



## CIÊNCIAS EXATAS E DA TERRA

**Previsão da concentração de material particulado inalável, através de modelos estatísticos de séries temporais para o município de Canoas, Rio Grande do Sul**

*Title Prediction of the concentration of inhalable particulate material, through statistical models of time series for the municipality of Canoas, Rio Grande do Sul*

Cleiton Tibulo<sup>1</sup>, Simone Erotildes Teleginski Ferraz<sup>2</sup>,  
Vaneza De Carli Tibulo<sup>3</sup>, Roselaine Ruviano Zanini<sup>4</sup>,  
Nathalie Tissot Boiaski<sup>5</sup>

**RESUMO**

O presente trabalho tem por objetivo modelar e prever a concentração de Material Particulado Inalável (PM<sub>10</sub>), a fim de analisar a capacidade preditiva dos modelos de séries temporais: ARMAX, ARIMA, Alisamento Exponencial Simples, Biparamétrico e Holt-Winters. Os dados que compõem as séries analisadas pertencem a Fundação Estadual de Proteção Ambiental (FEPAM), referente ao município de Canoas-RS, no ano de 2014. Os modelos foram ajustados com auxílio do *software* RStudio. Como resultados, constatou-se que o modelo ARMAX, com a inclusão de variáveis exógenas (Monóxido de carbono (CO) e Dióxido de enxofre (SO<sub>2</sub>)), produziu melhores resultados de previsão em relação aos seus concorrentes para a variável PM<sub>10</sub>. Também, verificou-se que modelos de séries temporais da Classe ARMA e ARMAX podem ser utilizados para previsão da PM<sub>10</sub>, entretanto os modelos da classe de Alisamento Exponencial não são recomendados, pois não se ajustaram adequadamente as séries em estudo.

**Palavras-chave:** Poluição; previsão; séries temporais.

**ABSTRACT**

*This work aims to model and predict the concentration of Inhalable Particulate Material (PM10), in order to analyze the predictive capacity of time series models: ARMAX, ARIMA, Simple Exponential Smoothing, Biparametric and Holt-Winters. The data that make up the analyzed series belong to the State Environmental Protection Foundation (FEPAM), referring to the municipality of Canoas-RS, in 2014. The models were adjusted with the aid of RStudio software. As a result, it was found that the ARMAX model, with the inclusion of exogenous variables (carbon monoxide (CO) and sulfur dioxide (SO2)), produced better forecasting results in relation*

<sup>1</sup> Universidade Federal de Santa Maria – UFSM, Santa Maria/RS – Brasil. E-mail: [tibulo\\_cleiton@hotmail.com](mailto:tibulo_cleiton@hotmail.com)

<sup>2</sup> E-mail: [simonetfe@gmail.com](mailto:simonetfe@gmail.com)

<sup>3</sup> E-mail: [vaneza\\_dc@yahoo.com.br](mailto:vaneza_dc@yahoo.com.br)

<sup>4</sup> E-mail: [rrzanini63@gmail.com](mailto:rrzanini63@gmail.com)

<sup>5</sup> E-mail: [ntboiaski@gmail.com](mailto:ntboiaski@gmail.com)



to its competitors for the PM<sub>10</sub> variable. Also, it was found that time series models of the ARMA and ARMAX Class can be used to predict the PM<sub>10</sub>, however the models of the Exponential Smoothing class are not recommended, as the series under study have not been properly adjusted.

**Keywords:** Pollution; forecast; time series.

## 1. INTRODUÇÃO

Nas últimas décadas a poluição do ar tem recebido destacada atenção, devido aos elevados índices de contaminação do ar e as consequências na saúde dos seres vivos e alterações no clima global, segundo Silva e Vieira (2017). Para Santos (2016), a qualidade do ar é motivo de preocupação, principalmente, em grandes centros urbanos onde os altos níveis de poluição causam diversos impactos negativos à saúde da população e ao meio ambiente.

Lopes *et al.* (2017) e Dapper *et al.* (2016), a poluição do ar, impulsionada pelo aumento demográfico e a concentração de população em áreas urbanas, vem constantemente se agravando, sendo de fundamental importância compreender os seus impactos para saúde humana. Mendes (2019) descreve que no Brasil, as mortes em decorrência da poluição atmosférica aumentaram 14% em dez anos e segundo a Organização Mundial da Saúde (OMS), calcula-se que ocorram anualmente 4,2 milhões de mortes prematuras atribuídas à poluição do ar no mundo. Para Freitas *et al.* (2013), a poluição atmosférica está diretamente associada ao número de mortes e baixas hospitalares ligadas a problema respiratórios e cardiovasculares.

Lovatto (2011) ressalta que o homem desencadeou uma série de processos que resultaram, dentre outros, na poluição dos ecossistemas, na concentração elevada de gases tóxicos na atmosfera e nas alterações climáticas. Dapper (2016), ressalta que a poluição atmosférica pode ser decorrente de diversos meios de produção, como também de fenômenos naturais. Independentemente do fator gerador, a poluição exerce influência no meio ambiente e na qualidade de vida dos seres humanos.

Para Santos (2016), Leite *et al.* (2011) e Barcellos *et al.* (2009), em condições climáticas favoráveis de altas temperaturas e baixa umidade o tempo de resistência de poluentes na atmosfera aumenta e podem ser transportados a longas distâncias, afetando a saúde de populações distantes das fontes geradoras. Desta forma, Pinto *et al.* (2018) destacam que a previsão da qualidade do ar pode ser utilizada como ferramenta de alerta sobre a concentração de poluentes na atmosfera e permitir a tomada de decisão, em tempo hábil, quanto à adaptação de comportamento da população de grupos de risco, como crianças, idosos e pessoas com doenças respiratórias.

Nesse contexto, o estudo tem por objetivo modelar e prever a concentração de PM<sub>10</sub> no município de Canoas-RS, comparando o desempenho dos modelos de séries temporais das seguintes classes: Alisamento Exponencial, Autorregressivos Integrados e de Médias Móveis (ARIMA) e modelos Autorregressivos Integrados de Médias Móveis que permitam a inclusão de variáveis exógenas (ARIMAX), a fim de analisar a capacidade preditiva dos modelos propostos.



## 2. REFERENCIAL TEÓRICO

Este tópico apresenta as fundamentações teóricas que sustentam o presente estudo. Para tanto inicialmente são apresentadas as discussões sobre as consequências dos poluentes do ar para a saúde, previsões de poluentes do ar, modelos ARIMA, ARIMAX, Alisamento exponencial e critérios de seleção/comparação de modelos.

### 2.1. POLUENTES DO AR E AS CONSEQUÊNCIAS PARA SAÚDE

A contínua emissão de poluentes para atmosfera prejudica gravemente a saúde humana. Os efeitos da poluição dependem do indivíduo (criança, adulto, idoso), tempo de exposição, dos níveis de poluição e das vias de contato (inalação, ingestão, contato cutâneo) (Lopes *et al.*, 2017). De acordo com Machín (2017) a exposição aos poluentes do ar causa efeitos negativos a saúde e tem sido uma das principais causas de internação por doenças do aparelho respiratório. Almeida *et al.* (2015) citam alguns impactos que os principais poluentes do ar podem causar como: diminuição nas funções pulmonares, doenças respiratórias e cardiovasculares.

Para melhor visualizar as características, principais fontes dos poluentes, seus efeitos para a saúde e meio ambiente, os dados coletados da Fundação Estadual de Proteção Ambiental (FEPAM) foram sintetizados no Quadro 1. Observa-se, no Quadro 1, que, dentre as principais fontes geradoras de poluição do ar, estão as emissões veiculares e as atividades industriais. Os efeitos à saúde humana estão relacionados diretamente ao sistema nervoso e respiratório, indiretamente pela contaminação do solo e dos alimentos.

