

Previsão da concentração de material particulado inalável, na Região da Grande Vitória, ES, Brasil, utilizando o modelo SARIMAX

Inhalable particulate matter concentration forecast, in the Greater Vitória Region, ES, Brazil, using the SARIMAX model

Wanderson de Paula Pinto¹, Valdério Anselmo Reisen², Edson Zambon Monte³

RESUMO

Este trabalho objetivou modelar e prever a concentração média diária de material particulado inalável (PM_{10}), na Região da Grande Vitória (RGV), Espírito Santo, 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 PM_{10} 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), sendo escolhida a estação da Enseada do Suá para fazer o estudo de predição e previsão. Baseando-se em indicadores de desempenho de modelagem, verificou-se que o modelo SARIMAX (1,0,2) (0,1,1), é o mais acurado entre os estudados, objetivando fazer predições e previsões da qualidade do ar na RGV. Em comparação com os modelos ARMA, o desempenho estatístico do modelo SARIMAX foi superior, no que diz respeito à predição de eventos de qualidade do ar regular. Dentre as variáveis meteorológicas avaliadas, a velocidade do vento e a precipitação pluvial foram significativas e melhoraram o ajuste do modelo. Em termos de previsão da qualidade do ar, os modelos de séries temporais mostraram resultados satisfatórios.

Palavras-chave: poluição do ar; PM_{10} ; séries temporais; SARIMAX.

ABSTRACT

This study aimed to model and forecast the average daily concentration of inhalable particulate matter (PM_{10}), in the Greater Vitoria Region (GVR), Espírito Santo, Brazil, using the SARIMAX model, for the period from January 1st, 2012 to April 30th, 2015. Data set from the State Environmental Institute was used. The Enseada do Suá station was chosen for purposes of prediction and forecasting. Some meteorological parameters (wind speed, relative humidity, rainfall and temperature) measured at the GVR were taken as explanatory variables of PM_{10} concentrations. Based on modelling performance indicators, it was verified the SARIMAX model (1,0,2) (0,1,1), is the most accurate between the ones studied, purposing to predict and forecast the air quality in the GVR. The statistical performance of the SARIMAX model was better than the ARMA model, with regard to prediction of regular air quality events. Among the evaluated meteorological variables, wind speed and rainfall were significant and improved the model estimated. Regarding to air quality forecasting, the time series models showed satisfactory results.

Keywords: air pollution; PM_{10} ; time series; SARIMAX.

INTRODUÇÃO

A poluição atmosférica caracteriza-se basicamente pela presença de gases tóxicos e partículas sólidas no ar (SEINFELD & PANDIS, 2006). Suas fontes são classificadas em antropogênicas, como, por exemplo, emissões provenientes de indústrias e escapamentos de veículos; e naturais, como as decorrentes de erupções vulcânicas.

Segundo Holgate *et al.* (1999), um nível elevado dos poluentes pode ocasionar desde irritação nos olhos, nariz e garganta, bronquite e pneumonia, até doenças respiratórias crônicas, câncer de pulmão, problemas cardíacos etc. Diversos estudos epidemiológicos têm demonstrado associações significativas entre a exposição às concentrações elevadas de poluentes atmosféricos e problemas de saúde (OSTRO *et al.*, 1996; MARTINS *et al.*, 2002; GOUVEIA *et al.*, 2003;

¹Doutorando em Engenharia Ambiental pela Universidade Federal do Espírito Santo (UFES) - Vitória (ES), Brasil.

²Professor do Departamento de Estatística, do Programa de Pós-Graduação em Engenharia Ambiental e do Programa de Pós-Graduação em Economia da UFES - Vitória (ES), Brasil.

³Professor do Departamento de Economia e do Programa de Pós-Graduação em Economia e membro do Grupo de Pesquisa em Econometria da UFES - Vitória (ES), Brasil.

*Autor correspondente: wandersonpp@gmail.com

Recebido: 02/09/16 - Aceito: 23/02/17 - Reg. ABES: 168758

ALMEIDA, 2006; NASCIMENTO *et al.*, 2006; CURTIS *et al.*, 2006; BRAGA *et al.*, 2007; SOUZA *et al.*, 2014).

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 quanto à adaptação de comportamento da população de grupos de risco, como crianças, idosos e pessoas com doenças respiratórias. Além disso, também pode servir para as autoridades competentes como informação para a preparação de planos para redução de emissões e gerenciamento da qualidade do ar (GOMES, 2009).

Na literatura, existem vários trabalhos que utilizaram modelos estatísticos para modelar e fazer previsão da qualidade do ar. Agirre-Basurko, Ibarra-Berastegui e Madariaga (2006) utilizaram três modelos, um de regressão linear múltipla e dois de rede neural para modelar e prever a qualidade do ar da cidade de Bilbao, Espanha. O fluxo de veículos e as variáveis meteorológicas foram usados como modelos de dados de entrada – temperatura, umidade relativa, pressão, radiação, gradiente de temperatura, direção do vento e velocidade do vento – no período de 1993 a 1994. Como saída prevista para modelos, adotou-se as concentrações de O_3 e NO_2 , com horizonte de previsão de oito horas à frente. Os resultados mostraram que os modelos de rede neural obtiveram resultados melhores para a previsão das concentrações de O_3 e NO_2 quando comparados ao modelo de regressão linear múltipla. Quanto ao desempenho dos modelos de rede neural, o que mais se destacou foi o modelo que considerou a sazonalidade da série das concentrações de O_3 e NO_2 .

Goyal, Chan e Jaiswal (2006) realizaram um estudo com três modelos estatísticos aplicados à média diária de concentração de MP_{10} , medido nas cidades de Delhi e Hong Kong. O trabalho objetivou desenvolver um modelo estatístico de previsão das concentrações de MP_{10} e promover um estudo comparativo através do desempenho dos modelos, a saber:

1. modelo de regressão linear múltipla (modelo 1);
2. modelo de séries temporais ARIMA (modelo 2);
3. combinação entre os modelos 1 e 2 (modelo 3).

Além do MP_{10} , alguns parâmetros meteorológicos foram adotados, como a velocidade do vento, a temperatura, a radiação solar e a umidade relativa do ar, medidos no período de junho de 2000 a junho de 2001. Na comparação entre os modelos, as medidas de erro mostraram que o modelo 3 foi o que obteve o melhor desempenho. O estudo de previsão ocorreu apenas para a cidade de Delhi, e compreendeu o período de junho de 2001 a junho de 2002. O modelo 3 foi utilizado, e os resultados da previsão foram satisfatórios.

Gomes (2009) realizou um estudo de previsão de índices de qualidade do ar da Região da Grande Vitória (RGV), Espírito Santo,

Brasil, utilizando o modelo autorregressivo de valores inteiros INAR(p). O período de análise foi de 01/01/07 a 19/03/07, sendo as previsões datadas de 20/03/07 a 25/03/07. Os poluentes investigados foram: monóxido de carbono (CO), dióxido de nitrogênio (NO_x), dióxido de enxofre (SO_2) e ozônio (O_3). Para a escolha do modelo mais adequado, o autor utilizou o critério de seleção automática para modelos INAR(p), o $AICC_{INAR}$, que seleciona a melhor ordem p para cada modelo. Os resultados mostraram que todas as previsões para os índices de qualidade do ar foram classificadas como boa, conforme a Resolução do Conselho Nacional do Meio Ambiente (CONAMA, 1990) nº 03. Porém, baseados nas diretrizes da World Health Organization (WHO, 2005), a previsão do poluente SO_2 no dia 20/03/07, estação do Centro de Vila Velha, excedeu o valor de $20 \mu g.m^{-3}$ para média de 24 horas, ou seja, mesmo estando dentro o limite do padrão nacional essa concentração é prejudicial para a saúde humana.

Gripa *et al.* (2012) compararam dois modelos: um de séries temporais e um de regressão linear múltipla, para modelagem e previsão das concentrações médias de MP_{10} , monitorados na RGV, com a incorporação de fatores meteorológicos. Ambos os modelos evidenciaram resultados semelhantes. No entanto, o modelo de regressão apresentou medidas de previsão das concentrações médias de MP_{10} um pouco melhores do que o modelo de séries temporais. Reisen *et al.* (2014) modelaram a média diária de concentração de MP_{10} , na cidade de Cariacica, Espírito Santo, Brasil, utilizando um processo integrado fracionado sazonal, com volatilidade. Os autores concluíram que o modelo, ajustado com erros heterocedásticos, captou bem a dinâmica da série e foi capaz de prever os períodos de maior volatilidade.

