

## **X-027 - MODELAGEM E PREVISÃO DA CONCENTRAÇÃO DE PM<sub>10</sub>, NA CIDADE DE VITÓRIA, ESPÍRITO SANTO, BRASIL**

### **Wanderson de Paula Pinto<sup>(1)</sup>**

Graduado em Matemática pela Faculdade da Região Serrana - FARESE. Mestre em Engenharia Ambiental pela Universidade Federal do Espírito Santo (UFES). Doutorando em Engenharia Ambiental no Programa de Pós-Graduação em Engenharia Ambiental da Universidade Federal do Espírito Santo (PPGEA/UFES). Professor do Curso de Engenharia Ambiental da FARESE e do Centro Universitário do Espírito Santo - UNESC.

### **Valdério Anselmo Reisen**

Doutor em Estatística pela University of Manchester Institute of Science And Technology. Professor do Departamento de Estatística e do Programa de Pós-graduação em Engenharia Ambiental, da UFES.

### **Gemael Barbosa Lima**

Engenheiro Ambiental pela FAESA - Faculdades Integradas Espírito-Santenses. Mestre em Engenharia Ambiental pela Universidade Federal do Espírito Santo (UFES). Professor do Curso de Engenharia Ambiental da FARESE e do Centro Universitário do Espírito Santo - UNESC.

**Endereço<sup>(1)</sup>:** Rua Jequitibá, 121 - Centro - Santa Maria de Jetibá – ES – CEP: 29645-000 - Brasil - Tel: (27) 3263-2010 - e-mail: **wandersonp@gmail.com**.

### **RESUMO**

Este trabalho objetivou foi modelar e realizar previsões da série de concentrações médias diárias de Material Particulado Inalável (PM<sub>10</sub>) monitorado no bairro de Jardim Camburi, Vitória, E.S., Brasil. As Concentrações do poluente foram obtidas junto ao Instituto Estadual de Meio Ambiente e Recursos Hídricos (IEMA). Em particular, foram apresentados os resultados dos ajustes, assim como as previsões obtidas utilizando os modelos ARMA(5,1) e ARMA(2,1). De acordo com os resultados o modelo ARMA(5, 1) foi considerado mais adequado para fazer previsões de PM<sub>10</sub>, além disso os resultados evidenciam que a Transformada de Fourier é uma opção para tratar a propriedade de sazonalidade de séries temporais.

**PALAVRAS-CHAVE:** Material Particulado, Sazonalidade, Séries Temporais, Poluição do ar.

### **INTRODUÇÃO**

O material particulado tem sido intensamente estudado por ocasionar alterações no clima, causar danos à saúde dos animais e, especialmente, na saúde dos humanos, principalmente, em ambientes urbanos industrializados (GODISH, 1997). O impacto sobre a saúde está diretamente relacionado ao tamanho, as forma e as composições das partículas. Segundo Trindade (2009) partículas pequenas adentram mais profundamente no sistema respiratório, sendo que partículas com formas pontiagudas e compostas com elementos químicos tóxicos podem lesar os tecidos do trato respiratório.

Segundo Holgate et al., (1999) um nível elevado dos poluentes pode ocasionar desde irritação dos olhos, nariz e da garganta, bronquite e pneumonia até doenças respiratórias crônicas, câncer de pulmão, problemas cardíacos, etc. Diversos estudos epidemiológicos têm demonstrado associações significativas entre a exposição à concentrações elevadas de PM<sub>10</sub> na atmosfera a problemas de saúde (OSTRO et al., 1996; NASCIMENTO et al., 2006; MARTINS et al., 2004; GOUVEIA et al., 2006).

Com a difusão de metodologias para análises de séries temporais na literatura científica, diversos trabalhos foram publicados com análise de séries temporais de dados ambientais. No âmbito da poluição atmosférica, pode-se citar diversos autores que usaram a metodologia de séries temporais para estudar e analisar séries de concentrações de poluentes atmosféricos como Shively (1990), Robeson e Steyn (1990) e Goyal et al., (2006). Vários métodos de previsão das séries temporais estão disponíveis na literatura, como os de médias móveis (MA), regressão linear com o tempo, suavização exponencial de Holt-Winters e os de Modelos ARIMA (Modelo Autorregressivo Integrado de Média Móvel). Os modelos ARIMA proporcionam previsões probabilísticas satisfatórias. Esses modelos apresentam diversas vantagens em relação aos outros modelos,

