ruído é submetido a processo de *bootstrap*, isto é, ele sofre um processo de reamostragem, neste caso 150 vezes. Todavia, como o termo aleatório de uma série temporal pode ser correlacionado no tempo, a reamostragem não é feita por observações, mas em blocos – *moving blocked bootstrap*. Por fim, estes novos componentes aleatórios, obtidos via *bootstrap*, agregados pela média (*bagging*), são reintroduzidos aos componentes de tendência e sazonalidade da série original, formando novas séries que são variantes dela.

7) TBATS: modelagem do tipo espaço-estado segue o método de previsão proposto por Livera, Hyndman e Snyder (2011), elaborado para séries temporais com padrão sazonal complexo (por exemplo, múltiplos períodos sazonais). Este método incorpora transformação *box-cox*, filtro Arma para correção de resíduo e componente sazonal trigonométrico. As principais vantagens do TBATS, apontadas em Livera, Hyndman e Snyder (2011) sobre outros métodos, são os parâmetros calculados por máxima verossimilhança e a capacidade de lidar com padrões complexos de sazonalidade, nos quais existem múltiplos ciclos e o tamanho dos ciclos não é número inteiro.

8) Bats: similar ao TBATS este método não incorpora componente sazonal trigonométrico.

4. O desempenho preditivo apresentado por outros critérios de informação foi inferior, por brevidade, os resultados foram omitidos.

9) Arfima: modelo Arfima (p, d, q) selecionado e estimado pelo algoritmo proposto por Hyndman e Khandakar (2008) para selecionar *p* e *q* e o algoritmo de Haslett e Raftery (1989) para estimar os parâmetros incluindo *d*.

10) NNETAR: modelo de rede neural *feed forward* de três camadas envolvendo funções de combinação linear e ativação logística. Denotado por NNAR (p, P, k, em que *p* informa a ordem autorregressiva não sazonal e *P* a ordem sazonal das variáveis de entrada defasadas, selecionadas pelo critério de informação AIC e *k* representa o número de nós na camada oculta única, selecionado k = (p + P + 1)/2, seguindo Hyndman e Athanasopoulos (2018).

11) Struct: modelo de componente não observado no tipo de modelo estrutural básico de série temporal, ou seja, modelo de tendência linear local acrescido de componente sazonal. Trata-se de modelo Gaussiano linear na forma espaço-estado, baseado na decomposição dos componentes da série temporal (Harvey, 1990). O modelo assim expresso permite utilização de filtro de Kalman (1960) na estimação do vetor de estado e das previsões requeridas – parâmetros desconhecidos estimados por máxima verossimilhança.

12) ELM: conhecido por *extreme learning machine*. Trata-se de rede neural do tipo *feedfoward* de camada oculta única, proposta em Huang, Zhu e Siew (2006). Neste caso, em vez de utilizar métodos de aprendizado *gradiente-based*, os parâmetros de entrada (pesos dos *inputs* e *hidden layer bias*) não requerem ajuste. Esses são aleatoriamente designados, permanecendo constantes durante o treinamento e a previsão. Pesos de saída são analiticamente determinados. Diante de parâmetros de entrada aleatoriamente escolhidos e fixos, o treinamento equivale a encontrar solução de mínimos quadrados ordinários de sistema de equações lineares (Salaken *et al.*, 2017; Wan *et al.*, 2014). Em seguida, os pesos de saída são determinados por regressão do tipo Lasso, removendo unidades irrelevantes e obtendo rede de tamanho ótimo. Finalmente, as redes são combinadas pela mediana, à luz dos resultados obtidos em Kourentzes, Barrow e Crone (2014).

13) Thief: método de previsão pela hierarquia temporal. Isto é, realizam-se previsões individuais das agregações temporais da série temporal, em seguida, as previsões são reconciliadas pelo algoritmo reconciliatório hierárquico (Athanasopoulos *et al.*, 2017). Neste caso, a série homicídios mensais é agregada em bimestre, trimestre, quadrimestre, semestre e anualmente. Assim, as previsões das séries agregadas são realizadas pelos algoritmos propostos por Hyndman e Khandakar (2008) (Thiefa) e do algoritmo de suavização exponencial de Hyndman *et al.* (2002) (Thiefe). Em seguida, continuando os melhores resultados obtidos por Athanasopoulos *et al.* (2017), as previsões agregadas por tempo são individualmente reconciliadas pelo algoritmo *structural scaling*.

Além desse conjunto de modelos univariados, incluímos dois esquemas de combinações, isto é, média e mediana das previsões univariadas. Seguindo os melhores resultados obtidos na literatura – por exemplo, Samuels e Sekkel (2017) –, desconsideramos modelos de baixa capacidade preditiva e incluímos nas combinações somente modelos univariados de RMSE médio inferior ao RMSE médio dos modelos univariados.

No caso dos modelos com variáveis antecedentes, o método de identificação proposto por Pankratz (1991c) sugere modelo com uma defasagem da variável dependente e uma defasagem da variável independente (no nosso caso, o número de internações por agressão), igualmente ao modelo sugerido no critério de informação BIC e HQ. Por sua vez, o AIC sugere ser o modelo de melhor ajuste àquele de cinco defasagens da variável internações e uma defasagem da variável dependente. Ao contrário do critério AICc, que propõe quatro defasagens da variável internações e uma defasagem da variável dependente.

A metodologia de avaliação implementada – MCS – apresenta em orientação decrescente a *performance* preditiva fora da amostra do conjunto de melhores modelos dado certo nível de confiança. Esta flexibilidade permite avaliar conjuntamente o desempenho de todos os modelos estimados. Portando, à luz das diferentes especificações sugeridas pelos critérios de informação e como forma de avaliar o maior conjunto de informações, também foram estimados modelos com especificações distintas das sugeridas pelos critérios de identificação tradicionais.

Considere o seguinte modelo:

$$y_{t+h} = f_n(x_t) + \mu_{t+h}, h = 1,..., 12, \quad (1)$$

O número de homicídios no mês t + h será representado por y_{t+h}, $x_t = (x_{1t},..., x_{nt})'$ é vetor de variáveis explicativas, possivelmente contendo defasagens da variável dependente, $f_n(\cdot)$ relaciona variáveis antecedentes e homicídios futuros. O primeiro método aplicado a $f_n(\cdot)$ foi o extreme learning machine (Elmx), neste caso, o método considera as variáveis antecedentes ao estimar o modelo ótimo. Além disso, diferentes especificações foram estimadas por outras três abordagens. As duas primeiras não contêm variável dependente defasada no conjunto de variáveis antecedentes. Neste caso, aplicamos o modelo com defasagens distribuídas (TslmX) e modelo com erros ArimaX Arima,[5] de acordo com algoritmo de Hyndman e Khandakar (2008). A equação de previsão estimada tem a seguinte forma:

5. Metodologia sugerida por Pankratz (1991a; 1991b) indica não existir efeito de *feedback*.

$$\widehat{y}_{t+h} = \widehat{\alpha}_0 + \sum_{i=0}^{5} \widehat{\beta}_0' x_{t-i} + \widehat{\mu}_{t+h}, \text{h} = 1,2 \ldots, 12, \quad (2)$$

Em que $\widehat{y}_{t+h}$ é o número de homicídios previsto no horizonte de previsão h; $\widehat{\alpha}_0$, a estimação do termo constante; x_t , o vetor de variáveis previsoras; e , o termo de erro. O modelo 0 considera relação contemporânea entre homicídios e variável antecedente (Tslm0ts), em seguida, o modelo 1 acrescenta a primeira defasagem de internações (Tslm1ts). Em função das diferentes especificações sugeridas pelos tradicionais critérios de informação, da flexibilidade da metodologia de avaliação adotada e como forma de verificar desempenho preditivo de distintas especificações em variados horizontes de previsão, este processo é repetido até a quinta defasagem de internações por agressão (Tslm5ts). Em seguida, esse processo é aplicado ao modelo autorregressivo de defasagens distribuídas, ou seja, ao conjunto de variáveis antecedentes, acrescentamos uma defasagem da variável dependente, especificando o seguinte modelo:

$$\widehat{Y}_{t+h} = \widehat{\alpha}_0 + \widehat{\gamma}_1 \gamma_{t-1} = + \sum_{i=0}^{5} \widehat{\beta}_0' x_{t-i} + \widehat{\mu}_{t+h}, \text{h} = 1,2 \ldots, 12, |\gamma_1| < 1 \quad (3)$$

Novamente, o modelo 0 considera relação contemporânea entre homicídios e internações (Tslm0tsy). Similar ao caso anterior, o método é estimado até a quinta defasagem de internações por agressão (Tslm5tsy).

Ao avaliarmos o desempenho previsor dos modelos, utilizamos duas abordagens. No primeiro método de avaliação, empregamos estatísticas de precisão ao comparar desempenho preditivo, isto é, utilizamos RMSE, MAE e MAD. As estatísticas RMSE e MAE foram escolhidas em função da fácil interpretação dos resultados e ampla utilização na literatura (Hyndman, 2015). A estatística MAD é apresentada por ser robusta a assimetrias e *outliers*.[6] Em cada estatística, calculamos razão entre modelo considerado e modelo *random walk*. Resultado inferior à unidade sugere que os modelos considerados apresentam melhor desempenho preditivo relativamente ao *random walk*.

Apesar de ampla utilização na literatura, as estatísticas de precisão não consideram a significância estatística da diferença de desempenho preditivo dos modelos. Portanto, comparamos desempenho previsor dos diversos modelos pela abordagem MCS, proposta por Hansen, Asger e James (2011).

6. Descrição das estatísticas na tabela A.2 do apêndice.

O objetivo do MCS é determinar um conjunto, M^*, que consista nos melhores modelos do conjunto inicial de modelos, M_0, em que melhor é definido por critérios pré-definidos, isto é, pelo nível de significância, pela função de perda e pelas estatísticas de teste. O MCS elabora um conjunto de modelos de confiança, $\widehat{M}^*$, ou seja, desenvolve um conjunto de modelos que contém os melhores modelos dado certo nível de confiança. Os modelos em $\widehat{M}^*$ são avaliados por informação amostral acerca da *performance* relativa dos modelos em M_0. Assim, o MCS é aleatório conjunto de modelos *data dependence* que inclui o(s) melhor(es) modelo(s) de previsão. Uma das vantagens do MCS é não requerer modelo *benchmark* e permitir mais de um modelo ser o melhor, ou seja, M^* talvez contenha mais de um modelo.

