

## PREVISÃO DA CONCENTRAÇÃO DE OZÔNIO NA REGIÃO DA GRANDE VITÓRIA, ESPÍRITO SANTO, BRASIL, UTILIZANDO O MODELO ARMAX-GARCH

EDSON ZAMBON MONTE<sup>1</sup>, TACIANA TOLEDO DE ALMEIDA ALBUQUERQUE<sup>2</sup>, VALDÉRIO  
ANSELMO REISEN<sup>3</sup>

<sup>1</sup>Universidade Federal do Espírito Santo (UFES), Departamento de Economia e Programa de Pós-Graduação em Engenharia Ambiental, Vitória, ES, Brasil.

<sup>2</sup>Universidade Federal de Minas Gerais (UFMG), Departamento de Engenharia Sanitária e Ambiental, Belo Horizonte, MG, Brasil e UFES, Programa de Pós-Graduação em Engenharia Ambiental, Vitória, ES, Brasil.

<sup>3</sup>UFES, Departamento de Estatística e Programa de Pós-Graduação em Engenharia Ambiental, Vitória, ES, Brasil.

edsonzambon@yahoo.com.br, ticianatoledo26@gmail.com, valderioanselmoreisen@gmail.com

Recebido Abril de 2014 - Aceito Janeiro de 2015

### RESUMO

O objetivo deste trabalho foi estimar e prever a concentração horária de ozônio na Região da Grande Vitória, Espírito Santo, Brasil, utilizando um modelo 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), sendo escolhidas três estações: Laranjeiras, Enseada do Suá e Cariacica. Adotou-se alguns parâmetros medidos nas estações como variáveis explicativas da concentração de ozônio, a saber: temperatura, umidade relativa, velocidade do vento e concentração de dióxido de nitrogênio. Estas foram significativas e melhoraram a estimativa do modelo ajustado. As previsões horárias para o dia 31/12/2011 revelaram-se muito próximas dos valores observados, sendo que as estimativas, em geral, seguiram a trajetória diária da concentração de ozônio. No mais, em comparação aos modelos ARMA e ARMAX, o modelo ARMAX-GARCH revelou-se mais eficaz na predição de episódios de poluição de ozônio (concentração horária superior a  $80 \mu\text{g}/\text{m}^3$ ), reduziu o número de falsos alarmes estimados e apresentou menor taxa de ocorrência de episódios não detectados.

**Palavras-chave:** ozônio; poluição do ar; séries temporais, ARMAX; GARCH.

### ABSTRACT: OZONE CONCENTRATION FORECAST IN THE REGION OF GRANDE VITÓRIA, ESPÍRITO SANTO, BRAZIL, USING THE ARMAX-GARCH MODEL

The objective of this study was to estimate and forecast the hourly ozone concentration in the Region of Grande Vitória, Espírito Santo, Brazil, using the ARMAX-GARCH model, for the period from 2011/01/01 to 2011/12/31. Data set from the State Environmental Institute (IEMA) was used. The models has run for three local stations: Laranjeiras, Enseada do Suá and Cariacica. Some parameters measured at the stations were taken as explanatory variables of ozone concentration. These variables significantly improved the model estimated. The hourly forecasts for 2011/12/31 (chosen to verify the model accuracy) were very close to the observed values and the estimated ones generally followed the path of daily ozone concentration. When compared with the ARMA and ARMAX models, ARMAX-GARCH model proved to be more effective in the prediction of ozone pollution episodes (hourly concentration higher than  $80 \mu\text{g}/\text{m}^3$ ), reduction on the number of false alarms and lowering the rate of undetected episodes.

**Keywords:** ozone; air pollution; time series, ARMAX; GARCH.

## 1. INTRODUÇÃO

A intensificação do processo de industrialização ocorrida no século XIX, aliado ao crescimento populacional, especialmente, o crescimento da população urbana em detrimento da rural, vem aumentando as preocupações relacionadas à proteção do meio ambiente. Logo, cada vez mais tem se dado atenção para os efeitos adversos que a poluição atmosférica pode causar na saúde humana, tais como: irritação dos olhos, problemas pulmonares, alergias, etc. Para Liu et al. (2013), os dois poluentes atmosféricos que mais preocupam em relação à saúde humana são o ozônio ( $O_3$ )<sup>1</sup> e o material particulado.

Quanto ao ozônio, esse é um poluente secundário formado na troposfera, por meio de reações fotoquímicas sobre os óxidos de nitrogênio e os compostos orgânicos voláteis, sendo a radiação solar um dos forçantes que contribui para sua formação. Conforme Seinfeld e Pandis (2006), a formação do ozônio depende de diversos fatores químicos e físicos, que variam no espaço e no tempo de forma não linear. Destaca-se que os hidrocarbonetos são emitidos por fontes naturais e antropogênicas, sendo que estas últimas incluem fontes móveis (automóveis) e estacionárias (usos industriais). Assim, a concentração de  $O_3$  tende a ser maior nos grandes centros urbanos, onde se concentram as grandes indústrias e o maior volume de automóveis. Além disso, como a formação de ozônio requer radiação ultravioleta, bem como a presença de precursores como óxidos de nitrogênio e compostos orgânicos voláteis, a concentração de  $O_3$  alcança o máximo durante os meses de verão (Ryan et al., 1999).

Ressalta-se que, de acordo com Moreira et al. (2008), as condições meteorológicas desempenham um papel importantíssimo na dispersão ou acumulação de poluentes. Liu e Johnson (2002) descreveram que a poluição do ar, particularmente a concentração de ozônio, é altamente correlacionada no tempo, estando associada, geralmente, a fatores como temperatura, umidade relativa, velocidade e direção do vento, dentre outros.

Cabe mencionar que, para o ozônio, por ser um poluente secundário, tornam-se mais difíceis as modelagens e as previsões a respeito de sua formação (Carvalho, 2006). No entanto, a natureza da concentração de ozônio em ambientes, principalmente urbanos, tem sido objetivo de diversos estudos estatísticos, especialmente no que tange a predição e a previsão das concentrações<sup>2</sup>. Segundo Liu et al. (2013), alguns métodos

utilizados são: análise de regressão e classificação em árvore (CART), modelos de redes neurais, modelos de séries temporais ARIMA (Box-Jenkins) e modelos de regressão.