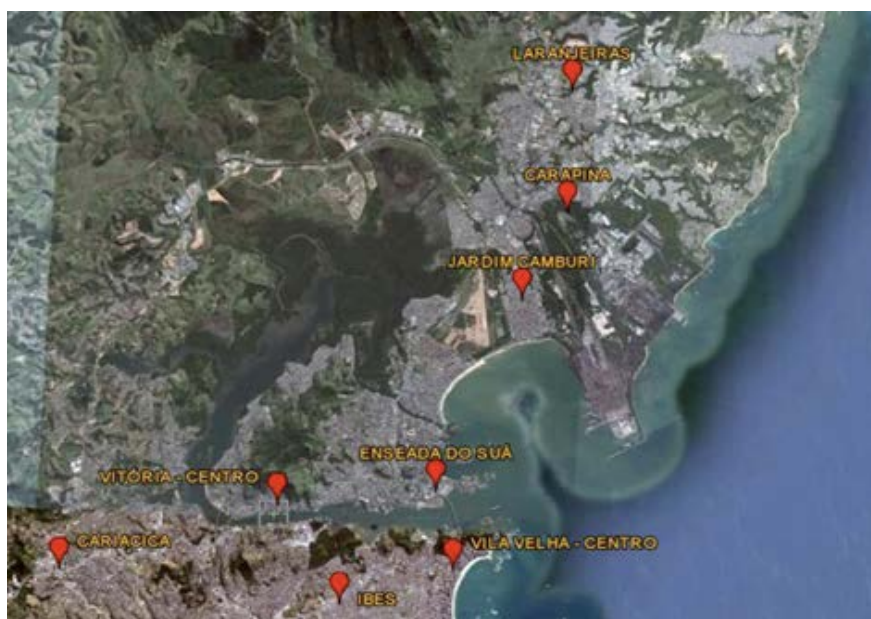
como alisamento exponencial, em particular em sua capacidade de previsão e a sua grande quantidade de informações sobre mudanças relacionadas ao tempo (MISHA e DESAI, 2005).

Assim, este trabalho objetivou modelar e realizar previsões das concentrações médias diárias de Material Particulado Inalável (PM<sub>10</sub>), na cidade de Vitória, Espírito Santo, Brasil, utilizando a abordagem de séries temporais.

## MATERIAIS E MÉTODOS

Esse trabalho foi realizado na Região da Grande Vitória (RGV), constituída pelos municípios de Vitória, Vila Velha, Cariacica, Serra e Viana, Espírito Santo. A região sofre com diversos tipos de problemas ambientais, dentre os quais está à deterioração da qualidade do ar, devido às emissões atmosféricas por indústrias e pela frota veicular.

A RGV possui uma Rede Automática de Monitoramento da Qualidade do Ar (RAMQAR) inaugurada em julho de 2000, de propriedade do Instituto Estadual de Meio Ambiente e Recursos Hídricos (IEMA). A referida rede é distribuída em oito estações localizadas nos municípios que compõem a RGV, da seguinte forma: o município Serra com duas estações localizadas nas regiões de Laranjeiras e Carapina; o município Vitória com três estações localizadas nas regiões de Jardim Camburi, Enseada do Suá e Centro de Vitória. O município de Vila Velha apresenta duas estações localizadas nas regiões do Ibes e Centro de Vila Velha e o município de Cariacica com uma estação em Cariacica. A localização espacial das estações de monitoramento da RAMQAR está ilustrada na Figura 1.



**Figura 1: Localização espacial das estações de monitoramento da qualidade do ar da RGV.**  
**Fonte: Google Earth (2015).**

A RAMQAR monitora os seguintes poluentes: Partículas Totais em Suspensão (PTS); Partículas Inaláveis (PM<sub>10</sub>); Ozônio (O<sub>3</sub>); Óxido de Nitrogênio (NO<sub>x</sub>); Monóxido de Carbono (CO) e Hidrocarbonetos (HC). E ainda, realiza-se o monitoramento dos seguintes parâmetros meteorológicos: Direção dos ventos (DV); Velocidade dos ventos (VV); Precipitação pluviométrica (PP); Umidade relativa do ar (UR); Temperatura (T); Pressão atmosférica (P) e Radiação solar (I).