Monte, Albuquerque e Reisen (2015) realizaram um estudo para estimar e prever a concentração horária de ozônio na RGV, utilizando um modelo ARMAX-GARCH no período de 01/01/2011 a 31/12/2011. O estudo utilizou dados cedidos pelo Instituto Estadual de Meio Ambiente e Recursos Hídricos (IEMA). De acordo com os resultados apresentados, as previsões horárias para o dia 31/12/2011 foram muito próximas dos valores observados. Observou-se também que as estimativas seguiram a trajetória diária da concentração de ozônio. Os autores concluíram que o modelo ARMAX-GARCH é mais eficaz na predição de episódios de poluição de ozônio, em comparação aos modelos autorregressivos e de médias móveis (ARMA) e autorregressivos e de médias móveis com variáveis explicativas (ARMAX).

Vale ressaltar que uma série temporal é composta, em geral, por três componentes não observáveis, a saber: tendência, sazonalidade e aleatoriedade. A sazonalidade é uma componente difícil de ser modelada, pois é necessário compatibilizar a questão física do problema em estudo com a questão estatística. Define-se um fenômeno sazonal como aquele que ocorre regularmente em períodos fixos de tempo (LATORRE & CARDOSO, 2001). Na análise de séries temporais, os

modelos ARMA podem ser empregados quando a série em estudo está livre de tendência e de sazonalidade, os modelos ARIMA são utilizados quando há tendência e, para incorporar o componente de sazonalidade, utilizam-se os modelos autorregressivos integrados e de médias móveis sazonal multiplicativo (SARIMA) (LATORRE & CARDOSO, 2001; MORETTIN & TOLOI, 2006). Conforme Reisen *et al.* (2014), os modelos que descrevem de forma adequada o comportamento físico dos dados são essenciais para a previsão precisa em qualquer área de aplicação, pois a sazonalidade é um fenômeno característico do poluente MP_{10} .

Nesse contexto, este trabalho objetivou modelar e prever a concentração média diária de MP_{10} na RGV, utilizando o modelo SARIMAX, no período 01/01/2012 a 30/04/2015. Mesmo tendo ultrapassado apenas 3 vezes os padrões primário e secundário ($150 \mu\text{g}\cdot\text{m}^{-3}$), média de 24 horas, estabelecidos pela Resolução CONAMA nº 03, de 28/06/1990 (CONAMA, 1990), no período de estudo, em 144 dias, de um total de 1.216, as concentrações monitoradas na estação da Enseada do Suá excederam o valor de $50 \mu\text{g}\cdot\text{m}^{-3}$, o que vai de encontro com as diretrizes estabelecidas pela WHO (2005) para esse poluente. O IEMA estabelece uma qualidade do ar boa, para concentrações de MP_{10} entre 0 e $45 \mu\text{g}\cdot\text{m}^{-3}$; regular para aquelas entre 46 e $120 \mu\text{g}\cdot\text{m}^{-3}$; inadequada, entre 121 e $250 \mu\text{g}\cdot\text{m}^{-3}$ (IEMA, 2013). No período de estudo, em 450 ocasiões a qualidade do ar foi classificada como regular; e em 6, como inadequadas. Logo, a importância desta pesquisa é justificada, principalmente no que diz respeito à formulação de medidas preventivas por parte das autoridades competentes, uma vez que a concentração de MP_{10} vem atingindo níveis que são prejudiciais à saúde na região de estudo.

MATERIAL E MÉTODOS

Área de estudo e variáveis analisadas

Para a realização desse estudo, utilizou-se séries temporais de concentrações de poluentes atmosféricos e de variáveis meteorológicas monitorados na RGV, que é constituída pelos municípios de Vitória, Vila Velha, Cariacica, Serra e Viana. Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2010), a RGV abrange uma área de 1.461 km^2 , com aproximadamente 1.475.332 habitantes, sendo um dos principais polos de desenvolvimento urbano e industrial do estado. A região sofre com diversos tipos de problemas ambientais, entre os quais está a deterioração da qualidade do ar, devido às emissões atmosféricas por indústrias, pela frota veicular e ressuspensão do solo causada pelo vento e tráfego veicular.

Vale ressaltar que a RGV possui uma Rede Automática de Monitoramento da Qualidade do Ar (RAMQAR), inaugurada em julho de 2000, de propriedade do IEMA. Essa rede é distribuída em oito estações, a saber:

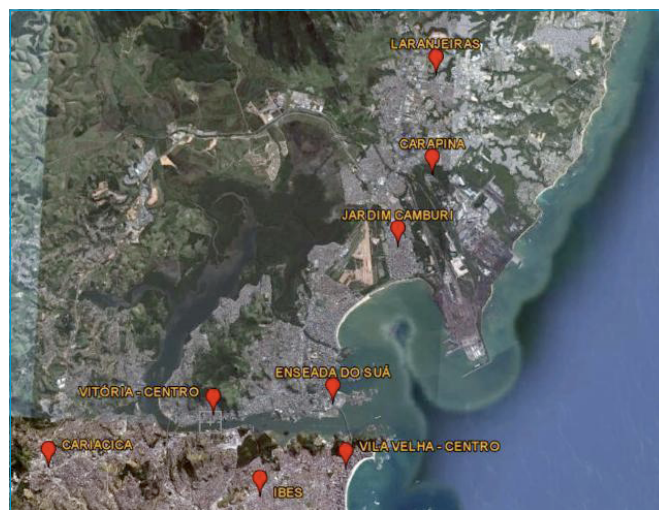
1. o município Serra; com duas estações localizadas nas regiões de Laranjeiras e Carapina;
2. o município Vitória, com três estações localizadas nas regiões de Jardim Camburi, Enseada do Suá e Centro de Vitória;
3. o município de Vila Velha apresenta duas estações localizadas nas regiões do Ibes e Centro de Vila Velha;
4. o município de Cariacica, com uma estação em Cariacica.

A localização espacial das estações de monitoramento da RAMQAR está ilustrada na Figura 1.

A RAMQAR monitora os seguintes poluentes: partículas totais em suspensão (PTS); MP_{10} ; O_3 ; NO_x ; CO e hidrocarbonetos (HC). E realiza, ainda, o monitoramento dos seguintes parâmetros meteorológicos: direção dos ventos (DV); velocidade dos ventos (VV); precipitação pluviométrica (PP); umidade relativa do ar (UR); precipitação pluviométrica (PP); temperatura (T); pressão atmosférica (P); e radiação solar (I). Os poluentes e parâmetros meteorológicos monitorados em cada estação RAMQAR encontram-se no Quadro 1. Na análise deste trabalho, as variáveis MP_{10} , VV, UR, PP e T foram utilizadas conforme descrição do Quadro 2.

Modelo SARIMAX

Uma série temporal é um conjunto de observações de qualquer fenômeno aleatório, ordenadas no tempo. A análise de séries temporais consiste em procurar alguma relação de dependência existente temporalmente nos dados, identificando o mecanismo gerador da série com objetivo de extrair periodicidades relevantes nas observações, descrever o seu comportamento e fazer previsões (MORETTIN & TOLOI, 2006; BAYER & SOUZA, 2010).



Fonte: Google Earth.

Figura 1 - Localização espacial das estações de monitoramento da qualidade do ar da RGV.

Seja Y_t ($t = 1, 2, 3, \dots$) um processo linear com representação dada pela Equação 1:

$$(1 - B)^d(1 - B^s)^p\Phi(B^s)\Theta(B)Y_t = \Theta(B^s)\theta(B)\varepsilon_t \tag{1}$$

Em que

s é chamado período sazonal do processo e,

ε_t é ruído branco (RB), definido como uma sequência de variáveis aleatórias não correlacionadas com média zero e variância constante ao longo do tempo (WEI, 2006), ou seja, $\varepsilon_t \sim RB(0, \sigma_\varepsilon^2)$.