O MCS é construído a partir do conjunto de modelos competidores, M_0, e critério de avaliação empírica. O procedimento é baseado em teste de equivalência, γ_M, e regra de eliminação e_M. Isto é, o procedimento consiste em sequência de testes, nos quais a hipótese nula de igualdade na capacidade preditiva não é rejeitada dado nível de significância. O teste de equivalência é aplicado ao conjunto de modelos $M = M_0$. Se γ_M for rejeitada, a evidência sugere que modelos em M não são igualmente bons e e_M é utilizada na eliminação de modelo de baixo desempenho em M. O procedimento é repetido até não ser rejeitada dado nível de significância, então, os modelos sobreviventes constituem conjunto de melhores modelos. Isto é, o MCS contém os melhores modelos do conjunto inicial M_0, dado certo nível de confiança $1-\alpha$.

O nível de significância, α é empregado em todos os testes, garantindo assintoticamente $P(M^* \subset \widehat{M}^*_{1-\alpha}) \geq 1-\alpha$. Os modelos serão ranqueados por seus p-valores. Modelos de p-valor inferior a α não estarão no subconjunto de melhores modelos M^*. Portanto, dado modelo $i \in M_0$, o *p-valor do MCS,* $\hat{p}_i$, é limiar no qual $i \in \widehat{M}^*_{1-\alpha}$ se e somente se $\hat{p}_i \geq \alpha$. Portanto, baixo $\hat{p}_i$ indica pouca chance de modelo i pertencer a conjunto de melhores modelos. Neste trabalho, os modelos em $\widehat{M}^*_{1-\alpha}$ serão selecionados considerando desempenho preditivo passados dos modelos em M_0.

Detalhadamente, considere conjunto, M_0, que contém número finito de modelos, $i=1,...,m_0$. Os modelos são avaliados ao longo da amostra $t = 1,..., n,$ em termos de função de perda do modelo i no período t, $L_{i,t}$. Definimos a *performance* relativa entre modelos como $d_{ij,t} = L_{i,t} - L_{j,t} \forall i,j \in M_0$. Logo, determinamos estatísticas de perdas amostrais relativas:

$$\bar{d}_{ij} = \frac{1}{n} \sum_{t=1}^{n} d_{ij,t} \qquad (4)$$

Representada perda amostral média do modelo *i* relativamente ao modelo *j*.

$$\bar{d}_{i.} = \frac{1}{m} \sum_{j \in M}^{n} \bar{d}_{ij.} \tag{5}$$

Representada perda amostral do modelo *i* comparada à perda média entre os modelos em M_0. Então, assumimos: $\mu_{ij} = E(d_{ij})$ e $\mu_{i.} = E(d_{i.})$ é finito e não depende de $t \forall i,j \in M_0$. Elaboramos *rank* de desempenho em termos da função de perda esperada, ou seja, modelo *i* é preferido ao *j* se $\mu_{ij} < 0$. Logo, definimos conjunto de melhores modelos:

$$M^* = \{i \in M^0 : \mu_{ij} \leq 0 \ \forall \ j \in M^0\} \tag{6}$$

A eliminação de modelos inferiores em M^0 é realizada verificando hipótese de igualdade na capacidade preditiva. Esta pode ser formulada de duas formas:

- $H_{0,M}: \mu_{ij} = 0 \ \forall \ i,j \in M$
- $H_{A,M}: \mu_{ij} \neq 0$ para algum $i,j \in M$

Ou:

- $H_{0,M}: \mu_{i} = 0 \ \forall \ i,j \in M$
- $H_{A,M}: \mu_{i} \neq 0$ para algum $i,j \in M$

Onde $M \subset M^0$. A partir destas estatísticas, construímos, respectivamente, as seguintes estatísticas *t*:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{var}(\bar{d}_{ij})}} \quad \text{e} \quad t_{i.} = \frac{\bar{d}_{i.}}{\sqrt{\widehat{var}(\bar{d}_{i.})}} \quad \text{para } \forall \ i,j \in M \tag{7}$$

Em que $\widehat{var}(\bar{d}_{ij})$ e $\widehat{var}(\bar{d}_{i.})$ são variâncias estimadas de $\bar{d}_{ij}$ e $\bar{d}_{i.}$, respectivamente. Os autores propõem duas estatísticas de teste utilizadas ao verificar hipótese nula de equivalência na capacidade preditiva dos modelos. Sendo a estatística t_{ij} avaliada pela seguinte estatística de teste: $T_{R,M} = \max_{i,j \in M} |tij|$.

A estatística $T_{R,M}$ assegura que o resultado da hipótese nula depende do modelo de maior perda relativa. Neste caso, o modelo eliminado obedece à seguinte regra de eliminação: $e_{R,M} = \arg \max_{i} \left(\sup_{j \in M} \frac{\bar{d}_{ij}}{\sqrt{\widehat{var}(\bar{d}_{ij})}} \right)$, ou seja, a regra de eliminação irá excluir o

modelo de maior estatística de teste *t*. Por sua vez, a estatística t_i será avaliada pela seguinte estatística de teste: $T_{max,\,M} = \underset{i \in M}{max}\, t_i$. Neste caso, o resultado da hipótese nula depende do maior desvio em relação à média da função de perda dos modelos investigados. Por fim, a regra de eliminação será $e_{max,M} = \arg \underset{i \in M}{max} \frac{\bar{d}_i}{\sqrt{\widehat{var}(\bar{d}_i)}}$, isto é, será excluído modelo de maior contribuição para a estatística de teste. Este modelo apresenta o pior desempenho relativamente à média dos modelos em *M*. A distribuição assintótica destas estatísticas de teste não é padronizada, portanto, o MCS utiliza método *bootstrap* ao estimar a distribuição destas estatísticas, detalhes em Hansen, Lunde e Nason (2010).

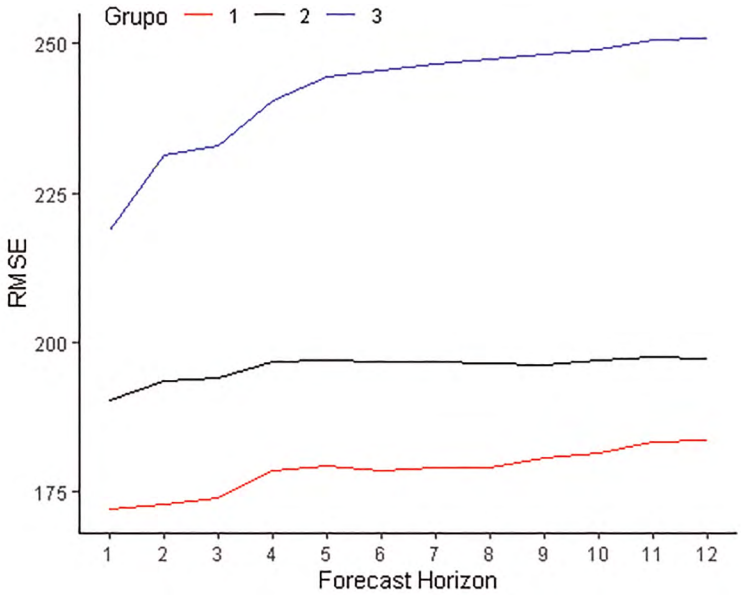
3 PRINCIPAIS RESULTADOS

Esta seção apresenta os resultados do exercício de previsão fora da amostra de homicídios. Para cada modelo estimado, computamos uma sequência de doze previsões fora da amostra, utilizando o método *rolling window* com janela de estimação de tamanho $T_s = 84$ observações mensais. Dessa forma, minimizam-se potenciais *outliers* e quebras estruturais, além de permitir implementar testes de capacidade preditiva em modelos aninhados (Giacomini e White, 2006). Em cada janela, os modelos são estimados e previsões de um a doze passos à frente são realizadas. Assim, a janela avança um mês e repete o processo anterior até o fim da amostra – similar a Medeiros, Vasconcelos e Freitas (2015) e Stock e Watson (2012). Isto resulta em 143 – *h* erros de previsões em cada horizonte de previsão *h*, os quais serão utilizados na avaliação de *performance* dos modelos estimados, selecionado como vencedor aquele de melhor desempenho fora da amostra. A seção 4 mostra que os principais resultados são estáveis para $T_s = 96$.

Ao analisar evidências qualitativas dos modelos estimados, o gráfico 1 expõe RMSE dos modelos estimados ao longo dos horizontes avaliados. O mesmo gráfico apresenta grupos de modelos agregados por similar desempenho preditivo. É possível perceber três padrões de desempenho preditivo. O grupo 1, composto por modelos (autorregressivos) de defasagem distribuída, apresenta ao longo de todos os horizontes de previsão o melhor desempenho preditivo, embora mostre gradual deterioração da capacidade preditiva ao longo dos horizontes de previsão. O conjunto dois, composto majoritariamente por modelos univariados, embora de desempenho inferior ao grupo 1, exibe estabilidade na capacidade preditiva. Por fim, o grupo 3, constituído sobretudo por modelos Arima com variáveis antecedentes, expõe desempenho preditivo inferior aos

demais grupos em todos horizontes de previsão e deterioração da capacidade preditiva ao longo dos horizontes de previsão.

GRÁFICO 1
RMSE por grupos preditivos e horizonte de previsão



Elaboração dos autores.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Seguindo Medeiros *et al*. (2019), a tabela 1 apresenta sumário de estatística do desempenho preditivo ao considerarmos todos os horizontes de previsão. As colunas (1), (2) e (3) apresentam RMSE, MAE e MAD médios, respectivamente. As colunas (4), (5) e (6) apresentam, respectivamente, RMSE, MAE e MAD máximos. Por sua vez, as colunas (7), (8) e (9) apresentam RMSE, MAE e MAD mínimos. Todas as estatísticas estão normalizadas relativamente ao modelo *random walk*. As colunas (10), (11) e (12) apresentam o número de vezes – ao longo do horizonte de previsão – cada modelo alcançou o menor RMSE, MAE e MAD, respectivamente. Na última linha, estão os valores apurados no modelo *random walk* de *benchmark*. A tabela 1 apresenta os resultados apurados de acordo com *rank* decrescente do RMSE médio.