Foram consideradas 731 observações compreendidas entre 01 de janeiro de 2003 e 31 de dezembro de 2004, de médias diárias das concentrações de material particulado (PM<sub>10</sub>), medidas em  $\mu\text{g} \cdot \text{m}^{-3}$  na estação de Jardim Camburi, Vitória-ES. Como os equipamentos de medição das concentrações de contaminantes na atmosfera podem apresentar falhas que impossibilitem seu funcionamento por algum tempo, ocasionando perda de

dados, as médias diárias foram calculadas com as amostras existentes. Isto justifica a escolha desta série, pois a mesma contém uma porcentagem menor que 5% de dados faltantes, viabilizando o estudo. Toda a análise estatística foi realizada no software R 2.15.1 (software livre).

### Modelo Autorregressivo Integrado Média Móvel [ARIMA (p,d,q)]

Uma série temporal  $\{X_t\}$  é dita um processo  $ARIMA(p, d, q)$  se,

$$\phi_p(B)(1 - B)^d X_t = \theta_0 + \theta_q(B)\varepsilon_t, \quad (1)$$

Em que  $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$  e  $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$  são respectivamente polinômios autorregressivo estacionário e médias móveis invertível. Portanto,  $\{X_t\}$  é processo não estacionário que depois de diferenciado  $\nabla^d X_t$  ( $d \geq 1$  e  $\nabla = 1 - B$ ) se transforma em processo estacionário e invertível  $ARMA(p, q)$ .

Para  $d = 0$ , o modelo  $ARIMA(p, d, q)$  pode ser escrito na forma,

$$\Phi(B)U_t = \Theta_q(B)\varepsilon_t \quad (2)$$

em que  $U_t = (1 - B)^d X_t$  é o processo estacionário  $ARMA(p, q)$ , quando  $d \geq 1$ , o processo  $X_t$  não é um processo estacionário. Para detalhes ver em Wei (2006) e Brockwell e Davis (2002).

### Metodologia de modelagem

A metodologia Box-Jenkins aplicada neste trabalho esta dividida nas seguintes etapas:

1. Identificação:
  - Comportamento geral da série
  - Transformação dos dados para estabilizar a variância;
  - Diferenciação dos dados para obter a série estacionária;
  - Seleção de um modelo a partir da observação da FAC e FACP.
2. Estimação:
  - Estimar os parâmetros do modelo;
  - Selecionar o melhor modelo através do critério de Akaike (AIC);
  - Avaliação do diagnóstico: análise residual.
3. Previsão.

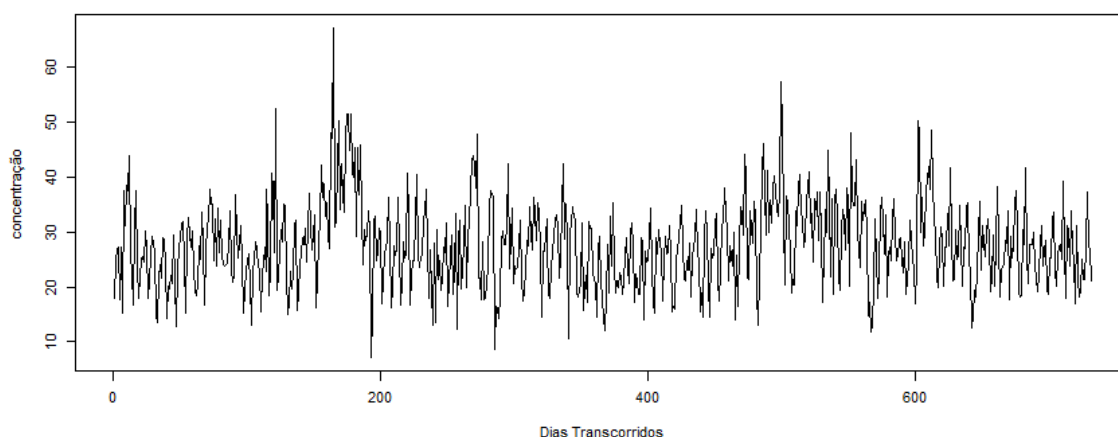
## RESULTADOS E DISCUSSÃO

### Ajuste dos Modelos

Para melhor compreensão das variáveis estudadas a Tabela 1 apresenta algumas medidas descritivas da série. Verifica-se que a média diária de  $PM_{10}$  é de, aproximadamente,  $27 \mu g.m^{-3}$ , sendo que a maior concentração foi de  $67,8 \mu g.m^{-3}$  e a menor foi  $7 \mu g.m^{-3}$ . Outra forma de analisar descritivamente a série estudada é por meio da análise gráfica. A Figura 2 apresenta graficamente a série temporal de  $PM_{10}$ , destaca-se que a série apresenta sazonalidade bem definida.

