

PREVISÃO DE HOMICÍDIOS NO ESTADO DO RIO DE JANEIRO: UM *BENCHMARK*¹

Gabriel de Oliveira Accioly Lins²

1 INTRODUÇÃO

As melhores práticas na segurança pública requerem antecipado conhecimento da dinâmica criminal, auxiliando na preparação de políticas de segurança e organização da força policial, permitindo à polícia superar papel reativo ao crime, tornando-se mais proativa (Roeder *et al.*, 2015). A elaboração de previsões de variáveis criminais serve a diferentes objetivos, tais como: desdobramento tático de recursos policiais, alocação de pessoal e planejamento estratégico (Meijer e Wessels, 2019). A finalidade das previsões é fornecer um instrumento capaz de aumentar a consciência tática e estratégica das polícias, permitindo planejamento de policiamento mais proativo e eficaz.

À luz destas possibilidades teóricas, a aplicação de métodos preditivos antecipando tendências, identificando locais e indivíduos propensos à futura vitimização é prática disseminada entre departamentos de polícia dos Estados Unidos (Perry *et al.*, 2013). Assim, no caso de previsões espaciais de curto prazo, Gorr, Olligschlaeger e Thompson (2003) sugerem que métodos simples de previsão (por exemplo, suavização exponencial) apresentam desempenho preditivo superior a técnicas heurísticas até então adotadas pelas polícias. Por sua vez, Shoemith (2013) propõe modelo capaz de antecipar tendências criminais de longo prazo nos Estados Unidos, permitindo planejamento estratégico de políticas de segurança. Assim, de acordo com Levine *et al.* (2017), em Nova Iorque, a adoção de sistema de vigilância capaz de antecipar ocorrência de crimes aumentou a eficiência alocativa de pessoal e reduziu em US\$ 50 milhões despesas operacionais, sendo parcialmente responsável por redução de 6% no índice geral de crimes.

Apesar da prática ter sucesso internacional, na literatura nacional são escassos os trabalhos propondo modelos capazes de antecipar tendências criminais. Silva *et al.* (2017) propõem o CrimeVis, ferramenta de visualização de dados capaz de relacionar informações socioeconômicas e criminais, enquanto outros estudos utilizam técnicas de previsão e métodos de avaliação modestos (Provenza, Costa e Silva, 2015), em relação ao praticado na literatura internacional (Tashman, 2000).

Este trabalho tem como objetivo estabelecer um modelo preditivo do número de homicídios dolosos no estado do Rio de Janeiro, trazendo evidências de longo prazo e auxiliando o *policymaker* da segurança pública no planejamento da política de segurança, desta forma, aumentando o bem-estar social, além de estabelecer modelo *benchmark* aos interessados na previsão da dinâmica criminal.

A partir do método *rolling window*, um conjunto de modelos univariados foi estimado, e a precisão fora da amostra, em diversos horizontes de previsão, foi avaliada por meio de estatísticas de precisão de previsões e com a abordagem *model confidence set*, proposta por Hansen, Asger e James (2011). Assim, escolhendo como melhor modelo aqueles de melhor desempenho preditivo fora da amostra. Este trabalho inova ao implementar técnicas de previsão e avaliação inéditas na literatura nacional de previsão de homicídios – alinhadas às melhores práticas da literatura.

1. DOI: <http://dx.doi.org/10.38116/bapi35art5>

2. Pesquisador do Subprograma de Pesquisa para o Desenvolvimento Nacional (PNPD) na Diretoria de Estudos e Políticas do Estado, das Instituições e da Democracia do Instituto de Pesquisa Econômica Aplicada (Diest/Ipea); e mestre em economia pela Universidade Federal Fluminense (UFF).

As evidências sugerem grupo composto por suavização exponencial, *extreme learning machine*, modelo do tipo espaço-estado e combinações de previsões – média e mediana – apresentam melhor desempenho preditivo em todos os horizontes de avaliação considerados. Na média, ganhos preditivos relativos ao *random walk* de *benchmark*, em termos de raiz do erro quadrático médio (*root mean squared error* – RMSE), erro médio absoluto (*mean absolute error* – MAE) e desvio absoluto médio (*mean absolute deviation* – MAD) são da ordem de 40%. Entretanto, não foram encontradas evidências de modelo superior aos demais em todos os horizontes de previsão, em linha com outros resultados na literatura de previsão (Medeiros *et al.*, 2019).

2 DADOS E METODOLOGIA

A série temporal de interesse será o número mensal de homicídios dolosos no estado do Rio de Janeiro, entre 1992 e 2018, com dados procedentes do Instituto de Segurança Pública do Rio de Janeiro (ISP/RJ). Além da tradicional utilização da literatura sobre criminalidade (Montes e Lins, 2018), no período analisado constatou-se que o homicídio doloso representa 87,4% de todas as mortes por causas violentas (agregação de latrocínio, homicídio doloso, lesão corporal e homicídio por intervenção policial) registradas pelas polícias. Ou seja, o homicídio doloso é o crime violento de maior importância no acompanhamento da dinâmica criminal.

