

Previsão de homicídios no Brasil: proposta de variável antecedente^a

Homicide Prediction in Brazil: Proposal for an Antecedent Variable

Daniel Ricardo de Castro Cerqueira^b 

Instituto de Pesquisa Econômica Aplicada, Diretoria de Estudos e Políticas do Estado,
Brasília (DF), Brasil

Gabriel de Oliveira Accioly Lins^c 

Instituto de Pesquisa Econômica Aplicada, Diretoria de Estudos e Políticas do Estado,
Brasília (DF), Brasil

Danilo Santa Cruz Coelho^d 

Instituto de Pesquisa Econômica Aplicada, Diretoria de Estudos e Políticas do Estado,
Brasília (DF), Brasil

Resumo: Neste artigo, investigamos a capacidade de variáveis antecedentes, entre elas, internações por agressão, na previsão do número de homicídios no Brasil. O objetivo principal é suprimir a lacuna referente à defasagem na divulgação de informações sobre homicídios no país, permitindo análises conjunturais atualizadas. Para tanto, por meio de esquema *rolling window* e abordagem Model Confidence Set (MCS), investigamos se modelos de variáveis antecedentes apresentam desempenho preditivo superior a um 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 de as 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).

Editor responsável: Diogo de Prince

^a Submissão: 09/02/2021 | Aprovação: 25/09/2021 | DOI: 10.22456/2176-5456.111440

^b daniel.cerqueira@ipea.gov.br

^c gablins@yahoo.com.br

^d danilo.coelho@ipea.gov.br

Todos os autores foram responsáveis pela concepção, pesquisa de dados e/ou documentos, análise dos dados e/ou documentos, participação ativa na discussão dos resultados e revisão e aprovação da versão final.



Esta publicação está licenciada sob os termos de
Creative Commons Atribuição 4.0 Internacional

Palavras-chave: Previsão. Previsão de homicídios. Conjunto de Modelos Confiáveis. 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.

JEL: C53. K42. C22.

1. Introdução

No Brasil, o número de homicídios tem apresentado tendência de crescimento ao longo dos últimos vinte anos, com taxa média de crescimento na ordem de 8,0% ao ano. À luz desse fenômeno, diversas análises investigaram as causas e consequências da criminalidade (Cerqueira, 2014; Cerqueira; Soares, 2014). A ausência de homogeneidade metodológica entre as Unidades da Federação ao se 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.

Criado com o objetivo de produção regular de informações sobre mortalidade no Brasil para fins de análise epidemiológica, o SIM é referência em diversas áreas da gestão pública e pesquisa acadêmica, como no campo da criminologia (Cerqueira, 2012; Cerqueira *et al.*, 2018; Cerqueira *et al.*, 2018; Murray; Cerqueira; Kahn, 2013). Apesar da ampla utilização, as informações disponibilizadas pelo SIM são divulgadas com defasagem. Existe, portanto, 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 políticas de segurança pública.

Este artigo explora a lacuna na disponibilidade de informações atualizadas e tem como objetivo mitigar a escassez informacional sobre criminalidade violenta, demonstrando os ganhos preditivos ao utilizar um conjunto de variáveis antecedentes, entre elas, internações por agressão, em modelos preditivos de homicídios, e construindo um modelo predictor multivariado de desempenho preditivo superior ao observado em modelos univariados. Dessa forma, pretende-se contribuir na produção de evidências preliminares acerca da dinâmica criminal brasileira, permitindo aos *policymakers* da segurança pública ajustar o impacto de políticas nacionais de segurança, aumentando o bem-estar social, além de estabelecer modelo predictor *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; Thompson, 2003). Sendo um campo pouco explorado na literatura nacional, estudos considerando

séries temporais utilizam técnicas de previsão e métodos de avaliação que não permitem conclusões robustas (De Souza, *et al.*, 2007; Provenza *et al.*, 2015; Provenza, Serra Costa; Silva, 2015); por sua vez, a literatura internacional utiliza crescente sofisticação técnica nas pesquisas (Gorr; Harries, 2003; Gorr; Olligschlaeger; Thompson, 2003; Klepinger; Weis, 1985; Pepper, 2008; Shoesmith, 2013). No entanto, a maioria desses trabalhos utiliza abordagem univariada. Portanto, inovamos ao introduzir inédita variável antecedente capaz de superar capacidade previsora de modelos univariados e 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 diferentes estatísticas de avaliação, funções de perda e larguras de janelas de estimação.

As evidências encontradas sugerem que a inclusão de variáveis antecedentes gera modelos de previsão com melhor desempenho fora da amostra àqueles do conjunto de modelos univariados. Especificamente, modelos de variáveis antecedentes apresentam 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 artigo está organizado da seguinte forma além desta introdução: na seção 2, abordamos conceitos e exemplos de previsão de informações criminais; na seção 3, apresentamos os dados utilizados no trabalho, as evidências iniciais e a metodologia; na seção 4, apresentamos os principais resultados; na seção 4, realizamos exercício de robustez; e, por fim, na seção 6, apresentamos as considerações finais.

2. Previsão de informações criminais

A prática de *crime forecasting* é a aplicação de técnicas – usualmente métodos quantitativos – capazes de identificar prováveis alvos (vítimas/ofensores) de intervenção policial e prevenir ou solucionar crimes passados, independentemente do nível de agregação da variável prevista (vítima/agressor, bairro, cidade, estado ou país) ou unidade de tempo. Apesar de ocorrer em maior frequência, a previsão de informações criminais não está restrita ao *predictive policing*, isto é, o uso de dados históricos na elaboração de previsão espaço-temporal de áreas de criminalidade ou pontos críticos de crime que serão a base para as decisões de alocação de recursos da polícia com a expectativa de que ter policiais no local e horário propostos deterá ou detectará atividades criminosas. O objetivo de previsão de informações criminais é antecipar tendências e padrões criminais, utilizando essa informação na elaboração da estratégia de prevenção de crimes (Bennett Moses; Chan, 2018).

A aplicação de métodos preditivos antecipando tendências, identificando locais e/ou indivíduos propensos à futura vitimização/agressão é prática disseminada entre departamentos de polícia dos Estados Unidos (Perry *et al.*, 2013) e, crescente campo de pesquisa na literatura especializada, por exemplo (Caplan; Kennedy; Piza, 2013; Cohen; Gorr; Olligschlaeger, 2007; Kounadi *et al.*, 2020).