As evidências sugerem que nas estatísticas de avaliação consideradas, majoritariamente os modelos competidores superam na média o modelo *random walk* – exceto o modelo *snaive*. Em geral, ganhos preditivos relativos ao *random walk*, em termos de RMSE e MAE são da ordem de 50% e aproximadamente 40% na estatística MAD.

Por sua vez, os melhores modelos de variáveis antecedentes apresentam melhorias relativas ao modelo *random walk* em torno de 60% em termos de RMSE, MAE e MAD. Além disso, modelos (autorregressivos) de defasagem distribuída com variáveis antecedentes sistematicamente obtêm melhores resultados preditivos relativamente aos demais modelos competidores. De acordo com as evidências encontradas, ao considerar o RMSE médio, o modelo univariado de melhor colocação (ETS) aparece somente na 11ª posição e, nas demais estatísticas de avaliação, modelos univariados não apresentam melhores resultados relativamente a modelos de variáveis antecedentes, pois, em nenhuma oportunidade modelos univariados mostram a menor estatística de desempenho preditivo. Ainda de acordo com as evidências, nenhum modelo apresenta desempenho sistematicamente superior aos demais, ou seja, não há modelo com melhor desempenho em todos os critérios de avaliação.

TABELA 1
Sumário estatísticas de previsão

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Model	Avg. RMSE	Avg. MAE	Avg. MAD	Max. RMSE	Max. MAE	Max. MAD	Min RMSE	Min. MAE	Min. MAD	#min RMSE	#min. MAE	#min. MAD
tslm1ts	0,41	0,39	0,36	0,64	0,62	0,51	0,36	0,33	0,30	6	3	0
tslm1tsy	0,42	0,38	0,27	0,62	0,59	0,40	0,36	0,33	0,22	6	9	8
tslm2ts	0,42	0,42	0,41	0,67	0,65	0,59	0,36	0,34	0,33	0	0	0
tslm2tsy	0,43	0,40	0,29	0,63	0,61	0,38	0,38	0,35	0,24	0	0	2
tslm3tsy	0,43	0,41	0,30	0,65	0,62	0,41	0,38	0,35	0,26	0	0	2
tslm3ts	0,43	0,42	0,40	0,68	0,66	0,56	0,37	0,35	0,33	0	0	0
tslm0tsy	0,44	0,42	0,40	0,66	0,64	0,55	0,39	0,36	0,31	0	0	0
tslm4tsy	0,44	0,41	0,34	0,66	0,62	0,48	0,39	0,35	0,28	0	0	0
tslm4ts	0,45	0,42	0,37	0,70	0,67	0,54	0,38	0,35	0,29	0	0	0
tslm5tsy	0,45	0,42	0,30	0,67	0,63	0,41	0,40	0,36	0,26	0	0	0
Ets	0,46	0,44	0,37	0,72	0,69	0,55	0,39	0,36	0,29	0	0	0
Tbats	0,46	0,44	0,38	0,71	0,67	0,52	0,40	0,36	0,30	0	0	0
Thiefe	0,46	0,44	0,38	0,71	0,68	0,54	0,40	0,37	0,30	0	0	0
Mediana	0,46	0,44	0,40	0,70	0,68	0,55	0,40	0,36	0,32	0	0	0
Stlme	0,46	0,44	0,39	0,73	0,70	0,57	0,40	0,36	0,29	0	0	0
BootStrap	0,46	0,44	0,36	0,71	0,68	0,53	0,40	0,37	0,28	0	0	0
Stlma	0,46	0,44	0,37	0,69	0,66	0,49	0,40	0,36	0,29	0	0	0
tslm0ts	0,46	0,45	0,40	0,73	0,71	0,55	0,40	0,37	0,31	0	0	0
Média	0,47	0,45	0,42	0,70	0,68	0,59	0,40	0,37	0,33	0	0	0
tslm5ts	0,47	0,44	0,38	0,73	0,68	0,53	0,40	0,36	0,30	0	0	0
Bats	0,47	0,45	0,40	0,74	0,71	0,58	0,41	0,38	0,34	0	0	0
arima2x	0,49	0,48	0,44	0,75	0,74	0,63	0,42	0,40	0,35	0	0	0

(Continua)

(Continuação)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Model	Avg. RMSE	Avg. MAE	Avg. MAD	Max. RMSE	Max. MAE	Max. MAD	Min RMSE	Min. MAE	Min. MAD	#min RMSE	#min. MAE	#min. MAD
arima1x	0,49	0,48	0,42	0,76	0,74	0,59	0,43	0,40	0,32	0	0	0
arima3x	0,50	0,48	0,41	0,77	0,74	0,57	0,43	0,40	0,33	0	0	0
arima2xy	0,50	0,48	0,37	0,76	0,75	0,51	0,44	0,40	0,30	0	0	0
arima4x	0,52	0,50	0,48	0,78	0,75	0,59	0,46	0,42	0,38	0	0	0
Thiefa	0,52	0,51	0,47	0,77	0,75	0,64	0,46	0,43	0,39	0	0	0
Arima	0,53	0,52	0,49	0,80	0,77	0,66	0,46	0,43	0,40	0	0	0
arima5x	0,54	0,51	0,48	0,81	0,76	0,61	0,47	0,43	0,38	0	0	0
arima1xy	0,54	0,51	0,36	0,76	0,73	0,46	0,49	0,44	0,29	0	0	0
arima5xy	0,54	0,53	0,35	0,80	0,78	0,46	0,49	0,45	0,29	0	0	0
arima0x	0,55	0,54	0,52	0,78	0,78	0,73	0,48	0,45	0,42	0	0	0
Struc	0,55	0,54	0,51	0,84	0,83	0,67	0,49	0,45	0,45	0	0	0
Nnetar	0,56	0,54	0,50	0,83	0,82	0,71	0,50	0,45	0,40	0	0	0
Elm	0,57	0,53	0,46	0,74	0,72	0,62	0,61	0,55	0,46	0	0	0
arima4xy	0,58	0,54	0,37	0,78	0,76	0,50	0,53	0,46	0,31	0	0	0
Elmx	0,61	0,61	0,62	0,94	0,95	0,85	0,53	0,51	0,57	0	0	0
arima3xy	0,63	0,56	0,38	0,80	0,77	0,53	0,58	0,49	0,31	0	0	0
Snaive	0,63	0,63	0,68	1,00	1,00	1,00	0,54	0,51	0,52	0	0	0
arima0xy	0,67	0,64	0,44	0,74	0,72	0,47	0,65	0,57	0,42	0	0	0
Arfima	0,69	0,68	0,65	0,96	0,93	0,85	0,61	0,58	0,54	0	0	0
RW	417,00	338,45	367,16	490,21	416,24	505,57	262,26	211,01	235,73			

Elaboração dos autores.
Obs.: Última linha consta valor apurado no *benchmark* RW.

A tabela 2 apresenta RMSE normalizado pelo modelo *random walk* de dois modelos com variáveis antecedentes de melhor RMSE médio e, também, de todos os modelos univariados e combinações, ao longo dos horizontes de previsão considerados. Os modelos de variáveis antecedentes são em todos os horizontes de previsão os que apresentaram maior ganho relativo ao *random walk*. Considerados todos os horizontes de previsão, na média, modelos de variáveis antecedentes são 13% melhores que o conjunto univariado e as combinações. Resultados similares são observados nas estatísticas MAE e MAD, apresentadas, no apêndice, nas tabelas A.4 e A.5, respectivamente.

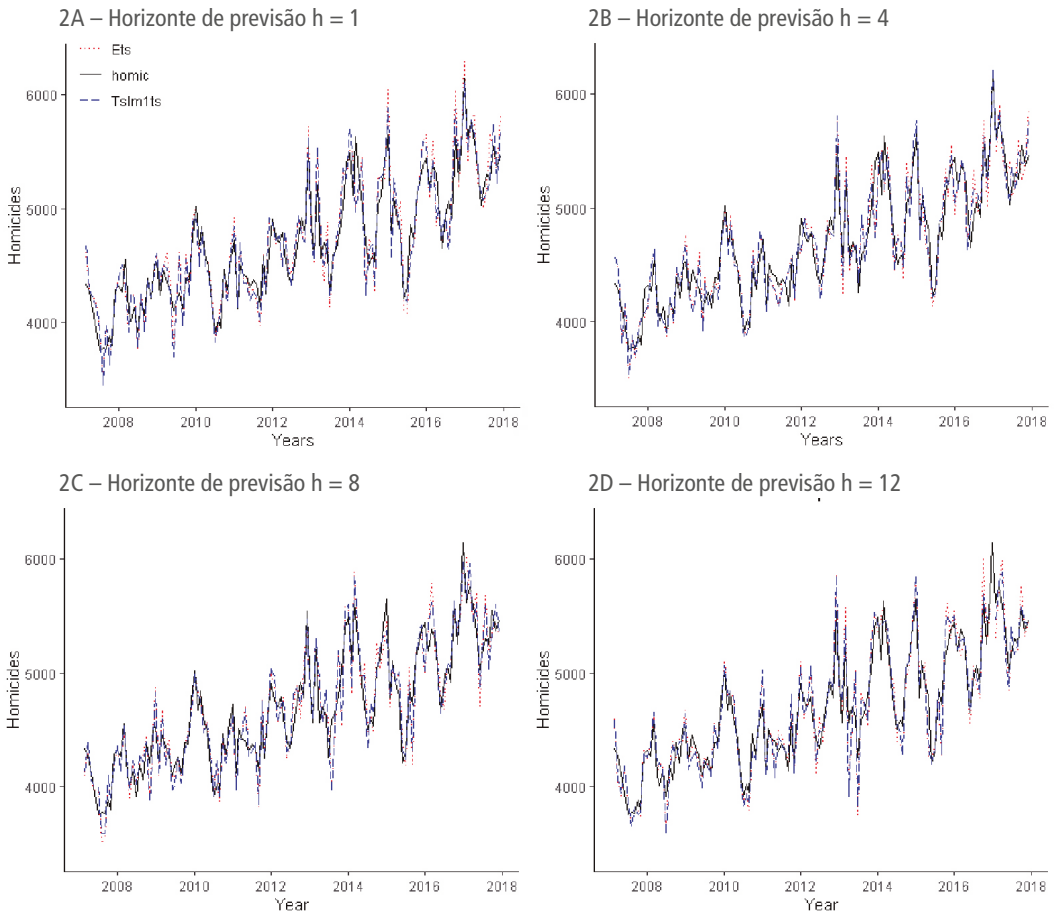
TABELA 2
RMSE normalizado por horizonte de previsão