A existência de raiz unitária na série temporal foi verificada a partir dos testes *augmented* Dickey-Fuller (ADF), Phillips-Perron (PP) e Kwiatkowski-Phillips-Schmidt-Shin (KPSS). Os resultados indicam homicídio como série I (0).³

Como forma de avaliar o maior conjunto de informações possíveis, e por causa da flexibilidade dos métodos de avaliação, foram estimados diversos modelos univariados. Os métodos empregados foram os seguintes.

- 1) *Random walk* – 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, $\hat{Y}_{T+h|t} = Y_{T+h-1|t}$. Este modelo será utilizado como *benchmark*.
- 2) *Snaive* – neste caso, o número de homicídios é igual ao número de doze meses antes $\hat{Y}_{T+h|t} = Y_{T+h-12|t}$.
- 3) 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 temporal a partir de classificação proposta em (Hyndman *et al.*, 2008).
- 4) *Autoregressive integrated moving average* (Arima) – aplicação de algoritmo proposto por Hyndman e Khandakar (2008), trata-se de regressão com erros Arima (p, d, q), estimada a partir de máxima verossimilhança. O algoritmo seleciona a ordem do termo autorregressivo p e média móvel q de acordo com critério de informação AICc.
- 5) *Seasonal and trend decomposition using Loess* (STL) – neste caso aplica-se a decomposição do tipo STL, proposto por Cleveland, Cleveland e Terpenning (1990). Isto é, decomparamos a série de homicídios em componente sazonal, tendência e resíduos e realizamos separadamente ajuste e previsão destes componentes. Ajustam-se séries livres da sazonalidade (tendência e termo de erro) por meio de modelo ETS (STLe) e Arima (STLa). Em seguida, é feita a previsão da série ajustada sazonalmente, então os resultados são ressazonalizados. Como o padrão sazonal se repete ao longo da série analisada,

3. As estatísticas descritivas e o resultado do teste de raiz unitária estão disponíveis nas tabelas A.1 e A.2 do apêndice.

adicionamos à série ajustada sazonalmente a estimativa do componente sazonal do ano anterior, ou seja, a previsão do componente sazonal é realizada a partir do método *random walk*.

- 6) *Bootstrap Aggregating* – método proposto por Bergmeir, Hyndman e Benitez (2016). A partir da série temporal original são criadas 150 amostras de treinamento. Em cada amostra o algoritmo de Hyndman e Khandakar (2008) é ajustado. Ao final, a previsão é elaborada a partir do valor médio das previsões.
- 7) *Tbats* – proposto por Livera, Hyndman e Snyder (2011), o método *Tbats* permite modelar diferentes padrões de sazonalidade e tendência. Utiliza abordagem de decomposição ao modelar a série temporal em seus componentes de tendência, sazonalidade e ruído. A sazonalidade é modelada usando funções trigonométricas, permitindo que o modelo capture padrões sazonais complexos.
- 8) *Bats* – similar ao *Tbats*, este método não incorpora componente sazonal trigonométrico.
- 9) *Arfima* – modelo *Arfima* (p, d, q) selecionado e estimado a partir 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 .
- 10) *Neural network autoregressive* (*Nnetar*) – o método *Nnetar* é um modelo de previsão de séries temporais que combina redes neurais artificiais com modelos autorregressivos. O modelo usa uma rede neural na captura de padrões não lineares na série temporal, enquanto o modelo autorregressivo é usado para capturar a dependência temporal.
- 11) *Struct* – modelo estrutural básico de série temporal, ou seja, modelo de tendência linear local acrescido de componente sazonal. O modelo assim expresso permite utilização de filtro de Kalman (1960) na estimação do vetor de estado e das previsões requeridas – parâmetros desconhecidos estimados a partir de máxima verossimilhança.
- 12) *Extreme learning machine* (*ELM*) – trata-se de rede neural do tipo *feedforward* de camada oculta única, proposta em (Huang, Zhu e Siew, 2006). O *ELM* é um modelo de rede neural de uma única camada oculta, em que os pesos das conexões entre a camada de entrada e a camada oculta são inicializados aleatoriamente, e a matriz de pesos para a camada de saída é calculada por uma solução de mínimos quadrados.
- 13) *Thief* – método de previsão a partir de hierarquia temporal. Neste caso, a série homicídio mensal é agregada de modo bimestral, trimestral, quadrimestral, semestral e anual. Então, as previsões das séries agregadas são realizadas por meio dos algoritmos propostos por Hyndman e Khandakar (2008) (*Thief*) e do algoritmo de suavização exponencial de Hyndman *et al.* (2002) (*Thiefe*). Em seguida, em consonância com os melhores resultados obtidos por Athanasopoulos *et al.* (2017), as previsões temporalmente agregadas são individualmente reconciliadas a partir do algoritmo *structural scaling*.

