

# Uso do modelo SARIMA na predição das concentrações de Ozônio ( $O_3$ ) na Região da Grande Vitória E.S.

Pezzin, G. C.<sup>1\*</sup>; Barbosa, G. C.<sup>2</sup>

1 Graduando, Faculdade do Centro Leste, Serra, ES, Brasil.

2 Professor, Faculdade do Centro Leste, Serra, ES, Brasil.

\* e-mail: [georgecp@ucl.br](mailto:georgecp@ucl.br) ; [geovanebarbosa@ucl.br](mailto:geovanebarbosa@ucl.br)

## Resumo

Este estudo teve como objetivo modelar as concentrações de Ozônio ( $O_3$ ) para o município de Cariacica, Região da Grande Vitória, Brasil por meio dos modelos de séries temporais. Os dados de concentração do poluente foram obtidos por uma das redes de monitoramento da qualidade do ar na Região da Grande Vitória no período de 2008 a 2010. Inicialmente ajustou-se um modelo SARIMA  $(1,1,2) \times (1,1,2)_7$  e comparou-se os valores preditos encontrados para dezembro de 2010 com as concentrações obtidas no mesmo período. Os resultados mostraram que o modelo Sazonal ARIMA prevê as concentrações de forma muito eficaz, sendo uma importante ferramenta para previsão da qualidade do ar para a região.

## Abstract

This study aimed to model the concentrations of ozone ( $O_3$ ) to the city of Cariacica, Region of Greater Vitória, Brazil through the time series models. The data of the pollutant concentration were obtained by one of the air quality in the Greater Victoria Region monitoring networks from 2008 to 2010. Initially set a model SARIMA  $(1,1,2) \times (1,1,2)_7$  and compared the predicted values found for December 2010 with the concentrations obtained in the same period. The results showed that the Seasonal ARIMA model predicts the concentrations very effectively, being an important tool for air quality forecast for the region.

Keywords (Palavras chaves): Air pollution (Poluição do ar), Time series (Séries temporais), Environment (Meio ambiente), VI Scientific meeting of applied physics (VI Encontro científico de física aplicada).

## 1. Introdução

A poluição do ar nos grandes centros urbanos tem registrado níveis cada vez menos seguros para a sua população, sendo tais efeitos nem sempre visíveis devido ao tamanho das partículas com inúmeras fontes de emissão tais como os meios de transporte a combustível fóssil e a atividade industrial que contribuem com uma grande parcela das emissões.

A qualidade do ar urbano tem causado sérios problemas às condições de vida das pessoas, principalmente crianças e idosos, dois grupos susceptíveis aos efeitos da poluição atmosférica. Os efeitos adversos dos poluentes atmosféricos na saúde humana têm sido uma preocupação dos órgãos

regulamentadores e de saúde pública a longo dos últimos 50 anos. Os efeitos nocivos da poluição têm sido observados no aumento da mortalidade e da morbidade por doenças respiratórias e cardiovasculares [1-5].

As características de dados de séries temporais ambientais exigem métodos estatísticos complexos e suficientemente sensíveis, no caso de estudos sobre a poluição do ar e saúde existem estruturas de tendências e flutuações sazonais, que interferem na dinâmica das variáveis observadas tais como uso de veículos, dias da semana, feriados, número de indústrias, entre outras. Ferramentas estatísticas têm sido exaustivamente utilizadas para a predição e

previsão de concentração de poluentes atmosféricos, principalmente os modelos SARIMA [6].

Neste artigo estudou-se o desempenho do modelo sazonal ARIMA para modelar as concentrações do poluente atmosférico conhecido como Ozônio ( $O_3$ ) no município de Cariacica, região da Grande Vitória, no período de janeiro de 2008 a novembro de 2010. Ajustou-se um modelo SARIMA para previsões diárias das concentrações deste poluente.