	Horizonte de Previsão											
Modelo	1	2	3	4	5	6	7	8	9	10	11	12
tslm1ts	0,37	0,43	0,46	0,36	0,42	0,35	0,42	0,37	0,43	0,42	0,37	0,65
tslm1tsy	0,34	0,40	0,44	0,36	0,42	0,35	0,42	0,37	0,44	0,43	0,38	0,67
Arima	0,44	0,55	0,59	0,47	0,55	0,45	0,54	0,48	0,57	0,56	0,49	0,85
BootStrap	0,39	0,48	0,52	0,40	0,48	0,39	0,48	0,42	0,49	0,49	0,42	0,74
Bats	0,41	0,49	0,54	0,41	0,48	0,39	0,49	0,43	0,49	0,49	0,43	0,75
Tbats	0,39	0,46	0,52	0,40	0,47	0,39	0,47	0,42	0,48	0,48	0,42	0,74
Stlme	0,41	0,47	0,52	0,41	0,47	0,39	0,48	0,42	0,48	0,48	0,42	0,73
Stlma	0,38	0,47	0,52	0,41	0,47	0,40	0,48	0,42	0,49	0,49	0,43	0,75
Ets	0,40	0,47	0,51	0,40	0,47	0,39	0,47	0,41	0,48	0,47	0,42	0,73
Nnetar	0,49	0,56	0,59	0,49	0,57	0,50	0,58	0,51	0,57	0,61	0,51	0,90
Snaive	0,55	0,64	0,71	0,55	0,64	0,53	0,65	0,57	0,67	0,65	0,58	1,00
Arfima	0,53	0,70	0,75	0,59	0,74	0,59	0,74	0,64	0,75	0,74	0,65	1,12
Struc	0,46	0,56	0,63	0,50	0,57	0,47	0,58	0,50	0,60	0,58	0,51	0,86
Thiefa	0,42	0,55	0,57	0,46	0,54	0,45	0,55	0,48	0,56	0,54	0,49	0,85
Thiefe	0,40	0,47	0,51	0,41	0,47	0,39	0,47	0,42	0,47	0,48	0,42	0,74
Mlp	0,41	0,51	0,51	0,40	0,61	0,51	0,74	0,66	0,84	0,82	0,80	1,46
Elm	0,41	0,49	0,54	0,42	0,54	0,45	0,62	0,56	0,66	0,67	0,64	1,13
Média	0,39	0,47	0,52	0,41	0,48	0,40	0,48	0,43	0,49	0,49	0,43	0,75
Mediana	0,38	0,47	0,51	0,40	0,47	0,39	0,48	0,42	0,49	0,48	0,42	0,74
RW	478,07	406,06	371,41	476,74	406,13	490,21	404,08	458,72	394,10	400,43	455,82	262,26

Elaboração dos autores.
Obs.: Última linha consta valor apurado no *benchmark* RW.

O gráfico 2 apresenta os valores realizados do número de homicídios ao lado das previsões do modelo univariado (ETS) e de variáveis antecedentes (Tslm1ts) de menor RMSE médio. Observamos que apesar do bom ajuste visual, o modelo ETS apresenta sistematicamente erros maiores àqueles observados no modelo de variáveis antecedentes, especialmente em horizontes de previsão mais longos.

GRÁFICO 2
Valores realizados de homicídios e previsões selecionadas estimadas por *rolling window*
ts = 84



Elaboração dos autores.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Seguindo Marcellino, Stock e Watson (2006) e Tarassow (2019), reportamos resultados de 1, 4, 8 e 12 horizontes de previsão. A tabela 3 apresenta o MCS ao considerar todos os modelos estimados desde 1999m01, assumindo função de perda quadrática. Como sumário, listamos modelos de melhor classificação pelo MCS.

TABELA 3
Top 10 MCS resultados baseados na *rolling window* de largura ts = 84 para homicídios (jan. 1999-dez. 2017)

Forecast Horizon h = 1				Forecast Horizon h = 4			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	26588,20	1,000	tslm1tsy	1	28817,67	1,000
tslm2tsy	2	27711,11	1,000				
tslm3tsy	3	28817,38	1,000				
tslm4tsy	4	29855,85	1,000				
tslm0tsy	5	29795,31	0,999				
tslm1ts	6	30923,70	0,977				
Forecast Horizon h = 8				Forecast Horizon h = 12			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1ts	1	28608,10	1,000	tslm1ts	1	28639,28	1,000

Elaboração dos autores.
Nota: [1] *Loss* refere-se ao valor médio da função de perda quadrática. Foram utilizadas 5 mil iterações *block-bootstrap*. O tamanho do *block bootstrap* foi determinado pelo número máximo de parâmetros significativos após ajustar processo AR (p) em todas as funções diferenças, como sugerido por Hansen *et al*. (2011). Ainda seguindo os autores do método, escolhemos nível de significância $\alpha = 10\%$ na elaboração do MCS ($\hat{M}^{*}_{90}$).
Obs.: Estatística t_{max} e função de perda quadrática[1].

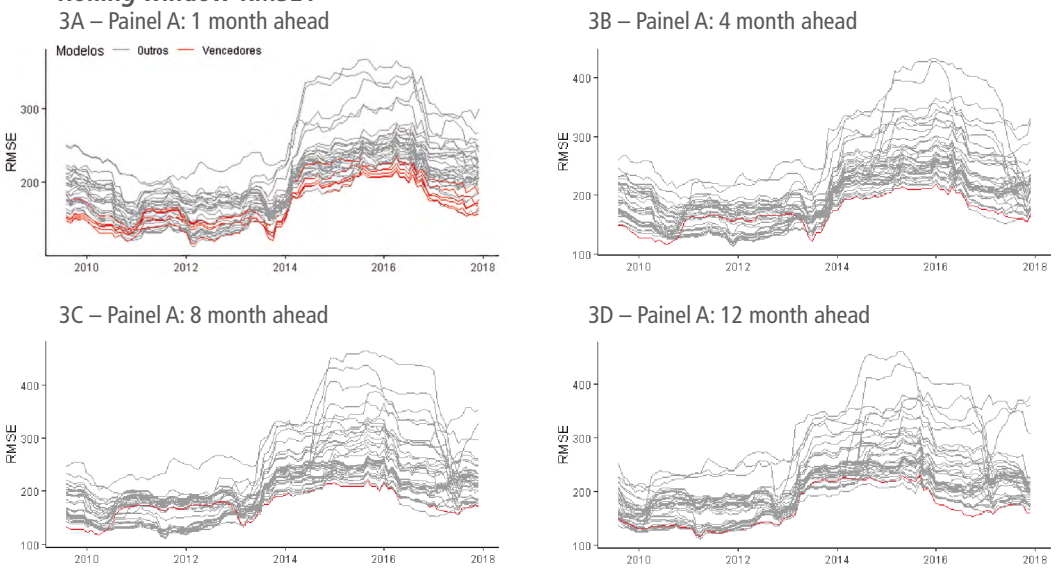
Alinhada às evidências anteriores, a tabela 3 sugere domínio dos modelos de variáveis antecedentes ao longo dos horizontes de previsão considerados. Especificamente, diferentes especificações do modelo (autorregressivo) de defasagem distribuída com até quatro defasagens estão no topo das previsões em t + 1, com diferença de 16% na função quadrática de perda estimada entre o modelo vencedor e o último colocado. Nos horizontes de previsão seguintes, o MCS selecionou somente dois modelos. Nomeadamente, no caso t + 4, modelo sugerido por método de identificação proposto por Pankratz (1999) e critérios BIC e HQ, apresentou melhor desempenho preditivo fora da amostra, enquanto em t + 8 e t + 12, modelos especificados por critérios tradicionais de identificação não apresentaram resultado preditivo vencedor, justificando a inclusão de outras especificações além das sugeridas por critérios de informação. Merece destaque a ausência no MCS de modelos com erros Arima estimados pelo algoritmo de Hyndman e Khandakar (2008) em todos os horizontes apresentados. Por fim, similar ao tradicionalmente observado na literatura de *forecasting*, o resultado sugere não existir modelo superior aos demais em todos os horizontes de previsão.

Uma fragilidade na avaliação do desempenho preditivo reside no fato de a precisão fora da amostra depender do período utilizado na avaliação das previsões. Isto é, modelos serão escolhidos pelo melhor desempenho preditivo ao considerar estimativa da função perda ao longo de todo o período de avaliação fora da amostra, no entanto, modelos concorrentes podem exibir, por exemplo, menor RMSE em determinado período da

amostra de avaliação. Para visualizar a estabilidade na capacidade preditiva dos modelos, observamos a variação temporal do RMSE, ao executar *rolling window* de tamanho $T_e = 30$ observações mensais.

O gráfico 3 apresenta a variação temporal do RMSE de todos os modelos estimados e destaca aqueles selecionados pelo MCS. Em todos os horizontes de previsão analisados, observamos gradual degradação na qualidade das previsões ao longo da amostra. No entanto, os modelos escolhidos pelo MCS, exceto em curtos períodos, apresentam melhor desempenho preditivo ao longo de toda a amostra. A evidência sugere estabilidade no desempenho preditivo dos modelos e, portanto, reforça os achados do MCS.[7]

GRÁFICO 3
***Rolling window* RMSE1**



Elaboração dos autores.
Nota: [1] Modelos escolhidos pelo MCS.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

As evidências apresentadas no gráfico 3 confirmam relevância das variáveis antecedentes utilizadas em termos de ganho na *performance* previsora. Modelos univariados são sistematicamente superados nos diferentes horizontes de previsão apresentados. Além disso, os resultados apontam alteração da especificação ótima, de acordo com o horizonte

7. No apêndice A, a figura A.1 apresenta esse exercício de acordo com a categoria do modelo – univariado, Arima ou (autorregressivo) de defasagens distribuídas. Novamente, as evidências estão alinhadas aos achados do MCS, isto é, modelos de variáveis antecedentes com defasagem distribuída apresentam os melhores resultados.

de previsão analisado. Considerando estatística de teste t_{rm}, os modelos selecionados pelo MCS são subconjunto do caso t_{max}, independentemente do tamanho da janela de estimação, como visto no apêndice A, tabelas A.6 e A.7, reforçando a validade dos achados.