Quadro 1 - Parâmetros meteorológicos e poluentes monitorados em cada estação da Rede Automática de Monitoramento da Qualidade do Ar.

Estação	PTS	MP ₁₀	SO ₂	CO	NO _x	HC	O ₃	Meteorologia
Laranjeiras	X	X	X	X	X		X	
Carapina	X	X						DV, VV, UR, PP, P, T, I
Jardim Camburi	X	X	X		X			
Enseada do Suá	X	X	X	X	X	X	X	DV, VV
Vitória Centro	X	X	X	X	X	X		
Ibes	X	X	X	X	X	X	X	DV, VV
Vila Velha		X	X					
Cariacica	X	X	X	X	X		X	DV, VV, T

Fonte: adaptado de IEMA, 2013.

PTS: partículas totais em suspensão; MP₁₀: material particulado inalável; SO₂: dióxido de enxofre; CO: monóxido de carbono; NO_x: dióxido de nitrogênio; HC: hidrocarbonetos; O₃: ozônio; DV: direção dos ventos; VV: velocidade dos ventos; UR: umidade relativa do ar; PP: precipitação pluviométrica; P: pressão atmosférica; T: temperatura; I: radiação solar.

Quadro 2 - Descrição das variáveis material particulado inalável, velocidade do vento, umidade, precipitação e temperatura.

Variáveis	Unidades	Descrição
MP ₁₀	µg.m ³	Existem medições para todas as oito estações da RAMQAR. Entretanto, como existem muitos dados faltantes para as demais estações, optou-se por trabalhar apenas com as médias diárias da estação da Enseada do Suá. Além disso, essa estação ultrapassou algumas vezes o padrão do CONAMA para este poluente.
Velocidade do vento	m.s ⁻¹	Valores medidos na estação de Carapina. Foi escolhida essa estação por apresentar a menor porcentagem de dados faltantes no período de estudo.
Umidade relativa	%	Como existem muitos dados faltantes para a estação de Cariacica, adotou-se o valor de Carapina.
Precipitação	mm	Valores medidos na estação de Carapina, única que possui medição para tal variável.
Temperatura	°C	Média aritmética entre as estações de Carapina e Cariacica, únicas que possuem medições para tal variável.

MP₁₀: material particulado inalável; RAMQAR: Rede Automática de Monitoramento da Qualidade do Ar; CONAMA: Conselho Nacional do Meio Ambiente.

Em $(1 - B)^d(1 - B^s)^p$, d e D são números inteiros não negativos e representam o número de diferenças simples e sazonais, respectivamente, aplicadas sobre o processo Y_t .

Tem-se que, B é o operador de defasagem definido como $B^k Y_t = Y_{t-k}$, $k \in \mathbb{N}$, $\phi(z) = 1 - \sum_{k=1}^p \phi_k z^k$, $\Phi(z^s) = 1 - \sum_{k=1}^P \Phi_k z^{sk}$, $\theta(z) = 1 - \sum_{k=1}^q \theta_k z^k$ e $\Theta(z^s) = 1 - \sum_{k=1}^Q \Theta_k z^{sk}$ são polinômios de ordem P , p , Q , $q \in \mathbb{N}$ respectivamente, com $z \in \mathbb{C}$, em que \mathbb{C} representa o conjunto dos números complexos e $\{\phi_k\}$, $\{\theta_k\}$, $\{\Phi_k\}$, $\{\Theta_k\}$ são sequências de números reais. O processo Y_t com representação dada na Equação 1 é denominado modelo autorregressivo integrado e de médias móveis sazonal multiplicativo (SARIMA), de ordem $(p, d, q) \times (P, D, Q)_s$. O processo Y_t é estacionário e invertível se $d = D = 0$ e as raízes de $\phi(z)$, $\Phi(z^s)$, $\theta(z)$ e $\Theta(z^s)$ são não comuns e encontram-se fora do círculo unitário (WEI, 2006). As ordens p, d, q, P, D, Q e s devem ser identificadas seguindo a metodologia de Box e Jenkins (1970).

Nesse estudo, foi considerado o modelo SARIMAX (modelo autorregressivo integrado e de médias móveis sazonal multiplicativo com variáveis explicativas), que é uma extensão do modelo SARIMA, utilizando outras séries temporais como variáveis explicativas. O modelo SARIMAX explica a variável dependente por meio de: variáveis explicativas; defasagens das variáveis explicativas; e defasagens da variável dependente (MOURA; MONTINI; CASTRO, 2011).

Metodologia de modelagem

A comprovação da sazonalidade foi verificada pela análise espectral e pelo teste G de Fisher. As hipóteses testadas foram as seguintes:

1. H_0 : não existe sazonalidade;
2. H_1 : existe sazonalidade.

A estatística do teste é dada pela Equação 2:

$$G = \frac{\max [I_p(f_i)]}{\sum_{i=1}^{\left(\frac{N}{2}\right)} I_p(f_i)} \tag{2}$$

Em que:

I_p = periodograma no período p ; e

N = número de observações da série.

Segundo Barbosa *et al.* (2015), o periodograma consiste na decomposição da série temporal em uma série de Fourier. No eixo das ordenadas ficam localizadas as frequências da série (f_i) e no eixo das abscissas as respectivas intensidades de cada frequência $I_p(f_i)$, definidas pela Equação 3:

$$I_p(f_i) = \frac{2}{\left(\frac{N}{2}\right)} \left\{ \sum_{t=1}^n \varepsilon_t \cos \left[\frac{2\pi i}{\left(\frac{N}{2}\right)} t \right] + \sum_{t=1}^n \varepsilon_t \sin \left[\frac{2\pi i}{\left(\frac{N}{2}\right)} t \right] \right\}^2 \tag{3}$$

Em que

ϵ_t representa o componente estocástico da série temporal associado ao tempo t e N é o número de observações da série. A distribuição exata de G é dada por $Z_\alpha = 1 - \left(\frac{\alpha}{c}\right)^{\frac{1}{c-1}}$ sendo $c = \left(\frac{N}{2}\right)$ e α o nível de significância adotado. Se $G > Z_\alpha$, a hipótese H_0 é rejeitada e a série apresenta periodicidade no período i . Para determinar o período sazonal (s) verifica-se a qual frequência está associado o maior valor $I_p(f_i)$ e, então, divide 1 por esse valor de frequência, isto é, $s = \left(\frac{1}{f_i}\right)$. Para detalhes, consultar Morettin e Tolo (2006).

A metodologia de Box e Jenkins aplicada neste trabalho está dividida nas seguintes etapas:

1. identificação;
2. estimação;
3. diagnóstico;
4. previsão.

A identificação do modelo a ser ajustado aos dados é uma das fases mais complicadas. Para isso, foram utilizados os critérios de seleção de modelos.

O *Akaike Information Criterion* (AIC) (AKAIKE, 1973) é um critério de seleção comumente utilizado. Utilizando os estimadores de máxima verossimilhança para os parâmetros do modelo, em que $\hat{\xi}$ é a função de log-verossimilhança maximizada e k é o número de parâmetros do modelo, o AIC é dado por: $AIC = -2l(\hat{\xi}) + 2k$. Em uma perspectiva bayesiana, Schwarz (1978) e Akaike (1978) introduziram o *Bayesian Information Criterion* (BIC), dado por: $BIC = -2l(\hat{\xi}) + k \log(n)$, em que n é o número de observações da amostra. Com base nesses critérios de informação, pode-se ajustar diversos modelos e escolher aquele que obtiver o menor valor para o critério de informação.

Além dos critérios de seleção de modelos, outras duas medidas de qualidade de ajuste foram utilizadas para auxiliar na seleção do modelo que melhor se ajusta à série temporal em estudo e, principalmente, na avaliação da qualidade do ajuste. As medidas utilizadas neste trabalho foram a raiz do erro quadrático médio (REQM) e erro absoluto médio (EAM), definidas pelas Equações 4 e 5:

$$REQM = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2} \quad (4)$$

$$EAM = \frac{1}{N} \sum_{i=1}^N |Y_i - \hat{Y}_i| \quad (5)$$

Em que

Y_i e \hat{Y}_i são, respectivamente, os valores observados e previstos no instante i . Um dos principais objetivos deste trabalho foi a comparação entre os modelos quanto ao seu desempenho no processo de ajustamento aos dados em estudo e uma das formas de verificar essa qualidade é o estudo de previsão.