Ryan (1995), por exemplo, realizou previsões para as altas concentrações diárias de ozônio (episódios), na região de Baltimore, Estados Unidos, utilizando diversas abordagens, dentre elas: o método CART e a análise de regressão. Jorquera et al. (1998) realizaram previsões para o nível máximo de concentração de ozônio diário, na cidade de Santiago, Chile, utilizando modelos de séries temporais (ARMAX), de redes neurais e o modelo fuzzy. Já Liu e Johnson (2002) fizeram previsões para picos diários de concentração de ozônio, em Milwaukee, Estados Unidos, no período de 1987 a 1993, por meio do modelo regressão com erros de séries temporais (RTSE), com a inclusão, dentre as variáveis exógenas, do que os autores denominaram de principal componente (PC) com gatilho. Liu e Johnson (2003) estudaram os picos diários de concentração de ozônio, para Milwaukee, Estados Unidos, no período de 1999 a 2002, utilizando o PC com gatilho na abordagem de Box-Jenkins com RTSE.

Nota-se que, independente do método econométrico de análise, alguns destes estudos, ao fazerem previsões ou predições da concentração de ozônio, têm ignorado a questão da heterocedasticidade (volatilidade)<sup>3</sup> temporal (Kumar e Ridder, 2010). Assim, esses autores estimaram um modelo de heterocedasticidade condicional autorregressivo generalizado (GARCH) associado com o método FFT-ARIMA (transformada rápida de Fourier–autorregressivo integrado de média móvel), para prever os episódios de concentração de ozônio em duas cidades europeias, Bruxelas e Londres. Os resultados revelaram que modelar a concentração de ozônio por meio do modelo GARCH, além de melhorar os intervalos de confiança das previsões de curto prazo, também proporcionou maior acurácia na probabilidade de previsão de episódios críticos de  $O_3$ . Reisen et al. (2014) modelaram a média diária de concentração de material particulado inalável ( $PM_{10}$ ), na cidade de Cariacica, Espírito Santo, Brasil, utilizando um processo integrado fracionado sazonal, com volatilidade.

Neste contexto, este trabalho objetivou estimar e prever a concentração de ozônio horária na Região da Grande Vitória (RGV), Espírito Santo, Brasil, utilizando o modelo ARMAX-GARCH, para o período 01/01/2011 a 31/12/2011. Mesmo não tendo ultrapassado os padrões primário e secundário ( $160 \mu g/m^3$ ) estabelecidos pela Resolução CONAMA 03, de 1990 (CONAMA, 1990), no período de estudo, em diversos momentos a concentração ultrapassou o valor de  $80 \mu g/m^3$  e, até

1. Vale ressaltar que os males causados pelo ozônio ocorrem na faixa de ar perto da superfície terrestre. Conforme Seinfeld e Pandis (2006), o ozônio apresenta um duplo paradoxo na atmosfera, pois, o mesmo tem um papel benéfico na estratosfera e maléfico na troposfera.

2. Em estatística, predições são realizadas dentro da amostra considerada e, previsões, fora da amostra considerada.

3. Conforme Matteson e Tsay (2011), a volatilidade pode ser entendida como o desvio padrão condicional da série. Séries com alta variabilidade ao longo do tempo tendem a apresentar esta característica. A concentração horária de ozônio é um exemplo de série volátil.

mesmo, o de  $100 \mu\text{g}/\text{m}^3$ , sendo que as maiores concentrações ocorreram no verão. Estes padrões estabelecem uma qualidade do ar regular (entre  $80$  e  $160 \mu\text{g}/\text{m}^3$ ), em termos de efeitos prejudiciais sobre a saúde, principalmente para população mais sensível, como idosos e crianças (Companhia Ambiental do Estado de São Paulo – CETESB, 2013). Lembrando que esta qualidade do ar regular é baseada no Índice de Qualidade do AR (IQA), da CETESB. Logo, esta pesquisa torna-se importante no que diz respeito, especialmente, à formulação de medidas preventivas por parte dos órgãos competentes, uma vez que a concentração de ozônio, na Grande Vitória, embora não tenha atingido níveis alarmantes, tem-se elevado nos últimos anos.

O presente artigo está estruturado da seguinte forma. Além desta introdução, a seção 2 traz uma descrição da região de estudo, as variáveis utilizadas e os modelos estatísticos adotados. Na seção 3 apresentam-se as estimativas do modelo ARMAX-GARCH, os testes de diagnóstico e as previsões para a concentração de ozônio. Por fim, as conclusões são apresentadas na seção 4.

## 2. MATERIAL E MÉTODOS

### 2.1 Região de estudo e apresentação das variáveis

Os dados deste estudo foram do tipo séries temporais, abrangendo variáveis relacionadas à poluição atmosférica (concentração de ozônio e dióxido de nitrogênio) e às condições meteorológicas (temperatura, umidade relativa e velocidade do vento), para a Região da Grande Vitória, Espírito Santo, Brasil. A RGV é composta por cinco municípios, localizando-se na costa

sul do oceano Atlântico (latitude  $20^\circ 19\text{S}$ , longitude  $40^\circ 20\text{W}$ ). O clima é tropical quente, com temperaturas médias variando entre  $24^\circ \text{C}$  e  $30^\circ \text{C}$ .

O período de análise foi de janeiro a dezembro de 2011, sendo os dados tomados de forma horária (8.760 observações) e coletados através do banco de dados do IEMA. Uma vez que alguns dados horários não estavam disponíveis, algumas observações foram inseridas utilizando o “pacote mtsdi” (*multivariate time series data imputation*) (Junger e Leon, 2012), presente no *software* R 3.0.2. Na Tabela 1 são apresentadas as variáveis utilizadas, unidades, siglas e fontes.

Vale ressaltar que, atualmente, a Região da Grande Vitória possui oito estações de monitoramento de qualidade do ar, a saber: Laranjeiras; Carapina; Jardim Camburi; Enseada do Suá; Vitória – Centro; Vila Velha – Ibes; Vila Velha – Centro; e, Cariacica (ver Figura 1). A concentração de ozônio é medida nas estações de Laranjeiras, Enseada do Suá, Vila Velha – Ibes e Cariacica. Dado o grande percentual de dados faltantes e a alta porcentagem de dados invalidados (seja pelo sistema, equipamento, usuário, etc.), na estação Vila Velha – IBES, trabalhou-se, somente, com os dados de concentração de ozônio das estações de Laranjeiras, Enseada do Suá e Cariacica (foi estimado um modelo para cada estação). Conseqüentemente, para a concentração de dióxido de nitrogênio, também se adotou os dados destas três estações.