A previsão de variáveis criminais é parte de estratégia de atuação policial chamado *prediction-led policing*, isto é, o gerenciamento dos recursos policiais através do *forecasting* de variáveis criminais. O objetivo das previsões é aumentar a consciência situacional no nível tático e estratégico, desenvolvendo policiamento proativo e mais eficiente. Assim, a previsão de variáveis criminais atende objetivos com horizontes preditivos diversos, tais como: desdobramento tático de recursos policiais, alocação de recursos e planejamento estratégico (Gorr; Olligschlaeger; Thompson, 2003).

Em nível tático, previsões criminais são geralmente associadas à previsão de *hot-spots*, ou seja, previsão de áreas com elevada ocorrência de crimes. Área de pesquisa em franca expansão, trabalhos nessa área usualmente utilizam técnicas de *machine learning*, *deep learning* ou estimação de densidade de *kernel*, realizando previsões considerando o padrão histórico do crime (Kounadi *et al.*, 2020). Previsões de curto e médio prazo não estão restritas a séries temporais agregadas ou a métodos de análise espaço-temporal que realizem previsão do número de crimes em determinada área, sendo utilizadas na previsão espacial de diversas informações

relacionadas ao crime, por exemplo, potenciais vítimas e/ou ofensores (Perry *et al.*, 2013).

De acordo com Levine *et al.* (2017), na cidade de Nova Iorque, a adoção de sistema de vigilância, composto por câmeras e sensores auditivos que abastecem modelos preditivos com diferentes objetivos (i.e., previsão de *hot-spots*, identificação de indivíduos propensos à vitimização) foi capaz de antecipar tendência de crimes e permitir o planejamento de pessoal, assim, aumentando a eficiência alocativa de pessoal e reduzindo em US\$ 50 milhões as despesas operacionais, sendo parcialmente responsável por redução de 6% no índice de geral de crimes.

O Estado do Rio de Janeiro utiliza o ISPGEO, na identificação de *hotposts* e de horários de maior probabilidade de ocorrência de crimes, e a ferramenta CrimeRadar, capaz de mapear ocorrências criminais em tempo real e, utilizando *machine-learning*, realizar previsão de horários e dias da semana com maior probabilidade de ocorrências policiais, em quadrantes de 500 por 500 metros da cidade. Abordagem similar foi proposta utilizando dados de Natal (Araujo *et al.*, 2017; 2018).

Embora úteis, especialmente no auxílio ao policiamento preventivo, abordagens utilizadas em nível tático atendem finalidades diferentes às requeridas neste trabalho. Além disso, o SIM, única base com informações nacionais sobre homicídios, não possui a granularidade necessária à aplicação dessas técnicas; portanto, essas abordagens são inadequadas ao problema apresentado, isto é, a inexistência de número contemporâneo dos homicídios no Brasil.

Do ponto de vista estratégico e, portanto, em longo prazo, os trabalhos estão concentrados na previsão de séries temporais agregadas. Assim, em contexto semelhante ao brasileiro, isto é, na ausência de números contemporâneos de homicídios, Shoenmith (2013) propõe o modelo *space-time autoregressive* (ST-AR) univariado, capaz de antecipar tendências criminais de longo prazo nos Estados Unidos, seus estados e regiões, permitindo o planejamento estratégico de políticas de segurança.

A proposta do trabalho é seguir a prática consolidada na literatura internacional de previsão de dados criminais e amenizar a inexistência de números contemporâneos de homicídios no Brasil, empregando modelo preditivo que, ao

utilizar a variável antecedente internação por agressão, vai além da prática usual de modelos univariados.

3. 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 (*e.g.* 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 este gerido pela Secretaria de Vigilância à Saúde, do Ministério da Saúde, em conjunto com as Secretarias Estaduais e Municipais de Saúde. As Secretarias de Saúde coletam as Declarações de Óbitos dos cartórios e as informações nelas contidas entram no SIM, o qual foi criado com objetivo de disseminar regularmente informações sobre mortalidade no país. O SIM possui abrangência universal.

Uma das informações 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.

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. Falecimento 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 através de septicemia. Nesse 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

¹ Veja, no Quadro 1 do Apêndice, fonte e descrição sumária das variáveis utilizadas.

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 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 informa 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 ocorrência, distorcendo o nível de violência capturado na variável homicídio. No entanto, tal possibilidade parece remota, pois, entre 2006 e 2017, dos homicídios com local de óbito determinado, 72% ocorreram fora de ambiente hospitalar, inexistindo possibilidade de internação de longa duração. Além disso, de acordo com os dados de internação, o tempo médio de internação daqueles mortos pela causa base agressão ou intervenção legal é 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, destinada 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.

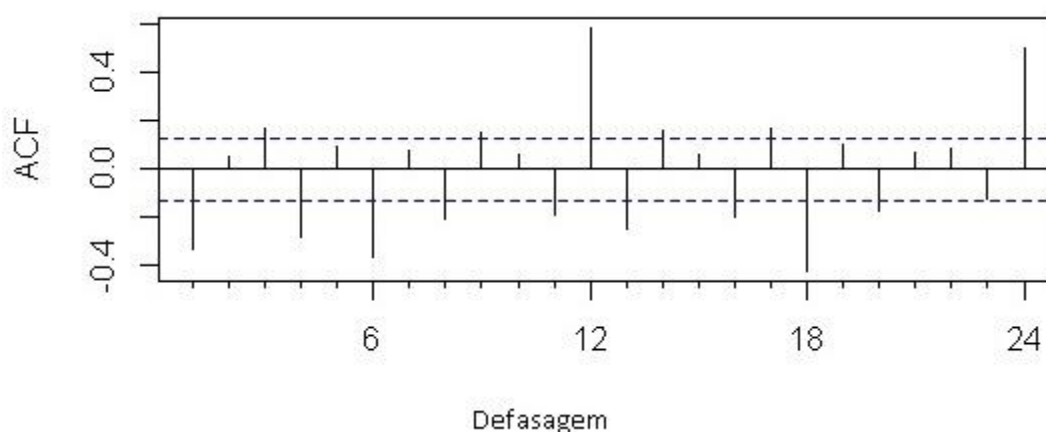
A variável internação por agressão foi construída através de microdados do SIH. Ocorre agregação mensal de novas internações cuja causa primária ou secundária seja codificada pelos códigos CID-10 X85 até Y09, Y35, Y36 e W50. Ou seja, agressões, intervenções legais e operações de guerra e exposição a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

Como usual em análises de série temporal é necessário verificar existência de raiz unitária, evitando a possibilidade de regressão espúria. Com esse objetivo foram realizados os testes ADF (*Augmented Dickey-Fuller*), PP (*Phillips-Perron*) e

KPSS (*Kwiatkowski-Phillips-Schmidt-Shin*).² Os resultados indicam que apenas interações é série $I(0)$.