Ao repetirmos o exercício de previsão, encontramos evidências de estabilidade nos resultados gerais ao utilizar no MCS função de perda absoluta em vez de função de perda quadrática. De acordo com as evidências apresentadas na tabela 4, novamente observamos o domínio dos modelos de variáveis antecedentes ao longo dos horizontes de previsão considerados. Similar ao observado na tabela 3, o modelo sugerido nos critérios de informação HQ e BIC e método de identificação de Pankratz (1999) aparece como o de melhor desempenho preditivo fora da amostra, nos horizontes de previsão t +1 e t + 4. Ainda nessa linha, os resultados em t + 8 e t + 12 repetem aqueles da tabela 3, ao não selecionar modelo indicador pelos métodos de identificação utilizados. Também, similar ao observado no exercício anterior, nenhum modelo supera os concorrentes em todos os horizontes de previsão. Estes resultados são qualitativamente iguais aos apurados, considerando estatística de teste t_{rm} e função de perda absoluta, como observado na tabela A.7 do apêndice.

TABELA 4
Top 10 MCS resultados baseados na rolling window de largura ts = 84 para mortes violentas intencionais[1,2] (jan. 1999-dez. 2016)

Forecast Horizon h = 1				Forecast Horizon h = 4			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	124,49	1,000	tslm1tsy	1	129,97	1,000

Forecast Horizon h = 8				Forecast Horizon h = 12			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	128,38	1,000	tslm1ts	1	130,36	1,000

Elaboração dos autores.
Nota: [1] Mais detalhes na tabela 3.
[2] Estatística $\langle t \rangle$_max e função de perda absoluta.

4 *CHECK* DE ROBUSTEZ

Na presença de quebra estrutural, os erros de previsão do método *rolling window* estão sujeitos à largura da janela de estimação (Pesaran e Timmermann, 2007). Contornamos este problema ampliando a largura da janela de estimação para Ts = 96 ao invés de Ts = 84 observações mensais.

A tabela 5 apresenta os achados desse exercício. Em primeiro lugar, os modelos vencedores apresentam maiores erros de previsão que aqueles reportados na tabela 3, indicando que as previsões são mais precisas na média ao escolhermos a janela Ts = 84. Apesar disso, os resultados gerais são similares aos observados anteriormente. Modelos com variáveis antecedentes superam sistematicamente os modelos univariados em todos os horizontes de previsão considerados. Estes achados são robustos à estatística de avaliação t_{rm} e função de perda, como observado nas tabelas A.6 e A.7 do apêndice – em que realizamos este exercício considerando estatística e função de perda absoluta – salientando os ganhos de previsão alcançados ao utilizar variáveis antecedentes.

TABELA 5
Top 10 MCS resultados baseados na *rolling window* de largura ts = 96 para mortes violentas intencionais[1] (jan. 1999-dez. 2016)

Forecast Horizon h = 1				Forecast Horizon h = 4			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	27376,13	1,000	tslm1tsy	1	29103,78	1,000
				tslm1ts	2	29511,11	0,923
Forecast Horizon h = 8				Forecast Horizon h = 12			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	29918,16	1,000	tslm1ts	1	30492,44	1,000
tslm1ts	2	30052,73	0,972				

Elaboração dos autores.
Nota: [1] Estatística t_{max} e função de perda quadrática.
Obs.: Para mais detalhes, ver tabela 3.

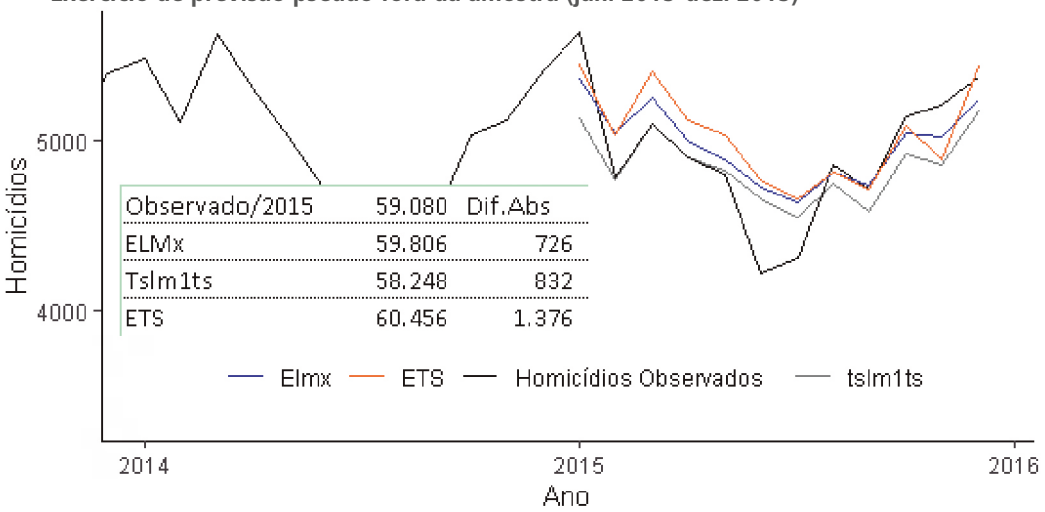
5 EXERCÍCIO DE PREVISÃO PSEUDO FORA DA AMOSTRA

Tradicionalmente, ao avaliar a capacidade preditiva de modelos previsores, a abordagem *rolling window* aparece como preferida (Medeiros *et al.*, 2019; Medeiros, Vasconcelos e Freitas, 2015; Panagiotelis *et al.*, 2019; Tarassow, 2019), pois atenua problemas relacionados a *outliers* e quebra estrutural. Além disso, ela produz série de erros preditivos em diferentes horizontes de previsão, utilizados na avaliação empírica dos modelos considerados, similar ao observado nas seções anteriores. No entanto, em função da estreita janela de estimação imposta nesta abordagem, métodos cujo desempenho preditivo acompanha o tamanho do conjunto de informações inseridas no modelo irão apresentar baixo desempenho preditivo. Em nossa análise, o método ELMX, ao contrário do esperado, não pertence ao grupo de melhores modelos escolhidos pelo MCS em nenhum

horizonte de previsão considerado, resultado similar ao observado em outros trabalhos (Haykin, 2009; Ploeg, Austin e Steyerberg, 2014).

A abordagem considerada nesta seção traz evidências qualitativas do desempenho preditivo ao ampliarmos a janela de estimação. Nesta seção, utilizamos a abordagem *fixed origin*, ou seja, ao fim do período de ajuste (T), são realizadas previsões referentes a T + 1, T + 2,..., T + 12. Buscando capturar desempenho preditivo em diversos ciclos, realizamos este exercício a partir de 2014, 2015 e 2016, sendo a amostra de teste os doze meses seguintes ao período de ajuste. Consideramos método univariado e de defasagens distribuídas de menor RMSE médio, além do método ELMX.[8]

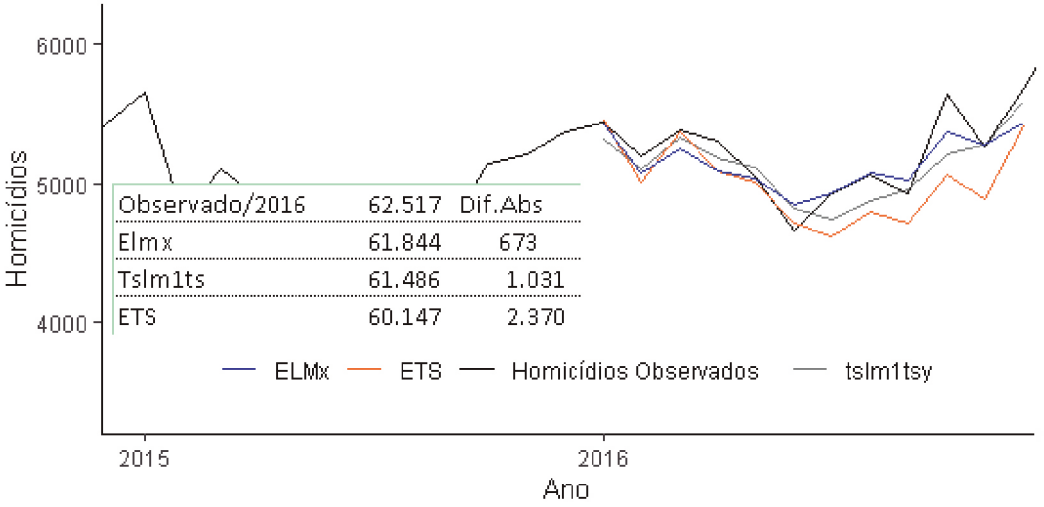
GRÁFICO 4
Exercício de previsão pseudo fora da amostra (jan. 2015-dez. 2015)



Elaboração dos autores.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

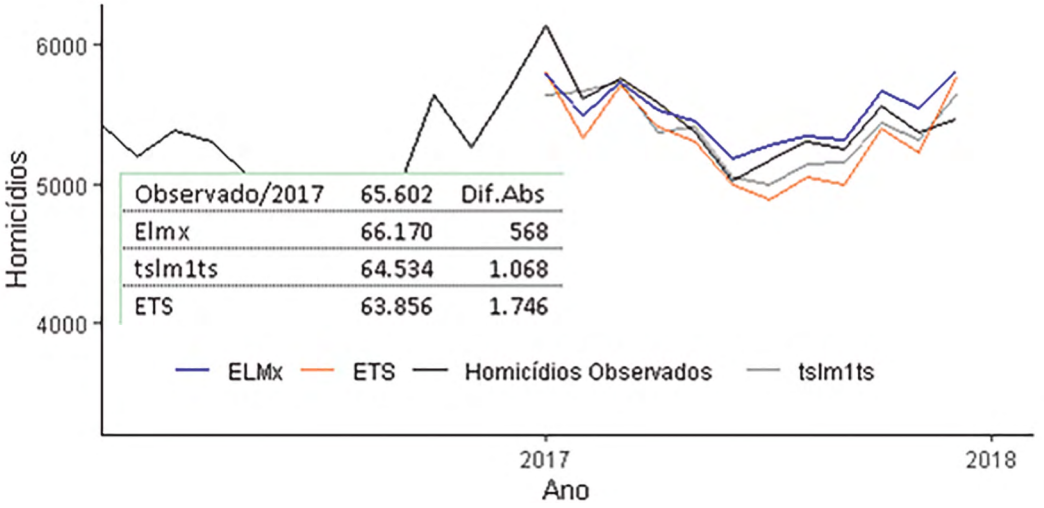
8. Estatísticas de precisão serão omitidas, pois não são capazes de separar erros em diferentes horizontes de previsão.