As variáveis temperatura, umidade relativa e velocidade do vento foram utilizadas conforme descrição da Tabela 2. Cabe mencionar aqui que foram testadas outras variáveis exógenas, como radiação solar e direção do vento. Porém, estas não se revelaram significativas ou comprometeram o ajuste do modelo.

**Tabela 1:** Variáveis, unidades, siglas e fontes.

Variáveis	Unidades	Siglas	Fontes
Concentração de ozônio em Laranjeiras – Frequência horária com amostra de uma hora a três metros.	$\mu\text{g}/\text{m}^3$	O3LAR	IEMA
Concentração de ozônio em Enseada do Suá – Frequência horária com amostra de uma hora a três metros.	$\mu\text{g}/\text{m}^3$	O3SUA	IEMA
Concentração de ozônio em Cariacica – Frequência horária com amostra de uma hora a três metros.	$\mu\text{g}/\text{m}^3$	O3CAR	IEMA
Concentração de dióxido de nitrogênio em Laranjeiras – Frequência horária com amostra de uma hora a três metros.	$\mu\text{g}/\text{m}^3$	NO2LAR	IEMA
Concentração de dióxido de nitrogênio em Enseada do Suá – Frequência horária com amostra de uma hora a três metros.	$\mu\text{g}/\text{m}^3$	NO2SUA	IEMA
Concentração de dióxido de nitrogênio em Cariacica – Frequência horária com amostra de uma hora a três metros.	$\mu\text{g}/\text{m}^3$	NO2CAR	IEMA
Temperatura média – Frequência horária com amostra de uma hora a três metros.	$^\circ\text{C}$	T	IEMA
Umidade relativa – Frequência horária com amostra de uma hora a três metros.	%	UR	IEMA
Velocidade escalar média do vento – Frequência horária com amostra de 1 hora a 10 metros.	m/s	VV	IEMA



**Figura 1:** Estações de monitoramento da qualidade do ar na Grande Vitória.  
 Fonte: Google Earth (2014).

**Tabela 2:** Descrição das variáveis temperatura, umidade relativa e velocidade do vento.

Variáveis	Descrição
Temperatura	Média entre as estações de Carapina e Cariacica, únicas que possuem medições para tal variável.
Umidade Relativa	Existem medições para as estações de Carapina e Cariacica. Entretanto, como existem muitos dados faltantes para a estação de Cariacica, optou-se por trabalhar apenas com a umidade da estação de Carapina. Mensurada nas estações de Carapina, Enseada do Suá, Vila Velha – Ibes e Cariacica.
Velocidade do Vento	Dado que a estação de Carapina não atende a alguns padrões importantes para captação correta da velocidade do vento, adotou-se uma média entre as estações da Enseada do Suá, de Vila Velha – Ibes e de Cariacica.

## 2.2 Modelo ARMAX-GARCH<sup>4</sup>

Um modelo ARMA(p,q) é uma combinação de um processo autorregressivo (AR) e um processo de médias móveis (MA) e pode ser expresso por:

$$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \tag{1}$$

$$(1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p) Y_t = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t, \tag{2}$$

em que  $Y_t$  é processo estocástico a ser modelado;  $\varphi_1, \varphi_2, \dots, \varphi_p$ , coeficientes do processo autorregressivo;  $\theta_1, \theta_2, \dots, \theta_q$ , coeficientes do processo de médias móveis;  $L$ , operador de defasagem; e,  $\varepsilon_t \sim RB(0, \sigma^2)$ . Caso as raízes de  $1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p = 0$  estejam fora do círculo unitário, o processo estocástico é dito estacionário.

Cabe ressaltar que as ordens de  $p$  e  $q$  podem ser determinadas, respectivamente, pela função de autocorrelação

4. Este item está baseado em Hamilton (1994).

parcial (FACP) e pela função de autocorrelação (FAC). Adicionalmente, critérios mais objetivos podem ser utilizados para identificar as ordens corretas de  $p$  e  $q$ , a saber: Critério de Informação de Akaike (AIC); Critério de Informação de Schwarz (SC); Critério de Informação de Hannan-Quinn (HQ); e, Erro de Predição Final (FPE) (Brockwell e Davis, 2002).

Já o modelo ARMAX é uma extensão do modelo ARMA, utilizando outras séries temporais como variáveis de entrada. Tal modelo pode ser descrito como:

$$Y_t = c + \sum_{i=1}^p \varphi_i Y_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{k=1}^n \lambda_k X(t, k), \tag{3}$$

em que  $\lambda$  é o vetor de coeficientes da matriz de componentes regressivos ( $X$ ); e,  $\varepsilon_t \sim RB(0, \sigma^2)$ . Na prática, os coeficientes  $\varphi_i, \theta_j, \lambda_k$  podem ser estimados pelo método de máxima verossimilhança (MV).

Neste contexto, a estimação do modelo ARMAX (Equação 3) requer que o termo de erro  $\varepsilon_t$  seja homocedástico (ausência de volatilidade estocástica). No entanto, nas situações onde a distribuição condicional difere da distribuição

incondicional, a suposição de variância do erro constante pode não ser verificada. Dessa forma, os modelos autorregressivos de heterocedasticidade condicional (ARCH) surgiram no início da década de 1980 (ver Engle, 1982), com o intuito de modelar, temporalmente, a variância condicional. Estes modelos foram generalizados por Bollerslev (1986), dando origem aos modelos autorregressivos de heterocedasticidade condicional generalizados (GARCH).

Para representar o modelo GARCH, toma-se  $\varepsilon_t$  com um processo estocástico real em tempo discreto. Neste estudo,  $\varepsilon_t$  são as inovações do processo ARMAX (Equação 3). Engle (1982) definiu um processo ARCH onde todos os  $\varepsilon_t$  são da forma,

$$\varepsilon_t = z_t \sigma_t, \quad (4)$$

em que  $z_t$  é um processo distribuído independente e identicamente com média zero e variância unitária. Embora  $\varepsilon_t$  seja serialmente não correlacionado por definição, sua variância condicional  $\sigma_t^2$  pode ser autocorrelacionada e, portanto, mudar ao longo do tempo.