RESULTADOS E DISCUSSÕES

Análise dos dados e ajuste dos modelos

Para um entendimento preliminar das variáveis em estudo são apresentadas algumas medidas descritivas (Tabela 1). A análise corresponde às variáveis apresentadas no Quadro 2, para o período de 01/01/2012 a 30/04/2015, perfazendo um total de 1.216 observações. Todas as análises foram feitas utilizando o *software* livre R (R CORE TEAM, 2016). A presença de dados faltantes nas séries motivam o uso da metodologia de imputação via algoritmo *expectation-maximisation* (EM), proposta por Junger e Leon (2015) e implementada na biblioteca R *multivariate time-series data imputation* (MTSDI).

A concentração média de MP_{10} foi de, aproximadamente, $43,0 \mu\text{g}\cdot\text{m}^{-3}$ com desvio padrão de $15,3 \mu\text{g}\cdot\text{m}^{-3}$. Nota-se que, em média, as concentrações não ultrapassaram o valor de $50,0 \mu\text{g}\cdot\text{m}^{-3}$, porém, o desvio padrão e o coeficiente de variação alto indicam uma média pouco representativa. Além disso, os resultados mostram que o valor máximo foi mais do que o triplo do valor médio, demonstrando a grande variabilidade das concentrações de MP_{10} na RGV. O valor mínimo foi observado em março de 2013, na estação chuvosa da região, que se inicia a partir do mês de outubro e vai até meados de abril (IEMA, 2013). Nessa época do ano, a atuação dos sistemas frontais e de zonas de convergência de umidade favorecem o aumento de precipitação. Nesse período, as concentrações de MP_{10} são menores, principalmente pela intensificação da eficiência dos processos de remoção por deposição úmida. Em relação às variáveis

Tabela 1 - Medidas descritivas das variáveis sob estudo.

Medidas descritivas	MP_{10}	Velocidade	Umidade	Precipitação	Temperatura
Média	43,8853	2,4918	76,9636	0,1217	23,1228
Mediana	41,7917	2,3683	76,7184	0,0000	23,7253
Desvio padrão	15,3020	0,7980	6,6742	0,5198	3,9298
Coeficiente de variação	34,8681	32,0233	8,6719	427,0784	16,9954
Valor máximo	159,1250	5,5412	98,2983	12,7750	30,7085
Valor mínimo	8,9167	1,0392	29,7162	0,0000	5,5700

MP_{10} : material particulado inalável.

meteorológicas, nota-se que as mesmas mostraram grande variabilidade em termos estatísticos.

A Tabela 2 mostra as correlações calculadas entre as concentrações de médias diárias de MP₁₀, de VV, da UR, de PP e T. Observa-se que as variáveis meteorológicas apresentam relação linear com as concentrações de MP₁₀, na estação de Enseada do Suá. Conforme esperado, os índices de MP₁₀ estão associados às mudanças dessas variáveis, ou seja, infere-se dos resultados (Tabela 2) que o aumento da VV e da T na RGV acarreta aumento nas concentrações de MP₁₀ e que o aumento da PP e UR acarreta diminuição das concentrações de MP₁₀.

A Figura 2 apresenta a série de MP₁₀ e sua decomposição em componentes de tendência, sazonalidade e aleatoriedade. Nota-se a presença da propriedade de sazonalidade e a inexistência de tendência nos dados, uma vez que o gráfico da componente tendência não apresenta um comportamento de crescimento ou decrescimento ao

longo do tempo. Segundo Wei (2006), o primeiro passo na análise de séries temporais é verificar se as mesmas apresentam média, variância e covariância constantes ao longo do tempo, ou seja, se são estacionárias. Caso elas não forem estacionárias, aplica-se a primeira diferença na mesma para tentar estacionarizá-las. O resultado para o teste de Dickey-Fuller Aumentado (ADF) (DICKEY & FULLER, 1981) revela que as séries em estudo são estacionárias na média, o que confirma que a série não apresenta tendência.

Como mencionado anteriormente, há indícios de sazonalidade nos dados. Portanto, para confirmar tal hipótese e verificar a necessidade de aplicação de diferenças sazonais, foi realizada a decomposição espectral da série MP₁₀, conforme mostra a Figura 3.

O periodograma demonstra que o maior pico está associado à frequência 0,14333; o que implica em $s = \frac{1}{0,14333} = 6,97$, ou seja, uma componente sazonal com periodicidade de sete dias. De acordo com

Tabela 2 - Matriz de correlação entre as variáveis sob estudo.

	MP ₁₀	Velocidade	Umidade	Precipitação	Temperatura
MP ₁₀	1,0000				
Velocidade do vento	0,1149	1,0000			
Umidade	-0,1377	-0,4653	1,0000		
Precipitação	-0,1332	-0,0979	0,2680	1,0000	
Temperatura	0,1794	0,2871	-0,2125	-0,0653	1,0000

MP₁₀: material particulado inalável.

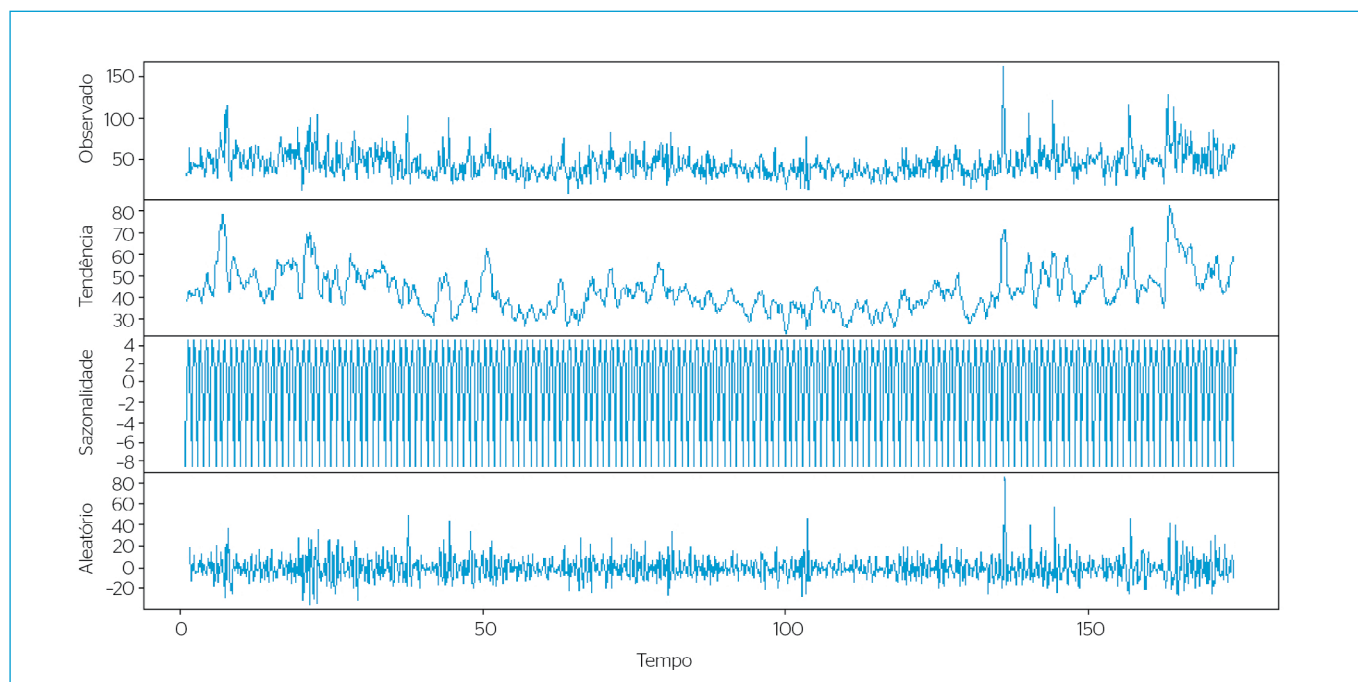


Figura 2 - Decomposição da série temporal em componentes de sazonalidade, de tendência e de aleatoriedade.

o teste de Fisher, foram obtidos os valores das estatísticas $G=0,1167$ e $Z=0,0156$. Como o valor de G é superior ao de Z , rejeita-se a hipótese H_0 , confirmando a existência da componente sazonalidade para períodos de 7 dias ao nível nominal de significância de 5%. A sazonalidade em séries de MP_{10} é esperada, visto que segundo o IEMA (2013), a principal fonte emissora de partículas na RGV são veículos automotores, representando mais de 60% das emissões de partículas que estão ligados à ressuspensão de partículas em vias. Dessa forma, existe variação entre as concentrações medidas nos dias úteis e finais de semana, uma vez que o fluxo de veículos é maior durante os dias

da semana. Logo, uma diferença sazonal de ordem $s=7$ foi aplicada para eliminar a sazonalidade presente na série. A Figura 4 mostra a função de autocorrelação (FAC) (a) da série e a FAC (b) da série diferenciada de ordem 7.