## 2. Materiais e Métodos

### 2.1 Descrição dos dados

As séries diárias das concentrações do poluente Ozônio, para o período de janeiro de 2008 a novembro de 2010, foram obtidos junto ao Instituto Estadual de Meio Ambiente e Recursos Hídricos (IEMA), que atualmente possui 8 estações automáticas de monitoramento da qualidade do ar na Região da Grande Vitória, sendo uma delas na região de Cariacica/ES, região de abrangência do estudo. Para esse poluente foi calculada a média aritmética diária sendo consideradas representativas segundo Braga et al [9]. Todas as análises foram realizadas com o Software R 3.1.3 [8] e foi adotado um nível de 5% de significância.

### 2.2 Modelo SARIMA

Os processos encontrados na prática, além de raramente serem estacionários, apresentam em suas estruturas componentes sazonais. Como muitas séries temporais, os dados de poluição do ar apresentam um componente sazonal em sua estrutura, tornando-se fundamental a alocação dessa componente sazonal na estimativa dos modelos ARIMA(p,d,q). O modelo denominado SARIMA passa a ser considerado pelas componentes sazonais SAR(p) e média móvel SMA(Q) de definido como SARIMA(p,d,q) x (P,D,Q), onde D é o termo da diferenciação sazonal do modelo. A definição matemática do modelo é:

$$\phi(B)\Phi(B^s)\Delta^d\Delta_s^D X_t = \theta(B)\Theta(B^s)\varepsilon_t \quad (1)$$

Onde:

$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  é o operador autorregressivo estacionário de ordem p;

$\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$  é o operador de médias móveis invertível de ordem q;

$\Phi(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{sP}$  é o operador autorregressivo sazonal de ordem P, estacionário;

$\Theta(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_Q B^{sQ}$  é o operador de médias móveis sazonal de ordem Q, invertível;

$\Delta = (1 - B)$  é o operador diferença;

$\Delta^d = (1 - B)^d$ , d indicando o número de diferenças;

$\Delta_s = (1 - B^s)^s$  é o operador diferença sazonal;

$\Delta_s^D = (1 - B^s)^D$ , D indicando o número de “diferenças sazonais”;

$\varepsilon_t$  representa o ruído branco com média zero.

## 3. Resultados

Para compreender melhor o comportamento ao longo do tempo estudado dos dados, foi gerado o gráfico da série temporal das concentrações de ozônio no período de jan/2008 a nov/2010 conforme figura (1). Observa-se que a distribuição da série contém uma grande variabilidade (picos) e sazonalidade evidenciada ao longo do período estudado com média e desvio padrão respectivamente iguais a 26,64 e 9.34. Outro dado importante é de que nenhum valor coletado ultrapassou o limite crítico de 160  $\mu\text{g}/\text{m}^3$  estabelecido pelo CONAMA 99/03.

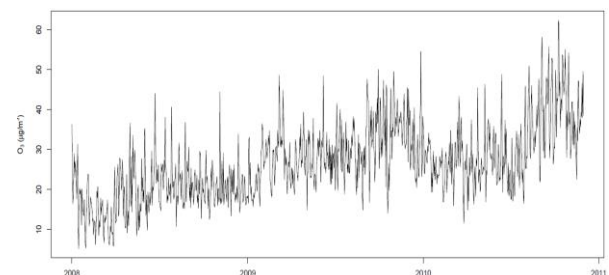


FIGURA 1 – Série temporal estacionária das concentrações diárias de ozônio ( $O_3$ ) na cidade de Cariacica/ES, no período de Jan/2008 a Nov/2010

Para a identificação do modelo foram utilizadas as funções de Autocorrelação (ACF) e Autocorrelação Parcial (PACF).

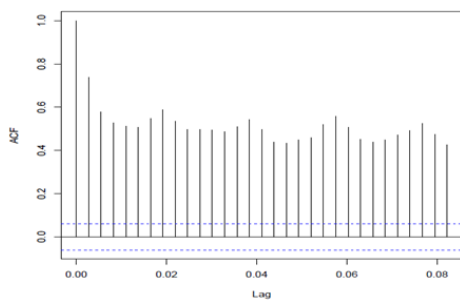


FIGURA 2 – Gráfico das correlações amostrais (ACF) das concentrações de ozônio no período de Jan/2008 a Nov/2010.