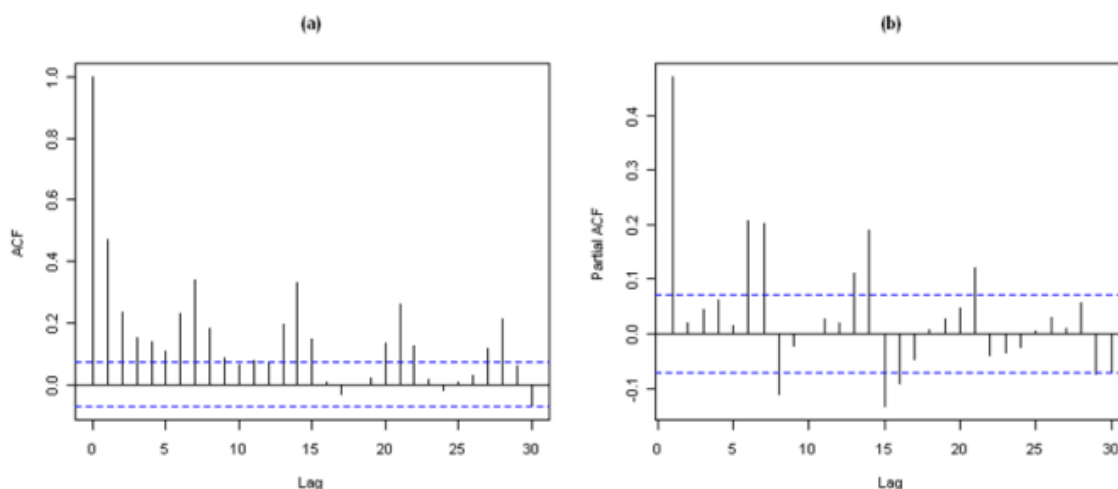
**Tabela 1: Estatísticas descritivas da série de concentração das médias diárias  $PM_{10}$ .**

	Média	Desvio	Min	Max	Mediana
Série $PM_{10}$	27.593	7.816	7.083	67.830	27.333



**Figura 2: Série Temporal das concentrações de PM<sub>10</sub> na RGV.**

Segundo Box e Jenkins (1970) para fazer a identificação dos modelos é essencial estudar o comportamento da função de autocorrelação (FAC) e da função de autocorrelação parcial (FACP). A Figura 3 mostra a FAC (a) e FACP (b) da série de concentração de PM<sub>10</sub>. Observa-se que a FAC apresenta correlações significativas com um decaimento lento para os *lags* múltiplos de 7, indicando a sazonalidade ( $s = 7$ ), confirmando assim a variação entre as concentrações medidas nos dias úteis e finais de semana.



**Figura 3: Função de autocorrelação (a) e função de autocorrelação parcial (b) da série temporal de PM<sub>10</sub>.**

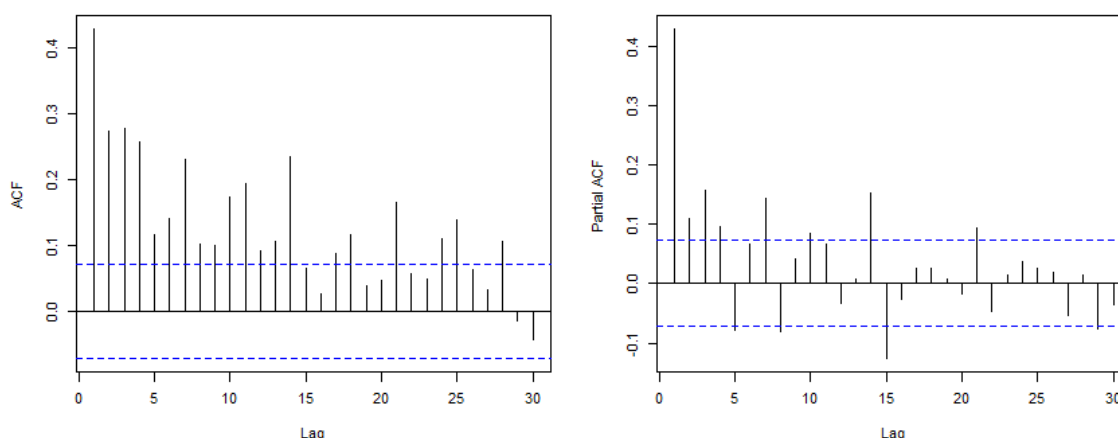
O uso do teste de Dickey-Fuller (ADF) fornece uma estatística adequada para definir se a série apresenta raiz unitária, sendo caracterizada, portanto como uma série não estacionária. A partir do resultado do teste ADF realizado pode-se concluir que a série de PM<sub>10</sub> é caracterizada como um processo estocástico estacionário na média ao nível de 5%.