A identificação preliminar das ordens, p e q e P e Q , do modelo SARIMAX a ser estimado ocorreu pela análise do comportamento da FAC e da função de autocorrelação parcial (FACP), para detalhes ver Wei (2006). Dessa forma, com os valores de p e q variando de 1 a 3, e P e Q de 0 a 3, intervalos escolhidos pela indicações do correlograma, testaram-se alguns modelos com essas combinações. Como a

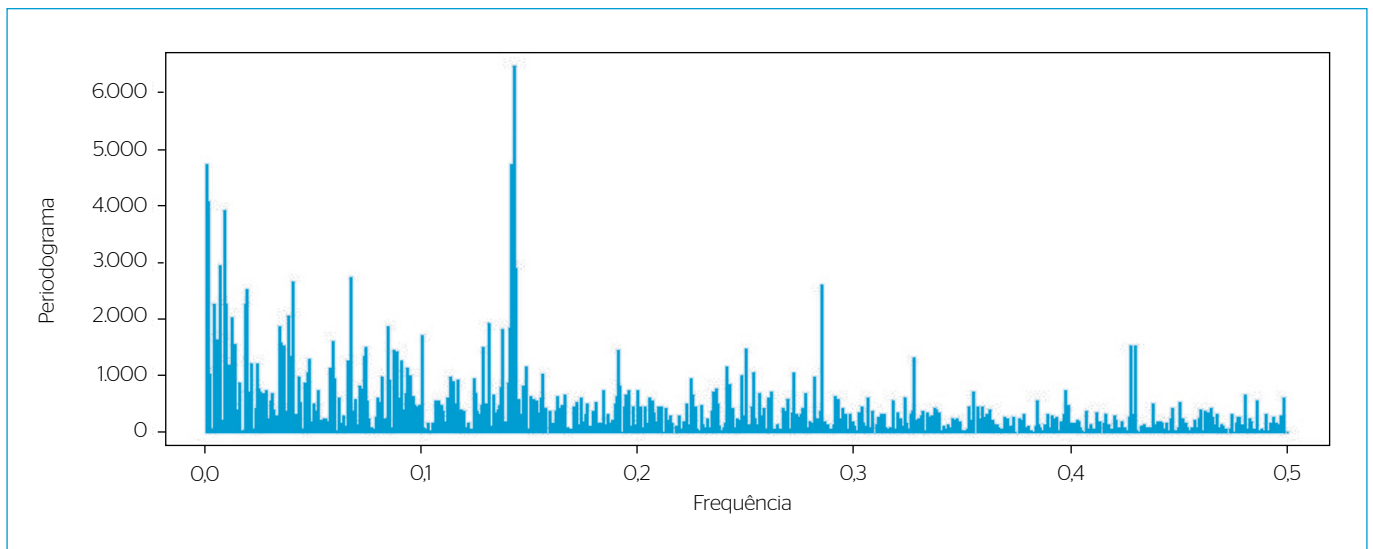


Figura 3 - Periodograma da série de material particulado inalável.

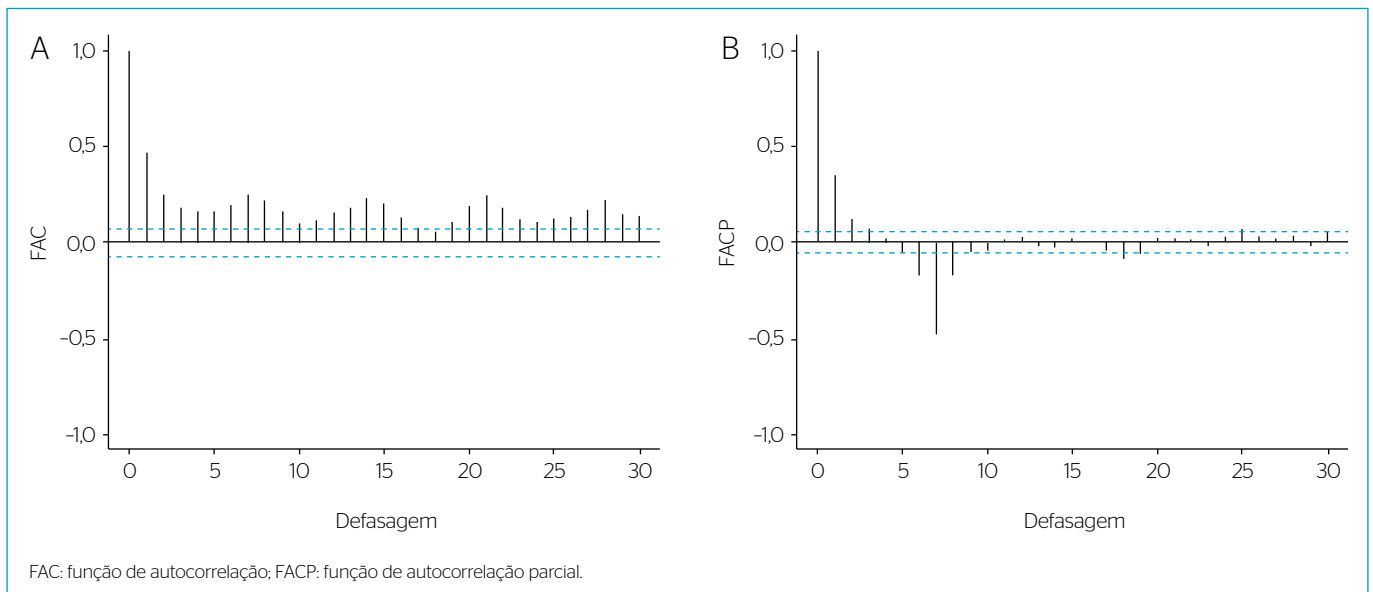


Figura 4 - Função de autocorrelação da série (A); e função de autocorrelação da série (B) após aplicação da diferença sazonal.

série é estacionária, não houve a necessidade de aplicar diferenças de ordem 1, então, tem-se que $d=0$. Da mesma forma, como uma diferença sazonal foi aplicada, tem-se $D=1$. Feito isso, a identificação das ordens autorregressivas e de médias móveis, do melhor modelo para representar os dados, baseou-se nos critérios de informação AIC, BIC e nos valores da REQM e EAM. Dentre todos o modelos estimados, o modelo com melhor ajuste foi o SARIMAX(1,0,2)(0,1,1)₇, por ter apresentado os menores valores calculados para os critérios de informação e medidas de qualidade. Sendo assim, esse torna-se o modelo de interesse.

Para verificar a adequação do modelo, a teoria estatística estabelece suposições das propriedades do mesmo, tais como não autocorrelação e normalidade dos resíduos (WEI, 2006). Na Tabela 3 são apresentadas as estatísticas para os testes Shapiro e Wilk (1965), Bera e Jarque (1981), Ljung e Box (1978) e Box e Pierce (1970). Observa-se que os resíduos não são normalmente distribuídos, resultado que já era esperado por se tratar de uma variável ambiental. Porém, assumiu-se, pela teoria assintótica sobre a média das distribuições de probabilidade, a suposição de que os resíduos são normalmente distribuídos. O valor p dos testes de correlação são superiores ao nível de significância de 5%, indicando que não se rejeita a hipótese

nula de erros não autocorrelacionados. Nenhuma autocorrelação apresentou-se significativamente diferente de zero no correlograma residual apresentado na Figura 5, o que vai ao encontro dos resultados apresentados na Tabela 3 e com a hipótese de homogeneidade dos resíduos.

