

Juliana Bottoni de Souza^IValdério Anselmo Reisen^{II}Jane Méri Santos^IGlaura Conceição Franco^{III}

Componentes principais e modelagem linear generalizada na associação entre atendimento hospitalar e poluição do ar

Principal components and generalized linear modeling in the correlation between hospital admissions and air pollution

RESUMO

OBJETIVO: Analisar a associação entre concentrações dos poluentes atmosféricos e atendimentos diários por causas respiratórias em crianças.

MÉTODOS: Estudo ecológico de série temporal. Foram analisadas as contagens diárias de admissões hospitalares de crianças < 6 anos e as concentrações diárias de poluentes atmosféricos (PM₁₀, SO₂, NO₂, O₃ e CO), na Região da Grande Vitória, ES, de janeiro de 2005 a dezembro de 2010. Foram combinadas duas técnicas para a análise estatística: modelo de regressão de Poisson em modelos aditivos generalizados e análise de componentes principais. Essas técnicas complementaram-se e forneceram estimativas mais expressivas na estimação do risco relativo. Os modelos foram ajustados para efeitos da tendência temporal, sazonalidade, dias da semana, fatores meteorológicos e autocorrelação. No ajuste final do modelo, foi necessária a inclusão de modelos do tipo *Autoregressive Moving Average Models* (p,q) nos resíduos, para eliminar as estruturas de autocorrelação presente nas componentes.

RESULTADOS: O aumento de 10.49 µg/m³ (intervalo interquartilico) nos níveis do poluente PM₁₀ resultou num aumento de 3,0% do valor do risco relativo estimado por meio do modelo aditivo generalizado – análise de componentes principais-sazonal autorregressivo –, enquanto no modelo aditivo generalizado usual a estimativa foi de 2,0%.

CONCLUSÕES: Em comparação ao modelo aditivo generalizado usual, em geral, a vertente proposta do modelo aditivo generalizado – análise de componentes principais apresentou melhores resultados na estimativa do risco relativo e na qualidade do ajuste.

DESCRITORES: Poluição do Ar, efeitos adversos. Admissão do Paciente. Hospitalização. Doenças Respiratórias, epidemiologia. Estudos de Séries Temporais. Estudos Ecológicos.

^I Programa de Pós-Graduação em Engenharia Ambiental. Universidade Federal do Espírito Santo. Vitória, ES, Brasil

^{II} Departamento de Estatística. Universidade Federal do Espírito Santo. Vitória, ES, Brasil

^{III} Departamento de Estatística. Universidade Federal de Minas Gerais. Belo Horizonte, MG, Brasil

Correspondência | Correspondence:

Juliana Bottoni de Souza
Rua Manoel Coutinho, 175 Porto de Santana
29153-020 Cariacica, ES, Brasil
E-mail: juliana_bottoni@yahoo.com.br

Recebido: 5/8/2013
Aprovado: 14/2/2014

Artigo disponível em português e inglês em:
www.scielo.br/rsp

ABSTRACT

OBJECTIVE: To analyze the association between concentrations of air pollutants and admissions for respiratory causes in children.

METHODS: Ecological time series study. Daily figures for hospital admissions of children aged < 6, and daily concentrations of air pollutants (PM₁₀, SO₂, NO₂, O₃ and CO) were analyzed in the *Região da Grande Vitória*, ES, Southeastern Brazil, from January 2005 to December 2010. For statistical analysis, two techniques were combined: Poisson regression with generalized additive models and principal model component analysis. Those analysis techniques complemented each other and provided more significant estimates in the estimation of relative risk. The models were adjusted for temporal trend, seasonality, day of the week, meteorological factors and autocorrelation. In the final adjustment of the model, it was necessary to include models of the Autoregressive Moving Average Models (p, q) type in the residuals in order to eliminate the autocorrelation structures present in the components.

RESULTS: For every 10:49 µg/m³ increase (interquartile range) in levels of the pollutant PM₁₀ there was a 3.0% increase in the relative risk estimated using the generalized additive model analysis of main components-seasonal autoregressive – while in the usual generalized additive model, the estimate was 2.0%.

CONCLUSIONS: Compared to the usual generalized additive model, in general, the proposed aspect of generalized additive model - principal component analysis, showed better results in estimating relative risk and quality of fit.

DESCRIPTORS: Air Pollution, adverse effects. Patient Admission. Hospitalization. Respiratory Tract Diseases, epidemiology. Time Series Studies. Ecological Studies.