**Quadro 1** - Fontes e características de alguns poluentes do ar.

Poluente	Características	Principais Fontes	Efeitos sobre o meio ambiente	Efeitos sobre a Saúde
Partículas inaláveis (PM <sub>10</sub> )	Partículas de material sólido ou líquido que ficam suspensas no ar, na forma de poeira, neblina, aerossol, fumaça, fuligem, etc. Tamanho < 10 micra	Processos de combustão (indústrias e veículos automotores), aerossol secundário (formado na atmosfera), Pólen, aerossol marinho e solo.	Danos à vegetação, redução da visibilidade e contaminação do solo.	Aumento de atendimentos hospitalares e mortes prematuras. Insuficiências respiratórias pela deposição deste poluente nos pulmões.
Dióxido de Enxofre (SO <sub>2</sub> )	Gás incolor, com forte odor, altamente solúvel. Na presença de vapor d'água pode ser transformado a SO <sub>3</sub> passando rapidamente a H <sub>2</sub> SO <sub>4</sub> , sendo um dos principais constituintes da chuva ácida. É um importante precursor dos	Combustão de combustíveis fósseis (carvão), queima de óleo combustível, refinaria de petróleo, veículos a diesel, vulcões, emissões de reações	Pode levar à formação de chuva ácida, causar corrosão aos materiais e danos à vegetação.	Desconforto na respiração, doenças respiratórias, agravamento de doenças respiratórias e cardiovasculares já existentes. Pessoas com asma, doenças crônicas de coração e pulmão são mais sensíveis ao SO <sub>2</sub> .



	sulfatos, um dos principais componentes das partículas inaláveis. No verão, através dos processos fotoquímicos, as reações do SO <sub>2</sub> são mais rápidas.	biológicas.		Irritação ocular.
Monóxido de Carbono (CO)	Gás incolor, inodoro e insípido.	Combustão incompleta em geral, principalmente em veículos automotores, queimadas e reações fotoquímicas.	-	Causa efeito danoso no sistema nervoso central, com perda de consciência e visão. Exposições mais curtas podem também provocar dores de cabeça e tonturas.

Fonte: FEPAM. Disponível em: <http://www.fepam.rs.gov.br/qualidade/poluente.asp>.  
Acesso em: 11 nov. 2019

## 2.2. PREVISÕES DE POLUENTES DO AR

Pinto *et al.* (2018) desenvolveram um estudo para prever a concentração média diária de material particulado inalável (MP<sub>10</sub>), na Região da Grande Vitória-ES, Brasil, utilizando o modelo SARIMAX para o período de 01/01/2012 a 30/04/2015. Os dados deste estudo foram do tipo séries temporais de concentrações de MP<sub>10</sub> e de variáveis meteorológicas (velocidade do vento, umidade relativa, precipitação pluvial e temperatura), obtidas junto ao Instituto Estadual de Meio Ambiente e Recursos Hídricos (IEMA). Na comparação dos modelos ARMA, com os modelos SARIMAX, os autores relatam um desempenho estatístico superior do modelo SARIMAX, no que diz respeito à previsão da qualidade do ar.

Souza *et al.* (2017), utilizando as técnicas de análise de componentes principais e regressão linear múltipla, desenvolveram um modelo para prever as concentrações de ozônio em função das variáveis climáticas. As medidas de concentração do ozônio foram realizadas na Universidade Federal de Mato Grosso do Sul (UFMS), utilizando dados do período de janeiro a dezembro de 2014. As variáveis preditoras referentes ao clima (umidade relativa, velocidade do vento, chuva, aerossóis, índice de claridade, radiação solar global e temperatura do ar) foram cedidas pelo Centro de Monitoramento e do Clima do Mato Grosso do Sul (CEMTEC-MS). Os autores concluíram que o modelo desenvolvido é capaz de explicar 99% da variabilidade nas concentrações de ozônio de superfície com um erro de 0,2%.

Monte, Albuquerque e Reisen (2015) realizaram um estudo com objetivo de estimar e prever a concentração horária de ozônio na Região da Grande Vitória - ES, utilizando um modelo de previsão de Série Temporal ARMAX-GARCH, para o período 01/01/2011 a 31/12/2011. Foram utilizados dados da rede de monitoramento do Instituto Estadual de Meio Ambiente e Recursos Hídricos (IEMA). Os autores constataram que os resultados de previsão foram muito próximas dos valores observados. Na comparação dos modelos ARMA e ARMAX, o modelo ARMAX-GARCH revelou-se mais eficaz na predição de episódios de poluição de ozônio.



Barbosa et al. (2014) modelaram as concentrações de material particulado inalável PM<sub>10</sub> para realizar previsões no município de Serra, Região da Grande Vitória - ES, por meio dos modelos de séries temporais SARIMA. Os dados de poluente correspondem ao período de janeiro de 2010 a novembro de 2013. Os resultados mostraram que o modelo SARIMA prevê as concentrações de forma muito eficaz, sendo uma importante ferramenta para previsão da qualidade do ar para a região.

Gripa *et al.* (2012) discutem técnicas de previsão da qualidade do ar para concentrações de PM<sub>10</sub>, através de modelos de regressão linear múltipla (envolvendo fatores meteorológicos) e modelos da classe SARIMA para cidade de Carapina - ES. Os resultados evidenciaram ajustes semelhantes para ambos os modelos, no entanto, o modelo de regressão linear múltipla obteve melhor desempenho na análise de previsão das concentrações médias de PM<sub>10</sub>.

Observa-se, em uma revisão bibliográfica nas principais plataformas de pesquisas online, que há poucos estudos publicados de aplicação de modelos de previsão de séries temporais, com a proposta metodológica dos modelos da classe de Alisamento Exponencial, SARIMA (autorregressivos integrados e de médias móveis) e SARIMAX (autorregressivos integrados e de médias móveis com entrada de variáveis exógenas) a séries históricas do poluente do ar PM<sub>10</sub>.

### 2.3. MODELO AUTORREGRESSIVO, INTEGRADO E DE MÉDIAS MÓVEIS (ARIMA)

Esses modelos, genericamente conhecidos por ARIMA ou Autorregressivos Integrados de Médias Móveis são modelos estatísticos que visam captar o comportamento da correlação seriada ou autocorrelação entre os valores da série temporal e, com base nesse comportamento, realizar previsões futuras, segundo Morettin e Toloí (2004) e Danfá (2009). De acordo com Morettin e Toloí (2004), essa classe de modelos vem sendo largamente utilizada. Sendo assim, o modelo ARIMA (p,d,q), é representado pela equação:

$$\Delta^d \tilde{Z}_t = \underbrace{\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p}}_{\text{autorregressivo (p)}} + \varepsilon_t - \underbrace{\theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}}_{\text{médias móveis (q)}} \quad (1)$$

em que os parâmetros  $\phi_1, \dots, \phi_p$  são os parâmetros referentes à parte autorregressiva,  $\theta_1, \dots, \theta_q$  são os parâmetros de médias móveis,  $\varepsilon_t$  é o erro que não pode ser estimado pelo modelo, e d representa o número de diferenças aplicadas na série.

### 2.4. MODELO AUTORREGRESSIVO, INTEGRADO E DE MÉDIAS MÓVEIS COM A INCLUSÃO DE VARIÁVEIS EXÓGENAS (ARIMAX)

O modelo possibilita a análise de mais de uma variável correlacionada a previsão. Vários modelos podem derivar a partir do modelo ARIMAX (p,d,q,r[]), como: ARIX (p,d,0,r[]), IMAX (0,d,p,r[]) ARMAX, ARX (p,0,0,r[]), MAX (0,0,q,r[]), SARIMAX (p,d,q,r[]) (P,D,Q) com a diferença que, além da estimação das ordens dos coeficientes dos modelos, é preciso estimar os coeficientes das variáveis exógenas. Essa estrutura pode considerar a inclusão de mais variáveis explicativas (exógenas), de acordo com Andrade (2009), Silva (2011), Chechi e Sanches (2013) e Lizzi (2012).



