

Estudo temporal dos acidentes de trânsito em Rodovias Federais no Estado de Santa Catarina

Marco Aurélio Boehme dos Santos¹
Andrea Cristina Konrath²
Luiz Ricardo Nakamura³
Rodrigo Gabriel de Miranda⁴
Paulo Siga Thomaz⁵
Ana Julia Righetto⁶

Resumo

O presente trabalho tem como objetivo realizar um estudo temporal dos acidentes de trânsito em Rodovias Federais do Estado de Santa Catarina. Os dados utilizados foram obtidos no banco de dados públicos e oficial da Polícia Rodoviária Federal (PRF), no qual um boletim de acidente de trânsito (BAT) é emitido. A série de dados corresponde ao período de 2007 até outubro de 2019, divididos em três categorias: acidentes leves, graves e fatais. Em decorrência de algumas características específicas observadas nas séries em estudo, o modelo autorregressivo integrado de média móvel sazonal (SARIMA) foi utilizado. O melhor modelo ajustado (selecionado com base no valor de U-Theil e no critério de informação de Akaike) foi utilizado para realizar previsões de acidentes, para cada uma das categorias em estudo, para os meses de novembro de 2019 até abril de 2020. Essas previsões foram posteriormente comparadas com os valores reais observados, a fim de se verificar a acurácia do modelo. Os resultados mostram que o modelo SARIMA ajustado é uma boa alternativa para o estudo das séries temporais analisadas.

Palavras-chaves: Modelo SARIMA; Previsão de acidentes; Técnicas Estatísticas.

Abstract

The main aim of this paper was to perform a temporal analysis of traffic accidents in federal highways in Santa Catarina, Brazil. The data was obtained from the official public database of the Polícia Rodoviária Federal (PRF), where a traffic accident bulletin is issued. The data series corresponds of a period from 2007 to 2019, and divided into three groups: minor, serious and fatal accidents. Due to some specific characteristics observed in all series, the seasonal autoregressive integrated moving average (SARIMA) model was considered. The best fitted

¹ Graduando do Curso de Ciências Econômicas – UFSC. E-mail:

boehme.marco@outlook.com.br

² Departamento de Informática e Estatística - UFSC

³ Departamento de Informática e Estatística - UFSC

⁴ Membro do Núcleo de Normalização e Qualimetria – NNQ - UFSC

⁵ Mestre em Modelagem Computacional - FURG

⁶ Alvez Agritech, Londrina, Paraná

model (based on U-Theil and Akaike information criterion values) was then used to predict the number of accidents for each categories from November 2019 to April 2020. These forecasts were compared with the real observed values in order to assess the model adequability. Results showed that the SARIMA model is a reasonable alternative to investigate the analysed time series and perform the forecasts.

Keywords: SARIMA model; Accident prediction; Statistical techniques

1. Introdução

Em maio de 2011, foi lançada pela Organização das Nações Unidas (ONU) a chamada Década de Ação pela Segurança no Trânsito 2011-2020, tendo como o objetivo que os governantes se comprometessem a aderir a novas medidas de prevenção contra os acidentes no trânsito (ONUBR, 2011). A criação desta ação pode ser justificada pelo elevado número de óbitos decorrentes dos acidentes de trânsito, que figura entre as principais causas de mortalidade de jovens do mundo (WHO, 2018). Adicionalmente, conforme a Organização Pan-americana da Saúde (OPAS, 2019), para a maioria dos países, os acidentes de trânsito custam aproximadamente 3% de seu Produto Interno Bruto (PIB).

O estudo realizado pelo Ministério de Transportes, Portos e Aviação Civil (EBELING et al., 2018), com base nos dados da Polícia Rodoviária Federal (PRF, 2019a, 2019b), identificou que as principais causas de acidentes nas rodovias federais, no período de 2007 e 2016 foram: o desrespeito às normas de trânsito (30,3%) e as condições de adormecimento e falta de atenção do condutor (23,4%), que estão relacionadas ao fator humano. Essas pesquisas são de suma importância, uma vez que, segundo Figueira et al. (2017), a identificação correta das prováveis causas dos acidentes pode fomentar a criação de medidas preventivas e reduzir o número de ocorrências. Corroborando com essa afirmação, temos os resultados obtidos na maioria dos estudos sobre acidentes de trânsito em países desenvolvidos (ABDEL-ATY, 2003; AGUERO-VALVERDE; JOVANIS, 2006; BARROSO JUNIOR et al., 2019).

Vários trabalhos vêm sendo desenvolvidos nesta temática, e o emprego de técnicas estatísticas, desde as mais simples até as mais sofisticadas, assume

papel de destaque, uma vez que fornecem o aporte formal para desenvolver políticas de prevenção mais específicas (BARROSO JUNIOR et al., 2019).

Dentre as técnicas mais sofisticadas, destacamos aqui aquelas referentes à análise de séries temporais, aplicadas nas mais diversas áreas, como por exemplo, na modelagem da demanda de energia elétrica (IRFFI et al., 2009); SILVEIRA, 2017; NUNES, 2019); previsão de incidência de doenças (EARNEST et al., 2005; AAKHTAR; ROZI, 2009; ALZHRANI; ALJAMAAN; AL-FAKIH, 2020; BENVENUTO et al., 2020); e também no estudo de acidentes de trânsito, foco deste trabalho. Alguns exemplos recentes de aplicações são ilustrados a seguir.

Vitorino (2019) utilizou o modelo dinâmico de Poisson para ajustar dados de acidentes automobilísticos nas Rodovias Federais (BRs) no Estado da Paraíba, entre os anos de 2013 a 2018. Dentre as técnicas foi utilizada a análise de séries temporais via inferência bayesiana. A autora conclui que a modelagem foi satisfatória, com um bom ajuste aos dados notificados, o qual pode ser empregada pelo poder público para diminuir e prever o número de acidentes de trânsito no estado.