INTRODUÇÃO

O impacto da poluição atmosférica no bem-estar humano é a principal motivação para o seu estudo e controle. A poluição atmosférica afeta a saúde da população mesmo quando seus níveis encontram-se abaixo do que determina a legislação vigente.⁶ Estudos apontam associações significativas entre os níveis diários de concentração de poluentes e atendimentos por causas respiratórias ou cardiovasculares^{2,4,10,11} entre outros. Os pulmões constituem o principal alvo de ataque dos poluentes do ar, sobretudo, os principais: material particulado (PM₁₀), dióxido de enxofre (SO₂) e ozônio (O₃).

Pelas características da variável de desfecho em saúde (atendimentos), o modelo de regressão não linear aditivo generalizado com distribuição marginal de Poisson é a ferramenta estatística para medir e quantificar a associação entre os poluentes atmosféricos e os efeitos adversos à saúde. A inclusão das covariáveis (poluentes) no modelo de regressão ocorre individualmente nessa metodologia, entretanto os poluentes são correlacionados. A utilização da análise de componentes principais (ACP) da matriz de covariância dos

poluentes é uma alternativa a essa questão. Avaliar os efeitos adversos à saúde por meio de uma combinação de poluentes pode ser mais interpretável e viável do que isolar os efeitos de um poluente apenas. Essa vertente de pesquisa é atual e relevante, já explorada anteriormente.^{7,11} Roberts & Martins⁷ (2006) avaliaram a associação dos poluentes PM₁₀, O₃, SO₂, NO₂ e CO e seus efeitos à saúde. O problema da multicolinearidade (correlação entre os poluentes) foi resolvido por meio de ACP. Essa é uma técnica estatística de análise múltipla utilizada para reduzir a dimensionalidade de um conjunto de dados com a preservação do máximo de variabilidade nas covariáveis,⁴ que permitiu agrupar os poluentes no modelo de regressão. Os autores sugerem um método descrito como “supervisionado de ACP”, em que a relação entre as covariáveis (poluentes) e os efeitos deletérios à saúde são observados antes da inserção desses no modelo de regressão. A associação conjunta dos poluentes na mortalidade diária foi analisada por Wang & Pham¹¹ (2011), utilizando a ACP e um método robusto. Os resultados apresentaram estimativas do risco relativo (RR) mais expressivas quando

aplicada a técnica de análise múltipla de ACP evidenciando associação mais forte entre as variáveis.

A aplicação dessa técnica requer que os dados sejam obtidos por meio de replicações independentes. Entretanto, seu uso para realizar inferências estatísticas necessita da devida atenção em covariáveis correlacionadas no tempo, como os poluentes atmosféricos.

A propriedade de estacionariedade deve ser satisfeita. As componentes principais são combinações lineares das covariáveis, cujas propriedades serão transferidas de forma linear para as componentes.

O efeito das correlações no contexto inferencial foi estudado por Zamprogn^a e mostra, teórica e empiricamente, o efeito da correlação temporal das covariáveis, quando essa é negligenciada no procedimento. Na Proposição 1, o autor mostra que as componentes são correlacionadas temporalmente. O uso das ACP em regressão deve ser feito de forma cautelosa, pois carregam a estrutura temporal das covariáveis.

O presente estudo teve como objetivo analisar a associação entre concentrações dos poluentes atmosféricos e atendimentos diários por causas respiratórias em crianças.

MÉTODOS

Estudo ecológico de série temporal. Foram analisados 59.353 atendimentos hospitalares de crianças < 6 anos na Grande Vitória, ES, de 1 de janeiro de 2005 a 31 de dezembro de 2010. Os dados foram obtidos no sistema de registro de atendimentos do hospital Infantil Nossa Senhora da Glória (HINSG), onde se encontra o principal pronto-socorro infantil da Região da Grande Vitória. Utilizou-se a Classificação Internacional de Doenças (CID-10) para as doenças aéreas respiratórias.

Informações sobre os níveis diários dos poluentes atmosféricos – material particulado (PM₁₀), dióxido de enxofre (SO₂), monóxido de nitrogênio (NO₂), ozônio (O₃) e monóxido de carbono (CO) – e as variáveis meteorológicas foram obtidas no Instituto Estadual de Meio Ambiente e Recursos Hídricos (IEMA), medidas e coletadas nas oito estações de monitoramento da rede automática de monitoramento da qualidade do ar (RAMQAr).

A coleta de dados compreendeu o período de 24h para todos os poluentes e foi iniciada na primeira meia hora do dia. Foi considerada a média de 24h para os poluentes PM₁₀ e o SO₂, a média móvel de 8h para o CO e o O₃, e a máxima de 24h para o NO₂ para cada estação. A média diária dessas variáveis entre as estações foi as covariáveis utilizadas no modelo aditivo generalizado

(MAG) e na sua extensão, MAG-ACP. As variáveis atmosféricas foram medidas em µg/m³ e as meteorológicas (temperatura e umidade relativa) foram medidas de acordo com suas unidades (°C e %, respectivamente).