Além deste conjunto de modelos univariados, são incluídos dois esquemas de combinações de modelos, isto é, a média e mediana das previsões univariadas. De acordo com os melhores resultados obtidos na literatura (Samuels e Sekkel, 2017), modelos de baixa capacidade preditiva são desconsiderados e, em cada horizonte de previsão as combinações incluem somente modelos univariados de RMSE médio inferior ao RMSE médio dos modelos univariados naquele horizonte.

Ao avaliar o desempenho preditivo são utilizadas duas abordagens. O primeiro método emprega estatísticas de precisão, ou seja, são utilizadas a RMSE, o MAE e o MAD. Em cada estatística, a razão entre modelo considerado e modelo *random walk* foi calculada. Resultado inferior à unidade sugere que os modelos considerados apresentam melhor desempenho preditivo relativo ao *random walk*.

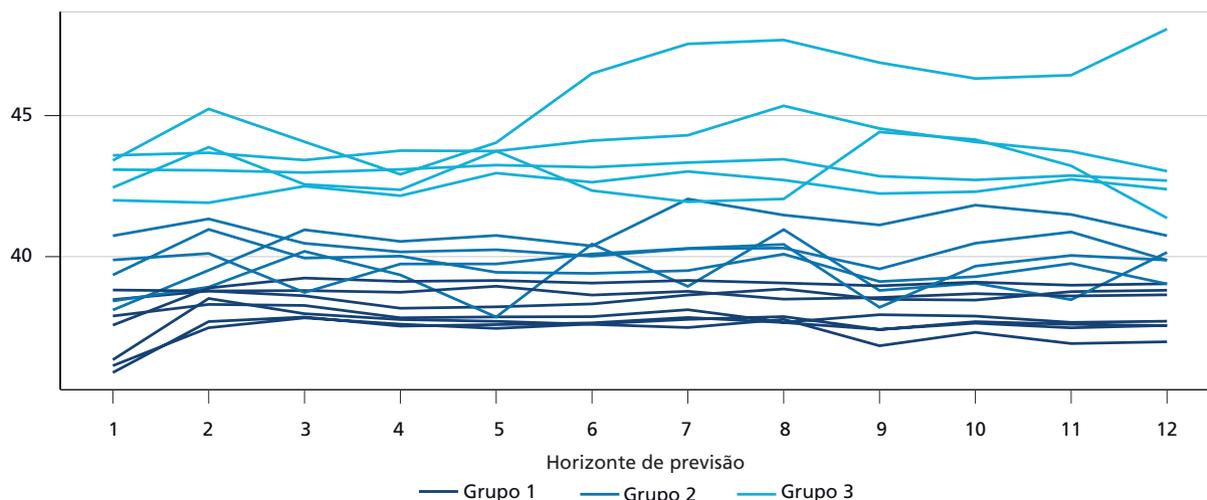
Apesar de ampla utilização na literatura, as estatísticas de precisão não consideram a significância estatística da diferença de desempenho preditivo dos modelos. Portanto, o desempenho preditivo dos diversos modelos também será comparado utilizando abordagem *model confidence set* (MCS) de (Hansen, Asger e James, 2011). O MCS consiste em sequência de testes de equidade de habilidade preditiva entre todos os modelos preditivos considerados, selecionando, assim, um conjunto de modelos estatisticamente indistinguíveis em termos de capacidade de previsão. Ao final, são selecionados os modelos com melhor capacidade preditiva (Medeiros, 2022).

3 PRINCIPAIS RESULTADOS

Em cada modelo são realizadas previsões fora da amostra, $\hat{y}_{i,T+h|T}$, utilizando método *rolling window* com janela de estimação de tamanho $w = 162$ observações mensais, assim, minimiza potenciais *outliers* e quebras estruturais. Em cada janela, os modelos são estimados e previsões de um a doze 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). Isto resulta em $162 - h$ previsões fora da amostra em cada horizonte de previsão h , as quais serão utilizadas na avaliação de desempenho dos modelos, selecionando como vencedor o modelo de melhor desempenho fora da amostra. A seção 4 mostra que os principais resultados são estáveis em janela de estimação $T_s = 184$.

Ao analisar evidências qualitativas, a figura 1 apresenta o RMSE de três grupos preditivos, agregados de acordo com a capacidade preditiva, dos modelos estimados ao longo dos horizontes avaliados. É possível perceber três padrões de desempenho preditivo. O grupo um, composto pela média, mediana e modelos Tbats, Stlme, Stlma, ETS e ELM, apresenta, ao longo de todos os horizontes de previsão, o menor RMSE e, portanto, o melhor desempenho preditivo e relativa estabilidade na capacidade preditiva. O segundo grupo, composto pelo modelo de erros Arima, Bats, Thiefa e Thiefe, embora de desempenho inferior ao grupo um, apresenta gradual melhoria ao longo dos horizontes de previsão. Por fim, o grupo três, constituído pelos demais modelos, apresenta desempenho preditivo inferior aos demais grupos em todos os horizontes de previsão e deterioração da capacidade preditiva.

FIGURA 1
RMSE por grupos preditivos e horizonte de previsão



Elaboração do autor.