A primeira evidência, apresentada na Figura 1, sugere autocorrelação de primeira ordem na série histórica de homicídios brasileiros, evidência em linha aos achados de Santos (2009). Por essas 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; Freitas, 2015).³

Figura 1 – Função de autocorrelação: homicídio



Fonte: elaboração própria.

Como forma 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, foi elaborado um conjunto de modelos univariados. Os métodos aplicados foram:

a) *Random Walk* – conhecido com passeio aleatório, trata-se do modelo mais simples no conjunto dos modelos estimados. Neste caso, o número homicídios no mês é igual ao número do mês anterior, $\hat{y}_{T+h|t} = y_{T+h-1|t}$. Este modelo não apresenta desempenho preditivo adequado na presença de elevada aleatoriedade

² Ver os resultados na Tabela 6 e as estatísticas descritivas na Tabela 7.

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

ou forte tendência e sazonalidade. No entanto, será preditor adequado de séries com frequentes mudanças de padrão. Esse modelo será utilizado como *benchmark*.

b) Snaive - neste caso, o número de homicídios é igual ao número de 12 meses antes $\hat{y}_{T+h|t} = y_{T+h-12|t}$. Método utilizado por departamentos policiais na previsão do nível de crime considerando variações sazonais (Gorr; Olligshlaeger; Thompson, 2003).

c) ETS - modelos de suavização exponencial, aplica o algoritmo proposto em Hyndman *et al.* (2002) ao modelar o 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 (1995), deriva-se o modelo espaço estado com erros de origem única, equivalente ao modelo de suavização exponencial selecionado de acordo com critério de informação 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*.

d) Arima - aplicação de algoritmo proposto por Hyndman e Khandakar (2008). Trata-se de regressão com erros ARIMA (p,d,q), estimada pelo máximo verossimilhança. O algoritmo seleciona a ordem do termo autoregressivo p e média móvel q de acordo com critério de informação AICc.⁴

e) STL - neste caso, aplica-se a decomposição do tipo STL ("Seasonal and trend decomposition using Loess"), proposto por Cleveland, Cleveland, and Terpenning (1990). Isto é, a série é decomposta em componente sazonal, tendência e termo de erro e cada componente é separadamente ajustado e previsto. Ajusta-se a série sazonalmente ajustada (tendência e termo de erro) utilizando o modelo ETS (STLe) e ARIMA (STLa). Em seguida, é feita a previsão da série ajustada sazonalmente, então, os resultados são resazonalizados. Como o padrão sazonal repete-se ao longo da série analisada, adicionamos à série ajustada sazonalmente, a estimativa do componente sazonal do ano anterior, ou seja, a previsão do componente sazonal é realizada pelo método *random walk*.

f) Bootstrap Aggregating (*bagging*) - método proposto por Bergmeir, Hyndman e Benitez (2016), aplica-se o *bootstrap* aos resíduos da série simulando

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

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). O ruído é submetido a processo de *bootstrap*, isto é, ele sofre um processo de re-amostragem, nesse caso, 150 vezes. Todavia, como o termo aleatório de uma série temporal pode ser correlacionado no tempo, a re-amostragem não é feita por observações, mas em blocos — “moving blocked bootstrap”. Por fim, esses 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.

g) TbatS - modelagem do tipo espaço-estado segue o método de previsão proposto por De Livera, Hyndman e Snyder (2011), elaborado para séries temporais com padrão sazonal complexo (e.g., múltiplos períodos sazonais). Esse método incorpora transformação *Box-Cox*, filtro ARMA para correção de resíduo e componente sazonal trigonométrico. As principais vantagens do TBATS apontadas em De 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, em que existem múltiplos ciclos, e o tamanho dos ciclos não é número inteiro.

h) Bats - similar ao TbatS, este método não incorpora componente sazonal trigonométrico.

i) Arfima - modelo ARFIMA (p, d, q) selecionado e estimado através do 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.

j) Nnetar - modelo de rede neural “*feed forward*” de três camadas envolvendo função de combinação linear e função de ativação logística. Denotado por $NNAR(p,P,k)_m$ onde p informa a ordem autoregressiva não sazonal e P ordem sazonal das variáveis de entrada defasadas, selecionadas pelo critério de informação AIC, e k representa número de nós na camada oculta única, selecionado $k = (p+P+1)/2$, seguindo (Hyndman; Athanasopoulos, 2018).

k) Struct - modelo de componente não observado do tipo 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 pela máxima verossimilhança.

l) ELM - conhecida por *Extreme Learning Machine*, trata-se de rede neural do tipo *feedforward* de camada oculta única, proposta em Huang, Zhu e Siew (2006). Neste caso, ao invés de utilizar métodos de aprendizado *gradiente-based*, os parâmetros de entrada (pesos dos *inputs* e *hidden layer bias*) não requerem ajuste. Estes são aleatoriamente designados, permanecendo constantes durante o treinamento e previsão. Pesos de saída são analiticamente determinados. Dado 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).

m) Thief - Método de previsão através de hierarquia temporal. Isto é, realizam-se previsões individuais das agregações temporais da série temporal; em seguida, as previsões são reconciliadas através de algoritmo reconciliatório hierárquico (Athanasopoulos *et al.*, 2017). Neste caso, a série homicídios mensais é agregada em bimestre, trimestre, quadrimestre, semestre e anualmente. Então, previsões das séries agregadas são realizadas através dos algoritmos propostos por Hyndman e Khandakar (2008) (Thief_a) e do algoritmo de suavização exponencial de Hyndman *et al.* (2002) (Thief_e). Em seguida, seguindo os melhores resultados obtidos por Athanasopoulos *et al.* (2017), as previsões temporalmente agregadas são individualmente reconciliadas pelo algoritmo “*Structural scaling*”.

Além desse conjunto de modelos univariados, incluímos dois esquemas de combinações de modelos, isto é, a média e mediana das previsões univariadas. Seguindo os melhores resultados obtidos na literatura (Samuels; 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 (1999) sugere modelo com uma defasagem da variável homicídios e até uma defasagem da variável independente (em nosso caso, o

número de interações por agressão), igualmente ao modelo sugerido no critério de informação BIC e HQ. Por sua vez, o critério de informação AIC sugere ser o modelo de melhor ajuste àquele de cinco defasagens da variável interações e uma defasagem da variável dependente, ao contrário do critério AICc, que sugere quatro defasagens da variável interações e uma defasagem da variável dependente.