Rodrigues, Santos e Nery (2018) realizaram uma pesquisa que teve como objetivo analisar a tendência temporal da mortalidade de acidentes de trânsito envolvendo motocicletas em Sergipe, no período de 2000 a 2016. Os autores observaram que a tendência temporal foi crescente, apresentando crescimento médio anual de 1,3 óbito para cada 100.000 homens. Verificaram também que devido a diversos fatores é necessária a realização de ações intersetoriais de prevenção, na perspectiva de reduzir esses acidentes, e por consequência, os óbitos decorrentes deles.

Parvareh et al. (2018) realizaram um estudo para modelar os padrões de acidentes de trânsito que levam a lesão na província do Curdistão, Irã, no período de março de 2009 a fevereiro de 2015. Os autores utilizaram o modelo de média móvel integrada autorregressiva (ARIMA) e o modelo autorregressivo integrado de média móvel sazonal (SARIMA) para prever os acidentes até 24 meses depois (fevereiro de 2017).

Eze, Asogwa e Okonkwo (2018) utilizaram o modelo SARIMA para modelar

casos de acidentes em quatro estradas principais que levam à cidade do estado de Enugu, na Nigéria. Foram analisados casos de acidentes do período de janeiro de 2007 a dezembro de 2015. Segundo os autores, o modelo ajustado foi utilizado para prever casos mensais de acidentes em cada uma das estradas para o ano de 2016.

Sanusi, Adebola e Adegoke (2016) analisaram dados de acidentes na Nigéria entre o período de 1960 e 2013. Os autores classificaram os acidentes como: acidentes graves, acidentes fatais e acidentes leves, e utilizaram os modelos autorregressivos integrados de médias móveis (ARIMA) (BOX; JENKINS, 1976). De acordo com os autores, os modelos obtidos foram utilizados para prever os diferentes casos de acidentes para o período de 2014 a 2020, e a previsão apresentou um aumento médio no número de casos. Previamente, Balogun et al. (2015) analisaram dados mensais de acidentes de trânsito no período de janeiro de 2004 a dezembro 2011, também na Nigéria, utilizando novamente os modelos ARIMA.

Destarte, o presente trabalho tem como objetivo realizar um estudo temporal dos acidentes de trânsito em rodovias federais para o Estado de Santa Catarina, utilizando os modelos ARIMA e considerando uma componente sazonal.

2. Fundamentação teórica

Na literatura, existem diferentes modelos para previsões, desde o mais simples, o qual consiste em usar a observação mais recente como uma previsão (método ingênuo) até métodos mais complexos, como, por exemplo, redes neurais. Conforme Hyndman e Athanasopoulos (2018), apresentam as seguintes classificações:

1) Modelos transversais: são utilizados quando a variável a ser prevista apresenta uma relação com uma ou mais variáveis. Neste caso, temos os seguintes modelos: modelos de regressão, modelos aditivos e alguns tipos de redes neurais;

2) Previsão de séries temporais: utiliza apenas informações sobre a

variável a ser prevista e não incluem fatores que podem afetar seu comportamento, como os modelos ARIMA, suavização exponencial e modelos estruturais; e,

3) Modelos dinâmicos de regressão: combinam variáveis explicativas com séries temporais, sendo também conhecidos como modelos de dados em painel, modelos longitudinais, modelos de função de transferência e modelos de sistemas lineares.

Neste trabalho, será empregado o modelo SARIMA, pertencente ao segundo grupo supracitado, uma vez que a série de estudo é referente a dados mensais e, desta maneira, pode apresentar um comportamento sazonal nos dados.

2.1. Modelos Autorregressivos e de Médias Móveis - ARIMA

Os modelos de Box e Jenkins (1976), denominados autorregressivos integrados de médias móveis (ARIMA), visam descrever a autocorrelação presente nos dados. Os modelos se baseiam na hipótese que as séries temporais são estacionárias, ou seja, as características estatísticas, como média, variância e autocorrelação são constantes ao longo do tempo (MORETTIN e TOLOI (2006), GUJARATI e PORTER (2011)). Quando a série é considerada estacionária, sua modelagem pode ser realizada por meio de um processo autorregressivo (AR), processo de média móvel (MA) e processo autorregressivo de médias móveis (ARMA). Os modelos ARIMA são utilizados para séries não estacionárias, enquanto os modelos SARIMA consideram uma componente extra sazonal. Maiores detalhes podem ser encontrados em Box, Jenkins e Reinsel (1994), Morettin e ToloI (2006), Ehlers (2009), Gujarati e Porter (2011) e Hyndman e Athanasopoulos (2018).

De acordo com Box e Jenkins (1976), um modelo $AR(p)$ é dado por:

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + a_t, \quad (1)$$

Em que p é a ordem de AR; Z_t é a série temporal observada no momento

t , a_t é o ruído aleatório que segue uma distribuição normal com média zero e variância um; e $\varphi_1, \dots, \varphi_p$ são os parâmetros a serem estimados. Ao inserir o operador de defasagem B , que determina $Z_{t-1} = BZ_t$; e, logo, $Z_{t-p} = B^p Z_t$, a Equação (1) pode ser reescrita como:

$$\varphi_p(B)Z_t = a_t, \quad (2)$$

Em que $\varphi_p(B)$ é o operador de $AR(p)$, definido por $\varphi_p(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$. Enquanto que o modelo $MA(q)$, conforme Box e Jenkins (1976), é dado por:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (3)$$