GRÁFICO 5
Exercício de previsão pseudo fora da amostra (jan. 2016-dez. 2016)



Elaboração dos autores.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO 6
Exercício de previsão pseudo fora da amostra (jan. 2017-dez. 2017)



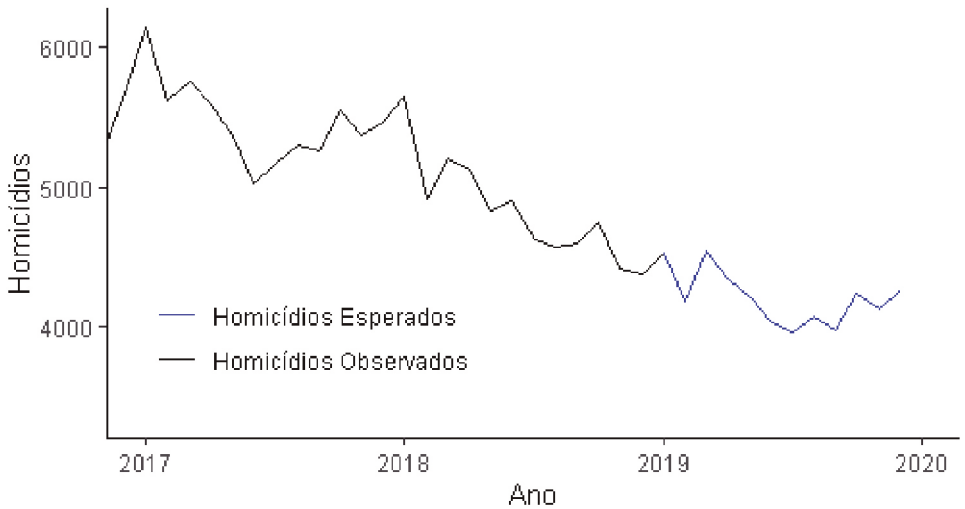
Elaboração dos autores.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Ao considerar a diferença absoluta entre o número acumulado de homicídios em 2015, 2016 e 2017 e o valor acumulado previsto nestes anos, são evidentes os ganhos de precisão pontual dos modelos multivariados e o desempenho superior do modelo ELMX nos diferentes períodos considerados.

5.1 Exercício de previsão de homicídios

Nesta seção, realizamos exercício de previsão fora da amostra, ou seja, elaboramos previsão do número de homicídios no Brasil de janeiro a dezembro de 2019 – período sem dados oficiais de homicídios publicados pelo SIM.

GRÁFICO 7
Trajetória prevista de homicídios (jan. 2019-dez. 2019)



Elaboração dos autores.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

De acordo com o resultado apurado, a expectativa do número total de homicídios no Brasil em 2019 é de 50.935, retração de 12,11% em relação ao observado no ano anterior.

6 CONCLUSÕES

A tendência crescente e os recordes em números absolutos de homicídios no Brasil despertaram o interesse de acadêmicos e *policymakers* nas causas e consequências da criminalidade violenta brasileira. Apesar do elevado custo econômico e social (Cerqueira e Soares, 2016) incorrido pela sociedade e da necessidade de acompanhamento conjuntural da criminalidade pelos interessados, as informações nacionais de criminalidade são divulgadas com defasagem.

Com o objetivo de produzir evidências preliminares acerca da criminalidade violenta nacional, este estudo inovou ao avaliar a capacidade de modelos com variáveis

antecedentes, entre elas, internações por agressão, apresentando desempenho preditivo do número de homicídios no Brasil superior a um conjunto de modelos univariados.

De acordo com os achados, modelos de variáveis antecedentes apresentam sistematicamente desempenho preditivo superior relativamente aos modelos univariados, embora não exista evidência de modelo multivariado superior aos demais em todos os horizontes de previsão. Na média, relativamente ao modelo *random walk*, os ganhos dos modelos de variáveis antecedentes e defasagens distribuídas em termos de RMSE, MAE e MAD são da ordem de 60%.

Os resultados gerais são robustos ao uso alternativo de estatística de avaliação, função de perda e largura de janela de estimação. Dessa forma, trazemos evidências da capacidade preditiva de internações por agressão aos pesquisadores interessados em aumentar a acurácia de suas previsões acerca da criminalidade violenta e estabelecemos *benchmark* a futuras tentativas de previsão de homicídios no Brasil.

Apesar das evidências, esta análise não esgota e, não buscou esgotar, possíveis abordagens de previsão. Análises futuras deverão considerar interação geográfica (Gorr, Olligschlaeger e Thompson, 2003; Shoesmith, 2013) ou ampliar o leque de variáveis antecedentes e aplicar métodos de regularização ou abordagem geral para o específico.

REFERÊNCIAS

ATHANASOPOULOS, G. *et al*. Forecasting with temporal hierarchies. **European Journal of Operational Research**, v. 262, n. 1, p. 60-74, Oct. 2017.

BERGMEIR, C.; HYNDMAN, R. J.; BENITEZ, J. M. Bagging Exponential Smoothing Methods using STL Decomposition and Box-Cox Transformation. **International Journal of Forecasting**, v. 32, n. 2, p. 303-312, 2016.

CANOVA, F.; GHYSELS, E. Changes in seasonal patterns. **Journal of Economic Dynamics and Control**, v. 18, n. 6, p. 1143-1171, Nov. 1994.

CERQUEIRA, D. R. C. Mortes violentas não esclarecidas e impunidade no Rio de Janeiro. **Economia Aplicada**, v. 16, n. 2, p. 201-235, 2012.

_____. **Causas e consequências do crime no Brasil**. Rio de Janeiro: BNDES, 2014.

CERQUEIRA, D. R. C.; SOARES, R. R. The welfare cost of homicides in brazil: accounting for heterogeneity in the willingness to pay for mortality reductions. **Health Economics**, v. 25, n. 3, p. 259-276, Mar. 2016.

CERQUEIRA, D. R. C. **Atlas da Violência 2018**. Rio de Janeiro: Ipea; FBSP, 2018.

CERQUEIRA, D. R. C. *et al*. **Uma análise da base de dados do Sistema de Informação Hospitalar entre 2001 e 2018**: dicionário dinâmico, disponibilidade dos dados e aspectos metodológicos para a produção de indicadores sobre violência. Rio de Janeiro: Ipea, 2019.

CLEVELAND, R. B.; CLEVELAND, W. S.; TERPENNING, I. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. **Journal of Official Statistics**, v. 6, n. 1, p. 3 73, 1990.

DAVIDSON, R.; MACKINNON, J. **Econometric theory and methods**. Cambridge, United Kingdon: Oxford University Press, 2004.

GIACOMINI, R.; WHITE, H. Tests of Conditional Predictive Ability. **Econometrica**, v. 74, n. 6, p. 1545-1578, 2006.

GORR, W.; HARRIES, R. Introduction to crime forecasting. **International Journal of Forecasting**, v. 19, n. 4, p. 551-555, Oct. 2003.

GORR, W.; OLLIGSCHLAEGER, A.; THOMPSON, Y. Short-term forecasting of crime. **International Journal of Forecasting**, v. 19, n. 4, p. 579-594, Oct. 2003.

HANSEN, P. R.; LUNDE, A.; NASON, J. M. The Model Confidence Sept. **SSRN Electronic Journal**, 2010.

HANSEN, P. R.; ASGER, L.; JAMES, M. N. The model confidence set. **Econometrica**, v. 79, n. 2, p. 453-497, 2011.

HARRIES, R.; SOLLIS, R. Testing for Unit Roots. *In*: HARRIES, R.; SOLLIS, R. **Applied time series modelling and forecasting**. 1. ed. Nova Jersey: John Wiley & Sons, 2003.

HARVEY, A. C. **Forecasting, structural time series models and the Kalman filter**. Cambridge: Cambridge University Press, 1990.

HASLETT, J.; RAFTERY, A. E. Space-time modelling with long-memory dependence: assessing ireland's wind power resource. **Applied Statistics**, v. 38, n. 1, p. 1, 1989.

HAYKIN, S. S. **Neural networks and learning machines**. 3. ed. London: Pearson Education, 2009.

HUANG, G.-B.; ZHU, Q.; SIEW, C. Extreme learning machine: theory and applications. **Neurocomputing**, v. 70, n. 1-3, p. 489-501, Dec. 2006.

HYNDMAN, R. J. *et al*. **Forecasting with exponential smoothing**: the state space approach. Berlin; Heidelberg: Springer-Verlag, 2008.

HYNDMAN, R. J. *et al*. A state space framework for automatic forecasting using exponential smoothing methods. **International Journal of Forecasting**, v. 18, n. 3, p. 439-454, July 2002.

HYNDMAN, R. J. Forecasting performance evaluation and reporting. *In*: TASHMAN, L.; GILLILAND, M.; SGLAVO, U. **Business forecasting**: the emerging role of artificial intelligence and machine learning. Hoboken, New Jersey: John Wiley & Sons, p. 177-184, 2015.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting**: principles and practice. 2. ed. Melbourne: 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.

KALMAN, R. E. A New approach to linear filtering and prediction problems. **Journal of Basic Engineering,** v. 82, n. 1, p. 35, 1960.

KLEPINGER, D. H.; WEIS, J. G. Projecting crime rates: an age, period, and cohort model using ARIMA techniques. **Journal of Quantitative Criminology**, v. 1, n. 4, p. 387-416, Dec. 1985.