De acordo com Medeiros *et al.* (2019), a tabela 1 apresenta sumário de estatística do desempenho preditivo médio ao considerarmos todos os horizontes de previsão. As colunas 1, 2 e 3 apresentam os RMSE, MAE e MAD médios, respectivamente. As colunas 4, 5 e 6 apresentam, respectivamente, o RMSE, MAE e MAD máximos, ou seja, a pior estatística preditiva do modelo. Por sua vez, as colunas 7, 8 e 9 apresentam o RMSE, MAE e MAD mínimos, a melhor estatística preditiva do modelo. 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) que cada modelo alcançou os menores valores de RMSE, MAE e MAD, respectivamente. Na última linha estão os valores apurados no modelo *random walk*. A tabela 1 apresenta *ranking* decrescente do RMSE médio.

As evidências sugerem que nas estatísticas de avaliação consideradas, em geral, os modelos competidores superam em todas as estatísticas o *benchmark random walk* – exceto o modelo *snaive*. Considerando todos os modelos, os ganhos preditivos relativos ao *random walk*, em termos de RMSE, MAE e MAD, são da ordem de 30%. Por sua vez, os cinco modelos de menor RMSE são, também, aqueles que geralmente obtêm os melhores resultados nas estatísticas MAE e MAD. Neste grupo de melhores modelos, as melhorias relativas ao *random walk* são de aproximadamente 40%, em termos de RMSE, MAE e MAD. Ainda de acordo com as evidências, nenhum modelo apresenta desempenho sistematicamente superior aos demais, ou seja, nenhum deles apresenta o melhor desempenho em todos os horizontes de previsão.

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 máximo	MAE máximo	MAD máximo	RMSE mínimo	MAE mínimo	MAD mínimo	RMSE mínimo ¹	MAE mínimo ¹	MAD mínimo ¹
Tbats	0,61	0,60	0,62	0,69	0,67	0,67	0,58	0,58	0,56	6	8	7
Mediana	0,61	0,61	0,66	0,68	0,66	0,74	0,57	0,58	0,58	3	1	1
Média	0,61	0,61	0,67	0,68	0,66	0,74	0,57	0,58	0,58	3	1	0
ETS	0,61	0,61	0,64	0,71	0,69	0,76	0,58	0,58	0,55	0	1	1
ELM	0,62	0,62	0,70	0,72	0,69	0,80	0,58	0,58	0,64	0	1	0
Stlme	0,63	0,64	0,68	0,73	0,72	0,80	0,59	0,60	0,59	0	0	0
Stlma	0,63	0,64	0,67	0,71	0,68	0,77	0,59	0,60	0,58	0	0	0
Thiefa	0,64	0,65	0,74	0,74	0,72	0,85	0,62	0,63	0,68	0	0	0
Thiefe	0,64	0,63	0,65	0,72	0,69	0,76	0,61	0,61	0,57	0	0	2
Arima	0,65	0,66	0,73	0,75	0,74	0,81	0,62	0,64	0,67	0	0	0
Bats	0,66	0,66	0,66	0,72	0,70	0,71	0,63	0,64	0,58	0	0	1
Struc	0,70	0,71	0,78	0,78	0,77	0,84	0,67	0,70	0,72	0	0	0
Nnetar	0,74	0,76	0,83	0,81	0,81	0,82	0,72	0,75	0,82	0	0	0
Snaive	0,86	0,88	0,84	1,00	0,99	0,95	0,81	0,82	0,76	0	0	0
RW	61,64	48,39	49,87	52,93	42,32	40,77	66,39	51,84	60,05	-	-	-

Elaboração do autor.

Nota: ¹ Representa o número de vezes (ao longo do horizonte de previsão) que cada modelo alcançou os menores valores.

Obs.: Última linha apresenta valor apurado no *benchmark random walk*.

A tabela 2 apresenta RMSE normalizado pelo modelo *random walk* de todos os modelos estimados ao longo dos horizontes de previsão. O grupo formado pela média, mediana e Tbats reúne modelos com as melhores previsões fora da amostra, embora comparativamente a outros modelos este desempenho superior é apenas marginal. Assim, a suavização exponencial (ETS), apesar de não constar no grupo de melhor desempenho, quase sempre aparece com desempenho marginalmente inferior ao modelo predictor vencedor.