Segundo Martínez e Zamprogno (2003), Silva (2011) e, Ozaki e Admi (2012), o modelo ARMAX (autorregressivo com médias móveis e entrada de variáveis exógenas) baseia-se em um modelo ARMA, em que é possível combiná-lo com uma estrutura de regressão (variáveis explicativas) na forma:

$$\begin{aligned} \tilde{Z}_t = & \underbrace{\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p}}_{\text{autorregressivo (p)}} + \underbrace{\beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_r Y_{t-r}}_{\text{variável independente}} + \\ & \underbrace{\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}}_{\text{médias móveis (q)}} \end{aligned} \quad (2)$$

onde  $Z_t$  é a variável dependente,  $Y_t$  é a variável independente,  $\varepsilon_t$  é ruído branco,  $\phi_i$  e  $\theta_i$  são os parâmetros desconhecidos,  $\beta_i$  é o coeficiente da variável exógena, e  $\varepsilon_t$  é a sequência de resíduos do modelo.

## 2.5. ALISAMENTO EXPONENCIAL

Uma das grandes classes de modelos de previsão que estuda o padrão de qualquer ruído contido em séries de tempo e que, através dele, prevê valores futuros é o de alisamento exponencial. A grande popularidade advém de sua simplicidade e eficiência, segundo Morettin e Toloí (1987). Neste trabalho serão utilizados os modelos Alisamento Exponencial Simples (indicado para séries sem tendência e/ou sazonalidade), Alisamento Exponencial Biparamétrico (indicado para séries com tendências lineares positivas ou negativas permitindo usar constantes com diferentes valores para o nível e tendência) e Alisamento Exponencial de Holt-Hinters (indicado para séries que possuem tendência e sazonalidade, derivando em dos modelos: aditivo para séries homogêneas e multiplicativo para séries com comportamento heterogêneas), descritos em Morettin e Toloí (2004) e Morettin e Toloí (1987).

## 2.6. CRITÉRIOS DE SELEÇÃO E COMPARAÇÃO DE MODELOS

Os critérios a serem utilizados para selecionar os melhores modelos tanto da classe ARIMA quanto do ARIMAX, são: AIC proposto por Akaike (1974) e BIC desenvolvido por Akaike (1978). Além desses critérios, a condição de ruído branco dos resíduos foi considerada, ou seja, faz-se necessário que os erros do modelo devem ter média zero e variância constante, além de serem descorrelacionados ao longo do tempo.

Para testar a eficácia dos modelos e comparar a qualidade das previsões entre as diferentes classes de modelos, foram utilizados os seguintes critérios: Erro Quadrático Médio (EQM), Soma dos Quadrados dos Erros (SSE), Média Absoluta dos Erros (MAD) e o Erro Absoluto Médio Percentual (MAPE), descritos em Tidre, Biase e Silva (2013), Esquivel, Senna e Gomes (2012) e Feliciano (2009). Quanto menores os valores encontrados nos critérios (EQM), (SSE), (MAD) e o (MAPE), mais próximos dos valores observados estão aos valores previstos pelos modelos ajustados.



### 3. MATERIAIS E MÉTODOS

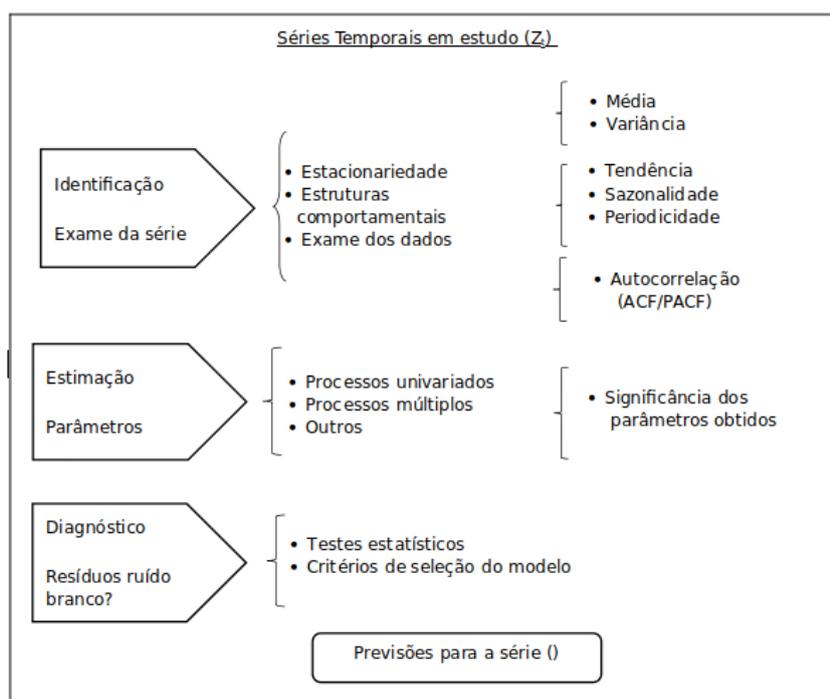
Esta pesquisa, quanto aos objetivos, segundo Gil (2008), é classificada como uma pesquisa exploratória e descritiva, visto que proporciona maior familiaridade aos estudos referentes a modelos estatísticos de previsão de séries temporais. Quanto aos procedimentos técnicos utilizados, é classificada como pesquisa bibliográfica (desenvolvida com base em material já elaborado constituído principalmente de livros e artigos científicos) e estudo de caso (estudo teórico aplicado a uma determinada situação, gerando possíveis hipóteses a estudos futuros) de acordo com Gil (2008).

Para realizar este estudo, foram obtidos dados da rede automática de monitoramento da FEPAM, do município de Canoas-RS. As medições são horárias dos seguintes poluentes do ar: material particulado inalável com diâmetro inferior a  $10 \mu\text{m}$  ( $\text{PM}_{10}$ ), dióxido de enxofre ( $\text{SO}_2$ ) e monóxido de carbono (CO), no período compreendido entre 01/01/2014 a 31/12/2014. As concentrações horárias de poluentes do ar foram transformadas em médias diárias.

Cabe ressaltar que outras séries históricas de poluentes do ar dos municípios de Porto Alegre-RS, Esteio-RS e próprio Canoas-RS, (para diferentes períodos) foram analisadas, entretanto o grande número de dados faltantes nas séries históricas impossibilitou o uso das mesmas.

Para ajuste dos modelos estatísticos de séries temporais, este trabalho segue a metodologia proposta por Box *et al.* (1994), que propõem um processo iterativo para a busca de um modelo estocástico que represente uma série temporal. O método consiste de três fases: identificação, estimação e diagnóstico, resumidas no Quadro 2.

**Quadro 2** - Processo iterativo de análise de séries temporais.



Fonte: (BARBIERO, 2003, p.38).



Para realizar a comparação dos modelos mais apropriado para previsão da  $PM_{10}$ , foram reservadas as 20 últimas observações das séries históricas em estudo, ou seja, dias não contabilizados no ajuste dos modelos e utilizados os critérios: MAPE, EQM, MAD e SSE, descritos no item 2.6. Como software de apoio utilizou-se o RStudio.

## 4. ANÁLISE DOS RESULTADOS

### 4.1. ESTATÍSTICA DESCRITIVA

Inicialmente são apresentadas as medidas descritivas dos poluentes do ar CO,  $PM_{10}$  e  $SO_2$  (Tabela 1). A concentração média de  $PM_{10}$  foi de aproximadamente  $31,43 \mu\text{g}/\text{m}^3$  com desvio padrão  $17,62 \mu\text{g}/\text{m}^3$ . Nota-se que em média as concentrações não ultrapassam o valor de  $50,0 \mu\text{g}/\text{m}^3$ , porém o desvio padrão e o coeficiente de variação (56,06) alto indicam uma média não representativa. Observa-se também que para as variáveis CO e  $SO_2$ , há um alto coeficiente de variação (105 e 130 respectivamente), indicando que os dados em estudo são heterogêneos (média não representativa), ou seja, demonstram um comportamento com variabilidade durante o ano, o que remete a necessidade de estudos dessas variáveis para entender os seus comportamentos.