A equação da variância do GARCH(r,l) pode ser representada por (Bollerslev, 1986; Brockwell e Davis, 2002):

$$z_t = D_\theta(0,1), \quad (5)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^l \beta_j \sigma_{t-j}^2, \quad (6)$$

em que  $D_\theta(0,1)$  é a função de densidade de probabilidade das inovações ou resíduos com média zero, variância unitária; e,

$$r \geq 0, l \geq 0; \alpha_0 > 0, \alpha_i \geq 0; i = 1, 2, \dots, r; \beta_j \geq 0, j = 1, 2, \dots, l.$$

Nota-se que, caso  $l = 0$ , o processo reduz-se para um ARCH(r). No mais, para  $r = l = 0$ , a variância condicional é constante, como em um modelo ARMA, e as inovações  $\varepsilon_t$  são reduzidas a ruídos brancos.

Bollerslev (1986) demonstrou que o processo GARCH(1,1) é estacionário com  $E(\varepsilon_t) = 0$ ,  $\text{var}(\varepsilon_t) = \alpha_0 / (1 - \alpha_1 - \beta_1)$  e  $\text{cov}(\varepsilon_t, \varepsilon_s) = 0$  para  $t \neq s$  se, e somente se,  $\alpha_1 + \beta_1 < 1$  (para mais detalhes, consultar Bollerslev, 1986). A estimação do modelo GARCH pode ser feita pelo método de máxima verossimilhança (MV).

### 3. RESULTADOS E DISCUSSÕES

#### 3.1 Estimativas por ARMAX-GARCH

O primeiro passo na análise de séries temporais é verificar se as mesmas são estacionárias<sup>5</sup>. Se elas não forem estacionárias em nível deve-se realizar algum procedimento para

5. Uma série temporal (processo estocástico) é considerada estacionária quando apresentar média, variância e covariância constantes ao longo do tempo.

estacionarizá-las (em geral, aplica-se a primeira diferença nas mesmas, dado que a maioria das séries é I(1), ou seja, integradas de primeira ordem). Os resultados dos testes *Augmented Dickey-Fuller – ADF* (Dickey e Fuller, 1981), *Phillips-Perron – PP* (Phillips e Perron, 1988) e *Kwiatkowski-Phillips-Schmidt-Shin – KPSS* (Kwiatkowski et al., 1992)<sup>6</sup> revelaram que todas as séries foram estacionárias em nível.

Para determinação do modelo ARMAX ideal, para cada estação de monitoramento, utilizou-se a FAC, a FACP e os critérios de informação de AIC, de SC e HQ. Foram estimados diversos modelos, sendo que o modelo com o melhor ajuste de cada estação encontra-se na Tabela 3. Importante mencionar que também foram estimados modelos ARMA, porém, pelos critérios de seleção, os mesmos foram desconsiderados.

No mais, todos os modelos estimados, constantes na Tabela 3, demonstraram resíduos (FAC) com características semelhantes à de um processo ruído branco, ou seja, não autocorrelacionados. Entretanto, quando se observou a FAC dos resíduos ao quadrado, verificou-se, para as três estações, um grande número de valores fora do intervalo de confiança. Logo, os resíduos ao quadrado não obedeceram à suposição de ruído branco, exibindo correlação na variância, ou seja, há heterocedasticidade condicional no processo. Para confirmar tal problema, realizou-se o teste de heterocedasticidade ARCH-LM<sup>7</sup> no modelo de cada estação (Tabela 3), e verificou-se que a variância condicional dos erros é autocorrelacionada. Vale lembrar que o teste foi realizado para diversos números de defasagens (a Tabela 3 apresenta o resultado para uma defasagem, ou seja, para defasagem de uma hora), sendo que, para todos, rejeitou-se a hipótese de ausência de volatilidade.

Dessa forma, adotou-se a técnica GARCH para modelar a volatilidade (variância condicional) da concentração de ozônio. Os modelos foram estimados considerando a suposição de que os erros do ARMAX seguem distribuição normal<sup>8</sup>. Foram testados vários modelos para o número de defasagens do GARCH, sendo que o melhor modelo para cada estação encontra-se na Tabela 4. Pelo teste ARCH-LM observou-se que não se rejeitou a hipótese nula de ausência de volatilidade, em todas

6. Também foram analisados os gráficos e as funções de autocorrelação das séries.

7. O teste ARCH-LM é utilizado para verificar se os resíduos (erros estimados) apresentam ou não volatilidade. A hipótese nula do teste é de ausência de volatilidade. Caso a hipótese nula não seja rejeitada, ocorre ausência de volatilidade e o modelo está adequado.

8. No que tange ao teste de normalidade dos resíduos, o teste de Jarque-Bera rejeitou a hipótese nula de que os resíduos são normais, para todas as estações de monitoramento. Entretanto, assumiu-se, pela teoria assintótica sobre a média das distribuições de probabilidade, a suposição de que os resíduos são normais, dando seguimento à análise do modelo. Esta suposição é importante para realização dos testes de hipóteses do modelo estimado.

**Tabela 3:** Estimativas das equações das médias condicionais.

	Laranjeiras		Enseada do Suá		Cariacica	
	Coef.	Ep.	Coef.	Ep.	Coef.	Ep.
Constante	31,45461*	3,678915	54,17850*	5,273380	14,14836*	4,277938
T	0,518867*	0,086809	0,240567**	0,122042	0,714503*	0,087138
UR	-0,041923*	0,017540	-0,166264*	0,030050	-0,047402*	0,018303
VV	1,104344*	0,152707	1,225437*	0,222257	1,396750*	0,145102
NO2	-0,675974*	0,007614	-1,000944*	0,014113	-0,459971*	0,008701
AR(1)	0,899618*	0,006371	0,824099*	0,012169	0,838127*	0,010965
AR(4)	-	-	0,029425*	0,009614	-0,031447*	0,008524
AR(12)	-	-	-	-	0,034332*	0,007915
AR(14)	-	-	-	-	-0,027510*	0,007968
AR(21)	0,047190*	0,010212	-	-	0,054735*	0,010831
AR(22)	0,029695**	0,014589	-	-	0,081962*	0,012935
AR(23)	0,060843*	0,013730	0,113795*	0,012975	-	-
AR(24)	-	-	0,092272*	0,016514	0,124145*	0,013372
AR(25)	-	-	-0,045601*	0,015871	-0,094428*	0,012322
AR(26)	-0,061539*	0,007720	-0,062426*	0,012344	-	-
MA(1)	0,173674*	0,016713	0,179530*	0,017111	0,161128*	0,015967
Teste de heterocedasticidade condicional						
ARCH-LM	529,010		231,581		578,191	
P-valor	0,00000		0,00000		0,00000	

Nota: \* Significativo a 1%; \*\* Significativo a 5%; Coef.: coeficiente; e, Ep.: erro-padrão.