TABELA 2
RMSE normalizado por horizonte de previsão

Modelo	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Arima	0,638	0,634	0,670	0,638	0,618	0,602	0,698	0,631	0,613	0,700	0,689	0,753
Bats	0,597	0,606	0,678	0,644	0,626	0,608	0,729	0,649	0,637	0,724	0,700	0,769
Tbats	0,569	0,591	0,629	0,600	0,579	0,566	0,649	0,591	0,570	0,645	0,622	0,698
Stlme	0,608	0,595	0,642	0,616	0,598	0,582	0,672	0,602	0,597	0,669	0,651	0,730
Stlma	0,588	0,597	0,650	0,622	0,601	0,588	0,678	0,611	0,603	0,676	0,657	0,737
ETS	0,594	0,588	0,634	0,601	0,581	0,570	0,660	0,589	0,587	0,655	0,635	0,712
Nnetar	0,680	0,695	0,730	0,682	0,677	0,700	0,824	0,747	0,726	0,802	0,784	0,909
Snaive	0,830	0,814	0,881	0,846	0,818	0,804	0,926	0,838	0,832	0,915	0,890	1,000
Theta	0,683	0,671	0,719	0,696	0,672	0,664	0,768	0,710	0,690	0,763	0,738	0,813
Arfima	0,658	0,643	0,704	0,670	0,660	0,642	0,746	0,668	0,654	0,732	0,721	0,801
Struc	0,665	0,674	0,705	0,674	0,672	0,638	0,727	0,658	0,688	0,764	0,729	0,781
Splinef	0,675	0,661	0,712	0,685	0,665	0,650	0,751	0,680	0,664	0,739	0,723	0,806
Thiefa	0,616	0,629	0,662	0,636	0,606	0,593	0,684	0,627	0,606	0,680	0,671	0,737
Thiefe	0,625	0,616	0,641	0,632	0,610	0,604	0,698	0,633	0,591	0,686	0,675	0,753
ELM	0,603	0,595	0,639	0,607	0,587	0,577	0,669	0,608	0,596	0,665	0,654	0,733
Média	0,566	0,575	0,626	0,598	0,575	0,566	0,654	0,592	0,579	0,652	0,634	0,709
Mediana	0,562	0,578	0,627	0,596	0,577	0,567	0,655	0,589	0,579	0,651	0,632	0,709
RW	63,8	65,1	60,3	62,9	65,1	66,4	57,7	63,9	64,6	57,8	59,3	52,9

Elaboração do autor.

Obs.: 1. Estatística do modelo de melhor desempenho destacado em negrito.

2. Última linha do valor apurado no *benchmark random walk*.

Segundo Tarassow (2019) são reportados os resultados de 1, 4, 8 e 12 como horizontes de previsão. A tabela 3 apresenta o *model confidence set* ao considerar todos os modelos estimados desde 1992, assumindo função de perda quadrática.

Em consonância com as evidências anteriores, a tabela 3 indica o domínio dos cinco modelos de melhor classificação no *ranking* da tabela 1 – média, mediana, Tbats, ETS e ELM. Especificamente, somente a mediana pertence ao grupo de melhores modelos em $t + 1$. Nos horizontes de previsão seguintes, o MCS selecionou diferentes combinações dos modelos de melhor qualificação no *ranking* apresentado. Nomeadamente, no caso $t + 4$ os cinco modelos foram incluídos no grupo de melhores modelos, com diferença de 3,4% na função de perda quadrática estimada entre o modelo vencedor e o último colocado. Em $t + 8$, dos cinco modelos de melhor qualificação, apenas ELM não pertence ao conjunto de melhores modelos. Por fim, no último horizonte de previsão considerado, $t + 12$, o conjunto de melhores modelos contém somente modelo Tbats.

Merece destaque a ausência no MCS do modelo com erros Arima estimado pelo algoritmo de Hyndman e Khandakar (2008) em todos os horizontes apresentados, usualmente utilizado em previsões. Além disso, o desempenho apresentado pelas combinações de previsões – média e mediana – justifica a inclusão destas na análise e ressalta a importância das combinações ao elaborar previsões. 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.

TABELA 3
MCS de *rolling window* com largura $ts = 162$: estatística t_{max} e função de perda quadrática

Horizonte de previsão h = 1			Horizonte de previsão h = 4		
Modelo	Classificação	Erro quadrático médio ¹	Modelo	Classificação	Erro quadrático médio ¹
Mediana	1	1285,04	Mediana	1	1406,45
			Média	2	1411,58
			Tbats	3	1424,02
			ETS	4	1428,67
			ELM	5	1454,50
Horizonte de previsão h = 8			Horizonte de previsão h = 12		
Modelo	Classificação	Erro quadrático médio ¹	Modelo	Classificação	Erro quadrático médio ¹
Mediana	1	1415,99	Tbats	1	1364,84
ETS	2	1416,68			
Tbats	3	1424,64			
Média	4	1431,96			

Elaboração do autor.

Nota: ¹ Refere-se ao valor médio da função de perda quadrática.

Obs.: 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, Asger e James (2011). Ainda de acordo com os autores do método, escolhemos nível de significância $\alpha = 10\%$ na elaboração do MCS ($M_{90\%}^*$).

Os resultados gerais são robustos ao uso alternativo de função de perda e estatística de avaliação. Como observado na tabela 4, ao utilizarmos função de perda absoluta, os modelos selecionados como melhores são os mesmos da função de perda quadrática.⁴ Desta forma, são apresentadas evidências da capacidade preditiva de conjunto de modelos univariados aos interessados em previsões do número de homicídios no estado do Rio de Janeiro e é estabelecido o *benchmark* aos interessados em se aprofundar nessa análise.