**Tabela 1** – Análise descritivas dos poluentes do ar.

Variáveis	Média	Desvio-padrão	Coeficiente de variação	Mínimo	Máximo
CO ( $\mu\text{g}/\text{m}^3$ )	0,06	0,063	105	0,03	0,42
$PM_{10}$ ( $\mu\text{g} / \text{m}^3$ )	31,43	17,62	56,06	4,34	133,37
$SO_2$ ( $\mu\text{g} / \text{m}^3$ )	1,33	1,73	130	0,65	21,05

\* CO: Monóxido de carbono;  $PM_{10}$ : Material particulado;  $SO_2$ : Dióxido de enxofre.

Fonte: Elaborada pelos autores.

Tendo em vista a variável  $PM_{10}$  ter o menor coeficiente de variação em relação ao CO e o  $SO_2$ , optou-se por realizar previsões para a variável  $PM_{10}$ . De acordo com Freitas et al. (2013), as populações expostas ao  $PM_{10}$  desenvolvem doenças respiratórias e cardiovasculares com maiores impactos. As demais variáveis consideradas neste trabalho CO e  $SO_2$ , serão utilizadas como variáveis exógenas no modelo ARIMAX.

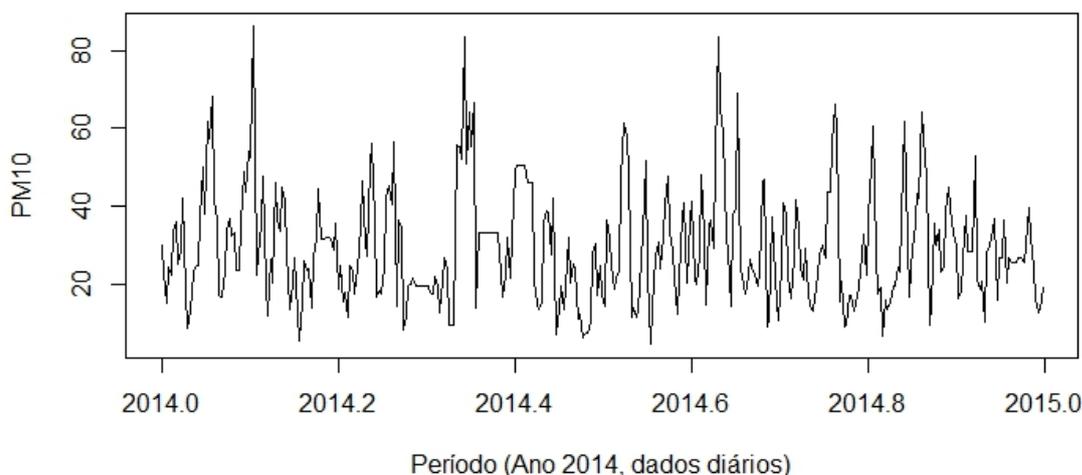
### 4.2. MODELOS DA CLASSE BOX-JENKINS

#### 4.2.1. Ajuste do modelo ARIMA para a variável $PM_{10}$

Para dar início ao ajuste do modelo, observa-se, na Figura 1, o comportamento da série original, buscando identificar suas componentes, como tendência e/ou sazonalidade. Para Box et al. (1994) na primeira fase verifica-se a condição de estacionariedade, estruturas comportamentais da série como tendência, sazonalidade, periodicidades, e ainda são observadas as autocorrelações.



**Figura 1** - Gráfico da série original  $PM_{10}$  em Canoas-RS.

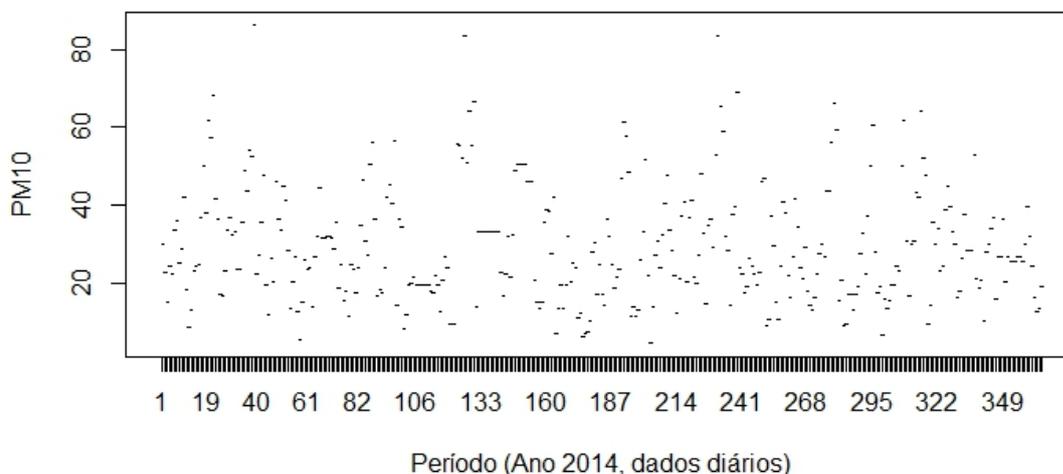


Fonte: Elaborado pelos autores.

Analisando a Figura 1, verifica-se que a série não possui indícios de tendência (crescente ou decrescente), sazonalidade e ciclos, não sendo necessário aplicar diferenciações na série original ou utilizar modelos que levem em conta a componente sazonal. Para confirmar a análise gráfica, foram aplicados testes estatísticos para verificar a condição de estacionariedade e sazonalidade.

Após a etapa inicial de inspeção visual (análise gráfica) dos dados, foram realizados o teste estatístico de sequências ( $p$ -valor  $> 0,05$ ), confirmando que série não possui indícios de tendência significativa ao nível de  $\alpha = 5\%$ , e os testes estatísticos para sazonalidade de Kruskal-Wallis e Friedman que rejeitaram a presença de sazonalidade ao nível de significância  $\alpha = 5\%$ . O resultado para o teste de Dickey-Fuller Aumentado (ADF), segundo Dickey e Fuller (1981) revela que as séries em estudo são estacionárias.

**Figura 2** - Gráfico da série  $PM_{10}$  em pontos.



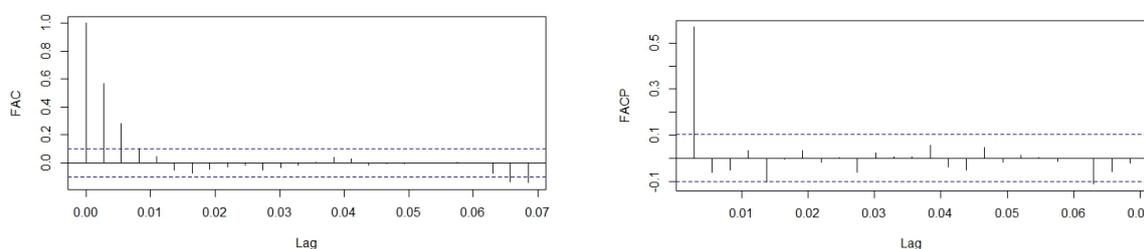
Fonte: Elaborado pelos autores.



No gráfico da Figura 2, não se identificam períodos com menor ou maior concentração de  $PM_{10}$ , sendo as observações heterogêneas durante o ano, o que indica que a média é pouco representativa. Ao término da etapa de análise inicial da série  $PM_{10}$ , não se identificou a presença de sazonalidade e tendência, a próxima etapa da fase de identificação de acordo com Box *et al.* (1994) é a construção dos gráficos de autocorrelação (FAC) e de autocorrelação parcial (FACP).

Os gráficos das funções de autocorrelação (FAC) e de autocorrelação parcial (FACP), expressos na Figura 3, demonstram as correlações seriais no domínio do tempo. Com base em suas frequências estimadas, é possível verificar sazonalidades e ciclos na série de acordo com Barros *et al.* (2018). Na Figura 3, apresentam-se os correlogramas das séries, sem diferença.

**Figura 3** - Função de Autocorrelação (FAC) e Autocorrelação Parcial (FACP).



Fonte: Elaborado pelos autores.