E pode ser reescrito como:

$$Z_t = \theta_q(B)a_t, \quad (4)$$

Em que q é a ordem de $MA(q)$; $\theta_1, \dots, \theta_q$ são os parâmetros do modelo $MA(q)$; e $\theta_q(B)$ é definido por $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$. As Equações (2) e (4) são combinadas, assim obtendo-se o modelo $ARMA(p, q)$, o qual possui processos $AR(p)$ e $MA(q)$, e pode ser representado por:

$$\varphi_p(B)Z_t = \theta_q(B)a_t. \quad (5)$$

Em situações em que a série é não estacionária, cabe um procedimento de transformação mediante processo de diferenciação dos dados. Neste caso é utilizada a componente de integração $I(d)$, que resulta no modelo $ARIMA(p, d, q)$, ou seja:

$$\varphi_p(B)\Delta^d Z_t = \theta_q(B)a_t, \quad (6)$$

Em que $\Delta^d Z_t = (1 - B)^d Z_t$ e d é a ordem de diferenciação.

Para situações em que as séries apresentam autocorrelação sazonal, Box e Jenkins (1976) definiram o modelo ARIMA sazonal multiplicativo, denominado como SARIMA(p,d,q)(P,D,Q) $_s$, dado por:

$$\varphi_p(B)\phi_p(B^S)\Delta^d\Delta S^D Z_t = \theta_q(B)\theta_Q(B^S)a_t, \quad (7)$$

Em que $\Delta S^D Z_t = (1 - B^S)^D Z_t$, D é a ordem de diferenciação sazonal; $\Phi_p(B^S)$ é operador sazonal AR(p), definido por $\Phi_p(B^S) = 1 - \Phi_1(B^S) - \Phi_2(B^{2S}) - \dots - \Phi_p(B^{pS})$; $\Theta_Q(B^S)$ é o operador sazonal MA(q) por $\Theta_Q(B^S) = 1 - \Theta_1(B^S) - \Theta_2(B^{2S}) - \dots - \Theta_Q(B^{qS})$; e Φ_1, \dots, Φ_p são parâmetros do modelo sazonal AR(p), enquanto $\Theta_1, \dots, \Theta_Q$ são parâmetros do modelo sazonal MA(q).

Em termos de recursos computacionais, existem *softwares* estatísticos que já possuem as funções implementadas para a aplicação dos modelos de Box e Jenkins, como, por exemplo, o *software* R (R CORE TEAM, 2020), cuja função *auto.arima* executa, automaticamente, o ajuste de vários modelos, comparando-os e selecionando o melhor modelo, de acordo com o critério de informação de Akaike (AKAIKE, 1973), apresentando sua ordem (p, d, q).

2.2. Avaliação do desempenho dos modelos

Conforme Moreira (2008), as previsões não são exatas e o erro associado é maior quanto maior for seu horizonte. Assim, o erro de previsão é a diferença entre o valor ocorrido e o valor previsto, dado por: $e_t = \hat{Y}_t - F_t$, em que e_t é o erro de previsão no período t ; \hat{Y}_t é o valor previsto no período t ; e F_t é o valor observado no período t , mas quando não é conhecido, utiliza-se o valor aceito.

De acordo com Souza, Samohyl e Miranda (2008), há diferentes índices que medem o erro de previsão, auxiliando assim na busca de quais modelos podem fornecer a melhor estimativa possível. Na Tabela 1 são apresentados esses índices; para maiores detalhes, verificar o trabalho de Souza, Samohyl e Miranda (2008).

Tabela 1: Principais índices que medem o erro de previsão.

Índice	Significado	Fórmula
ME	Erro Médio	$ME = \frac{1}{n} \sum_{t=1}^n e_t \cong 0$
MAE	Erro Absoluto Médio	$MAE = \frac{1}{n} \sum_{t=1}^n e_t $
MSE	Erro Quadrado Médio	$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2$
MPE	Erro Percentual Médio	$MPE = \frac{1}{n} \sum_{t=1}^n \frac{Y_t - F_t}{Y_t}$
MAPE	Erro Percentual Absoluto Médio	$MAPE = \frac{1}{n} \sum_{t=1}^n \left \frac{Y_t - F_t}{Y_t} \right $
U de Theil	Coefficiente de Desigualdade	$U^2 = \frac{\sum_{t=1}^{r-1} \left\{ \left(\frac{Y_{i+1} - Y_i}{Y_i} \right) - \left(\frac{F_{i+1} - Y_i}{Y_i} \right) \right\}^2}{\sum_{t=1}^{r-1} \left(\frac{Y_{i+1} - Y_i}{Y_i} \right)^2}$

Fonte: Souza, Samohyl e Miranda (2008).

Na estatística U-Theil, seus valores variam sempre entre zero e um, no qual as melhores previsões são aquelas que se aproximam de zero. Conforme Ehlers (2009), o erro superior a um indica que o modelo proposto possui erro superior ao passeio aleatório.

Além dos índices mencionados na Tabela 1, pode-se optar pela utilização de outros critérios amplamente utilizados, como, por exemplo, o critério de informação de Akaike (AIC), definido por $AIC = -2 \log(\hat{L}) + 2k$, em que k é o número de parâmetros do modelo e \hat{L} é a função de verossimilhança estimada do modelo. O modelo que apresenta menor AIC é considerado o de melhor ajuste.

3. Materiais e métodos

Os dados utilizados foram obtidos no banco de dados públicos oficial da Polícia Rodoviária Federal (PRF, 2019b), no qual um boletim de acidentes de trânsito (BAT) é emitido e, desta forma, os acidentes podem ser visualizados.