4. Considerando a estatística de teste t_{rm} , os modelos selecionados pelo MCS são subconjunto do caso t_{max} .

TABELA 4
MCS de *rolling window* com largura $ts = 162$: estatística t_{max} e função de perda absoluta

Horizonte de previsão $h = 1$			Horizonte de previsão $h = 4$		
Modelo	Classificação	Erro quadrático médio	Modelo	Classificação	Erro quadrático médio
Mediana	1	27,78	ELM	1	29,26
			ETS	2	29,37
			Tbats	3	29,42
			Média	4	29,43
			Mediana	5	29,47
Horizonte de previsão $h = 8$			Horizonte de previsão $h = 12$		
Modelo	Classificação	Erro quadrático médio	Modelo	Classificação	Erro quadrático médio
ETS	1	29,19	Tbats	1	28,78
Tbats	2	29,32	ETS	2	29,26
Mediana	3	29,54	Média	3	29,40
Média	4	29,58	Mediana	4	29,41

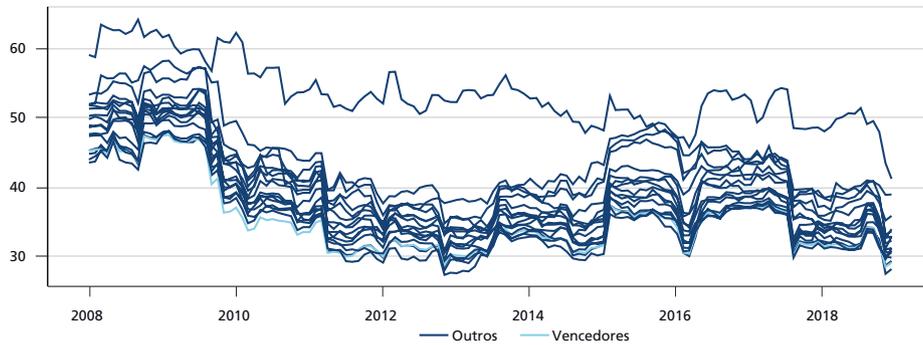
Elaboração do autor.
Obs.: Detalhes na tabela 3.

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 a partir do melhor desempenho preditivo ao considerar a 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.

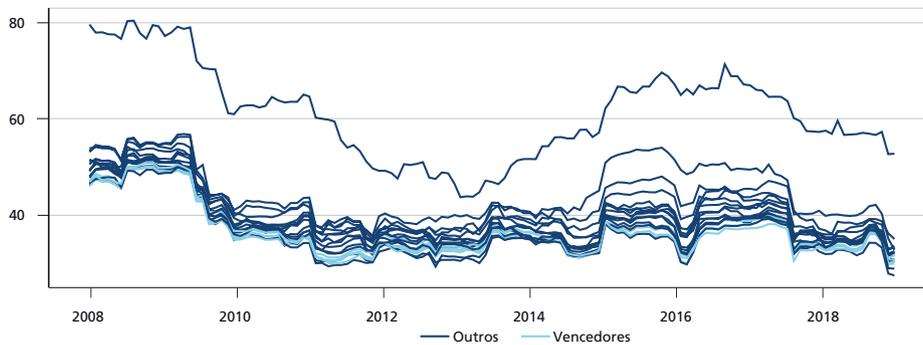
A figura 2 apresenta a variação temporal do RMSE de todos os modelos estimados e destaca aqueles selecionados pelo MCS. Em todos os horizontes de previsão analisados, observamos gradual melhora na qualidade das previsões ao longo da amostra. Ainda assim, os modelos escolhidos pelo MCS, exceto em curtos 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.