Ao se analisar os FAC e FACP, percebe-se características de um modelo estacionário, autorregressivo de ordem 1 ou 2 e de médias móveis de ordem 1. Os testes estatísticos e a análise gráfica indicam que um modelo do tipo ARMA deve ser considerado. A identificação é preliminar baseada no comportamento do FAC e FACP, para detalhes ver Morettin e Toloí (2004). Esta análise serve para dar início ao processo de modelagem, embora vários modelos de ordem diferentes foram considerados e testados.

Na Tabela 2, estão expostos alguns modelos concorrentes ajustados para a série a  $PM_{10}$ , com os respectivos valores de AIC e BIC, os parâmetros do modelo.

**Tabela 2** - Modelos concorrentes ARIMA e os respectivos critérios (variável  $PM_{10}$ ).

ARIMA (p,d,q)	Parâmetros	Critérios de seleção	
		AIC	BIC
(1,0,0)	$\phi_1 = 0,5861$ C = 29,6131	2850,11	2861,39
(2,0,0)	$\phi_1 = 0,6022$ $\phi_2 = -0,0608$ C = 29,2913	2863,33	2868,93
(1,1,1)	$\phi_1 = 0,5719$ $\theta_1 = -0,9999$	2852,35	2864,39

Fonte: Elaborado pelos autores.

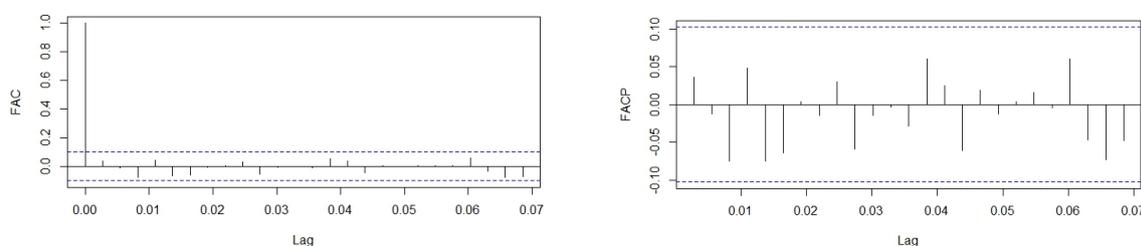


Considerando os três melhores modelos significativos ajustados, analisou-se a condição de ruído branco dos resíduos. Na análise dos resíduos, os modelos atendem aos pressupostos previstos.

Os testes de autocorrelação dos resíduos Ljung-Box (1978) ( $p$ -valor = 0,4884) e Box-Pierce (1970) ( $p$ -valor = 0,4903) do melhor modelo AR (1), de acordo com os critérios de seleção AIC e BIC, indicam que os erros do modelo ajustado são descorrelacionados ao longo do tempo.

Na Figura 4, gráfico de ACF e PACF dos resíduos, percebe-se que não há autocorrelação significativa diferente de zero em nenhuma defasagem residual, portanto os dados são independentes e demonstram um bom ajuste do modelo, corroborando com os resultados dos testes de Ljung-Box e Box-Pierce.

**Figura 4** – Função de autocorrelação (FAC) e autocorrelação parcial (FACP) dos resíduos.

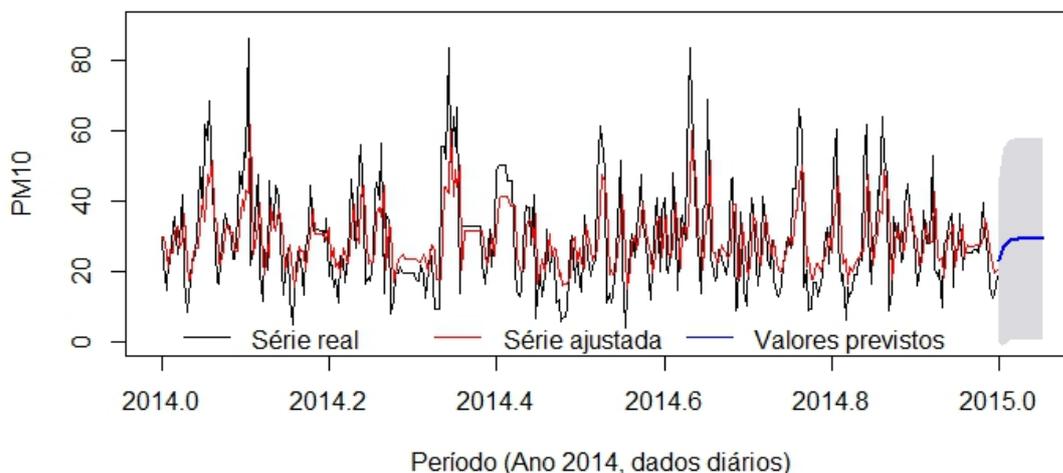


Fonte: Elaborado pelos autores.

Considerando a análise residual e os critérios AIC e BIC, o modelo AR(1), com um parâmetro autorregressivo, atende a todas as condições de ruído branco e normalidade dos resíduos e concluem a última etapa da escolha do modelo para realizar previsões futuras.

Nessa etapa do trabalho, foram obtidas as previsões da série  $PM_{10}$  para um horizonte de 20 passos à frente. Na Figura 5, é demonstrada a série real, ajustada, e os valores previstos para os limites de previsão com intervalo de confiança de 95%.

**Figura 5** – Modelo AR(1) ajustado à série observada (variável  $PM_{10}$ ).



Fonte: Elaborado pelos autores.



#### 4.2.2. Ajuste do modelo ARIMAX

Para o ajuste do modelo ARIMAX, a variável endógena a ser usada é a  $PM_{10}$  e as variáveis exógenas, serão utilizados o CO e o  $SO_2$ , que possuem um alto coeficiente de correlação com a  $PM_{10}$ . O intuito é que CO e o  $SO_2$ , possam auxiliar nas previsões da variável em  $PM_{10}$ .

Dentre os diversos modelos testados, na Tabela 3, estão apresentados os três modelos com melhor desempenho nos critérios de seleção AIC e BIC.

**Tabela 3** – Modelos concorrentes ARIMAX e os respectivos critérios.

ARIMAX (p,d,q,r[I])	Parâmetros	Critérios de seleção	
		AIC	BIC
(1,[1,1])	$\phi_1 = 0,5588$ C = 26,8307 SO <sub>2</sub> = 0,4484 CO = 34,38	2701,38	2720,60
(2,[1,1])	$\phi_1 = 0,5762$ $\phi_2 = -0,0296$ C = 26,9305 SO <sub>2</sub> = 0,4557 CO = 32,7264	2703,10	2726,16
(1,0,1[1,1])	$\phi_1 = 0,5289$ $\theta_1 = 0,0445$ C = 26,9303 SO <sub>2</sub> = 0,4544 CO = 33,002	2703,14	2726,2

Fonte: Elaborado pelos autores.

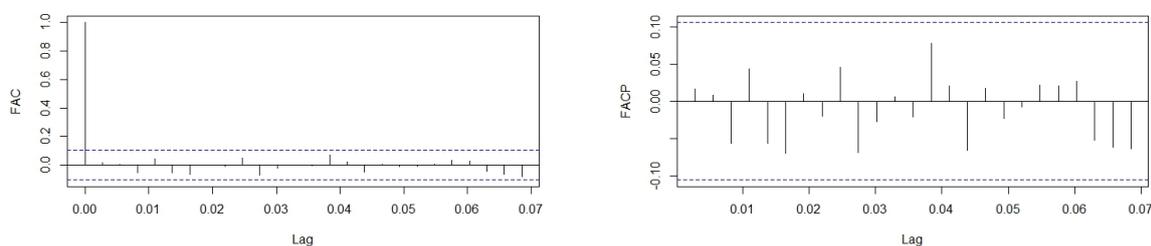
Considerando os três melhores modelos significativos ajustados, analisou-se a condição de ruído branco dos resíduos. Na análise dos resíduos os modelos atenderam a todos os pressupostos.