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

Considere o seguinte modelo:

$$Y_{t+h} = f_h(\mathbf{x}_t) + \mu_{t+h}, h = 1, \dots, 12, \quad t = 1, \dots, T \quad (1)$$

O número de homicídios no mês $t + h$ será representado por Y_{t+h} , $\mathbf{x}_t = (x_{1t}, \dots, x_{nt})'$ é vetor de variáveis explicativas possivelmente contendo defasagens da variável dependente, $f_h(\cdot)$ relaciona variáveis antecedentes e homicídios futuros. As duas primeiras abordagens não contêm variável homicídios defasada no conjunto de variáveis antecedentes. Nesse caso, aplicamos o modelo com defasagens distribuídas (Tslm) e modelo com erros ARIMA (Arima),⁵ de acordo com algoritmo de Hyndman e Khandakar (2008). A equação de previsão estimada tem a seguinte forma:

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

Onde $\hat{Y}_{t+h}$ é o número de homicídios previsto no horizonte de previsão h , $\hat{\alpha}_0$ estimação do termo constante, $\mathbf{x}_t$ vetor de variáveis previsoras e $\hat{\mu}_t$ termo de erro. O modelo 0 considera relação contemporânea entre homicídios previstos e

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

variável antecedente (Tslm0ts), em seguida, o modelo 1 acrescenta a primeira defasagem de interações (Tslm1ts). Em função das diferentes especificações sugeridas através dos 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, esse processo é repetido até a quinta defasagem de interações 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:

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

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

Ao avaliarmos desempenho preditor dos modelos, utilizamos duas abordagens. No primeiro método de avaliação, empregamos estatísticas de precisão ao comparar desempenho preditivo, isto é, utilizamos a raiz do erro quadrado médio (RMSE), erro absoluto médio (MAE) e desvio absoluto médio da média (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 o modelo considerado apresenta 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 preditor dos diversos modelos através da abordagem *model confidence set* (MCS), proposta por Hansen, Asger e James, (2011).

O objetivo do MCS é determinar conjunto, M^* , que consiste dos “melhores” modelos do conjunto inicial de modelos, M_0 , onde melhor é definido por critérios

⁶ Descrição das estatísticas no Quadro 2 do Apêndice.

pré-definidos, isto é, nível de significância, função de perda e estatísticas de teste. O MCS elabora conjunto de modelos de confiança, $\widehat{M}^*$, ou seja, elabora conjunto de modelos que contém os melhores modelos dado certo nível de confiança. Os modelos em $\widehat{M}^*$ são avaliados através de 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 $\mathbf{E}_M$. Isto é, o procedimento consiste em sequência de testes, onde 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 $\mathbf{E}_M$ é utilizada na eliminação de modelo de baixo desempenho em M . O procedimento é repetido até γ_M não ser rejeitada dado o 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 através dos 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 $d_{ij,t} = L_{i,t} - L_{j,t} \forall i, j \in M_0$. Então, definimos estatísticas de perdas amostrais relativas:

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

Representando perda amostral média do modelo i relativamente ao modelo j .

$$\bar{d}_i = \frac{1}{m} \sum_{j \in M} \bar{d}_{ij}.$$

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

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

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

$$H_{0,M} : \mu_{ij} = 0 \forall i, j \in M,$$

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

Ou

$$H_{0,M} : \mu_i = 0 \forall i, j \in M,$$

$$H_{A,M} : \mu_i \neq 0 \text{ para algum } i, j \in M$$

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

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

Onde $\widehat{var}(\bar{d}_{ij})$ 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 através da seguinte estatística de teste: $T_{R,M} = \max_{i,j \in M} |t_{ij}|$.

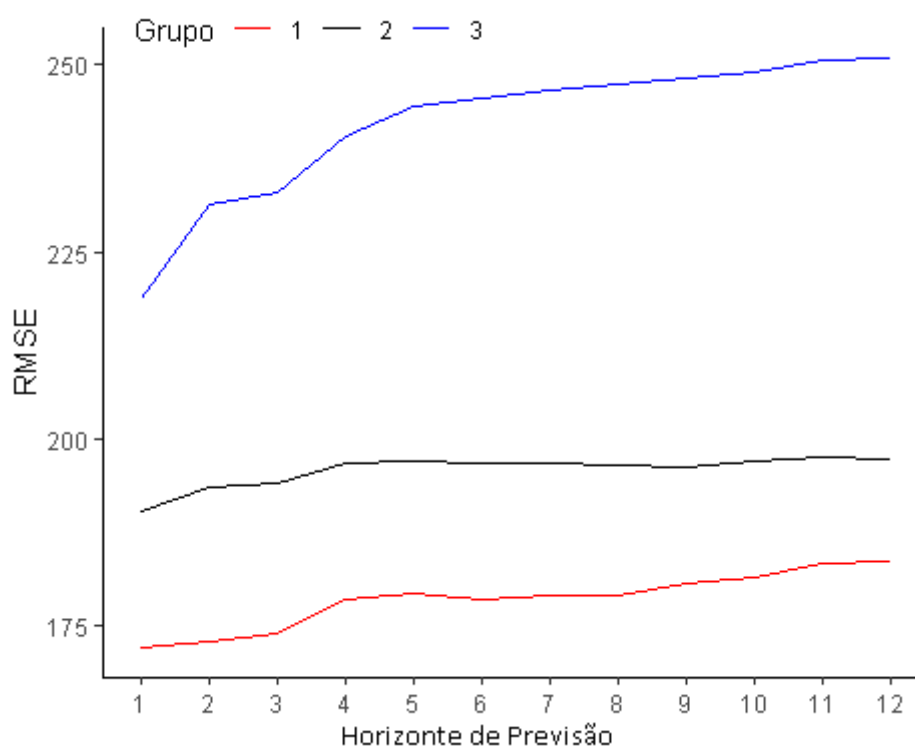
A estatística $T_{R,M}$ assegura que o resultado da hipótese nula depende do modelo de maior perda relativa. Nesse caso, o modelo eliminado obedece a seguinte regra de eliminação: $e_{R,M} = \arg \max_i \left(\sup_{j \in M} \frac{\bar{d}_{ij}}{\sqrt{\text{var}(\bar{d}_{ij})}} \right)$, ou seja, a regra de eliminação irá excluir o modelo de maior estatística de teste t . Por sua vez a estatística t_i será avaliada através da seguinte estatística de teste: $T_{\max,M} = \max_{i \in M} t_i$. Nesse 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 $e_{\max,M} = \arg \max_{i \in M} \frac{\bar{d}_i}{\sqrt{\text{var}(\bar{d}_i)}}$, isto é, será excluído modelo de maior contribuição para a estatística de teste. Esse modelo apresenta o pior desempenho relativamente à média dos modelos em M . A distribuição assintótica dessas estatísticas de teste não é padronizada, portanto, o MCS utiliza método *bootstrap* ao estimar a distribuição dessas estatísticas. (veja detalhes em Hansen, Lunde e Nason (2010)).