FIGURA 2
Rolling window RMSE: modelos escolhidos pelo MCS

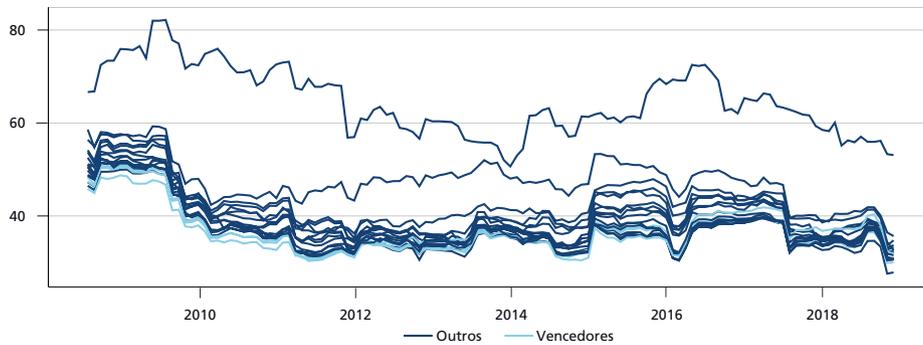
2A – Horizonte de previsão: $h = 1$



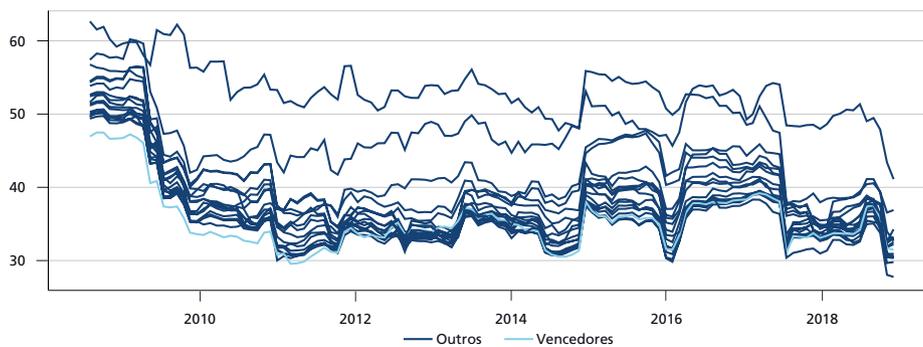
2B – Horizonte de previsão: $h = 4$



2C – Horizonte de previsão: $h = 8$



2D – Horizonte de previsão: $h = 12$



Elaboração do autor.

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). Este problema é contornado ampliando a largura da janela de estimação para $w = 184$ em vez de $w = 162$ observações mensais.

A tabela 5 apresenta os achados deste exercício. Em primeiro lugar, os modelos vencedores apresentam menores erros de previsão que aqueles reportados na tabela 3, indicando que as previsões são mais precisas na média ao escolhermos janela $w = 184$. Apesar disso, os resultados gerais são similares aos observados na janela de estimação mais curta. Isto é, grupo composto por suavização exponencial, *extreme learning machine*, modelo Tbat e combinações de previsões – média e mediana – apresentam melhor desempenho preditivo. Similar ao observado no exercício anterior, estes achados são robustos para a estatística de avaliação e função de perda. Portanto, são robustas as evidências de capacidade preditiva dos modelos analisados.

TABELA 5
MCS de *rolling window* com largura $w = 184$: estatística t_{max} e função de perda quadrática

Horizonte de previsão $h = 1$			Horizonte de previsão $h = 4$		
Modelo	Classificação	Erro quadrático médio	Modelo	Classificação	Erro quadrático médio
Média	1	1119,60	Média	1	1206,54
			Mediana	2	1221,06
			Elm	3	1218,70
			ETS	4	1236,28
			Tbats	5	1245,85
Horizonte de previsão $h = 8$			Horizonte de previsão $h = 12$		
Modelo	Classificação	Erro quadrático médio	Modelo	Classificação	Erro quadrático médio
Tbats	1	1190,59	Tbats	1	1152,56
Média	2	1190,52			
ETS	3	1196,77			
Mediana	4	1199,824			

Elaboração do autor.
Obs.: Detalhes na tabela 3.

5 CONCLUSÃO

A necessidade de antecipar a trajetória temporal da criminalidade violenta torna imperativa a elaboração de modelos previsores. Com o objetivo de produzir evidências preliminares acerca dos homicídios dolosos no Rio de Janeiro, este trabalho inovou ao investigar a capacidade preditiva de conjunto diverso de modelos univariados.

De acordo com as evidências encontradas, o grupo composto por suavização exponencial, *extreme learning machine*, modelo Tbat e combinações de previsões – média e mediana – apresenta melhor desempenho preditivo. Na média, ganhos preditivos relativos ao *benchmark random walk*, em termos de RMSE, MAE e MAD são da ordem de 40%. Entretanto, não existe evidência de modelo superior aos demais em todos os horizontes de previsão, em linha a outros resultados na literatura.

Os resultados gerais são robustos ao alternativo uso de estatística de avaliação, função de perda e largura de janela de estimação. Portanto, são apresentadas evidências da capacidade preditiva de modelos univariados aos interessados na previsão da dinâmica criminal fluminense e é estabelecido *benchmark* a futuras tentativas de previsão de homicídios no Rio de Janeiro.

REFERÊNCIAS

- ATHANASOPOULOS, G. *et al.* Forecasting with temporal hierarchies. **European Journal of Operational Research**, v. 262, n. 1, p. 60-74, 2017. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0377221717301911>>.
- 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. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169207015001120>>.
- 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. 373, 1990. Disponível em: <<https://search.proquest.com/openview/cc5001e8a0978a6c029ae9a41af00f21/1?pq-origsite=gscholar&cbl=105444>>.
- GORR, W.; OLLIGSCHLAEGER, A.; THOMPSON, Y. Short-term forecasting of crime. **International Journal of Forecasting**, v. 19, n. 4, p. 579-594, 2003. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S016920700300092X>>.
- HANSEN, R.; ASGER, L.; JAMES, M. N. The model confidence set. **Econometrica**, v. 79, n. 2, p. 453-497, 2011. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA5771>>.
- 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. Disponível em: <<https://www.jstor.org/stable/10.2307/2347679?origin=crossref>>.
- HUANG, G.-B.; ZHU, Q.; SIEW, C. Extreme learning machine: theory and applications. **Neurocomputing**, v. 70, n. 1-3, p. 489-501, 2006. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0925231206000385>>.
- 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. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0169207001001108>>.
- HYNDMAN, R. J. *et al.* **Forecasting with exponential smoothing: the state space approach**. [s.l.]: Springer-Verlag Berlin Heidelberg, 2008. Disponível em: <<https://www.springer.com/gp/book/9783540719168>>.
- 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. Disponível em: <<https://doi.org/10.18637/jss.v027.i03>>.