Os testes de autocorrelação dos resíduos Ljung-Box (1978) (p-valor = 0,7506) e Box-Pierce (1970) (p-valor = 0,7517), do melhor modelo ARX (1,[1,1]), de acordo com os critérios de seleção AIC e BIC, indicam que os erros do modelo ajustado são descorrelacionados ao longo do tempo.

Na análise gráfica (Figura 6), ACF e PACF dos resíduos do modelo ARX, percebe-se que não há autocorrelação significativa diferente de zero em nenhuma defasagem, portanto, os dados são independentes e demonstram um bom ajuste do modelo, confirmando os resultados dos testes de Ljung-Box e Box-Pierce.



**Figura 6** – Função de autocorrelação (FAC) e autocorrelação parcial (FACP) dos resíduos ARX (1,[1,1]).

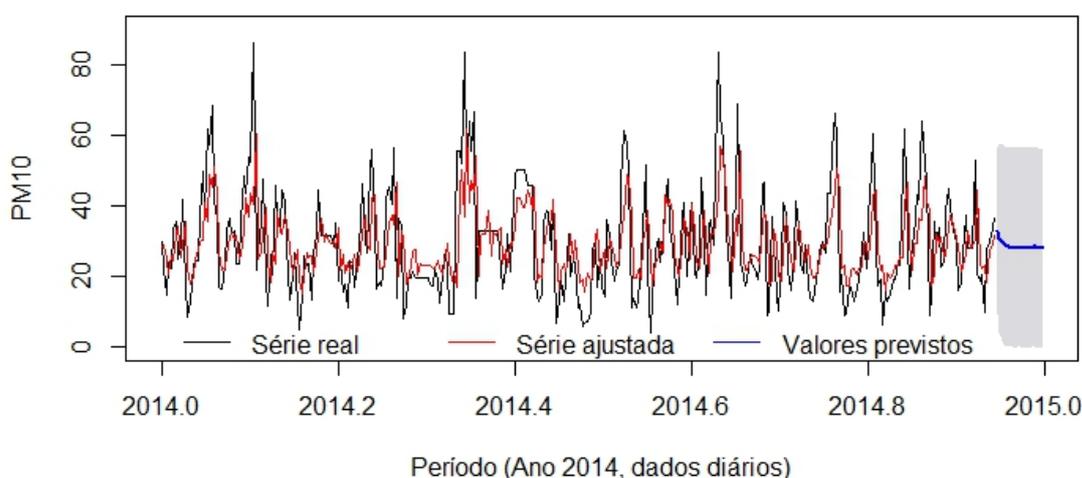


Fonte: Elaborado pelos autores.

Considerando a análise residual e os critérios AIC e BIC, o modelo ARX (1,[1,1]), com um parâmetro autorregressivo ( $p$ ) e um parâmetro das variáveis exógenas Monóxido de Carbono e Dióxido de Enxofre, o modelo atende a todas as condições de ruído branco e normalidade dos resíduos e concluem a última etapa da escolha do modelo para realizar previsões futuras.

Nessa etapa do trabalho, foram obtidas as previsões da série  $PM_{10}$  para um horizonte de 20 passos à frente. A Figura 7 demonstra a série ajustada e os valores previstos para os limites de previsão com intervalo de confiança de 95%.

**Figura 7** – Modelo ARX (1,[1,1]) ajustado à série observada (variável  $PM_{10}$ ).



Fonte: Elaborado pelos autores.

#### 4.3. AJUSTE DE MODELOS DA CLASSE DE ALISAMENTO EXPONENCIAL PARA VARIÁVEL $PM_{10}$

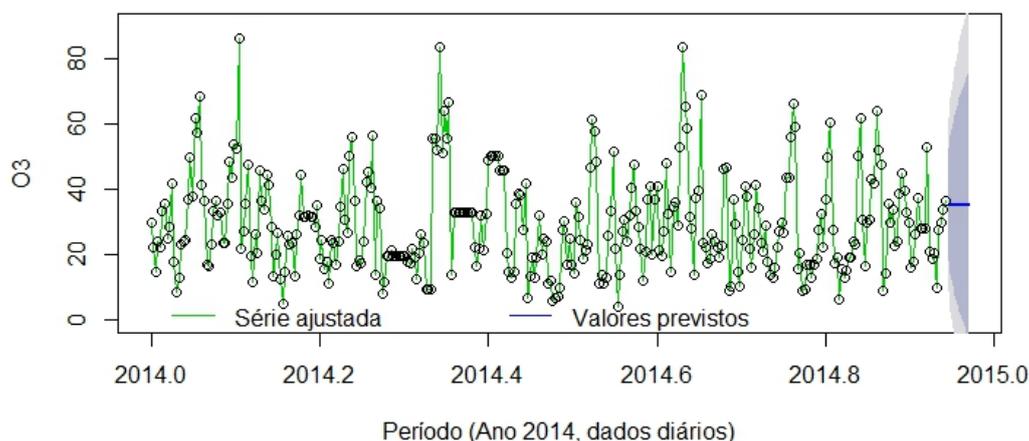
Nesta etapa do trabalho, busca-se ajustar os modelos de Alisamento Exponencial Simples, Alisamento Exponencial Biparamétrico e Alisamento Exponencial de Holt-Winters.



### 4.3.1. Alisamento exponencial simples

No ajuste do modelo de alisamento exponencial simples, encontrou-se o parâmetro de suavização alfa  $\alpha = 0,7176$ . Na Figura 8, é demonstrada a série ajustada e os valores previstos para os limites de previsão com intervalo de confiança de 95%.

**Figura 8** - Modelo de alisamento Exponencial Simples.



Fonte: Elaborado pelos autores.

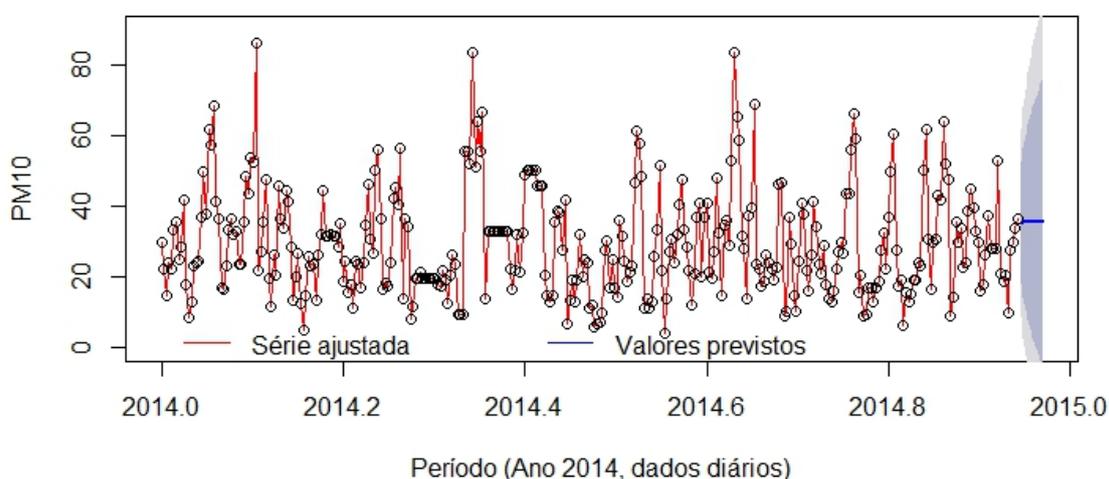
A série ajustada representada pela cor verde, as circunferências de cor preta é a série  $PM_{10}$  e, em azul, o valor encontrado para as previsões futuras.

### 4.3.2. Alisamento exponencial biparamétrico

Mesmo não sendo detectada tendência na série, o modelo foi testado. O valor da constante para o nível  $A = 0,7182$  e para constante da tendência  $C = 0,00001$ , baixo valor para constante já erra esperado tendo em vista a série não possuir tendências.

Na Figura 9, encontra-se a série ajustada e os valores previstos para os limites de previsão com intervalo de confiança de 95%.

**Figura 9** - Modelo de alisamento biparamétrico (variável  $PM_{10}$ ).



Fonte: Elaborado pelos autores.