KOURENTZES, N.; BARROW, D. K.; CRONE, S. F. Neural network ensemble operators for time series forecasting. **Expert Systems with Applications**, v. 41, n. 9, p. 4235-4244, July 2014.

LIVERA, A. M. de; HYNDMAN, R. J.; SNYDER, R. D. Forecasting time series with complex seasonal patterns using exponential smoothing. **Journal of the American Statistical Association**, v. 106, n. 496, p. 1513-1527, 2011.

MARCELLINO, M.; STOCK, J. H.; WATSON, M. W. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. **Journal of Econometrics**, v. 135, n. 1-2, p. 499-526, Nov. 2006.

MEDEIROS, M. C. *et al*. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. **Journal of Business & Economic Statistics**, p. 1-22, 19 Aug. 2019.

MEDEIROS, M. C.; VASCONCELOS, G.; FREITAS, E. Forecasting Brazilian Inflation with High-Dimensional Models. **Brazilian Review of Econometrics**, v. 99, n. 99, p. 223-254, 2015.

MURRAY, J.; CERQUEIRA, D. R. DE C.; KAHN, T. Crime and violence in Brazil: systematic review of time trends, prevalence rates and risk factors. **Aggression and Violent Behavior**, v. 18, n. 5, p. 471-483, Sept. 2013.

ORD, J. K.; KOEHLER, A. B.; SNYDER, R. D. Estimation and prediction for a class of dynamic nonlinear statistical models. **Journal of the American Statistical Association**, v. 92, n. 440, p. 1621-1629, Dec. 1997.

PANAGIOTELIS, A. *et al*. Macroeconomic forecasting for Australia using a large number of predictors. **International Journal of Forecasting**, v. 35, n. 2, p. 616-633, Apr. 2019.

PANKRATZ, A. A primer on ARIMA Models. *In*: PANKRATZ, A. **Forecasting with dynamic regression models**. 1. ed. London: John Wiley & Sons, p. 24-81, 1991a.

_____. Dynamic Regression Models in a Vector ARMA Framework. *In*: _____. **Forecasting with dynamic regression models**. 1. ed. London: John Wiley & Sons, p. 342-356, 1991b.

_____. Building Dynamic Regression Models: Model Identification *In:* PANKRATZ, A. **Forecasting with dynamic regression models**. [s.l: s.n.]. 1991c. p. 167-201.

PEPPER, J. V. Forecasting crime: a city-level analysis. *In*: ROSENFELD, Richard, GOLDBERGER, Arthur S. **Understanding crime trends**: workshop report. Washington, D.C.: National Academies Press, p. 177-210, 2008.

PESARAN, M. H.; TIMMERMANN, A. Selection of estimation window in the presence of breaks. **Journal of Econometrics**, v. 137, n. 1, p. 134-161, Mar. 2007.

PLOEG, T. Van Der; AUSTIN, P. C.; STEYERBERG, E. W. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. **BMC Medical Research Methodology**, v. 14, n. 1, p. 137, 22 Dec. 2014.

PROVENZA, M. M. *et al*. Análise estatística e previsão de séries temporais de roubos e furtos a transeuntes no município do Rio de Janeiro de 2009 a 2013. **Cadernos do IME – Série Estatística**, v. 39, n. 2, p. 17-34, 23 dez. 2015.

PROVENZA, M. M.; COSTA, J. F. da S.; SILVA, L. D. C. análise e previsão de séries temporais do homicídio doloso no Rio de Janeiro. **Revista Cadernos de Estudos Sociais e Políticos**, v. 4, n. 7, p. 63-85, 2015.

ROEDER, O. K. *et al.* What Caused the Crime Decline? **Available at SSRN 2566965**, p. 142, 2015.

SALAKEN, S. M. *et al*. Extreme learning machine based transfer learning algorithms: a survey. **Neurocomputing**, v. 267, p. 516-524, Dec. 2017.

SAMUELS, J. D.; SEKKEL, R. M. Model Confidence Sets and forecast combination. **International Journal of Forecasting**, v. 33, n. 1, p. 48-60, 2017.

SANTOS, M. J. dos. Dinâmica temporal da criminalidade: mais evidências sobre o "efeito inércia" nas taxas de crimes letais nos estados brasileiros. **Revista Economia**, v. 10, n. 1, p. 169-194, 2009.

SHOESMITH, G. L. Space-time autoregressive models and forecasting national, regional and state crime rates. **International Journal of Forecasting**, v. 29, n. 1, p. 191-201, Jan. 2013.

SOUZA, F. M. de *et al*. Reductions in firearm-related mortality and hospitalizations in brazil after gun control. **Health Affairs**, v. 26, n. 2, p. 575-584, 1º Mar. 2007.

STOCK, J. H.; WATSON, M. W. Generalized shrinkage methods for forecasting using many predictors. **Journal of Business & Economic Statistics**, v. 30, n. 4, p. 481-493, Oct. 2012.

TARASSOW, A. Forecasting U.S. money growth using economic uncertainty measures and regularisation techniques. **International Journal of Forecasting**, v. 35, n. 2, p. 443-457, 2019.

TASHMAN, L. J. Out-of-sample tests of forecasting accuracy: an analysis and review. **International Journal of Forecasting**, v. 16, n. 4, p. 437-450, Oct. 2000.

WAN, C. *et al*. Probabilistic forecasting of wind power generation using extreme learning machine. **IEEE Transactions on Power Systems**, v. 29, n. 3, p. 1033-1044, May 2014.

APÊNDICE A

QUADRO A.1
Descrição das variáveis e fonte dos dados

Variável	Descrição da Variável	Fonte	Referência.
Homicídios	Agregação de óbitos por causas externas e agressão por intervenção legal.	Sistema de Informações sobre Mortalidade – SIM. Ministério da Saúde.	Óbitos segundo as seguintes CIDs-10: X85-Y09 Agressões, Y35-Y36 Intervenções legais e operações de guerra.
Internações	Número de internações registradas no SIH provenientes de agressão.	Sistema de Informações Hospitalar - SIH . Ministério da Saúde	Internações cuja causal primária e/ou secundária remete ao código CID relativo ao capítulo XX da classificação CID-10 de doenças. (Cerqueira *et al*., 2019)

Elaboração dos autores.

TABELA A.1
Estatísticas descritivas

	Internações	Homicídios
Média	3.971	4.341
Mediana	4.048	4.272
Máximo	5.339	5.677
Mínimo	2.464	3.341
Desvio Padrão	6.408	5.020

Elaboração dos autores.

Estatísticas de avaliação da previsão:

Raiz do erro quadrado médio (RMSE)

$$RMSE_{m,h} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}\text{^}2_{t,m,h}}$$

Erro absoluto médio (MAE)

$$MAE_{m,h} = \frac{1}{n}\sum_{t=1}^{n} |\hat{\epsilon}_{t,m,h}$$

Desvio absoluto médio da média (MAD)

$$MAE_{m,h} = \text{média}\,(|\hat{\epsilon}_{t,m,h} - \text{média}\,(\hat{\epsilon}_{t,m,h})|)$$

Em que: $\hat{\epsilon}_{t,m,h} = y_t - \hat{y}_{t,m,h}$ e $\hat{y}_{t,m,h}$ é a previsão de mortes violentas intencionais no período t, modelo m relativo ao horizonte h e n é o número de erros de previsão.

TABELA A.2
Teste de raiz unitária[1]

				ADF			
Variables	Test	eq.	lag	t-stat	10%	5%	1%
Homicídio	Level	N	12	2.408	-1.615	-1.942	-2.575
	1st difference	C	11	-6.796	-2.573	-2.874	-3.459
Internações	Level	N	1	0.153	-1.615	-1.942	-2.575
	1st difference	N	0	-21.669	-1.615	-1.942	-2.575
				pp			
		eq.	Band	t-stat	10%	5%	1%
Homicídio	Level	C/T	0	-7.053	-3.138	-3.429	-3.998
	1st difference	N	10	-33.207	-1.615	-1.942	-2.575
Internações	Level	C/T	7	-4.199	-3.138	-3.429	-3.998
	1st difference						
				Kpss			
		eq.	Band	t-stat	10%	5%	1%
Homicídio	Level	C/T	12	0.297	0.119	0.146	0.216
	1st difference	C	30	0.212	0.347	0.463	0.739
Internações	Level	C/T	15	0.110	0.119	0.146	0.216
	1st difference						

Elaboração dos autores.
Nota: [1] Constante (C) e tendência (T) incluídas de acordo com critério de Shcwarz. ADF: escolha de defasagens baseada no critério de Schwarz. PP e KPSS: Bartlett Kernell é o método de estimação espectral utilizado e *Bandwidth Newey West* foi utilizado.

TABELA A.3
MAE normalizado por horizonte de previsão

	Horizonte de Previsão											
Model	1	2	3	4	5	6	7	8	9	10	11	12
tslm1ts	0,36	0,43	0,44	0,36	0,40	0,32	0,41	0,35	0,40	0,40	0,35	0,62
tslm1tsy	0,33	0,39	0,41	0,35	0,39	0,31	0,40	0,34	0,39	0,40	0,36	0,64
Arima	0,44	0,56	0,56	0,48	0,54	0,43	0,54	0,47	0,55	0,54	0,47	0,85
BootStrap	0,39	0,47	0,49	0,40	0,46	0,36	0,47	0,40	0,45	0,47	0,40	0,71
Bats	0,42	0,47	0,51	0,42	0,46	0,36	0,48	0,41	0,46	0,47	0,41	0,72
Tbats	0,39	0,46	0,48	0,38	0,45	0,36	0,45	0,39	0,44	0,46	0,39	0,72
Stlme	0,40	0,47	0,49	0,41	0,45	0,36	0,46	0,40	0,45	0,46	0,39	0,71
Stlma	0,37	0,47	0,48	0,40	0,45	0,36	0,46	0,40	0,45	0,46	0,40	0,71
Ets	0,40	0,47	0,48	0,40	0,45	0,35	0,45	0,39	0,44	0,45	0,39	0,70
Nnetar	0,48	0,56	0,56	0,48	0,54	0,45	0,56	0,49	0,54	0,58	0,48	0,86
Snaive	0,57	0,67	0,69	0,57	0,64	0,51	0,65	0,56	0,65	0,66	0,56	1,00
Theta	8,98	10,21	9,92	10,26	11,33	8,32	10,41	10,81	12,26	13,03	10,63	21,38
Arfima	0,53	0,71	0,73	0,58	0,72	0,55	0,75	0,62	0,73	0,73	0,63	1,15