4. Principais resultados

Esta seção apresenta os resultados do exercício de previsão fora da amostra de homicídios. Para cada modelo estimado, computamos sequência de 12 previsões fora da amostra, utilizando método *rolling window* com janela de estimação de tamanho $T_s = 84$ observações mensais, dessa forma, minimizando potenciais *outliers* e quebras estruturais, além de permitir implementar testes de capacidade preditiva em modelos aninhados (Giacomini; White, 2006). Em cada janela, os modelos são estimados e previsões de 1 a 12 passos à frente são realizadas. Então, a janela avança um mês e repete o processo anterior até o final da amostra – similar a Medeiros, Vasconcelos e Freitas (2015) e Stock e Watson (2012). Isso 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 vencedor àquele 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, a Figura 2 expõe RMSE dos modelos estimados ao longo dos horizontes avaliados. A figura apresenta grupos de modelos agregados por similar desempenho preditivo. É possível perceber três padrões de desempenho preditivo. O grupo 1, composto por modelos (autoregressivos) de defasagem distribuída, apresenta ao longo de todos os horizontes de previsão o melhor desempenho preditivo, embora apresente 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, apresenta desempenho preditivo inferior aos demais grupos em todos os horizontes de previsão e deterioração da capacidade preditiva ao longo dos horizontes de previsão.

Figura 2 - RMSE por grupos preditivos e horizonte de previsão



Fonte: elaboração própria.

Segundo 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 o RMSE médio, o MAE médio e MAD médio,

respectivamente. As colunas (4), (5), (6) apresentam, respectivamente, o RMSE, MAE e MAD máximo. Por sua vez, as colunas (7), (8) e (9) apresentam o RMSE, MAE e MAD mínimo. 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 o *rank* decrescente do RMSE médio.

As evidências sugerem que, nas estatísticas de avaliação consideradas, majoritariamente os modelos competidores superam na média modelo *random walk* – exceto modelo Snaive. Em geral, ganhos preditivos relativos ao *random walk*, em termos de RMSE e MAE, são da ordem de 50% e, na estatística MAD, de aproximadamente 40%. 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 (autoregressivo) 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 11^o 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 apresentam a menor estatística de desempenho preditivo. Ainda de acordo com as evidências, nenhum modelo apresenta desempenho sistematicamente superior aos demais, ou seja, apresenta o melhor desempenho em todos os critérios de avaliação.

Tabela 1 – Sumário das estatísticas de previsão

Modelo	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	RMSE Médio	MAE Médio	MAD Médio	RMSE Max.	MAE Max.	MAD Max.	RMSE Min.	MAE Min.	MAD Min.	#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
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			

Fonte: elaboração própria a partir dos dados da pesquisa.

Nota: última linha 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. Melhores resultados estão destacados em negrito. Os modelos de variáveis antecedentes são, em todos os horizontes de previsão considerados, aqueles de maior ganho relativo ao *random walk*, sugerindo estabilidade na superioridade preditiva dos modelos com variável antecedente. Considerado todos os horizontes de previsão, na média, modelos de variáveis antecedentes considerados são 13% melhores que conjunto univariado e combinações. Resultados similares são observados nas estatísticas MAE e MAD, apresentadas, nas Tabelas 8 e 9 do Apêndice, respectivamente.

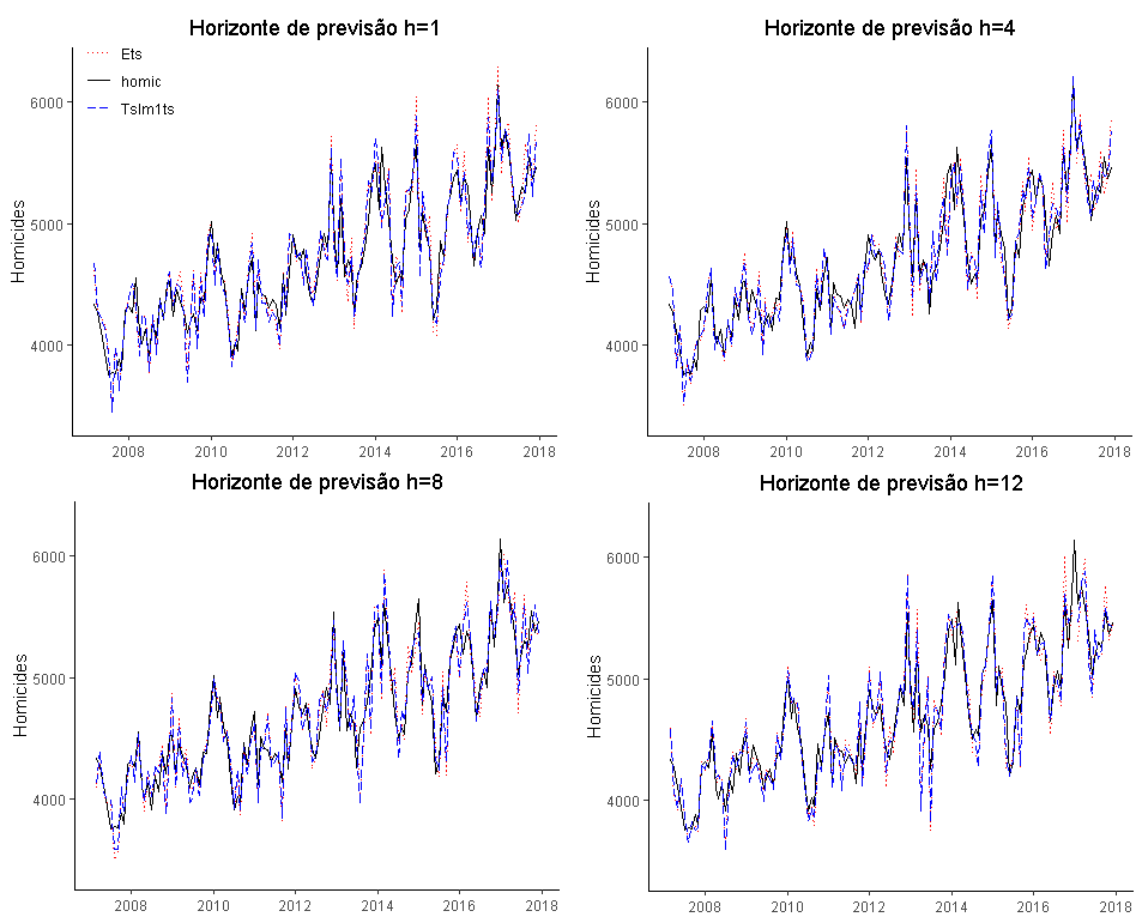
Tabela 2 - RMSE normalizado por horizonte de previsão

Modelo	Horizonte de Previsão											
	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
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

Fonte: elaboração própria a partir dos dados da pesquisa.
Nota: estatística do modelo de melhor desempenho destacado em negrito.