A cor vermelha representa a série ajustada, as circunferências de cor preta, a série PM10 e, em azul, o valor encontrado para as previsões futuras.

### 4.3.3. Alisamento exponencial de Holt-Winters

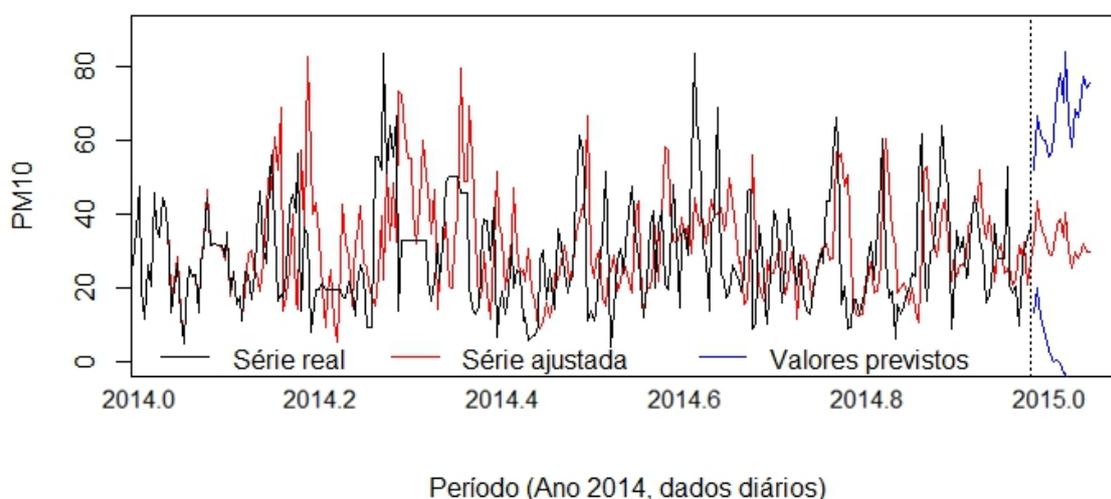
Para a série PM<sub>10</sub>, foram testadas as duas variações de ajuste desse modelo, o aditivo e o multiplicativo.

No ajuste dos modelos, os valores das constantes de suavização que melhor se ajustaram ao modelo aditivo foram:  $A=0,5447$  (constante de suavização do nível);  $C=0,0000$  (constante de suavização da tendência) e  $D=0,5980$  (constante de suavização da sazonalidade) e, para o multiplicativo, foi:  $A=0,2725$  (constante de suavização do nível);  $C=0,0000$  (constante de suavização da tendência) e  $D=0,3650$  (constante de suavização da sazonalidade).

Utilizando os critérios de seleção MAPE, EQM, MAD e SSE, o modelo multiplicativo ajustou-se de forma mais adequada à série em relação ao modelo aditivo. Esse resultado se deve ao comportamento explosivo dos dados, ou seja, a série de dados é heterogênea, conforme detectado na fase inicial da estatística descritiva.

Na Figura 10, podem ser observadas as séries com valores reais e previstos, além do intervalo de previsão ( $h=20$ ) do modelo Holt-Winters multiplicativo.

**Figura 10** - Holt-Winters multiplicativo (variável PM<sub>10</sub>).



Fonte: Elaborado pelos autores.

### 4.4. COMPARAÇÃO DAS PREVISÕES DOS MODELOS AJUSTADOS

Na Tabela 4, estão expostos os resultados de comparação encontrados para os modelos que melhor se ajustaram à série de dados PM<sub>10</sub>.



**Tabela 4** – Comparação dos modelos utilizando os critérios propostos.

Modelo	Variável	MAPE(%)	EQM	MAD	SSE
AR(1)	PM <sub>10</sub>	8,8%	6,22	2,38	595,63
ARX(1,[1,1])	PM <sub>10</sub>	5,5%	2,29	1,5	400,95
ALS	PM <sub>10</sub>	24,49%	54,60	7,38	1728,41
ALB	PM <sub>10</sub>	28,01%	61,63	7,85	1812,01
HW Multiplicativo	PM <sub>10</sub>	14,1%	14,34	3,75	1417,30

Fonte: Elaborada pelos autores.

Considerando os resultados dos critérios de comparação propostos, o modelo ARX demonstra melhor desempenho em todos os métodos utilizados em relação aos modelos concorrentes AR e da classe de alisamento exponencial, para PM<sub>10</sub>. Demonstrando que as previsões se aproximam de forma mais adequada à série observada quando se utilizam variáveis explicativas. Resultados similares são encontrados nos trabalhos de Pinto et al. (2018).

De maneira geral, os resultados de modelagem das médias diárias da PM<sub>10</sub> para o município de Canoas-RS, tanto para o modelo ARMA como para o modelo ARMAX, demonstram bons desempenhos de previsão com erro médio percentual de apenas 8,8% e 5,5% respectivamente, sendo os resultados satisfatórios estatisticamente.

## 5. CONSIDERAÇÕES FINAIS

Este trabalho objetivou modelar e prever a concentração média diária de PM<sub>10</sub> para o município de Canoas-RS, utilizando os modelos SARIMA, SARIMAX, Alisamento Exponencial Simples, Biparamétrico e de Holt-Winters, para o ano de 2014. Baseando-se em indicadores de desempenho de modelagem (MAPE, EQM, MAD e SSE), verificou-se que o modelo ARX(1,[1,1]) é o mais acurado, entre os modelos comparados, para fazer previsões e previsões da variável PM<sub>10</sub>.

Cabe ressaltar que as classes de modelos de previsão ajustadas e comparadas para variável PM<sub>10</sub> demonstraram desempenhos diferentes de previsão. Os modelos da classe de alisamento exponencial não se ajustaram adequadamente à série em estudo e não se recomenda sua aplicação para a região em estudo. Já os modelos da classe ARIMA e ARIMAX demonstraram desempenho satisfatório com erro médio percentual de apenas 8,8% e 5,5% respectivamente. Dessa forma, podem ser considerados bons, uma vez que tais resultados foram satisfatórios estatisticamente.

Sendo assim, a previsão da PM<sub>10</sub> pelos modelos ARIMA ou ARIMAX pode auxiliar no planejamento de diversos setores públicos e privados. Medidas preventivas podem ser tomadas com base em informações seguras, como: restringir fluxo de veículos, manter crianças e idosos em locais fechados ou com uso de máscaras, fiscalizar os principais geradores de poluição, entre outros pontos. Entretanto, é necessário que os dados coletados sejam precisos e em tempo hábil para se efetuar a modelagem necessária e para que se reproduzam cenários próximos ao real.



## 6. REFERÊNCIAS

- AKAIKE, H. A. Bayesian analysis of the minimum AIC procedure. **Annals of the Institute of Statistical Mathematics**, v.30, p.9-14, 1978.
- AKAIKE, H. A. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v.19, p.716-723, 1974.
- ALMEIRA, A. M.; RIOS, E. C. dos S. V.; OLIVEIRA, P. G. Saúde Humana e a Poluição do Ar. **Revista Conhecimento em Destaque**, v.4, n.10, 2015.
- ANDRADE, W. W. W. **Estimação de Modelos ARIMA/ARIMAX e Aplicação em Inferência de Perdas de Propano**. 2009. 127 f. Dissertação (Mestrado em Engenharia Elétrica e Computação) – Universidade Federal do Rio Grande do Norte, Natal 2009.
- BARBOSA, G. C.; NALLI, S.; GRIPA, W. R. Modelos de séries temporais aplicados à previsão de Material Particulado (PM10) na Região da Grande Vitória/ES. **Blucher Physics Proceedings**, v.1, n.2, p.34-36, 2014.
- BARCELLOS, C. *et al.* Mudanças climáticas e ambientais e as doenças infecciosas: cenários e incertezas para o Brasil. **Epidemiologia e Serviços de Saúde**, v.18, n.3, p.285-304, 2009.
- BARBIERO, C. C. M. **Séries temporais**: um estudo de previsão para a receita operacional da ECT – Empresa Brasileira de Correios e Telégrafos. 2003, 108f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2003.
- BARROS, A. C. *et al.* **Análise de Séries Temporais em R**: curso introdutório. Organização Pedro Guilherme Costa Ferreira. 1. Ed. Rio de Janeiro: Elsevier/FGV IBRE, 2018.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis**: forecasting and control. 3. ed. Prentice Hall: New Jersey, 1994.
- BOX, G. E. P.; PIERCE, D. A. Distribution of residual autocorrelations in autorregressive integrated moving average time series models. **Journal of the American Statistical Association**, v.65, n.332, p.1509-1526, 1970.
- CHECHI, L.; SANCHES, F. de O. Análise de uma série temporal de precipitação para Erechim (RS) e um possível método de previsão climática. **Ambiência Guarapuava**, v.9, p.43-45, 2013.
- DANFÁ, S. **Distribuição espacial da precipitação pluvial e sua erosividade para Guiné-Bissau**. 2009. 104 f. Dissertação (Mestrado em Engenharia Agrícola) – Universidade Federal de Lavras do Sul, Lavras, 2009.
- DAPPER, S. N. **Estudo da associação dos efeitos da poluição do ar na saúde da população de Canoas/RS**. 2016. 98 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Maria, Santa Maria, 2016.