**Tabela 4:** Estimativas dos modelos de volatilidade condicional.

	Laranjeiras		Enseada do Suá		Cariacica	
	Coef.	Ep.	Coef.	Ep.	Coef.	Ep.
Constante	2,713916*	0,304517	2,217166*	0,366801	1,111629*	0,161478
ARCH(1)	0,289816*	0,025081	0,297134*	0,024871	0,223160*	0,017258
GARCH(1)	0,522876*	0,032864	0,315804*	0,065489	0,736404*	0,016601
GARCH(2)	-	-	0,336062*	0,053538	-	-
Teste de heterocedasticidade condicional						
ARCH-LM	0,04678		0,81524		0,19865	
P-valor	0,82880		0,36360		0,65580	

Nota: \* Significativo a 1%; Coef.: coeficiente; e, Ep.: erro-padrão.

as estações, eliminando-se o problema da heterocedasticidade condicional. Novamente, o teste foi realizado para diversos números de defasagens (a Tabela 4 apresenta o resultado para uma defasagem), sendo que, para todos, não se rejeitou a hipótese de ausência de volatilidade.

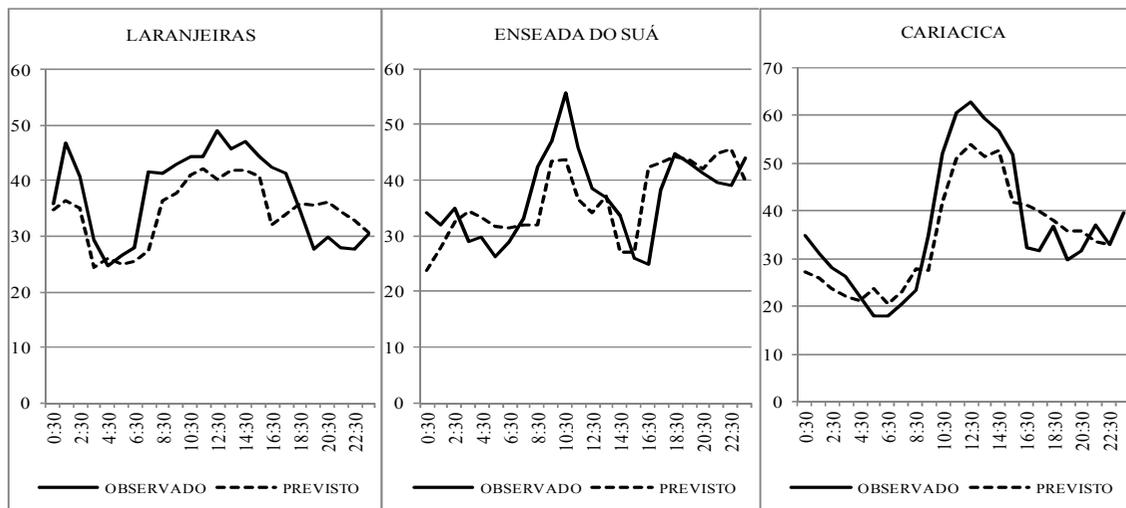
Para corroborar a utilização do modelo ARMAX-GARCH, realizou-se estimativas para os modelos ARMA, ARMAX e ARMAX-GARCH e verificou-se qual o melhor método para realizar previsões. Para isto, adotou-se o período de 0:30 horas de 01/01/2011 até as 23:30 horas de 30/12/2011 para fazer as estimativas e reservou-se as 24 horas do dia 31.12.2011 para se fazer as previsões. A Tabela 5 demonstra os valores estimados para o Erro Absoluto Médio (MAE), para a Raiz Quadrada do Erro Quadrático Médio (RMSE) e para o Erro Percentual Absoluto Médio (MAPE), para os modelos das estações de Laranjeiras, Enseada do Suá e Cariacica. Conforme observado, em todos os casos, o método ARMAX-GARCH apresentou os menores valores das estatísticas MAE, RSME e MAPE, sendo o mais adequado para realizar as previsões.

Ainda, para fins de comparação entre o pior modelo (ARMA) e o melhor modelo (ARMAX-GARCH) para realização de previsões, nas Figuras 2 e 3 encontram-se os valores observados e previstos para o dia 31/12/2011, para as estações de Laranjeiras, Enseada do Suá e Cariacica. Verifica-se a melhor adequação do modelo ARMAX-GARCH. O comportamento dos valores previstos é semelhante ao dos valores observados, inclusive no que diz respeito aos períodos de máxima concentração, que ocorrem durante as maiores temperaturas. Vale ressaltar que existem alguns pontos de subestimação e/ou superestimação, mas, as previsões seguiram a trajetória observada da concentração de ozônio durante o dia 31/12/2011.

Por fim, após estabelecer o melhor método para previsão, cabe dizer que os sinais estimados para as variáveis temperatura, umidade, velocidade do vento e dióxido de nitrogênio, estão de acordo com o esperado. Em relação à temperatura, o sinal positivo deve-se a correlação direta entre a radiação solar e a temperatura, uma vez que uma elevação da radiação solar implica em aumento da temperatura e, conseqüentemente, das reações fotoquímicas que contribuem para formação do ozônio.

**Tabela 5:** Critérios de seleção de modelo para previsão.

Estação	Modelo	MAE ( $\mu\text{g}/\text{m}^3$ )	RSME ( $\mu\text{g}/\text{m}^3$ )	MAPE (%)
Laranjeiras	ARMA	3,556287	5,053671	9,675806
	ARMAX	3,035573	3,903782	8,249631
	ARMAX-GARCH	2,999764	3,899851	8,119810
Enseada do Suá	ARMA	4,847291	5,763318	13,39339
	ARMAX	3,408568	3,894462	9,290900
	ARMAX-GARCH	3,269677	3,794599	8,861550
Cariacica	ARMA	4,335348	5,779029	12,09641
	ARMAX	3,188361	3,853879	9,525701
	ARMAX-GARCH	3,055559	3,763219	8,921103

**Figura 2:** Valores observados e previstos para a concentração horária de ozônio (em  $\mu\text{g}/\text{m}^3$ ), para o dia 31/12/2011, utilizando o modelo ARMA.