(Continua)

(Continuação)

	Horizonte de Previsão											
Model	1	2	3	4	5	6	7	8	9	10	11	12
Struc	0,47	0,57	0,61	0,51	0,54	0,44	0,55	0,48	0,57	0,56	0,49	0,83
Thiefa	0,42	0,56	0,53	0,45	0,52	0,43	0,54	0,46	0,52	0,53	0,47	0,84
Thiefe	0,40	0,47	0,48	0,41	0,45	0,36	0,45	0,40	0,44	0,46	0,39	0,71
Mlp	0,40	0,50	0,49	0,40	0,55	0,45	0,68	0,59	0,77	0,78	0,75	1,40
Elm	0,41	0,48	0,50	0,41	0,50	0,40	0,58	0,50	0,59	0,62	0,59	1,09
Média	0,39	0,47	0,48	0,40	0,46	0,37	0,47	0,40	0,46	0,47	0,41	0,73
Mediana	0,38	0,47	0,48	0,40	0,45	0,36	0,46	0,40	0,45	0,46	0,40	0,72
RW	374,31	317,22	309,54	373,57	331,69	416,24	323,90	376,80	328,71	323,01	375,46	211,01

Elaboração dos autores.
Obs.: Última linha consta valor apurado no *benchmark* RW.

TABELA A.4
MAD normalizado por horizonte de previsão

	Horizonte de Previsão											
Model	1	2	3	4	5	6	7	8	9	10	11	12
tslm1ts	0,41	0,44	0,39	0,40	0,38	0,26	0,32	0,36	0,33	0,41	0,28	0,55
tslm1tsy	0,30	0,32	0,24	0,28	0,26	0,21	0,27	0,27	0,25	0,31	0,23	0,47
Arima	0,47	0,63	0,51	0,55	0,51	0,40	0,45	0,47	0,46	0,52	0,40	0,73
BootStrap	0,40	0,42	0,36	0,35	0,37	0,25	0,35	0,39	0,36	0,42	0,30	0,53
Bats	0,52	0,45	0,43	0,39	0,37	0,28	0,41	0,39	0,39	0,45	0,33	0,58
Tbats	0,41	0,47	0,36	0,35	0,38	0,29	0,32	0,40	0,36	0,47	0,30	0,56
Stlme	0,44	0,45	0,38	0,38	0,37	0,27	0,37	0,39	0,36	0,47	0,33	0,59
Stlma	0,35	0,45	0,35	0,38	0,37	0,27	0,32	0,37	0,35	0,43	0,31	0,62
Ets	0,41	0,41	0,34	0,35	0,36	0,27	0,37	0,39	0,36	0,45	0,31	0,57
Nnetar	0,55	0,56	0,49	0,50	0,47	0,37	0,46	0,47	0,55	0,57	0,42	0,77
Snaive	0,75	0,75	0,65	0,69	0,68	0,52	0,64	0,74	0,67	0,79	0,53	1,00
Theta	2,17	2,08	1,65	1,78	1,61	1,22	1,71	2,03	1,69	2,15	1,41	2,83
Arfima	0,64	0,75	0,70	0,55	0,64	0,42	0,65	0,62	0,69	0,77	0,55	1,10
Struc	0,58	0,55	0,50	0,49	0,43	0,37	0,48	0,53	0,61	0,60	0,45	0,67
Thiefa	0,51	0,61	0,45	0,48	0,44	0,38	0,44	0,48	0,40	0,52	0,37	0,74
Thiefe	0,43	0,45	0,36	0,39	0,40	0,28	0,35	0,37	0,37	0,43	0,31	0,54
Mlp	0,44	0,50	0,40	0,41	0,43	0,34	0,58	0,69	0,74	0,87	0,69	1,28
Elm	0,44	0,49	0,42	0,42	0,38	0,32	0,43	0,47	0,46	0,58	0,43	0,98
Média	0,43	0,50	0,43	0,43	0,42	0,28	0,37	0,43	0,41	0,49	0,35	0,64
Mediana	0,39	0,47	0,41	0,38	0,38	0,27	0,37	0,40	0,39	0,48	0,36	0,62
RW	334,33	318,76	390,67	369,17	383,25	505,57	388,44	358,79	372,87	306,90	441,43	235,73

Elaboração dos autores.
Obs.: Última linha consta valor apurado no *benchmark* RW.

TABELA A.5
Top 10 MCS resultados baseados na *rolling window* de largura ts = 84 para mortes violentas intencionais (jan. 1999-dez. 2016)

Forecast Horizon h = 1				Forecast Horizon h = 4			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	26588,20	1,000	tslm1tsy	1	28817,67	1,000

Forecast Horizon h = 8				Forecast Horizon h = 12			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1ts	1	28608,10	1,000	tslm1ts	1	28202,07	1,000

Elaboração dos autores.
Obs.: Estatística $t_{,m}$ e função de perda quadrática. Para mais detalhes, ver tabela 3.

TABELA A.6
Top 10 MCS resultados baseados na *rolling window* de largura ts = 96 para mortes violentas intencionais (jan. 1999-dez. 2016)

Forecast Horizon h = 1				Forecast Horizon h = 4			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	27376,13	1,000	tslm1tsy	1	29103,78	1,000
				tslm1ts	2	29511,11	0,923

Forecast Horizon h = 8				Forecast Horizon h = 12			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	28608,10	1,000	tslm1ts	1	28639,28	1,000

Elaboração dos autores.
Obs.: Estatística e função de perda quadrática. Para mais detalhes, ver tabela 3.

TABELA A. 7
Top 10 MCS resultados baseados na *rolling window* de largura ts = 96 para mortes violentas intencionais (jan. 1999-dez. 2016)

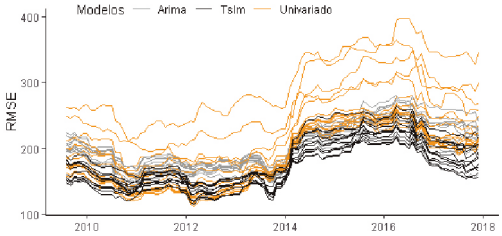
Forecast Horizon h = 1				Forecast Horizon h = 4			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	124,63	1,000	tslm1tsy	1	128,05	1,000

Forecast Horizon h = 8				Forecast Horizon h = 12			
Model	Rank	Loss	*p*mcs	Model	Rank	Loss	pmcs
tslm1tsy	1	129,73	1,000	tslm1ts	1	132,69	1,000
				tslm1tsy	1	133,11	1,000

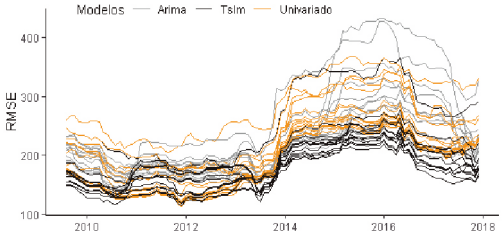
Elaboração dos autores.
Obs.: Estatística e função de perda absoluta. Para mais detalhes, ver tabela 3.

GRÁFICO A.1
Rolling RMSE: classe de modelos

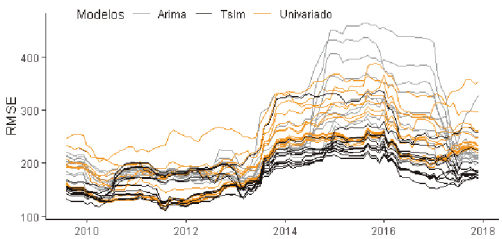
A.1A – Painel A: 1 month ahead



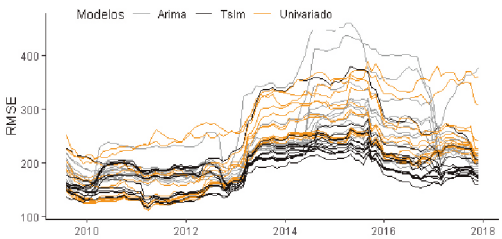
A.1B – Painel A: 4 month ahead



A.1C – Painel A: 8 month ahead



A.1D – Painel A: 12 month ahead



Elaboração dos autores.
Nota: [1] Modelos escolhidos pelo MCS.
Obs.: Figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Ipea – Instituto de Pesquisa Econômica Aplicada

Assessoria de Imprensa e Comunicação

EDITORIAL

Coordenação
Reginaldo da Silva Domingos

Assistente de Coordenação
Rafael Augusto Ferreira Cardoso

Supervisão
Camilla de Miranda Mariath Gomes
Everson da Silva Moura

Revisão
Amanda Ramos Marques
Ana Clara Escórcio Xavier
Clícia Silveira Rodrigues
Idalina Barbara de Castro
Luiz Gustavo Campos de Araújo Souza
Olavo Mesquita de Carvalho
Regina Marta de Aguiar
Hellen Pereira de Oliveira Fonseca (estagiária)
Ingrid Verena Sampaio Cerqueira Sodré (estagiária)

Editoração
Aeromilson Trajano de Mesquita
Cristiano Ferreira de Araújo
Danilo Leite de Macedo Tavares
Herllyson da Silva Souza
Jeovah Herculano Szervinsk Junior
Leonardo Hideki Higa

Capa
Danielle de Oliveira Ayres
Flaviane Dias de Sant'ana

Projeto Gráfico
Renato Rodrigues Bueno

The manuscripts in languages other than Portuguese published herein have not been proofread.

Livraria Ipea
SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo
70076-900 – Brasília – DF
Tel.: (61) 2026-5336
Correio eletrônico: livraria@ipea.gov.br

Composto em adobe garamond pro 12/16 (texto)
Frutiger 67 bold condensed (títulos, gráficos e tabelas)
Brasília-DF

Logotipo/slogan de governo removido em observancia ao disposto no art. 40 da Lei n 9.504/1997

Missão do Ipea
Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.

ipea Instituto de Pesquisa Econômica Aplicada

MINISTÉRIO DA
ECONOMIA

ISSN 1415-4765