A Figura 3 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.

Figura 3 – Valores realizados de homicídios e previsões selecionadas estimadas através de “rolling window” $ts=84$



Fonte: elaboração própria.

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 *model confidence set* ao considerar todos os modelos estimados desde 1999m01, assumindo função de perda quadrática. Como sumário, listamos modelos de melhor classificação pelo MCS.

Tabela 3 – Resultado do MCS baseado na rolling window de largura $ts = 84$ entre 1999m01 e 2017m12 para homicídios. Estatística t_{max} e função de perda quadrática

Horizonte de Previsão h = 1			Horizonte de Previsão h = 4		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	26.588	tslm1tsy	1	28.818
tslm2tsy	2	27.711			
tslm3tsy	3	28.817			
tslm4tsy	4	29.856			
tslm0tsy	5	29.795			
tslm1ts	6	30.924			
Horizonte de Previsão h = 8			Horizonte de Previsão h = 12		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1ts	1	28.608	tslm1ts	1	28.639

Fonte: elaboração própria a partir dos dados da pesquisa.

Nota: erro refere-se ao valor médio da função de perda quadrática. Foram utilizadas 5.000 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 ($M_{90\%}^*$).

Em linhas 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 (autoregressivo) 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

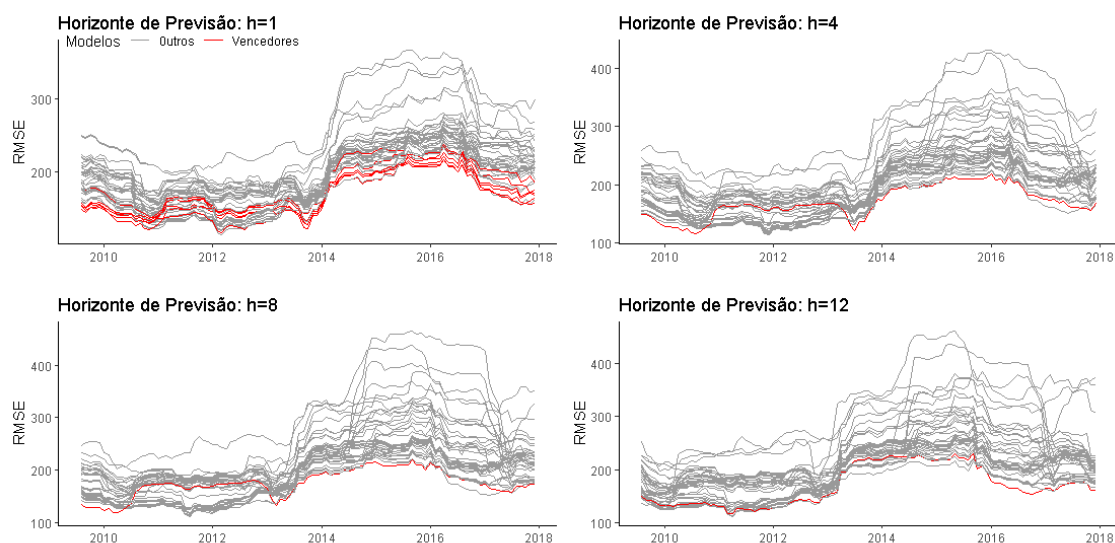
vencedor e o último colocado. Nos horizontes de previsão seguintes, o MCS selecionou somente dois modelos. Nomeadamente, no caso $t + 4$, modelo sugerido através do método de identificação proposto por Pankratz (1991b) e critérios BIC e HQ apresentou melhor desempenho preditivo fora da amostra, enquanto em $t + 8$ e $t + 12$, modelos especificados através de critérios tradicionais de identificação não apresentaram resultado preditivo vencedor, justificando a inclusão de outras especificações além das sugeridas através de 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 através do melhor desempenho preditivo ao considerar a estimativa da função perda ao longo de todo 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.

A Figura 4 apresenta a variação temporal do RMSE de todos os modelos estimados e destaca àqueles 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 amostra. A evidência sugere estabilidade no desempenho preditivo dos modelos e, portanto, reforça os achados do MCS.⁷

⁷ No Apêndice, a Figura 5 apresenta esse exercício de acordo com a categoria do modelo univariado, Arima ou (autoregressivo) de defasagens distribuídas. Novamente, as evidências estão em linhas aos achados do MCS, isto é, modelos de variáveis antecedentes com defasagem distribuída apresentam os melhores resultados.

Figura 4 – *Rolling window* RMSE. Modelos escolhidos pelo MCS



Fonte: elaboração própria.

As evidências apresentadas 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 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 nas Tabelas 10 e 11 do Apêndice, 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 ao invés 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 no caso da função de perda quadrática, o modelo sugerido nos critério de informação HQ e BIC e método de identificação de Pankratz (1991b) aparece como aquele de melhor desempenho preditivo fora da amostra, nos horizontes de previsão $t+1$ e $t+4$. Ainda nesta linha, os resultados em $t+8$ e $t+12$ repetem aqueles da Tabela 3, ao não selecionar modelo indicado pelos métodos de especificação utilizados. Também, similar ao observado no exercício anterior, nenhum modelo supera os concorrentes em todos os horizontes de previsão. Esses resultados são

qualitativamente iguais aos apurados considerando estatística de teste t_{rm} , e função de perda absoluta, como observado na Tabela 12 do Apêndice.

Tabela 4 – Resultado do MCS baseado na “*rolling window*” de largura $ts=84$ entre 1999m01 e 2016m12 para homicídios. Estatística t_{max} e função de perda absoluta

Horizonte de Previsão h = 1			Horizonte de Previsão h = 4		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	124,49	tslm1tsy	1	129,97
Horizonte de Previsão h = 8			Horizonte de Previsão h = 12		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	128,38	tslm1ts	1	130,36

Fonte: elaboração própria a partir dos dados da pesquisa.