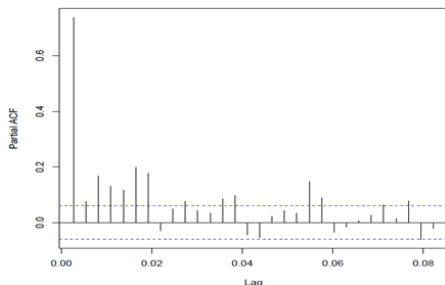


FIGURA 3 – Gráfico das correlações amostrais parciais (PACF) das concentrações de ozônio no período de Jan/2008 a Nov/2010.

Analisando a figura (2) e (3), nota-se que existe sazonalidade de período 7, possível ainda observar partes autorregressivas e partes com médias móveis não sazonais. Portanto, o modelo escolhido para ajuste desta série temporal foi o SARIMA(1,1,2)x(1,1,2)<sub>7</sub>, que se mostrou o melhor modelo.

Assim como comprovação da assertividade do modelo, na figura 4, é exposto os dados originais e seu modelo ajustado para uma melhor visualização.

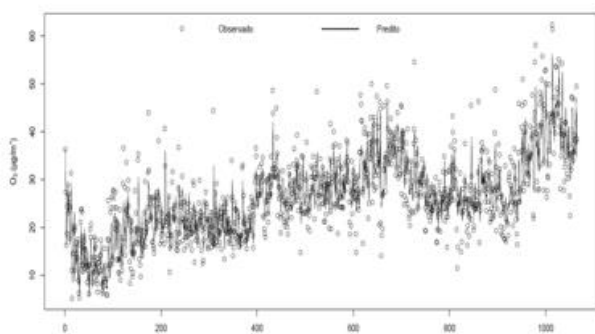


FIGURA 4 – Valores Observados versus ajuste do modelo SARIMA (1,1,2) x (1,1,2)<sub>7</sub> para concentrações de ozônio.

Foram refeitos os gráficos de ACF e PACF (figuras 5 e 6) e assim podemos constatar que em grande maioria, os valores permanecem dentro dos limites de controle. O Ajuste do modelo mostra uma boa captura da

variabilidade, bem como também a inclusão da tendência e sazonalidade do modelo.

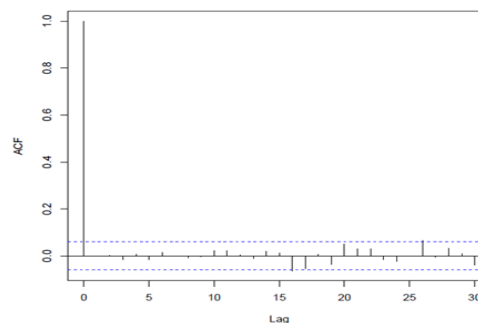


FIGURA 5 – ACF residuais ao nível de 5%.

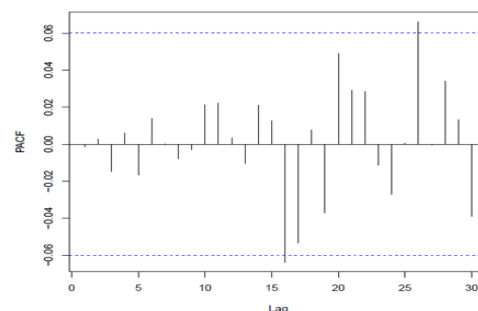


FIGURA 6 – PACF residuais ao nível de 5%.

Antes de se poder utilizar o modelo em previsões, deve-se testar a sua adequação. Em termos práticos, um modelo é considerado adequado se os resíduos não puderem ser usados para melhorar a previsão, apresentam um comportamento aleatório. A análise geral à adequação do modelo, é feita pelo teste do qui-quadrado baseado na estatística Q de Ljung-Box. Este teste analisa as autocorrelações dos resíduos como um grupo. Se o valor-*p* associado à estatística Q é pequeno (por exemplo, valor-*p* < 0,05), o modelo é considerado inadequado.