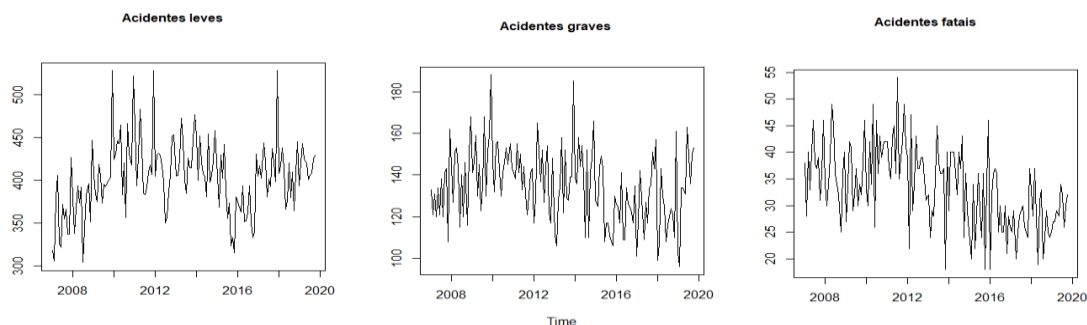
As séries originais dos dados correspondem ao período de 2007 até outubro de 2019, e foram divididas em três categorias: acidentes leves, graves e fatais, conforme abordagem adotada por Sanusi et al. (2016). As observações de novembro de 2019 até abril de 2020 também foram coletadas, para a realização de comparações com as previsões obtidas pelo melhor modelo ajustado para cada uma das categorias consideradas.

Todas as análises foram realizadas com auxílio do *software* R (R Core Team, 2020). Para as previsões foi utilizado o pacote *forecast* (HYNDMAN et al., 2020) e utilizou-se a função *auto.arima()* que ajusta, de forma automática, diversos modelos e seleciona o melhor, baseado em algum critério de qualidade de ajuste, como por exemplo o AIC, utilizado neste trabalho. Para escolher o modelo candidato foi utilizada a estatística de U-Theil e o critério de informação AIC.

Finalmente, a fim de satisfazer as premissas assumidas, isto é, para verificar se todas as pressuposições assumidas pelo modelo ajustado foram atendidas e, conseqüentemente, validar as previsões realizadas, com um nível de significância de 5%, foram utilizados: o teste Dickey-Fuller (DICKEY; FULLER, 1981), para verificação da estacionariedade da série; o correlograma e testes de Ljung-Box e Box-Pierce, para teste de correlação dos resíduos (LJUNG; BOX, 1978); e o teste Jarque-Bera (JARQUE; BERA, 1987), para verificar a normalidade dos resíduos.

4. Resultados e discussão

Para identificação dos modelos, inicialmente foi criado um gráfico com as séries de acidentes leves, acidentes graves e acidentes fatais (Figura 1). As séries utilizadas na modelagem totalizam 154 observações mensais cada, começando no ano de 2007 até outubro de 2019.

Figura 1: Séries de acidentes mensais: leves, graves e fatais.

Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Na Tabela 2 são apresentadas as medidas descritivas para cada uma das categorias. Pode-se observar que as médias de acidentes mensais das categorias leves, graves e fatais são, respectivamente, 402,99, 134,84 e 33,53. Ademais, todas as categorias apresentam coeficiente de assimetria próximo a zero e coeficiente de curtose próximo a três, o que fomenta a aplicação da modelagem escolhida neste trabalho, uma vez que a distribuição dos dados apresenta simetria, com grau de curtose próximo à da distribuição normal.

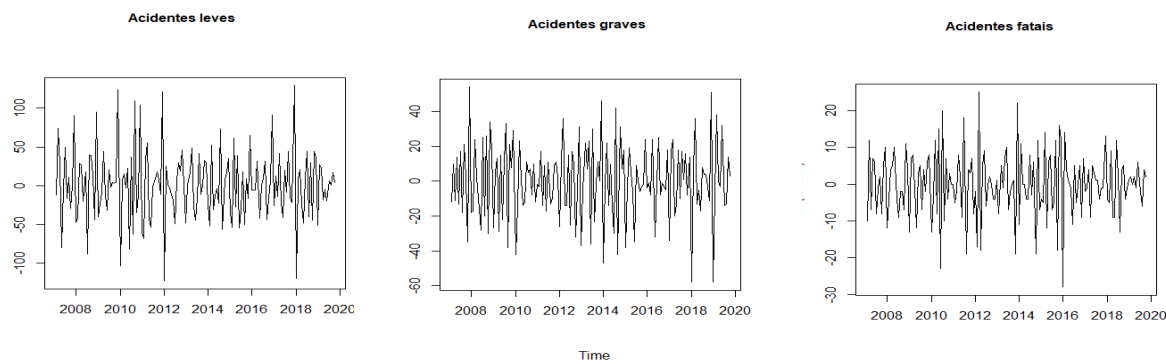
Tabela 2: Medidas descritivas.

	Leves	Graves	Fatais
Observações	154	154	154
Média	402,99	134,84	33,53
Mínimo	304	96	18
Q1	380	122	28
Mediana	405	135	34
Q3	425,75	146	39
Máximo	528	188	54
Desvio Padrão	41,78	17,31	7,38
Assimetria	0,27	0,14	0,08
Curtose	3,82	2,87	2,57

Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Todas as séries em estudo foram consideradas estacionárias, baseado no teste de Dickey Fuller (acidente leves: p -valor=0,02; acidentes graves: p -valor=0,01; e acidentes fatais: p -valor=0,03). Entretanto, com o intuito de atender todas as pressuposições básicas da aplicação de um modelo SARIMA, foi utilizada a primeira diferença das séries (Figura 2).

Figura 2: Primeira diferença das séries de acidentes mensais: leves, graves e fatais.



Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Com base na função *auto.arima()*, disponível no *software* R, foram ajustados, analisados e comparados a diferentes modelos SARIMA, a partir da primeira diferença da série. A Tabela 3 apresenta os três melhores modelos, baseado nos valores retornados de AIC, para cada uma das categorias de acidentes (leves, graves e fatais).

Tabela 3: Critérios de Informação para os três melhores modelos para cada categoria (Continua).

	Modelo SARIMA	AIC
Acidentes Leves	$(0,1,1)(1,1,1)_{12}$	1.367,974
	$(0,1,1)(0,1,1)_{12}$	1.368,069
	$(0,1,1)(0,1,2)_{12}$	1.368,314
Acidentes graves	$(0,1,1)(1,0,1)_{12}$	1.272,875
	$(3,1,1)(1,0,1)_{12}$	1.273,126
	$(1,1,1)(1,0,1)_{12}$	1.274,681

Tabela 3: Critérios de Informação para os três melhores modelos para cada categoria (Continuação).

	Modelo SARIMA	AIC
Acidentes fatais	$(0,1,1)(0,1,1)_{12}$	942,194
	$(2,1,1)(0,1,1)_{12}$	942,942
	$(3,1,2)(2,1,0)_{12}$	942,942

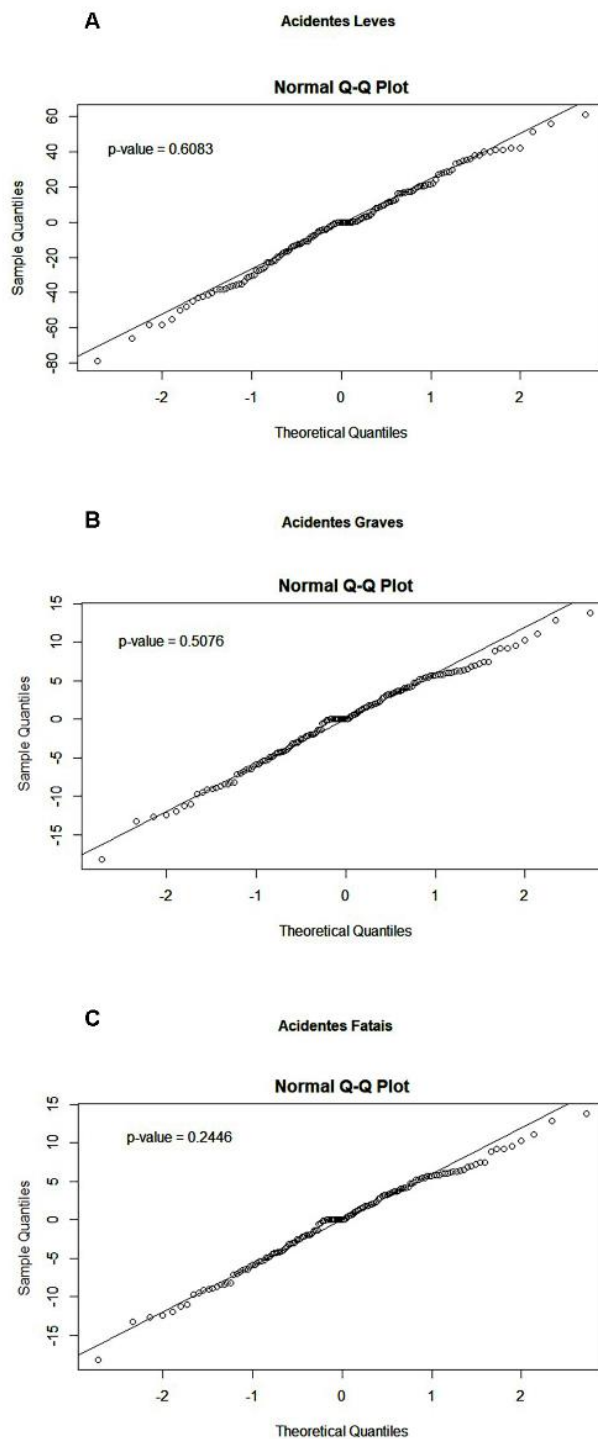
Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Observa-se, na Tabela 3, que o melhor modelo de cada categoria, baseado no valor do AIC, para cada uma das séries foi: $(0,1,1)(1,1,1)_{12}$, $(0,1,1)(1,0,1)_{12}$, $(0,1,1)(0,1,1)_{12}$, para leves (AIC igual a 1.367,974), graves (AIC igual a 1.272,875) e fatais (AIC igual a 942,194), respectivamente. Para verificar se os resíduos apresentam autocorrelação foi aplicado o teste de Box-Ljung, cujos p -valores encontrados para acidentes leves, graves e fatais foram 0,97, 0,93 e 0,99, respectivamente. Assim, não se rejeita a hipótese nula de que os erros não possuem autocorrelação, característica necessária para a aplicação e inferência do modelo SARIMA. Para a verificação da normalidade, foi construído um gráfico de probabilidade normal (Figura 4), e aplicado o teste Jarque-Bera, ao nível de significância 5%.

Baseado na Figura 4, constatou-se que todas as categorias de acidentes respeitam as pressuposições básicas de normalidade. Ademais, pode-se considerar que os resíduos das três categorias têm comportamento de um ruído branco e, logo, os modelos finais ajustados são adequados para descrever o padrão das séries em estudo.

Na Tabela 4 são apresentadas as medidas dos erros de previsão. Pode-se observar que os modelos ajustados para as três categorias têm bons valores de acurácia, e suas previsões são mais adequadas quando comparadas a um modelo *random walk*.

Figura 4: Gráfico dos quantis teóricos e teste Jarque-Bera para acidentes leves (A), graves (B) e fatais (C).



Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Tabela 4: Medidas dos erros de previsão.

	RMSE	MAE	MAPE	MASE	U-Theil
Leves	36,046	29,139	6,546	0,875	0,51
Graves	36,918	29,871	6,669	0,896	0,49
Fatais	3,625	3,431	14,404	0,497	0,51

Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Os parâmetros estimados para os modelos selecionados para cada uma das categorias são apresentados na Tabela 5. Observa-se que os erros-padrão associados aos parâmetros de todos os modelos são relativamente pequenos, o que sugere boa precisão das estimativas.

Tabela 5: Estimativa dos parâmetros dos modelos SARIMA.

		θ_1	ϕ_1	Θ_1
Leves	Estimativa	-0,707	0,134	-1,000
	Erro Padrão	0,057	0,091	0,143
Graves	Estimativa	-0,774	0,998	-0,962
	Erro Padrão	0,068	0,012	0,118
Fatais	Estimativa	-0,876	-	-1,000
	Erro Padrão	0,049	-	0,119

Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Nas tabelas 6, 7 e 8 são apresentados os resultados das previsões e limites, com 95% de confiança para cada categoria. Com exceção dos dois meses após o início da quarentena, decorrente da pandemia por COVID-19, e o valor previsto para novembro de 2019 na categoria de acidentes leves, os valores observados estão todos dentro do intervalo de confiança. É possível notar uma grande diferença entre o valor previsto e o valor observado nos meses de quarentena para as categorias acidentes leves e graves, porém o mesmo não acontece para a categoria de acidentes fatais, no qual as previsões continuam muito próximas do valor observado.

Tabela 6: Previsão de acidentes leves para os próximos 6 meses, com intervalo de confiança de 95%.

Período	Previsão	Limite Inferior	Limite Superior	Valor observado	Erro (%)
Nov/19	422,45	367,02	477,89	484	-12,7
Dec/19	493,45	435,68	551,22	491	0,5
Jan/20	435,37	375,41	495,32	412	5,7
Fev/20	427,68	365,57	489,79	399	7,2
Mar/20	451,05	386,85	515,24	363	24,2
Abr/20	459,59	393,38	525,79	273	68,3

Fonte: Elaborado pelos autores autor, a partir dos dados da pesquisa.

Tabela 7: Previsão de acidentes graves para os próximos 6 meses, com intervalo de confiança de 95%.

Período	Previsão	Limite Inferior	Limite Superior	Valor observado	Erro (%)
Nov/19	142,67	114,22	171,12	138	3,4
Dec/19	162,90	133,74	192,07	147	10,8
Jan/20	136,35	106,51	166,20	128	6,5
Fev/20	135,09	104,56	165,62	138	-2,1
Mar/20	152,97	121,77	184,17	82	86,5
Abr/20	148,17	116,31	180,02	92	61,1

Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Tabela 8: Previsão de acidentes fatais para os próximos 6 meses, com intervalo de confiança de 95%. (Continua).

Período	Previsão	Limite Inferior	Limite Superior	Valor observado	Erro (%)
Nov/19	27,60	15,43	39,78	25	10,4
Dec/19	32,27	20,01	44,54	30	7,6
Jan/20	25,29	12,96	37,63	20	26,45
Fev/20	23,44	11,02	35,87	27	-13,18

Tabela 8: Previsão de acidentes fatais para os próximos 6 meses, com intervalo de confiança de 95%. (Continuação).

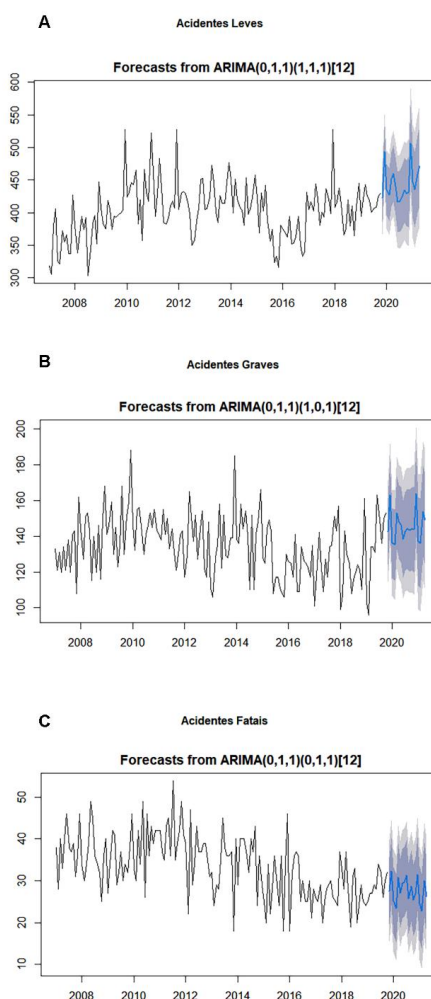
Período	Previsão	Limite Inferior	Limite Superior	Valor observado	Erro (%)
---------	----------	-----------------	-----------------	-----------------	----------

		Inferior	Superior	observado	
Mar/20	30,75	18,24	43,27	31	-0.8
Abr/20	26,98	14,38	39,59	25	7,9

Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

A Figura 6 apresenta os gráficos das previsões para cada categoria de acidente.

Figura 6: Gráfico das previsões de acidentes leves (A), graves (B) e fatais (C).



Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Na Tabela 9 estão os valores das previsões realizadas para cada categoria, no período de maio de 2020 até dezembro de 2020, baseadas nos modelos

ajustados para cada categoria, sem os dados observados. Ressalta-se que pode haver discrepâncias entre os valores previstos e observados, tendo em vista a execução da quarentena e de medidas de isolamento social devido à pandemia.