Nota: para mais detalhes, ver Tabela 3.

5. *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; Timmermann, 2007). Contornamos esse problema ampliando a largura da janela de estimação para $T_s = 96$ ao invés de $T_s = 84$ observações mensais.

A Tabela 5 apresenta os achados desse exercício. 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 janela $T_s = 84$. Apesar disso, os resultados gerais são similares aos observados anteriormente. Modelos com variáveis antecedentes superam sistematicamente os modelos univariados em todos os horizontes previsão considerados. Esses achados são robustos à estatística de avaliação e função de perda, como observado nas Tabelas

11 e 12 do Apêndice — onde realizamos este exercício considerando estatística t_{rm} , e função de perda absoluta — salientando os ganhos de previsão alcançados ao utilizar variáveis antecedentes.

Tabela 5 – Resultado do MCS baseado na “rolling window” de largura $ts=96$ entre 1999m01 e 2016m12 para homicídios. Estatística t_{max} e função de perda quadrática

Horizonte de Previsão h = 1			Horizonte de Previsão h = 4		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	27.376	tslm1tsy	1	29.104
			tslm1ts	2	29.511
Horizonte de Previsão h = 8			Horizonte de Previsão h = 12		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	29.918	tslm1ts	1	30.492
tslm1ts	2	30.053			

Fonte: elaboração própria a partir dos dados da pesquisa.

Nota: para mais detalhes, ver Tabela 3.

6. Considerações finais

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; Soares 2014) 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 artigo inovou ao avaliar a capacidade de modelos com variáveis antecedentes, entre elas, interações por agressão, e ao apresentar

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 interações por agressão aos pesquisadores interessados em aumentar a acurácias 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; Thompson, 2003; Shoesmith, 2013) ou ampliar o leque de variáveis antecedentes e aplicar métodos de regularização ou abordagem geral para o específico.

Referências

ARAUJO, A. *et al.* *A predictive policing application to support patrol planning in smart cities*. In: INTERNATIONAL SMART CITIES CONFERENCE (ISC2), 14., 2017, Wuxi, China. *Anais [...]* Wuxi: IEEE, set. 2017.

ARAUJO, A. *et al.* *Towards a crime hotspot detection framework for patrol planning*. In: INTERNATIONAL CONFERENCE ON SMART CITY, 16., 2018, Exeter. *Anais [...]* Exeter: IEEE, jun. 2018.

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

BENNETT MOSES, L.; CHAN, J. Algorithmic prediction in policing: assumptions, evaluation, and accountability. *Policing and Society*, v. 28, n. 7, p. 806-822, set. 2018.

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.

BRZOZA-BRZEZINA, M.; KOTŁOWSKI, J.; MIŚKOWIEC, A. How forward-looking are central banks? Some evidence from their forecasts. *Applied Economics Letters*, v. 20, n. 2, p. 142-146, fev. 2013.

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

CAPLAN, J. M.; KENNEDY, L. W.; PIZA, E. L. Joint utility of event-dependent and environmental crime analysis techniques for violent crime forecasting. *Crime & Delinquency*, v. 59, n. 2, p. 243-270, mar. 2013.

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

CERQUEIRA, D. *et al.* Guns and suicides. *The American Statistician*, v. 72, n. 3, p. 289-294, 3 jul. 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: Instituto de Pesquisa Econômica Aplicada, 2019.

CERQUEIRA, D. R. D. C. *Causas e consequências do crime no Brasil*. Rio de Janeiro: BNDES, 2014.

CERQUEIRA, D.; 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, dez. 2014.

CLEVELAND, R. B. *et al.* STL: a seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, v. 6, n. 1, p. 3-73, 1990.

COHEN, J.; GORR, W. L.; OLLIGSCHLAEGGER, A. M. Leading indicators and spatial interactions: a crime-forecasting model for proactive police deployment. *Geographical Analysis*, v. 39, n. 1, p. 105-127, jan. 2007.

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

DE LIVERA, A. M.; 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.

FÁTIMA MARINHO DE SOUZA, 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, mar.-abr. 2007.

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, out. 2003.

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

HANSEN, P. R.; LUNDE, A.; NASON, J. M. The model confidence set. *SSRN Electronic Journal*, mar. 2010.

HANSEN, 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: Applied time series modelling and forecasting*. 1. ed. Southern Gate: 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. 1-50, 1989.

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

HYNDMAN, R. *et al.* *Forecasting with exponential smoothing: the state space approach*. [S.l.]: Springer-Verlag Berlin Heidelberg, 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, 2002.

HYNDMAN, R. J. Forecasting performance evaluation and reporting. *In: Business Forecasting*. Hoboken: John Wiley & Sons, Inc., 2015. p. 177-184.

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, jul. 2008.

KALMAN, R. E. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, v. 82, p. 35-45, 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, dez. 1985.

KOUNADI, O. *et al.* A systematic review on spatial crime forecasting. *Crime Science*, v. 9, n. 7, 2020.

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

LEVINE, E. S. *et al.* The New York city police department's domain awareness system. *Interfaces*, v. 47, n. 1, 2017.

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, 2006.

MEDEIROS, M. C. *et al.* Forecasting inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics*, v. 39, n. 1, p. 1-45, 2019.

MEDEIROS, M. C.; VASCONCELOS, G.; FREITAS, E. Forecasting Brazilian inflation with high-dimensional models. *Brazilian Review of Econometrics*, v. 99, n. 99, 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, set. 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, feb. 1995.

PANKRATZ, A. A primer on ARIMA models. *In: Forecasting with dynamic regression models*. 1. ed. [S.l.]: Wiley-Interscience, 1991a.

PANKRATZ, A. Dynamic regression models in a vector ARMA framework. *In: Forecasting with dynamic regression models*. 1. ed. [S.l.]: Wiley-Interscience, 1991b.