Um dos métodos para resolver o problema da sazonalidade é fazer a transformada de Fourier. O espectro de um processo estacionário é a transformada de Fourier da função de autocovariância de um processo absolutamente somável e de forma geral, um processo estacionário e pode ser representado pela função de distribuição espectral (WEI, 2006).

Segundo Brockwell e Davis (2002) o espectro é a representação de uma série temporal estacionária  $\{X_t\}$  em uma soma de componentes senoidais com coeficientes aleatórios não correlacionados, em conjunto com a função de autocovariância de  $\{X_t\}$ . A decomposição espectral é análoga aos processos estacionários da

representação de Fourier de funções determinísticas. A análise dos processos estacionários por meio de sua representação espectral é muitas vezes referida como "análise de domínio de frequência" de séries temporais ou "análise espectral". É equivalente a análise no domínio do tempo que utiliza a função de autocovariância, mas fornece uma forma alternativa de visualização do processo, o que para algumas aplicações pode ser mais esclarecedora.

A Figura 4 apresenta a ACF e a PACF da série transformada. Através da análise da mesma fica evidente que o problema da sazonalidade foi parcialmente resolvido, ou seja, os fatores sazonais não vão contribuir de maneira forte para a variância total. Assim, podemos seguir para o próximo passo da modelagem que é a etapa de identificação.



**Figura 4: Correlograma da série temporal após transformada de Fourier.**

Para escolher o melhor modelo o critério de informação de Akaike (AIC) (AKAIKE, 1974), foi utilizado. De acordo com o critério de AIC e análise da ACF dos resíduos, o ARMA(5,1), denominado modelo 1, foi o modelo mais bem ajustado a série após aplicar a transformada de Fourier. Para fins de comparação escolheu-se também um modelo ARMA(2,1), denominado modelo 2, ajustado a série sem transformação. Nas Tabelas 2 e 3 contém as estimativas dos parâmetros e os respectivos erros padrão. A análise residual dos modelos validados para o passo seguinte.

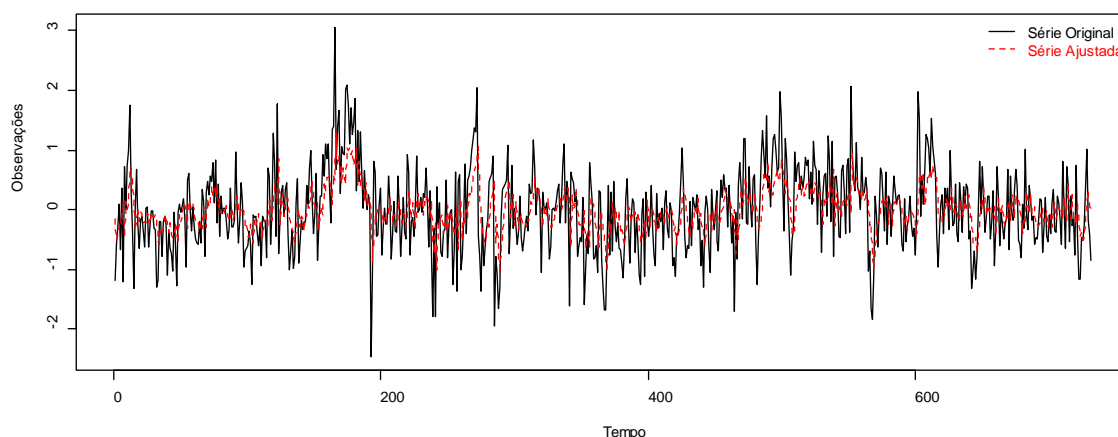
**Tabela 2: Estatísticas do modelo 1.**

Modelo 1	Parâmetro	Estimativa	Erro Padrão
ARMA(5,1)	$\phi_1$	0,2066	0,1907
	$\phi_2$	0,1092	0,0754
	$\phi_3$	0,1334	0,0381
	$\phi_4$	0,1430	0,0435
	$\phi_5$	-0,0684	0,0422
	$\theta_2$	0,1541	0,1882