Tabela 9: Previsões dos acidentes de maio/20 até dezembro/20.

Período	Leves	Graves	Fatais
Mai/20	440,82	146,32	29,29
Jun/20	416,89	138,40	29,75
Jul/20	422,72	142,62	31,29
Ago/20	434,45	144,11	25,75
Set/20	429,66	143,18	28,75
Out/20	431,09	144,19	25,45
Nov/20	506,77	143,60	26,88
Dez/20	447,74	163,80	31,54

Fonte: Elaborado pelos autores, a partir dos dados da pesquisa.

Conclusões

O presente trabalho teve como objetivo desenvolver um estudo utilizando técnicas estatísticas de séries temporais, a fim de identificar e entender os possíveis movimentos dos acidentes de trânsito nas Rodovias Federais do Estado de Santa Catarina. Os modelos de séries temporais, baseados na metodologia Box e Jenkins, podem ser uma boa alternativa na construção de modelos de previsão de acidentes, e auxiliar no desenvolvimento de estratégias para ajudar a prevenir e reduzir o número de acidentes de trânsito.

Os acidentes foram divididos em três categorias: leves; graves e fatais. Dada a periodicidade mensal das observações, foram utilizados modelos SARIMA para o ajuste de cada uma das séries, selecionados por meio do valor do critério de qualidade de ajuste AIC. Os modelos selecionados foram: SARIMA(0,1,1)(0,1,1)_[12]; SARIMA(0,1,1)(1,0,1)_[12]; e SARIMA(0,1,1)(1,1,1)_[12], para as categorias fatal, grave e leve, respectivamente. Foram realizadas as previsões de novembro de 2019 a abril de 2020. Entre os achados é possível

notar uma grande diferença entre o valor previsto e o valor observado nos meses de quarentena, para as categorias de acidentes leves e graves, porém o mesmo não acontece para a categoria de acidentes fatais, no qual as previsões continuam muito próximo do valor observado.

Seria interessante realizar um estudo para identificar a causa desta diferença. Com a quarentena, houve uma diminuição no tráfego de automóveis, o que possivelmente resultou em um menor número de acidentes. Contudo, o número de fatalidades não sofreu a mesma redução no período, o que leva à indagação se é eficiente a tentativa de reduzir o número de fatalidades via redução de acidentes. Ainda para trabalhos futuros sugere-se a construção de modelos que incorporam variáveis explicativas, como por exemplo o modelo autorregressivo integrado de médias móveis sazonal com entradas exógenas (SARIMAX).

Referências

AAKHTAR, Saeed; ROZI, Shafquat. An autoregressive integrated moving average model for short-term prediction of hepatitis C virus seropositivity among male volunteer blood donors in Karachi, Pakistan. *World Journal of Gastroenterology*, n. 15: 1607-1612, 2009.

ABDEL-ATY, M. Analysis of driver injury severity levels at multiple locations using ordered probit models. *J. Saf. Res.*, n. 34: 597–603, 2003.

AGUERO-VALVERDE, J.; JOVANIS, P.P. Spatial analysis of fatal and injury crashes in Pennsylvania. *Accid. Anal. Prev.*, n. 38: p. 618–625, 2006.

AKAIKE, H. Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, v. 60, n. 2: p. 255-265, 1973. Disponível em: http://www.jstor.org/stable/2334537?seq=1#page_scan_tab_contents. Acesso em: jan. 2016.

ALZHRANI, Saleh I.; ALJAMAAN, Ibrahim A.; AL-FAKIH, Ebrahim A. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *Journal of Infection and Public Health*, v. 13: 914–919, 2020.

BALOGUN, O.S. et al. On model selection of road accident data in Nigeria: a time series approach. *Am. J. Res. Commun.*, n. 3: p. 139–177, 2015.

BARROSO JUNIOR, Gilvan Teles; BERTHO, Ana Carolina Soares; VEIGA,

Alinne de Carvalho. A Letalidade dos acidentes de trânsito nas rodovias federais brasileiras em 2016. *R. Bras. Est. Pop.*, v. 36; p. 1-22, e0074, 2019. Disponível em: <http://dx.doi.org/10.20947/S0102-3098a0074>. Acesso em: 21 mar. 2020.

BENVENUTO, Domenico et al. Application of the *ARIMA* model on the COVID-2019 epidemic dataset. *Data in Brief*, v. 29: 1-5, 2020.

BOX, G.E.; JENKINS, G.M. *Time series analysis, control, and forecasting*. San Francisco: Holden Day, 1976.

BOX, G.E.P.; JENKINS, G.M.; REINSEL, G. C. *Time Series Analysis: Forecasting and Control*. 3. ed. New Jersey: Prentice Hall, 1994.

DICKEY, D.A.; FULLER, W.A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, v. 49, n. 4: p. 1057-1072, 1981. Disponível em: http://www.jstor.org/stable/1912517?seq=1#page_scan_tab_contents. Acesso em: jan. 2019.

EARNEST, Arul et al. Using autoregressive integrated moving average (*ARIMA*) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore. *BMC Health Services Research*, v. 5, n. 1: 36, 2005.

EBELING, Eimair Bottega et al. Avaliação das Políticas Públicas de Transportes – Segurança nas Rodovias Federais. [s.l.]: Ministério dos Transportes, Portos e Aviação Civil, 2018. Disponível em: <https://www.gov.br/infraestrutura/pt-br/assuntos/noticias/artigos/avaliacao-da-seguranca-viaria-nas-rodovias-federais>. Acesso em: mai. 2020.

EHLERS, R.S. *Análise de Séries Temporais*. 5. ed. 2009. Disponível em: <http://www.icmc.usp.br/~ehlers/stemp/stemp.pdf>. Acesso em: 15 fev. 2020.