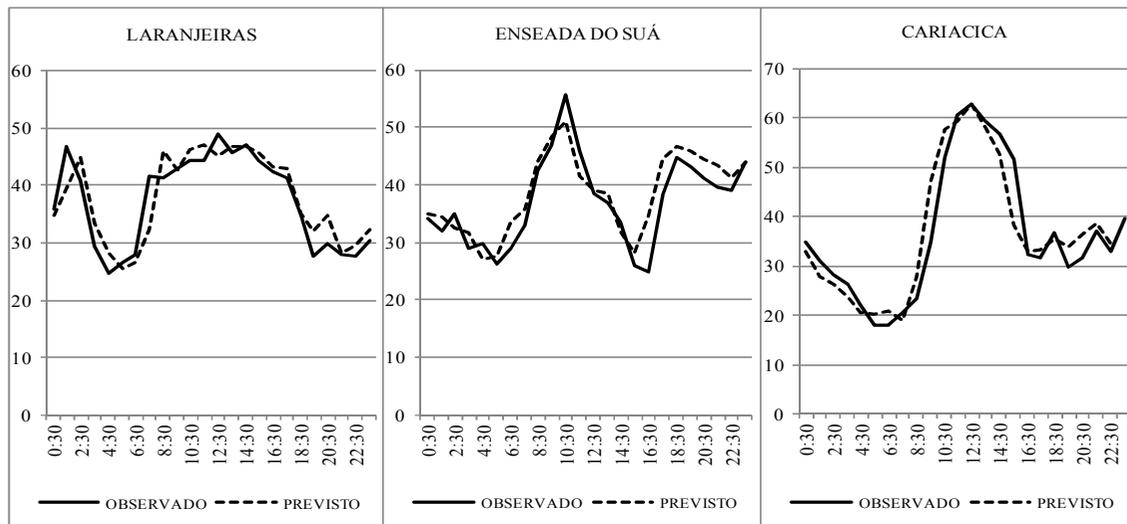
No que se refere à correlação negativa entre a concentração de ozônio e a umidade relativa, conforme Seinfeld e Pandis (2006), os diversos mecanismos cinéticos podem ocorrer segundo a presença ou não de compostos orgânicos voláteis (COV) na atmosfera. Um destes mecanismos seria redução da formação de ozônio, quando da ocorrência de alta umidade relativa. Conforme Carvalho et al. (2004), a relação negativa entre a concentração de ozônio e a umidade pode ser decorrência da radiação solar, uma vez que menores índices de umidade relativa podem estar relacionados a horários do dia com maior radiação solar e, assim, com maior formação de ozônio. Carvalho et al. (2004) salienta que “outro mecanismo cinético associado à presença dos COVs pode possibilitar a amplificação da produção do ozônio, devido à reação do radical peroxila com o óxido de nitrogênio para formar dióxido de nitrogênio, que numa segunda etapa, pode formar o poluente ozônio”.

O sinal estimado da variável velocidade do vento foi positivo, demonstrando que aumentos na velocidade do vento elevam a concentração de ozônio. Esta relação pode ser influenciada por outros fatores meteorológicos, visto que o

aumento da velocidade do vento na região pode surgir devido à atuação de sistemas sinóticos ou devido à circulação local que poderá ser gerada.

Independente da origem do aumento da velocidade do vento, existe um efeito direto entre o aumento da velocidade do vento e o aumento das concentrações de ozônio, devido à intensificação do transporte de gases precursores de outras regiões para as estações de Laranjeiras, Enseada do Suá e Cariacica, assim como o favorecimento do transporte do próprio ozônio já formado em outras regiões para as estações analisadas. Porém, em dias com atuação de sistemas sinóticos (por exemplo, frentes frias) existirá um aumento da velocidade do vento, mas a formação de ozônio ficará comprometida pela diminuição da incidência de radiação solar, consequentemente de temperatura do ar, aumento da umidade relativa do ar e nebulosidade e, a probabilidade de ocorrência de precipitação que contribuirá para a diluição dos poluentes precursores na atmosfera.

Em contrapartida, o aumento da velocidade do vento pode estar relacionado com a formação e intensificação de sistemas de circulação local, como a brisa marítima. Nestas



**Figura 3:** Valores observados e previstos para a concentração horária de ozônio (em  $\mu\text{g}/\text{m}^3$ ), para o dia 31/12/2011, utilizando o modelo ARMAX-GARCH.

situações, tem-se um aumento não somente da velocidade do vento, mas também da incidência de radiação solar, aumento da temperatura do ar, como consequência deste aumento, existirá uma intensificação do gradiente de pressão que é formado entre o mar e o continente, devido ao aquecimento diferenciado (mar-terra), surgindo com isso a circulação de brisa marítima. A existência da circulação de brisa marítima na RGV favorece o aumento da velocidade do vento local, logo o transporte de poluentes precursores é intensificado, assim como as condições meteorológicas são favoráveis para a ocorrência de reações fotoquímicas para a formação do ozônio.

Sob qualquer condição meteorológica que influencie a RGV, as horas do dia em que ocorrem as maiores concentrações de ozônio são próximas dos horários de máxima incidência de radiação solar e temperatura.

Já o dióxido de nitrogênio apresentou relação negativa com a concentração de ozônio. Por ser um poluente secundário, o ozônio é formado por reações fotoquímicas sobre os óxidos de nitrogênio e os compostos orgânicos voláteis. Assim, para formação do ozônio deve ocorrer o consumo de dióxido de nitrogênio, o que determina a relação negativa entre estes poluentes.

### 3.2 Desempenho estatístico do modelo ARMAX-GARCH

Além dos testes já realizados, baseando-se na metodologia de Ryan (1995) e Liu e Johnson (2003), calculou-se algumas estatísticas (Tabela 6), para comparar a capacidade de predição de episódios de concentração de ozônio do modelo ARMAX-GARCH, com os modelos ARMA e ARMAX. As estatísticas<sup>9</sup> mensuradas foram: a taxa de alarme falso (FAR), que mede 9. A descrição das estatísticas encontra-se no apêndice.

a tendência de a predição superestimar os episódios de concentração de ozônio; a probabilidade de detecção (POD), que mensura a probabilidade de o modelo estimar corretamente os episódios horários da concentração de ozônio, ou seja, de prever os episódios quando eles realmente ocorreram; e, a taxa de perda (MISS), que se refere à taxa na qual episódios de ozônio ocorreram, porém, não foram previstos.