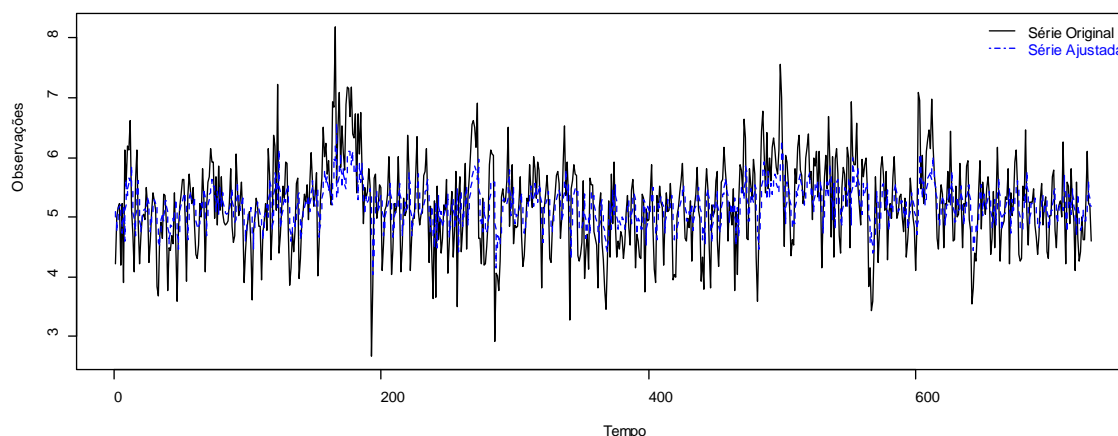
**Tabela 3: Estatísticas do modelo 2.**

Modelo 2	Parâmetro	Estimativa	Erro Padrão
ARMA(2,1)	$\phi_1$	0,2161	1,3337
	$\phi_2$	0,1054	0,6052
	$\theta_2$	0,2415	1,3354

Em seguida é apresentada a análise comparativa da qualidade dos modelos ajustados. As Figuras 5 e 6 mostram uma análise visual dos dados ajustados pelos modelos 1 e 2 a série de  $PM_{10}$ . Essa análise evidencia um bom desempenho dos modelos, uma vez que estes mostram representar bem a série.



**Figura 5: Valores de PM<sub>10</sub> observados e Preditos pelo modelo 1.**



**Figura 6: Valores de PM<sub>10</sub> observados e Preditos pelo modelo 2.**

### Análise de Previsão

Um dos principais objetivos do trabalho é fazer a comparação entre os modelos. Portanto, nesta seção, é apresentado o estudo das previsões um passo à frente, para avaliar o desempenho dos modelos ajustados. No intuito de quantificar os erros no ajuste, algumas medidas de erros de predição e previsão são apresentadas nas Tabelas seguintes.

**Tabela 4: Avaliação dos erros de predição dos modelos.**

Medidas de erro	Modelo 1	Modelo 2
Erro Quadrático Médio (MSE)	0.3691	0.4358
Erro Absoluto Médio (MAE)	0.4703	0.5124

**Tabela 5: Avaliação dos erros de previsões dos modelos.**

Medidas de erro	horizonte	Modelo 1	Modelo 2
Erro Quadrático Médio (MSE)	h=1	0,2595	0,3098
Erro Absoluto Médio (MAE)	h=1	0.4290	0.4390

De acordo com as medidas de qualidade apresentadas nas Tabelas 4 e 5, observa-se que o modelo 1 possui maior acurácia que o modelo 2, superioridade que pode ser verificada pelos menores valores de MSE e MAE, nos dois horizontes de previsão ( $h = 0$  e  $h = 1$ ). As conclusões comparativas sugerem a superioridade do modelo 1. Percebe-se que a adequacidade do modelo depende da série em estudo e de suas características. Por



meio da análise gráfica e de medidas de qualidade pode-se verificar que a Transformada de Fourier é uma opção para tratar o fenômeno sazonalidade, pois o modelo 1 foi superior ao modelo 2. Com isso, sugere-se o uso dessa metodologia como ferramenta para o estudo de séries temporais com esta característica.