KALMAN, R. E. A new approach to linear filtering and prediction problems. **Journal of Basic Engineering**, v. 82, n. 1, p. 35, 1960. Disponível em: <<http://fluidsengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1430402>>.

LEVINE, E. S. *et al.* The New York city police department's domain awareness system. **Informations: Journal on Applied Analytics**, v. 47, n. 1, p. 70-84, 2017. Disponível em: <<http://pubsonline.informs.org/doi/10.1287/inte.2016.0860>>.

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. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1198/jasa.2011.tm09771>>.

MEDEIROS, M. C. Forecasting with machine learning methods. *In*: FELIX, C.; MÁTYÁS, L. (Ed.). **Econometrics with machine learning**. 1st ed. [s.l.]: Springer, 2022. p. 111-149. Disponível em: <https://link.springer.com/10.1007/978-3-031-15149-1_4>.

MEDEIROS, M. C. *et al.* Forecasting inflation in a data-rich environment: the benefits of machine learning methods. **Journal of Business and Economic Statistics**, v. 39, n. 1, p. 1-22, 2019. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/07350015.2019.1637745>>.

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. Disponível em: <<https://doi.org/10.12660/bre.v99n992016.52273>>.

MEIJER, A.; WESSELS, M. Predictive policing: review of benefits and drawbacks. **International Journal of Public Administration**, v. 42, n. 12, p. 1031-1039, 2019. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/01900692.2019.1575664>>.

MONTES, G. C.; LINS, G. O. Deterrence effects, socio-economic development, police revenge and homicides in Rio de Janeiro. **International Journal of Social Economics**, v. 45, n. 10, p. 1406-1423, 2018. Disponível em: <<https://www.emerald.com/insight/content/doi/10.1108/IJSE-09-2017-0379/full/html>>.

PERRY, W. *et al.* **Predictive policing**: the role of crime forecasting in law enforcement operations. [s.l.]: Rand Corporation, 2013. Disponível em: <https://www.rand.org/pubs/research_reports/RR233.html>.

PESARAN, M. H.; TIMMERMANN, A. Selection of estimation window in the presence of breaks. **Journal of Econometrics**, v. 137, n. 1, p. 134-161, 2007. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0304407606000418>>.

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. Disponível em: <<https://www.e-publicacoes.uerj.br/index.php/CESP/article/view/18992/13810>>.

ROEDER, O. K. *et al.* **What caused the crime decline?** New York: Brennan Center for Justice, 2015. p. 142. Disponível em: <https://www.brennancenter.org/sites/default/files/publications/What_Caused_The_Crime_Decline.pdf>.

SAMUELS, J. D.; SEKKEL, R. M. Model confidence sets and forecast combination. **International Journal of Forecasting**, v. 33, n. 1, p. 48-60, 2017. Disponível em: <<http://dx.doi.org/10.1016/j.ijforecast.2016.07.004>>.

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, 2013. Disponível em: <<http://dx.doi.org/10.1016/j.ijforecast.2012.08.002>>.

SILVA, L. J. S. *et al.* CrimeVis: an interactive visualization system for analyzing crime data in the State of Rio de Janeiro. *In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS*, 19., 2017, Porto, Portugal. **Proceedings...** Porto: Iceis, 2017. p. 193-200. Disponível em: <<http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006258701930200>>.

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. Disponível em: <<https://doi.org/10.1016/j.ijforecast.2018.09.012>>.

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, 2000. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0169207000000650>>.

APÊNDICE

TABELA A.1
Estatísticas descritivas

	Homicídio
Média	504,1
Mediana	505,0
Máximo	831,0
Mínimo	272,0
Desvio-padrão	111,2

Elaboração do autor.

TABELA A.2
Teste de raiz unitária

		ADF					
		Equação	Número de defasagens	Estatística <i>t</i>	10%	5%	1%
Homicídios	Em nível	C/T	1	-5,547	-3,134	-3,423	-3,986
	Primeira diferença						
		PP					
		Equação	Parâmetro de defasagem	Estatística <i>t</i>	10%	5%	1%
Homicídios	Em nível	C/T	4	-6,688	-3,134	-3,423	-3,986
	Primeira diferença						
		KPSS					
		Equação	Parâmetro de defasagem	Estatística <i>t</i>	10%	5%	1%
Homicídios	Em nível	C/T	13	0,106	0,119	0,146	0,216
	Primeira diferença						

Elaboração do autor.

Obs.: 1. Constante (C), tendência (T), incluídas de acordo com o critério de Schwarz.

2. ADF: escolha de defasagens baseada no critério de Schwarz.

3. PP e KPSS: Bartlett Kernel é o método de estimação espectral utilizado e *bandwidth* Newey West foi utilizado.