DAPPER, S. N.; SPOHR, C.; ZANINI, R. R. Poluição do ar como fator de risco para a saúde: uma revisão sistemática no estado de São Paulo. **Estudos Avançados**, v.30, n.86, p.83-97, 2016.

DICKEY, D. A.; FULLER, W. A. Likelihood ratio statistics for autoregressive time series with a unit root. **Econométrica**, v.49, n.4, p.1057-1072, 1981.

ESQUIVEL, R. de M.; SENNA, V.; GOMES, G. S. S. Análise espectral singular: comparação de previsão em séries temporais. **Revista do Mestrado em Administração e Desenvolvimento Empresarial**, v.16, n.2, p.87-101, 2012.

FELICIANO, R. A. **Uma proposta de gerenciamento integrado de demanda e distribuição, utilizando sistemas de apoio à decisão (SAD) com Business Intelligence (BI)**. 2009. 179 f. Dissertação (Mestrado em Engenharia) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2009.

FEPAM. **Fundação Estadual de Proteção Ambiental Henrique Luiz Roessler**. RS. Disponível em: <http://www.fepam.rs.gov.br/>. Acesso em: 15 nov. 2016.

FREITAS, C. U. *et al.* Poluição do ar em cidades brasileiras: selecionando indicadores de impacto na saúde para fins de vigilância. **Epidemiologia e Serviços de Saúde**, Brasília, v.22, n.3, p.445-454, jul./set. 2013.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. Ed. São Paulo: Editora Atlas, 2008.

GRIPA, W.; REISEN, V.; REIS, N.; FAJARDO, F. Análise de predição e previsão das concentrações de material particulado inalável (PM10) na cidade de Cariacica, ES. **Revista Brasileira de Estatística**, v.73, n.237, p.37-57, 2012.

INMET. **Instituto Nacional de Meteorologia**. Disponível em: <http://www.inmet.gov.br/>. Acesso em: 15 nov. 2016.

LEITE, R. C. M.; GUIMARÃES, E. C.; LIMA, E. A. P.; BARROZO, M. A. de S.; TAVARES, M. Utilização de regressão logística simples na verificação da qualidade do ar atmosférico de Uberlândia. **Engenharia Sanitária e Ambiental**, v.16, n.1, 2011.

LIZZI, E. A. S. **Predição do número de casos mensal dos casos de dengue por modelos de séries temporais**. 2012. 131 f. Dissertação (Mestrado em Saúde na Comunidade) – Universidade de São Paulo, Ribeirão Preto, 2012.

LJUNG, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. **Biometrika**, v.65, n.2, p.297-303, 1978.

LOVATTO, P. B.; ALTEMBURG, S. N.; CASALINHO, H.; LOBO, E. A. Ecologia profunda: o despertar para uma educação ambiental complexa. **Revista REDES**, Santa Cruz do Sul, v.16, n.3, p.122-137, set./dez. 2011.

LOPES, M.; FERREIRA, J.; PIMENTEL, C.; BORREGO, C. Qualidade do ar, exposição e saúde, como indicadores de desenvolvimento e qualidade de vida nas cidades. In: PIRES, S. M.; ARAGÃO, A.; FIDÉLIS, T.; MENDES, I. **Indicadores de desenvolvimento sustentável: instrumentos estratégicos e inovadores para municípios sustentáveis | O caso de Estarreja**. Aveiro: Editora Instituto Jurídico/Universidade de Aveiro/OHMI/CNRS, 2017. p.258. Cap. 11.



MACHÍN, A. B. **Efeitos da exposição a poluentes do ar na saúde humana: internações por doenças respiratórias em Cuiabá, MT.** 2017, 74 f. Dissertação (Mestrado em Engenharia Mecânica) – Universidade Estadual Paulista, Guaratinguetá, 2017.

MARTÍNEZ, R. O.; ZAMPROGNO, B. Comparação de algumas técnicas de previsão em análise de séries temporais. **Revista Colombiana de Estatística**, v.26, n.2, p.129-157, 2003.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais.** São Paulo: Edgard Blucher, 2004.

MORETTIN, P. A.; TOLOI, C. M. C. **Previsão de Séries Temporais.** 2. ed. São Paulo: Editora Atual, 1987.

MENDES, A. **Mortes devido à poluição aumentam 14% em dez anos no Brasil.** 2019. Disponível em: <https://www.saude.gov.br/noticias/agencia-saude/45500-mortes-devido-a-poluicao-aumentam-14-em-dez-anos-no-brasil> . Acesso em: 04 jun. 2020.

MONTE, E. Z.; ALBUQUERQUE, T. T. A.; REISEN, V. A. Previsão da concentração de ozônio na Região da Grande Vitória, Espírito Santo, Brasil, utilizando o modelo ARMAX-GARCH. **Revista Brasileira de Meteorologia**, v.30, n.3, p.285-294, 2015.

OZAKI, V. A.; ADAMI, A. C. O. Modelagem estatística dos prêmios do seguro rural. **Revista Política Agrícola**, v.21, n.1, p.60-75, 2012.

PINTO, W. P.; REISEN, V. A.; MONTE, E. Z. Previsão da concentração de material particulado inalável, na Região da Grande Vitória, ES, Brasil, utilizando o modelo SARIMAX. **Engenharia Sanitária Ambiental**, v.23, n.2, p.307-318, 2018.

SANTOS, T. C. **Relação de Variáveis Atmosféricas com a qualidade do ar no Estado de São Paulo.** 2016. 88 f. Dissertação (Mestrado em Meio Ambiente e Recursos Hídricos) – Universidade Federal de Itajubá, Itajubá, 2016.

SILVA, L. C. **Utilizando o ARMAX para estimar a influência do PIB, SELIC e Inflação no faturamento de empresas do setor siderúrgico com ações na bolsa.** 2011. 82 f. Monografia (Bacharelado em Administração) – Universidade de Brasília, Brasília 2011.

SILVA, A. F.; VIEIRA, C. A. Aspectos da poluição atmosférica: uma reflexão sobre a qualidade do ar nas cidades brasileiras. **Revista Ciências e Sustentabilidade**, v.3, n.1, 2017.

SOUZA, Amaury *et al.* Impacto de fatores meteorológicos sobre as concentrações de ozônio modelados por análise de séries temporais e métodos estatísticos multivariados. **Revista HOLOS**, v.5, p.2-16, 2017.

TIDRE, P. V.; BIASE, N. G.; SILVA, M. I. S. Utilização dos modelos de séries temporais na previsão do consumo mensal de energia elétrica da região norte do Brasil. **Revista Eletrônica Matemática e Estatística em Foco**, v.1, n.1, p.57-63, 2013.

Submetido em: **11/11/2019**

Aceito em: **14/07/2020**