Realizando o teste Ljung-Box no *software* R foi encontrado um valor-*p* igual 0,9998. Sendo este valor-*p* maior que 0,05 pode-se dizer que existe um nível de significância de 5% ( $H_0$ : O ruído é branco) não rejeita-se a hipótese nula, portanto o ruído é branco [9].

Os resíduos do modelo estão completamente aleatórios e independentes, mostrando a boa performance do modelo propostos. Podemos concluir que o modelo é adequado, e que não existem problemas de autocorrelação residual. Previsões são extrapolações obtidas através das funções de previsão, para além do

período no qual o modelo foi estimado. A previsão abaixo estabelece os valores futuros da série fazendo uma relação dos valores observados e previstos. A previsão foi realizada para os próximos 15 dias após final da série original, conforme figura (7).

Na figura (7), também podemos verificar o comportamento da previsão e seus limites, compostos pelos seus respectivos erros padrões. Para finalizar, quando separamos os últimos 15 dias dos valores observados e comparamos com o modelo previsto pelo SARIMA  $(1,1,2) \times (1,1,2)_7$  observamos que o modelo acompanha a série original de forma satisfatória.

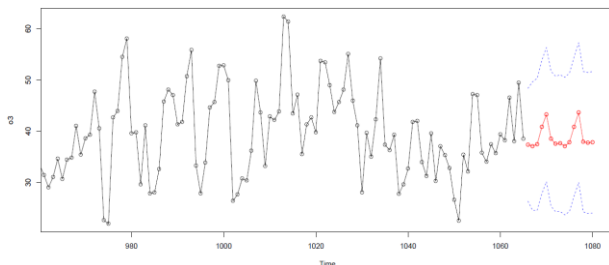


FIGURA 7 - Previsão da concentração de ozônio para 15 dias com seus respectivos limites de erros padrões .

## 4. Conclusão

Foi demonstrado como os modelos de séries temporais podem ser bastante úteis na previsão das concentrações de ozônio na região de Cariacica. O modelo SARIMA  $(1,1,2) \times (1,1,2)_7$  mostrou-se o mais adequado para previsão das concentrações do poluente em questão.

Destaca-se que os resultados obtidos neste estudo fornecem um bom indicativo da qualidade do ar na região estudada.

## 5. Agradecimentos

À Faculdade do Centro Leste (UCL).

## 6. Referências

- [1] SCHWARTZ, J. Particulate air pollution and daily mortality: a synthesis. *Public Health Revist* 19, p.39-60, 1992.
- [2] SCHWARTZ, J. Air pollution and hospital admissions for respiratory disease. *Epidemiology* 7, p.20-28, 1995.
- [3] BURNETT, R. T.; BROOK, J. R.; YUNG, W. T.; DALES, R. E.; KREWSKI, D. Association between ozone and hospitalization for respiratory diseases in 16 Canadian cities. *Environmental Respiratory* 72, p.24-31, 1997.
- [4] ATKINSON, R. W.; BREMNER, S. A.; ANDRESON, H. R.; STRACHAN, D. P.; BLAND, J. M.; LEON, A. P. Short-term associations between emergency hospital admissions for respiratory and cardiovascular disease and outdoor air pollution in London. *Arch. Environmental Health* 54, p.398-411, 1999.
- [5] SCHWARTZ, J. Air pollution and blood markers of cardiovascular risk. *Environmental Health Perspective* (Suppl. 3), p.405-409, 2001.
- [6] BOX, G.E.P.; JENKINS, G.W. and REINSEL, G.C. *Time Series Analysis, Forecasting and Control*. Third Edition. Prentice-Hall, Englewood Cliffs, N.J., 1994.
- [7] BRAGA, A.L.F.; ZANOBETTI, A.; SCHWARTZ, J. Do respiratory epidemics confound the association between air pollution and daily deaths. *European Respiratory Journal* 16, p.723-728, 2000.
- [8] R Development Core Team; R : A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, p.409, 2009.
- [9] MORETTIN, P.A.; TOLOI, C.M.C. *Análise de séries temporais*. 2ª ed. São Paulo: Editora Edgarg Blucher, 2004.