A Tabela 4 contém as estimativas dos parâmetros do modelo ajustado, considerando como variáveis explicativas a VV e a PP. Observa-se que todos os coeficientes são estatisticamente significativos de acordo com o teste z. Uma vez que as variáveis umidade e temperatura não foram estatisticamente significativas, as mesmas não foram incorporadas no ajuste do modelo final. Chaloulakou, Grivas, Spyrellis (2003) estimaram a média diária de MP_{10} em Atenas por meio da análise de regressão. Os autores observaram que a variável VV influencia na concentração de MP_{10} . Esses resultados vão ao encontro dos apresentados no presente trabalho. Lyra, Oda-Souza e Viola (2011) observaram que as variáveis UR e a PP foram os elementos meteorológicos significativos para explicar a variabilidade do MP_{10} na cidade do Rio de Janeiro. Os resultados para umidade encontrados pelos autores contrariam as estimativas deste trabalho, uma vez que a umidade não foi

Tabela 3 - Testes estatísticos de normalidade* e correlação dos resíduos do modelo escolhido.**

Parâmetro	Estimativa
Jarque-Bera*	<0,0001
Shapiro-Wilk*	<0,0001
Ljung-Box**	0,9095
Box-Pierce**	0,9114

Tabela 4 - Estimativa dos parâmetros do modelo ajustado SARIMAX(1,0,2)(0,1,1)₇.

Parâmetro	Estimativa	Erro padrão	Valor de z	Valor p
Veloc	2,4419	0,9544	2,5586	0,0105
precip	-2,4995	0,5955	-4,1971	<0,0001
ϕ_1	0,7168	0,0763	9,3974	<0,0001
θ_1	-0,3444	0,0823	-4,1831	<0,0001
θ_2	-0,1261	0,0418	-3,0155	0,0026
Θ_1	-0,9058	0,0124	-72,8414	<0,0001

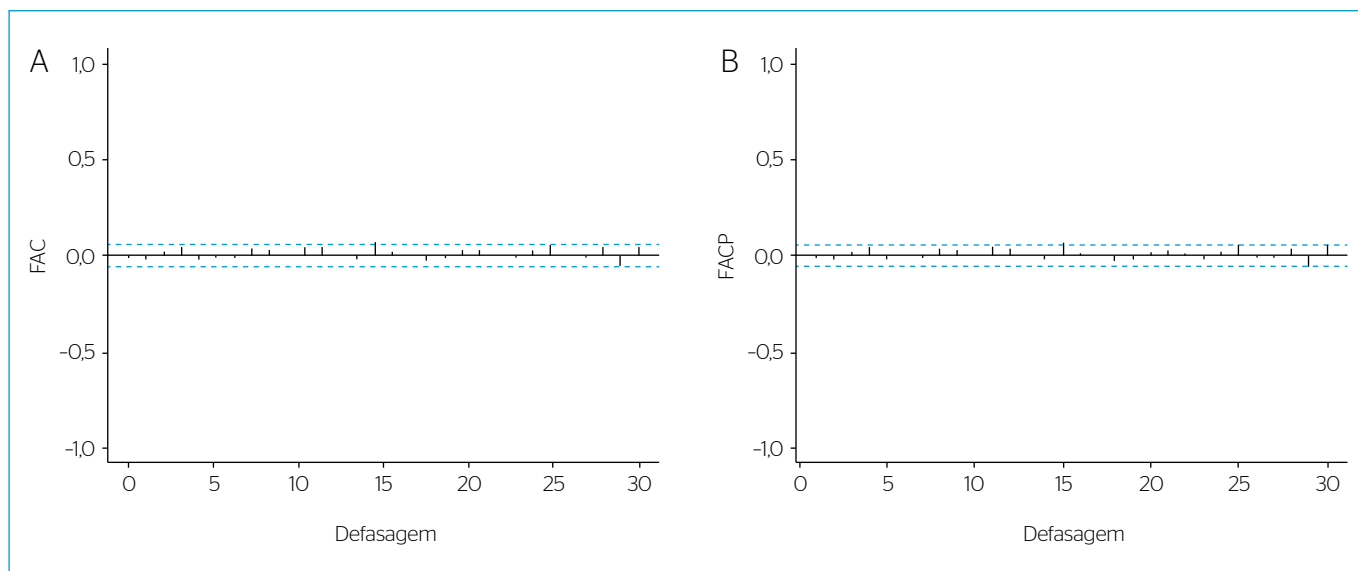


Figura 5 - Funções amostrais de autocorrelação (A) e autocorrelação parcial (B) dos resíduos do modelo SARIMAX(1,0,2)(0,1,1)₇.

um elemento significativo para explicar as concentrações de MP_{10} na RGV. Entretanto, no trabalho de Lyra, Oda-Souza e Viola (2011) a precipitação foi estatisticamente significativa, similar ao observado para o modelo SARIMAX ajustado.

O sinal positivo do coeficiente estimado da variável VV indica a existência de uma relação direta entre a concentração de MP_{10} e essa variável, demonstrando que o aumento da VV eleva a concentração de MP_{10} . Segundo o Relatório da Qualidade do Ar da Grande Vitória (IEMA, 2013), 69,3% das emissões de MP_{10} para a atmosfera da RGV estão ligadas à ressuspensão causada pelo vento e tráfego veicular, corroborando o resultado encontrado, uma vez que o aumento da VV tende a elevar a ressuspensão do solo. Já a PP apresentou relação negativa com a concentração de MP_{10} . O coeficiente negativo no modelo se deve ao processo de deposição úmida e atenuação da ressuspensão do solo (LYRA; ODA-SOUZA; VIOLA, 2011). Durante eventos de precipitação ocorre diminuição do MP_{10} , pois o solo úmido dificulta a ressuspensão do particulado no solo (VARDOULAKIS & KASSOMENOS, 2008).

Estudo de previsões

Para corroborar a utilização do modelo SARIMAX e para fins de comparação entre modelos, realizou-se estimativas para os modelos ARMA, SARMA, ARMAX e SARMAX e verificou-se qual é o melhor método para fazer previsões das concentrações de MP_{10} . Para isso, o conjunto de dados sob estudo foi dividido em dois subconjuntos; adotou-se o período de 01/01/2012 a 14/04/2015 para fazer as estimativas e reservou-se os dados compreendidos entre 15/04/2015 a 30/04/2015 para se fazer as previsões, ou seja, para que o cálculo das estatísticas de avaliação das previsões dê um passo à frente.

Na Tabela 5 são apresentados os resultados estimados para REQM e EAM para os modelos ajustados. Observa-se que o modelo SARIMAX apresentou um aumento na precisão das previsões nos horizontes (h) calculados. Os resultados da comparação das medidas de qualidade mostram que para $h=0$ e $h=1$ o modelo SARIMAX foi o que obteve melhor resultado, sendo o mais adequado para realizar as previsões.

Tabela 5 – Medidas de avaliação da qualidade de previsão obtidas a partir dos modelos ajustados.

Horizonte	Medida	Modelos				
		ARMA	SARMA	ARMAX	SARMAX	SARIMAX
h=0	REQM	13,2885	13,2540	13,0881	13,0504	12,8507
	EAM	9,6660	9,6456	9,4331	9,4220	9,3400
h=1	REQM	13,3421	13,3047	13,1487	13,1089	12,9212
	EAM	9,6610	9,6386	9,4356	9,4223	9,3564

REQM: raiz do erro quadrático médio; EAM: erro absoluto médio.