## CONCLUSÃO

As previsões de concentrações de poluentes atmosféricos podem ser utilizadas como ferramenta de alerta sobre futuros picos atípicos de concentrações e permitir a tomada de decisão quanto à adaptação de comportamento da população e grupos de risco, como crianças, idosos e pessoas com doenças respiratórias. Pode também servir para as autoridades competentes como informação para a preparação de planos para a redução de emissões e gerenciamento da qualidade do ar. A metodologia empregada neste trabalho se baseia na proposta por Box-Jenkins (1970). Essa metodologia foi aplicada para estudar o comportamento da série, verificar a estacionariedade e sazonalidade, ajustar o melhor modelo aos dados e fazer previsões. De acordo com as análises realizadas o modelo que melhor representou a série em estudo foi ARMA(5, 1).

A estrutura de correlação da série de PM<sub>10</sub>, analisada nesse trabalho, apresentou a propriedades de sazonalidade. Uma opção de metodologia para resolver tal problema foi a Transformada de Fourier, que segundo os resultados é uma opção razoável para o estudo de séries temporais. Tal conclusão justifica-se devido aos resultados obtidos com o modelo ARMA(5,1), ajustado a série transformada, no estudo de previsão. Considerando a importância da previsão de concentrações de PM<sub>10</sub>, para o planejamento de medidas de segurança, espera-se que os resultados encontrados neste trabalho propiciem incentivos para estudos futuros ligados ao planejamento dessas medidas.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. Akaike, H. A New Look at The Statistical Model Identification. IEEE Transactions Automatic Control, vol AC 19, Nº 6, 716-723, 1974.
2. BOX, G. E. P., JENKINS, G. M. Time Series Analysis, Forecasting and Control. Holden-Day, 1970.
3. Brockwell, P., Davis, R. Introduction to Time Series and Forecasting, second edn, Springer Verlag, 2002.
4. GODISH. T. Air quality. Boca Raton: CRC Press. LLC. 1997.
5. GOUVEIA. N. et al. Poluição do e efeitos na saúde nas populações de duas grandes metrópoles brasileiras. Epidemiol. Serv. Saúde, 12. 29-40, 2006.
6. GOYAL. P., CHAN. A. T., JAISWAL. N. Statistical models for the prediction of respirable suspended particulate matter in urban cities. Atmospheric Environment 40. 2068-2077, 2006.
7. HOLGATE. S. T., SAMET. J. M., KOREN. H. S., MAYNARD. R. L. Air Pollution and Health. San Diego. EUA. Academic Press. 1999.
8. MARTINS. L. C. et al. Poluição atmosférica e atendimentos por pneumonia e gripe em São Paulo. Brasil. Revista de Saúde Pública 36. 88-94, 2004.
9. MISHARA A. K., DESAI V. R. Drought forecasting using stochastic model.. Stochastic Environmental Research and Risk Assessment. 19. 326-339, 2005.
10. NASCIMENTO. L.F.C., PEREIRA. L.A.A., BRAGA. A.L.F., MÓDOLOA. M.C.C., CARVALHO. J.A.C. Efeitos da poluição atmosférica na saúde infantil em São José dos Campos. SP. Revista de Saúde Pública 40. 77-82, 2006.
11. OSTRO. B., SANCHES. J. M., ARANDA. C., ESKELAND. G. S. Air pollution and mortality: results from a study of Santiago. Chile. Journal of Exposure Analysis and Environmental Epidemiology 6. 97-114, 1996.
12. ROBESON S.M., STEYN. D.G. Evaluation and comparison of statistical forecast models for daily maximum ozone concentrations. Atmospheric Environment. Part B. 24(2). 303-312, 1990.
13. SHIVELY. T. S. An analysis of the long-term trend in ozone data from two Houston. Texas monitoring sites Atmospheric Environment. Part B. Urban Atmosphere 24(2). 293-301, 1990.
14. TRINDADE, C. C. Avaliação do uso de diferentes modelos receptores com dados de PM<sub>2,5</sub>: balanço químico de massa (BQM) e fatoração de matriz positiva (FMP). 143 f. Dissertação (Mestrado) Programa de Pós-Graduação em Engenharia Ambiental do Centro Tecnológico, UFES, Vitória, 2009.
15. WEI. W. Time Series Analysis: univariate and multivariate methods. Pearson. Boston. 2006. 2 ed.