Conforme descrito na introdução, pela avaliação dos dados das estações de Laranjeiras, Enseada do Suá e Cariacica, em diversas horas do ano de 2011, a concentração de ozônio esteve acima de  $80 \mu\text{g}/\text{m}^3$  e, até mesmo, superior a  $100 \mu\text{g}/\text{m}^3$ , o que pode ser considerada como uma qualidade do ar regular (entre  $80$  e  $160 \mu\text{g}/\text{m}^3$ ), em termos de efeitos prejudiciais sobre a saúde, segundo a CETESB (2013). Logo, as concentrações acima de  $80 \mu\text{g}/\text{m}^3$  foram escolhidas como episódios neste estudo. Vale ressaltar, que níveis iguais ou inferiores a  $80 \mu\text{g}/\text{m}^3$  não estão isentos de causar riscos à saúde.

Para fins de exemplificação das medidas de desempenho, verifica-se que, para a estação de Laranjeiras, o modelo ARMAX-GARCH apresentou uma FAR de 0,07, o que significa que ele prevê 7% de falsos alarmes, enquanto que para o modelo ARMAX a taxa foi de 10% e para o ARMA de 16%. No caso da POD, nota-se que a probabilidade de detectar episódios corretamente é maior para o modelo ARMAX-GARCH do que para os outros modelos. O modelo ARMAX-GARCH obteve probabilidade de 68% no que diz respeito a estimar corretamente concentrações acima de  $80 \mu\text{g}/\text{m}^3$  por hora. Além disso, a taxa (MISS) referente à ocorrência de episódios que não foram detectados foi menor para o ARMAX-GARCH.

Para as demais estações, o comportamento da FAR, da POD e da MISS foi semelhante ao encontrado para a estação

**Tabela 6:** Desempenho estatístico dos modelos ARMA, ARMAX e ARMAX-GARCH, quando a predição da concentração de ozônio excede  $80 \mu\text{g}/\text{m}^3$  por hora.

Estação	Modelo	FAR	POD	MISS
Laranjeiras	ARMA	0,16	0,38	0,62
	ARMAX	0,10	0,64	0,36
	ARMAX-GARCH	0,07	0,68	0,32
Enseada do Suá	ARMA	0,28	0,39	0,61
	ARMAX	0,21	0,49	0,51
	ARMAX-GARCH	0,18	0,53	0,47
Cariacica	ARMA	0,28	0,54	0,46
	ARMAX	0,25	0,63	0,38
	ARMAX-GARCH	0,21	0,65	0,35

Nota: FAR = Taxa de falso alarme; POD = Probabilidade de detecção; e, MISS = Taxa de perda.

de Laranjeiras, corroborando, novamente, que o desempenho estatístico do modelo ARMAX-GARCH foi superior aos modelos ARMA e ARMAX, no que diz respeito à predição de episódios de poluição de ozônio. Por fim, vale dizer que, em geral, os melhores desempenhos das estatísticas FAR, POD e MISS foram para as estações de Laranjeiras, Cariacica e Enseada do Suá, respectivamente. Isto pode ser decorrência de a estação da Enseada do Suá ter apresentado um número muito maior de picos de concentração observados, quando comparado com as outras duas estações.

#### 4. CONCLUSÕES

Este trabalho teve como objetivo estimar e prever a concentração horária de ozônio na Região da Grande Vitória, Espírito Santo, utilizando um modelo ARMAX-GARCH. Devido às limitações dos dados, as estimativas foram realizadas para as estações de Laranjeiras, Enseada do Suá e Cariacica. O modelo ARMAX apresentou volatilidade da variância dos resíduos, como já era esperado. Assim, estimou-se um ARMAX-GARCH para captar o efeito desta volatilidade. Ressalta-se que todas as séries utilizadas foram estacionárias em nível.

Neste contexto, destaca-se que, mesmo sendo um poluente de difícil modelagem estatística, em função da sua formação secundária, o modelo ARMAX-GARCH apresentou boas estimativas para as três estações de monitoramento. Os principais resultados foram:

1. As variáveis exógenas, temperatura, umidade, velocidade do vento e dióxido de nitrogênio, foram significativas e melhoram o ajuste do modelo final estimado;
2. As estatísticas MAE, RMSE e MAPE revelaram que o modelo ARMAX-GARCH é mais adequado para realização de previsões do que os modelos ARMA e ARMAX;
3. As previsões horárias do modelo ARMAX-GARCH, para o dia 31/12/2011, revelaram-se muito próximas dos

valores observados, sendo que as estimativas, em geral, seguiram a trajetória diária da concentração de ozônio;

4. Em comparação com os modelos ARMA e ARMAX, o modelo ARMAX-GARCH revelou-se mais eficaz na predição de episódios de poluição de ozônio (concentração horária superior a  $80 \mu\text{g}/\text{m}^3$ ), reduziu o número de falsos alarmes e apresentou menor taxa de ocorrência de episódios não detectados.

Ressalta-se que, estudos desta natureza são de grande importância, uma vez as preocupações quanto aos efeitos adversos que a poluição atmosférica pode causar na saúde humana tem aumento a cada dia. Sendo assim, estabelecer, por exemplo, como as condições meteorológicas afetam a concentração de ozônio e tentar prever os picos (episódios) de concentração deste poluente é fundamental, dado que pode auxiliar na tomada de decisões dos agentes públicos no que diz respeito ao combate à poluição, à prevenção de altas concentrações e a formulação de legislações para este fim.

Vale mencionar que este é um estudo preliminar. Para trabalhos futuros outras técnicas estatísticas podem ser utilizadas e comparadas com as adotadas nesta pesquisa, como os modelos de redes neurais, os modelos não lineares, e até mesmo os modelos vetoriais autorregressivos (VAR). Também podem ser utilizados outros métodos para modelar a variância condicional e tentar realizar estimativas separadas para as distintas estações do ano.