As variáveis de interesse foram modeladas utilizando técnicas de séries temporais, de modelos de regressão e de análise múltipla. O MAG e suas vertentes permitem funções paramétricas e não paramétricas para o ajuste da curva média dos dados. O desfecho foi modelado tendo como pressuposto básico de distribuição que as contagens de eventos de saúde (admissões hospitalares) seguem uma distribuição de Poisson. O número diário de atendimentos por doenças respiratórias foi considerado variável dependente e os níveis de concentrações diárias dos poluentes atmosféricos, variáveis independentes. Uma característica comum das variáveis foi a falta de observações (*missings*), seja por medições incorretas, por falhas nos equipamentos, seja por medições inválidas, entre outras. Essas variáveis foram ajustadas via imputação, de acordo com a metodologia descrita por Junger,^b cujas estimativas são obtidas pela correlação espacial entre os níveis do poluente e pela autocorrelação dos níveis desse poluente.

A análise e ajuste dos modelos consistiram em procedimentos implementados em etapas. Tratou-se a sazonalidade de curta duração com variáveis indicadoras para os dias da semana e os feriados. A função suavizadora *loess*² foi utilizada para a sazonalidade de longa duração. Ela permite controlar dependência não linear entre a variável de interesse (atendimentos) e a sazonalidade. As covariáveis de confusão (temperatura e umidade relativa) foram modeladas por curvas suavizadoras *splines*.^{2,10} As componentes principais foram calculadas por meio da matriz de covariância dos poluentes considerados.

Foi aplicada a técnica de análise múltipla de ACP para avaliar os efeitos dos poluentes conjuntamente, eliminando a correlação entre eles, e investigar o efeito simultâneo dos poluentes. O modelo de regressão utilizado foi o MAG, e sua extensão, o MAG-ACP. Os efeitos da poluição sobre a saúde foram quantificados por meio do RR, que expressa a probabilidade de um indivíduo desenvolver uma doença relativa à exposição a um fator de risco. A estimativa do RR foi utilizada para comparação entre os modelos propostos.

O RR foi obtido pela resolução de um sistema de equações, do modelo MAG e da aplicação de ACP. Os resultados consideraram as variações interquartílicas dos poluentes e foram calculados pela expressão %RR = (RR - 1)*100.

^a Zamprogn B. Análise de componentes principais no domínio do tempo e suas implicações em dados autocorrelacionados [tese de doutorado]. Vitória: Programa de Pós-Graduação em Engenharia Ambiental do Centro Tecnológico da UFES; 2013.

^b Junger WL. Análise, Imputação de dados e interfaces computacionais em estudos de séries temporais epidemiológicas [tese de doutorado]. Universidade Estadual do Rio de Janeiro; 2008.

O MAG¹³ com distribuição marginal de Poisson é usualmente referido nas análises de associação entre a variável desfecho e as covariáveis explicativas. Sua ampla utilização deve-se ao fato de a modelagem não linear descrever a relação entre as variáveis de interesse.^{1,2,8,9}

Seja $\{Y_t\}$, $t = 1, \dots, N$, uma série temporal de contagem formada por números inteiros não negativos. A densidade condicional de $\{Y_t\}$ dado o passado F_{t-1} , denotado por Y_t / F_{t-1} , possui distribuição de Poisson, com média μ_t , se satisfizer:

$$f(y_t; \mu_t / F_{t-1}) = \frac{e^{-\mu_t} \mu_t^{y_t}}{y_t!}, t = 1, \dots, N. \quad (1)$$

Seja $X = [x_1, \dots, x_p]^T$ (2)

o vetor de dimensão p de covariáveis que podem incluir os valores passados de Y_t e outras informações auxiliares, como os poluentes e as variáveis de confusão (tendência, sazonalidade, variáveis meteorológicas entre outras).⁵

A curva que descreve a relação entre Y_t e o vetor de covariáveis X é obtida por meio da transformação logarítmica de μ_t :

$$\log(\mu_t) = \sum_{j=0}^q \beta_j X_j + \sum_{j=q+1}^p f(x_j) \quad \text{com } q \leq p \quad (3)$$

em que β_j é o vetor dos coeficientes a serem estimados (covariáveis) e $f(x_j)$ são as funções suavizadoras para as variáveis de confusão (temperatura e umidade) e a sazonalidade de longa duração presente nos dados. β_0 corresponde ao intercepto da curva associado ao vetor de valores unitários.