- PANKRATZ, A. Building dynamic regression models: model identification. *In: Forecasting with dynamic regression models*. 1. ed. [S.l.]: Wiley-Interscience, 1991c.
- PEPPER, J. V. Forecasting crime: a city-level analysis. *In: Understanding crime trends*. Washington: National Academies Press, 2008.
- PERRY, W. *et al.* Predictive policing: the role of crime forecasting in law enforcement Operations. [S.l.]: RAND Corporation, 2013.
- PESARAN, M. H.; TIMMERMAN, A. Selection of estimation window in the presence of breaks. *Journal of Econometrics*, v. 137, n. 1, p. 134-161, mar. 2007.
- 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*, v. 39, n. 2, p. 17-34, dez. 2015. (Série Estatística).
- PROVENZA, M. M.; SERRA COSTA, J. F. DA; 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.
- SACHSIDA, A. *et al.* Inequality and criminality revisited: further evidence from Brazil. *Empirical Economics*, v. 39, n. 1, p. 93-109, aug. 2010.
- SALAKEN, S. M. *et al.* Extreme learning machine based transfer learning algorithms: a survey. *Neurocomputing*, v. 267, p. 516-524, jul. 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*, Elsevier, v. 29, n. 1, p. 191-201, 2013.
- 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, jun. 2012.
- TARASSOW, A. Forecasting U.S. money growth using economic uncertainty measures and regularisation techniques. *International Journal of Forecasting*, Elsevier, v. 35, n. 2, p. 443-457, 2019.

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, 2014.

Apêndice

Quadro 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) do 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) do 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 <i>et al.</i> , 2019).

Fonte: elaboração própria.

Tabela 6 – Teste de raiz unitária

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

Fonte: elaboração própria a partir dos dados da pesquisa.

Nota: constante (C), tendência (T) incluídas de acordo com critério de *Schwarz*. ADF: escolha de defasagens baseada no critério de Schwarz. PP e KPSS: *Bartlett Kernell* é método de estimação espectral utilizado e “*Bandwidth*” *Newey West* foi utilizado.

Tabela 7 - Estatísticas descritivas

	Homicídio	Internação
Média	3.971	4.341
Mediana	4.048	4.272
Máximo	5.339	5.677
Mínimo	2.464	3.341
Desv. Padrão	6.408	5.020

Fonte: elaboração própria a partir dos dados da pesquisa.

Quadro 2 - Estatísticas de avaliação da previsão

Raiz do erro quadrado médio (RMSE)

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

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)

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

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

Fonte: Hyndman (2015).

Tabela 8 - MAE normalizado por horizonte de previsão

Model	Horizonte de Previsão											
	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
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
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
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

Fonte: elaboração própria a partir dos dados da pesquisa.
Nota: para mais detalhes, ver Tabela 2.

Tabela 9 - MAD normalizado por horizonte de previsão

Model	Horizonte de Previsão											
	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
Affima	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
Splinef	0,74	0,78	0,59	0,62	0,61	0,46	0,60	0,67	0,66	0,76	0,55	1,00
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
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

Fonte: elaboração própria a partir dos dados da pesquisa.
Nota: para mais detalhes, ver Tabela 2.

Tabela 10 – Resultado do MCS baseado na “*rolling window*” de largura $ts=84$ entre 1999m01 e 2016m12 para homicídios. Estatística t_{rm} e função de perda quadrática

Horizonte de Previsão h = 1			Horizonte de Previsão h = 4		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	26.588	tslm1tsy	1	28.818
Horizonte de Previsão h = 8			Horizonte de Previsão h = 12		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1ts	1	28.608	tslm1ts	1	28.202

Fonte: elaboração própria a partir dos dados da pesquisa.

Nota: para mais detalhes, ver Tabela 3.

Tabela 11 – Resultado do MCS baseado na “*rolling window*” de largura $ts=96$ entre 1999m01 e 2016m12 para homicídios. Estatística t_{rm} e função de perda quadrática

Horizonte de Previsão h = 1			Horizonte de Previsão h = 4		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	27.376	tslm1tsy	1	29.104
			tslm1ts	2	29.511
Horizonte de Previsão h = 8			Horizonte de Previsão h = 12		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	28.608	tslm1ts	1	28.639

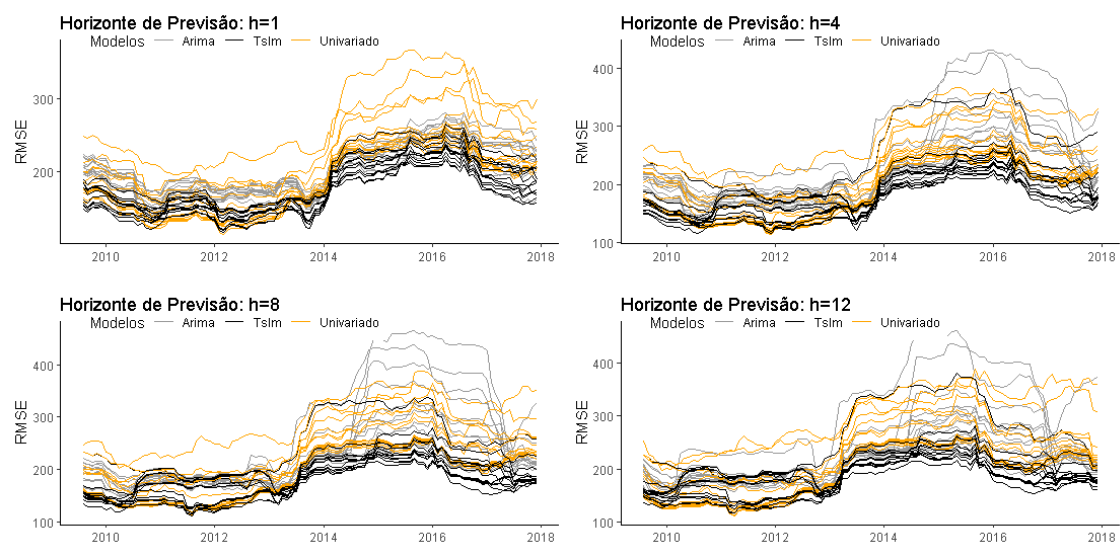
Fonte: elaboração própria a partir dos dados da pesquisa.

Nota: para mais detalhes, ver Tabela 3.

Tabela 12 - Resultado do MCS baseado na “rolling window” de largura $ts=96$ entre 1999m01 e 2016m12 para homicídios. Estatística t_{rm} e função de perda absoluta

Horizonte de Previsão h = 1			Horizonte de Previsão h = 4		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	125	tslm1tsy	1	128
Horizonte de Previsão h = 8			Horizonte de Previsão h = 12		
Modelo	Posição	Erro	Modelo	Posição	Erro
tslm1tsy	1	130	tslm1ts	1	133
			tslm1tsy	1	133

Fonte: elaboração própria a partir dos dados da pesquisa.
Nota: para mais detalhes, ver Tabela 3.

Figura 5 – *Rolling* RMSE - Classe de modelos

Fonte: elaboração própria a partir dos dados da pesquisa.