Avaliação do modelo SARIMAX para estimativa da qualidade do ar

Conforme apresentado na introdução, o IEMA classifica a qualidade do ar da RGV como boa, para concentrações de MP_{10} entre 0 e $45 \mu\text{g.m}^{-3}$, regular entre 46 e $120 \mu\text{g.m}^{-3}$, e inadequada entre 121 e $250 \mu\text{g.m}^{-3}$ (IEMA, 2013). A qualidade do ar pode, ainda, ser classificada como má, péssima ou crítica. Como no período de estudo a qualidade do ar, monitorada na estação da Enseada do Suá, foi classificada como regular em 450 ocasiões, avaliou-se apenas a eficiência do modelo em prever eventos classificados como regular. Para isso, baseando-se na metodologia de Ryan (1995) e Liu e Johnson (2003), calculou-se as seguintes estatísticas de avaliação: a probabilidade de detecção (POD), que mensura a probabilidade do modelo prever corretamente as concentrações maiores ou iguais a $46 \mu\text{g.m}^{-3}$ (qualidade do ar regular), quando realmente esse valor ocorreu; a razão de alarme falso (FAR), que mede a tendência da predição superestimar o valor observado da concentração de MP_{10} ; o escore de ameaças (EA), que representa a relação entre as predições corretas de eventos pelo total de eventos regular observado; e a taxa de perda (MISS), que mensura a taxa a qual os eventos de concentrações maiores ou iguais a $46 \mu\text{g.m}^{-3}$ observados não são preditos.

Para fins de comparação entre o pior modelo (ARMA) e o melhor modelo (SARIMAX) para fazer a predição de eventos regulares de qualidade do ar, a Tabela 6 contém a descrição e os resultados das estatísticas de avaliação. Observa-se que, das vezes que foram observados eventos de qualidade regular, o modelo SARIMAX estimou corretamente 64,39% desses eventos. Além disso, apresentou uma FAR de 0,3792, o que significa que em torno de 37,00% das predições do modelo os eventos de qualidade do ar foram classificados como regular, quando o observado foi boa; enquanto no modelo ARMA esse resultado foi em torno de 38,00%. O maior valor de EA mostra que o modelo SARIMAX foi superior ao modelo ARMA na predição correta de eventos de qualidade do ar regular. Além disso, a taxa

Tabela 6 – Descrição e resultados das estatísticas de avaliação da eficiência do modelo em prever eventos classificados como regular.

Estatísticas/equação ¹	Modelos	
	ARMA	SARIMAX
$POD = A/(A + B)$	0,6100	0,6439
$FAR = C/(C + A)$	0,3888	0,3792
$EA = A/(A + B + C)$	0,4395	0,4621
$MISS = 1 - POD$	0,3900	0,3561

¹A: número de eventos observados e preditos; B: número de eventos observados mas não preditos; C: número de eventos estimado mas não observados; POD: probabilidade de detecção; FAR: razão de alarme falso; EA: escore de ameaças; MISS: taxa de perda.

(MISS) referente à ocorrência de eventos que não foram detectados foi menor para o modelo SARIMAX, corroborando, novamente, que o desempenho estatístico do modelo SARIMAX foi superior ao modelo ARMA, no que diz respeito à predição de eventos de qualidades do ar regular.

De maneira geral, os resultados obtidos com a modelagem dos dados de concentrações de MP_{10} com o modelo SARIMAX, tanto no ajuste quanto na previsão e predição de eventos de qualidade do ar, podem ser considerados bons, uma vez que tais resultados foram satisfatórios em termos estatísticos.

CONCLUSÕES

Este trabalho objetivou modelar e prever a concentração média diária de MP_{10} na RGV, Espírito Santo, Brasil, utilizando o modelo

SARIMAX, no período de 01/01/2012 a 30/04/2015. Baseando-se em indicadores de desempenho de modelagem, verificou-se que o modelo SARIMAX (1,0,2) (0,1,1)₇ é o mais acurado, entre os estudados, para fazer predições e previsões da qualidade ar da RGV. Em comparação com os modelos ARMA, o desempenho estatístico do modelo SARIMAX foi superior, no que diz respeito à predição de eventos de qualidade do ar regular.

Entre as variáveis meteorológicas avaliadas, a VV e a PP foram significativas e melhoraram o ajuste do modelo. A UR e a T não foram estatisticamente significativas para explicar a variação das concentrações de MP_{10} . Em termos de previsão da qualidade do ar, os modelos de séries temporais mostraram resultados satisfatórios.

Como parte de um estudo posterior, sugere-se a utilização de um modelo de heterocedasticidade condicional autorregressivo generalizado (GARCH) juntamente aos modelos da classe SARIMAX.

REFERÊNCIAS

- AGIRRE-BASURKO, E.; IBARRA-BERASTEGUI, G.; MADARIAGA, I. (2006) Regression and multilayer perceptron-based models to forecast hourly O_3 and NO_2 levels in the Bilbao area. *Environmental Modelling and Software*, v. 21, p. 430-446. <http://dx.doi.org/10.1016/j.envsoft.2004.07>
- AKAIKE, H. (1973) Information theory and an extension of the maximum likelihood principle. In: PETROV, B.N.; CSAKI, F. (Eds.). *Proceedings of the Second International Symposium on Information Theory*. Budapeste: Akademiai Kiado. p. 267-281.
- _____. (1978) A Bayesian analysis of the minimum AIC procedure. *Annals of the Institute of Statistical Mathematics*, v. 30, n. 1, p. 9-14.
- ALMEIDA, M.A.I. (2006) *Modelo Aditivo Generalizado (MAG) no estudo da relação entre o número de atendimentos hospitalares por causas respiratórias e a qualidade do ar*. Dissertação (Mestrado em Engenharia Ambiental) - Programa de Pós-Graduação em Engenharia Ambiental, Centro Tecnológico, Universidade Federal do Espírito Santo, Vitória.
- BARBOSA, E.C.; SÁFADI, T.; NASCIMENTO, M.; NASCIMENTO, A.C.C.; SILVA, C.H.O.; MANULI, R.C. (2015) Metodologia Box & Jenkins para previsão de temperatura média mensal da cidade de Bauru (SP). *Revista Brasileira de Biometria*, v. 33, n. 1, p. 104-117.
- BAYER, F.M.; SOUZA, A.M. (2010) Wavelets e modelos tradicionais de previsão: um estudo comparativo. *Revista Brasileira de Biometria*, v. 28, n. 2, p. 40-61.
- BERA, A.K.; JARQUE, C.M. (1981) Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Economics Letters*, v. 7, n. 4, p. 313-318. [https://doi.org/10.1016/0165-1765\(81\)90035-5](https://doi.org/10.1016/0165-1765(81)90035-5)
- BOX, G.; JENKINS, G. (1970) *Time Series Analysis: Forecasting and Control*. São Francisco: Holden-Day.
- BOX, G.; PIERCE, D.A. (1970) 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. DOI: 10.2307/2284333
- BRAGA, A.L.F.; PEREIRA, L.A.A.; PROCÓPIO, M.; ANDRÉ, P.A.D.; SALDIVA, P.H.D.N. (2007) Associação entre poluição atmosférica e doenças respiratórias e cardiovasculares na cidade de Itabira, Minas Gerais, Brasil. *Cadernos de Saúde Pública*, v. 23, supl. 4, p. S570-S578. <http://dx.doi.org/10.1590/S0102-311X2007001600017>
- CHALOULAKOU, A.; GRIVAS, G.; SPYRELLIS, N. (2003) Neural network and multiple regression models for PM_{10} prediction in Athens: A comparative assessment. *Journal of the Air & Waste Management Association*, v. 53, n. 10, p. 1183-1190. <https://doi.org/10.1080/10473289.2003.10466276>
- CONAMA - Conselho Nacional de Meio Ambiente. (1990) Resolução nº 03, de 28 de junho de 1990. Dispõe sobre padrões de qualidade do ar, previstos no PRONAR. *Diário Oficial da República Federativa do Brasil*, Brasília, Seção 1, p. 15937-15939.