EZE, C.M.; ASOGWA, O.C.; OKONKWO, C.R. On the modelling of road traffic crashes: A case of SARIMA Models. *Journal of Advance Research in Mathematics and Statistics*, v. 5, n. 8: 15-35, 2018.

FIGUEIRA, Aurenice da Cruz et al. Identification of rules induced through decision tree algorithm for detection of traffic accidents with victims: A study case from Brazil. *Case Studies On Transport Policy*, v. 5, n. 2: p. 200-207, 2017. Disponível em: <http://dx.doi.org/10.1016/j.cstp.2017.02.004>. Acesso em: jun. 2020.

GUJARATI, D.N.; PORTER, D.C. *Econometria Básica*. 5. ed. Porto Alegre: AMGH, 2011.

HYNDMAN, R.J.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. 2. ed. Melbourne, Australia: OTexts, 2018. Disponível em: <https://www.otexts.org/fpp/>. Acesso em: 15 dez. 2019.

HYNDMAN, R, et al, Forecast: Forecasting functions for time series and linear models, R package version – 13. R Foundation for Statistical Computing: Vienna, Austria, 2020. Disponível em <https://pkg,robjhyndman,com/forecast/>, Acesso em: ago, 2020.

IRFFI, G. et al. Previsão da demanda por energia elétrica para classes de consumo na região Nordeste, usando OLS dinâmico e mudança de regime. *Economia Aplicada*, v. 13, n. 1: p. 69-98, 2009. Disponível em: <http://www.scielo.br/pdf/ecoa/v13n1/v13n1a04.pdf>. Acesso em: jan. 2020.

JARQUE, C.M.; BERA, A.K.A. Test for Normality of Observations and Regression Residuals. *International Statistical Review*, v. 55, n. 2: 163-172, 1987.

LJUNG, G.M.; BOX, G.E.P. On a measure of lack of fit in Time Series Models. *Biometrika*, n. 65: p. 297-303, 1978.

MOREIRA, D.A. *Administração da Produção e Operações*. 2. ed. São Paulo: Cengage Learning, 2008.

MORETTIN, P.A.; TOLOI, C.M.C. *Análise de séries temporais*. 2. ed. São Paulo: Edgard Blücher, 2006.

NUNES, G. *Estudo da demanda energética no Rio Grande do Sul*. 2019 Dissertação de Mestrado - Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal de Rio Grande (FURG), Rio Grande, RS, 2019.

ONUBR. Organização das Nações Unidas no Brasil. *Década de ação pela segurança no trânsito (2011-2020)*. ONUBR, Brasil, 2011. Disponível em: <https://nacoesunidas.org/campanha/seguranca-transito/>. Acesso em: 10 dez. 2019.

OPAS. Organização Pan-Americana da Saúde. *Folha informativa - Acidentes de trânsito*. OPAS, [s.l.], fev. 2019. Disponível em: https://www.paho.org/bra/index.php?option=com_content&view=article&id=5147:acidentes-de-transito-folha-informativa&Itemid=779. Acesso em: 01 dez. 2019.

PARVAREH, M., et al. Assessment and prediction of road accident injuries trend using time-series models in Kurdistan. *Burns & Trauma*, v. 6, n. 9: s41038–018–0111–6, 2018. Disponível em: <https://doi.org/10.1186/s41038-018-0111-6>.

PRF. Polícia Rodoviária Federal. *Estatísticas: Relatório 2007-2018*. PRF, [S.l.], 2019a. Disponível em: <https://www.prf.gov.br/portal/dados-abertos/acidentes>. Acesso em: 03 nov. 2019.

PRF. Polícia Rodoviária Federal. *Dados Abertos*. PRF, [s.l.]: 2019b. Disponível em: <https://www.prf.gov.br/portal/dados-abertos/>. Acesso em: 01 ago. 2019.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. Disponível em: <https://www.R-project.org>. Acesso em: 16 ago. 2020.

RODRIGUES, Victor Levi Rocha; SANTOS, Josevan Conceição; NERY, Felipe Souza. Mortes de motociclistas em Sergipe: tendência temporal de 2000 a 2016. In: CONGRESSO NACIONAL DE ENFERMAGEM (CONENF), 2018, Aracaju, Sergipe. **Anais...** Aracaju, Sergipe: CONENF, 2018. p. 1-5.

SANUSI, R. A.; ADEBOLA, F. B.; ADEGOKE, N. A. Cases of road traffic accident in Nigeria: a time series approach. *Mediterr. J. Soc. Sci.*, n. 7: p. 542–552, 2016.

SILVEIRA, A.G. *Estudo da demanda de energia elétrica no Brasil*. 2017. Dissertação de Mestrado - Universidade Federal do Rio Grande, Rio Grande, RS, 2017.

SOUZA, G.P.; SAMOBYL, R.W.; MIRANDA, R.G. *Métodos simplificados de previsão empresarial*. Rio de Janeiro: Ciência Moderna Ltda, 2008.

VITORINO, Rafaella Santos. *Modelo dinâmico de Poisson aplicado a dados de acidentes automobilísticos nas Rodovias Federais do Estado da Paraíba*. 2019. Monografia (Trabalho de Conclusão de Curso em Estatística) - Bacharelado em Estatística, Departamento de Estatística, Centro de Ciências e Tecnologia, Universidade Estadual da Paraíba, 2019.

WHO. World Health Organization. *Global status report on road safety 2018*. WHO, Geneva., 2018. Disponível em: http://www.who.int/violence_injury_prevention/road_safety_status/2015/Summary_GSRRS2015_POR.pdf?ua=1. Acesso em: 06 nov. 2019.