#### 5. AGRADECIMENTOS

Os autores agradecem aos pareceristas anônimos por seus comentários construtivos. Ao IEMA pelo fornecimento das concentrações dos poluentes atmosféricos e das variáveis meteorológicas. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Fundação de Amparo à Pesquisa do Espírito Santo (FAPES) pelo suporte financeiro parcial.

#### 6. APÊNDICE A

a) **Probabilidade de detecção (POD):** mensura a ocorrência de horários de alta concentração de ozônio, iguais ou superiores a  $100 \mu\text{g}/\text{m}^3$ , que foram preditas corretamente de acordo com os valores observados:

$$POD = \frac{A}{A + B} \quad (A1)$$

b) **Taxa de alarme falso (FAR):** mede a tendência da predição de ozônio superestimar o valor observado da concentração de ozônio:

$$FAR = \frac{C}{C + A} \quad (A2)$$

c) **Taxa de perda (MISS):** mensura a taxa a qual os episódios de ozônio observados não são preditos:

$$MISS = 1 - POD = \frac{B}{A + B} \quad (A3)$$

em que:

	Estimado – Sim	Estimado – Não
Observado – Sim	A	B
Observado – Não	C	D

Nota: as descrições, observado e estimado, referem-se aos valores característicos dos episódios.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, v. 31, p. 307-327, 1986.
- BROCKWEL, P. J.; DAVIS, R. A. **Introduction to time series and forecasting**. 2ª ed. New York: Springer, 2002, 437 p.
- CARVALHO, V. S. B.; CAVALCANTI, P. M. P. S.; CATALDI, M.; PIMENTEL, L. C. G. Avaliação da Concentração do Ozônio e de seus precursores na RMRJ e correlação deste com variáveis meteorológicas durante o ano de 2002. In: Congresso Brasileiro de Meteorologia, XIII, 2004, Fortaleza. **Anais eletrônicos...** Disponível em: <<http://www.cbmet.com/cbm-files/22-106a74513a8169304ab1ec402bddd658.doc>>. Acesso em: 05 nov. 2013.
- CARVALHO, V. S. B. **Meteorologia da qualidade do ar no que tange as concentrações de ozônio e dos óxidos de nitrogênio na região Metropolitana do Rio de Janeiro**. Rio de Janeiro, RJ. 2006. 134 f. Dissertação (Mestrado em Engenharia Mecânica). Programa de Pós-Graduação em Engenharia Mecânica, Universidade Federal do Rio de Janeiro, RJ.
- CETESB. **Relatório da qualidade do ar do estado de São Paulo 2012**. São Paulo: CETESB, 2013.
- CONSELHO NACIONAL DE MEIO AMBIENTE – CONAMA (Brasil). Resolução nº 08, de 6 de dezembro de 1990. **Diário Oficial [da] República Federativa do Brasil**, Brasília, 28 dez. 1990. Seção 1, p. 25539.
- DICKEY, D. A.; FULLER, W. A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, v. 49, n. 4, p. 1057-1073, 1981.
- ENGLE, R. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, v. 20, n.3, p. 339-350. 1982.
- GOOGLE EARTH. **Informações geográficas**. 2014. Disponível em: <<http://www.google.com.br/intl/pt-PT/earth/>>. Acesso em: 20 de mar. de 2014.
- HAMILTON, J. D. **Time series analysis**, New Jersey: Princeton University Press, 1994, 820 p.
- JORQUERA, H.; PÉREZ, R.; CIPRIANO, A.; ESPEJO, A.; LETELIER, M. V.; ACUÑA, G. Forecasting ozone daily maximum levels at Santiago, Chile. *Atmospheric Environment*. v. 32, n. 20, p. 3425-3424, 1998.
- JUNGER, W.; LEON, A. P. **mtsdi: Multivariate time series data imputation**. R package version 0.3.3. 2012. Disponível em: <http://CRAN.R-project.org/package=mtsdi>. Acessado em 25.03.2014.
- KUMAR, U.; RIDDER K. GARCH modeling in association with FFT-ARIMA to forecast ozone episodes. *Atmospheric Environment*, v. 44, p. 4252-4265, 2010.
- KWIATKOWSKI, D.; PHILLIPS, P. C. B.; SCHMIDT, P.; SHIN, Y. Testing the null hypothesis of stationarity against the alternative of unit root. *Journal of Econometrics*, v. 54, n. 1, p. 159-178, 1992.
- LIU, P. W. G.; JOHNSON, R. Forecasting peak daily ozone levels-I. A regression with time series errors model having a principal component trigger to fit 1991 ozone levels. *Journal of the Air & Waste Management Association*. v. 52, n. 9, p1064-1074, 2002.
- LIU, P. W. G.; JOHNSON, R. 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, 2003.
- LIU, P. W. G.; TSAI, J. H.; LAI, H. C.; TSAI, D. M.; LI, L. W. Establishing multiple regression models for ozone sensitivity analysis to temperature variation in Taiwan. *Atmospheric Environment*, v. 79, p. 225-235, 2013.
- MATTESON, D. S.; TSAY, R. S. Dynamic orthogonal components for multivariate time series. *Journal of the American Statistical Association*, v. 106, n. 496, p. 1450-1463, 2011.
- MOREIRA, D. M.; TIRABASSI, T.; MORAES, M. R. Meteorologia e poluição atmosférica. *Ambiente & Sociedade*. v. 11, n. 1, p. 1-13, 2008.
- PHILLIPS, P. C. B.; PERRON, P. Testing for unit roots in time series regression. *Biometrika*, v. 75, n. 3, p. 335-346, 1988.
- REISEN, V. A.; SARNAGLIA, A. J. Q.; REIS JUNIOR, N. C.; SANTOS, J. M. Modeling and forecasting daily average PM<sub>10</sub> concentrations by a seasonal long-memory model with volatility. *Environmental Modelling & Software*, v. 51, p. 286-295, 2014.
- RYAN, W. F. Forecasting severe ozone episodes in the Baltimore metropolitan area. *Atmospheric Environment*, v. 29, n. 17, p. 2387-2398, 1995.
- RYAN, W. F.; PIETY, C. A.; LUEBEHUSEN, E. D. Air quality forecasts in the Mid-Atlantic Region: current practice and benchmark skill. *Weather and Forecasting*, v. 15, n. 1, p. 46-60, 1999.
- SEINFELD, J. H.; PANDIS, S. N. **Atmospheric chemistry and physics: from air pollution to climate change**. J. Wiley, New York, 2006.