- CURTIS, L.; REA, W.; SMITH-WILLIS, P.; FENYVES, E.; PAN, Y. (2006) Adverse health effects of outdoor air pollutants. *Environment International*, v. 32, n. 6, p. 815-830. <https://doi.org/10.1016/j.envint.2006.03.012>
- DICKEY, D.A.; FULLER, W.A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, v. 49, n. 4, p. 1057-1072. DOI: 10.2307/1912517
- GOMES, K.S. (2009) *Modelagem INAR(p) para a previsão de índices de qualidade do ar*. 71 f. Dissertação (Mestrado em Engenharia Ambiental) - Programa de Pós-Graduação em Engenharia Ambiental, Centro Tecnológico, Universidade Federal do Espírito Santo, Vitória.
- GOOGLE EARTH. (2014) *Informações geográficas*. Disponível em: <<http://www.google.com.br/intl/pt-PT/earth/>>. Acesso em: 10 jun. 2016.
- GOUVEIA, N.; MENDONÇA, G.A.S.; LEON, A.P.; CORREIA, J.E.M.; JUNGER, W.L.; FREITAS, C.U.; DAUMAS, R.P.; MARTINS, L.C.; GIUSSEPE, L.; CONCEIÇÃO, G.M.S.; MANERICH, A.; CUNHA-CRUZ, J. (2003) Poluição do ar e efeitos na saúde nas populações de duas grandes metrópoles brasileiras. *Epidemiologia e Serviços de Saúde*, v. 12, p. 29-40. <http://dx.doi.org/10.5123/S1679-49742003000100004>
- GOYAL, P.; CHAN, A.T.; JAISWAL, N. (2006) Statistical models for the prediction of respirable suspended particulate matter in urban cities. *Atmospheric Environment*, v. 40, n. 11, p. 2068-2077. <https://doi.org/10.1016/j.atmosenv.2005.11.041>
- GRIPA, W.R.; REISEN, V.A.; FAJARDO, F.A.; REIS JUNIOR, N.C. (2012) Análise de predição e previsão das concentrações de material particulado inalável (PM₁₀) na cidade de Carapina, ES. *Revista Brasileira de Estatística*, v. 73, n. 237, p. 37-57.
- HOLGATE, S.T.; SAMET, J.M.; KOREN, H.S.; MAYNARD, R.L. (1999) *Air Pollution and Health*. San Diego: Academic Press.
- IBGE - Instituto Brasileiro de Geografia e Estatística. (2010) *Censo 2010*. Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/resultados_dou/ES2010.pdf>. Acesso em: 01 dez. 2015.
- IEMA - Instituto Estadual de Meio Ambiente e Recursos Hídricos do Estado do Espírito Santo. (2013) *Relatório da qualidade do ar da Região da Grande Vitória*. Vitória: IEMA. Disponível em: <https://iema.es.gov.br/Media/iema/Downloads/RAMQAR/Relat%C3%B3rio_Anual_de_Qualidade_do_Ar_2013.pdf>. Acesso em: 10 jun. 2016.
- JUNGER, W.L.; LEON, A.P. (2015) Imputation of missing data in time series for air pollutants. *Atmospheric Environment*, v. 102, p. 96-104. <https://doi.org/10.1016/j.atmosenv.2014.11.049>
- LATORRE, M.R.DO.; CARDOSO, M.R.A. (2001) Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. *Revista Brasileira de Epidemiologia*, v. 4, n. 3, p. 145-152. <http://dx.doi.org/10.1590/S1415-790X2001000300002>
- LIU, P.W.G.; JOHNSON, R. (2003) Forecasting peak daily ozone levels: part 2. A regression with time series errors model having a principal component trigger to forecast 1999 and 2002 ozone levels. *Journal of the Air & Waste Management Association*, v. 53, n. 12, p. 1472-1489. <https://doi.org/10.1080/10473289.2003.10466321>
- LJUNG, G.M.; BOX, G.E.P. (1978) On a measure of lack of fit in time series models. *Biometrika*, v. 65, n. 2, p. 297-303.
- LYRA, G.B.; ODA-SOUZA, M.; VIOLA, D.N. (2011) Modelos lineares aplicados à estimativa da concentração do material particulado (PM₁₀) na cidade do Rio de Janeiro, RJ. *Revista Brasileira de Meteorologia*, v. 26, n. 3, p. 392-400. DOI: 10.1590/s0102-77862011000300006
- MARTINS, L.C.; LATORRE, M.R.DO.; CARDOSO, M.R.A.; GONÇALVES, F.L.T.; SALDIVA, P.H.N.; BRAGA, A.L.F. (2002) Poluição atmosférica e atendimentos por pneumonia e gripe em São Paulo, Brasil. *Revista de Saúde Pública*, v. 36, n. 1, p. 88-94. <http://dx.doi.org/10.1590/S0034-89102002000100014>
- MONTE, E.Z.; ALBUQUERQUE, T.T.A.; REISEN, V.A. (2015) 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. <http://dx.doi.org/10.1590/0102-778620140006>
- MORETTIN, P.A.; TOLOI, C.M.C. (2006) *Análise de séries temporais*. 2. ed. São Paulo: Blucher.
- MOURA, F.A.; MONTINI, A.A.; CASTRO, J.B.B. (2011) Modelagem do Consumo de Energia Elétrica Residencial no Brasil Através de Modelos ARMAX. In: SEMINÁRIO DE ADMINISTRAÇÃO, 14., 2011. *Anais...*
- NASCIMENTO, L.F.C.; PEREIRA, L.A.A.; BRAGA, A.L.F.; MÓDOLO, M.C.C.; CARVALHO JR., J.A.C. (2006) Efeitos da poluição atmosférica na saúde infantil em São José dos Campos, SP. *Revista de Saúde Pública*, v. 40, p. 77-82. <http://dx.doi.org/10.1590/S0034-89102006000100013>
- OSTRO, B.; SANCHEZ, J.M.; ARANDA, C.; ESKELAND, G.S. (1996) Air pollution and mortality: results from a study of Santiago, Chile. *Journal of Exposure Analysis and Environmental Epidemiology*, v. 6, p. 97-114.
- R CORE TEAM (2016). *R: A language and environment for statistical computing*. Viena: R Foundation for Statistical Computing. Disponível em: <<http://www.R-project.org/>>.
- REISEN, V.A.; SARNAGLIA, A.J.Q.; REIS JUNIOR, N.C.; LÉVY-LEDUC, C.; SANTOS, J.M. (2014) Modeling and forecasting daily average PM₁₀ concentrations by a seasonal long-memory model with volatility. *Environmental Modelling & Software*, v. 51, p. 286-295. <https://doi.org/10.1016/j.envsoft.2013.09.027>
- RYAN, W.F. (1995) Forecasting severe ozone episodes in the Baltimore metropolitan area. *Atmospheric Environment*, v. 29, n. 17, p. 2387-2398. [https://doi.org/10.1016/1352-2310\(94\)00302-2](https://doi.org/10.1016/1352-2310(94)00302-2)
- SCHWARZ, G. (1978) Estimating the dimensional of a model. *The Annals of Statistics*, Hayward, v. 6, n. 2, p. 461-464.

SEINFELD, J.H.; PANDIS, S.N. (2006) *Atmospheric Chemistry and Physics*. 2. ed. Nova Jersey: John Wiley & Sons.

SHAPIRO, S.S.; WILK, M.B. (1965) An analysis of variance test for normality (complete samples). *Biometrika*, v. 52, n. 3-4, p. 591-611. DOI: 10.2307/2333709

SOUZA, J.B.; REISEN, V.A.; SANTOS, J.M.; FRANCO, G.C. (2014). Componentes principais e modelagem linear generalizada na associação entre atendimento hospitalar e poluição do ar. *Revista de Saúde Pública*, v. 48, n. 3, p. 451-458. DOI: 10.1590/S0034-8910.2014048005078

VARDOULAKIS, S.; KASSOMENOS, P. (2008) Sources and factors affecting PM_{10} levels in two European cities: implications for local air quality management. *Atmospheric Environment*, v. 42, n. 17, p. 3949-3963. <https://doi.org/10.1016/j.atmosenv.2006.12.021>

WEI, W. (2006) *Time Series Analysis: Univariate and Multivariate Methods*. Nova York: Addison Wesley.

WORLD HEALTH ORGANIZATION (WHO). (2005) *Air quality guidelines global update 2005*. Report on a working group meeting, Bonn/Alemanha: WHO. Disponível em: <http://www.euro.who.int/_data/assets/pdf_file/0008/147851/E87950.pdf>. Acesso em: 20 jan. 2016.