Medida muito utilizada em estudos epidemiológicos para mensurar o impacto da concentração de poluentes atmosféricos na saúde da população exposta é o RR. Pode-se defini-lo como a relação de probabilidade de um evento ocorrer após certa exposição a um fator de risco, que no estudo refere-se à exposição de níveis de concentração de poluentes. No caso do modelo MAG, com distribuição marginal de Poisson, $RR(x)$ é estimado pela fórmula:¹²

$$\overline{RR}(x = \zeta) = e^{(\zeta \hat{\beta})} \quad \text{com } i = 1, 2, \dots, p \quad (4)$$

em que ζ é a variação da concentração do poluente que pode, por exemplo, assumir o valor de 10 $\mu\text{g}/\text{m}^3$, da variação interquartilica, entre outros, e $\hat{\beta}$ é o coeficiente estimado associado ao poluente estudado. Para um nível de significância α , a hipótese a ser testada é definida como $H_0: RR(x) = 1$ contra $H_1: RR(x) > 1$. A não rejeição de H_0 implica estatisticamente que o poluente estudado não causa efeito adverso à saúde.

A ACP é uma técnica estatística de análise múltipla que visa reduzir a dimensionalidade do espaço matricial dos dados por meio de transformações lineares das

variáveis originais. A correlação existente entre as variáveis implica a ocorrência de multicolinearidade nos modelos de regressão. A aplicação da técnica de ACP neste artigo deve-se ao fato de solucionar o problema da correlação entre os poluentes. A explicação de toda a variabilidade do sistema determinado por k variáveis só pode ser efetuada por k componentes principais. No entanto, grande parte dessa variabilidade pode ser explicada por um número r menor de componentes, $r \leq k$.⁴

Sejam $(\lambda_1, \alpha_1), (\lambda_2, \alpha_2), \dots, (\lambda_k, \alpha_k)$ os pares de autovalores-autovetores, respectivamente, da matriz Σ de covariância do vetor X . A i -ésima componente principal de Σ é dada por

$$CP_i = \alpha_i' x = \alpha_{i1} x_1 + \alpha_{i2} x_2 + \dots + \alpha_{ik} x_k, i = 1, 2, \dots, k. \quad (5)$$

As covariáveis geradas pela aplicação de ACP, definidas aqui por CP, são combinações lineares das variáveis ambientais originais. A inclusão dessas novas covariáveis no modelo MAG é definida da seguinte forma:

$$\log(\mu_t) = \sum_{i=0}^q v_i CP_i + \sum_{i=q+1}^p f(x_i) \quad \text{com } q \leq p \quad (6)$$

em que v_i é o vetor estimado das componentes principais (CP) e $f(x_i)$ são as funções suavizadoras para as variáveis de confusão (temperatura e umidade). A estimativa do RR para o modelo 6 é dada por:

$$\overline{RR}(x = \zeta) = e^{(\zeta \hat{v})} \quad \text{com } i = 1, 2, \dots, p \quad (7)$$

em que ζ é a variação da concentração do poluente dada pela variação interquartilica. \hat{v}_i é representado pela expressão:

$$\hat{v}_i = \sum_{j=1}^k \alpha_{ij} \hat{v}_j, j = 1, \dots, k. \quad (8)$$

em que α_{ij} corresponde aos autovetores associados das covariáveis j ; \hat{v}_i é o coeficiente estimado da i -ésima componente principal e \hat{v}_j é facilmente obtido por meio das equações 5 e 6.

O coeficiente \hat{v}_i é obtido com a solução linear das equações 5 e 6, em que se extrai a contribuição individual de cada poluente da combinação linear de todos os poluentes (equação 5) por meio da equação 6.

RESULTADOS

O número médio de atendimentos diários foi de 27,1 com DP = 18,1 (Tabela 1).

As variáveis meteorológicas (temperatura e umidade relativa) foram da estação de monitoramento de Carapina. A média da temperatura máxima utilizada no modelo foi de 29,4°C (DP = 3,3°C) e a umidade relativa do ar obteve média de 77,4% (DP = 6,0%).

Tabela 1. Estatística descritiva dos atendimentos por doenças respiratórias nas áreas de abrangência de cada estação da rede de monitoramento da qualidade do ar. Grande Vitória, ES, janeiro de 2005 a dezembro de 2010.

Variável	Média	Desvio padrão	Mínimo	Percentis			
				25	50	75	Máximo
PM ₁₀ (µg/m ³)	33,5	8,8	9,0	27,9	32,7	38,4	86,7
SO ₂ (µg/m ³)	12,4	3,1	4,9	10,1	12,2	14,6	26,5
O ₃ (µg/m ³)	31,9	8,4	12,1	26,0	30,7	36,6	72,3
NO ₂ (µg/m ³)	24,8	6,9	9,0	19,6	24,1	29,4	62,6
CO (µg/m ³)	885,8	231,3	295,0	724,8	866,6	1031,1	2141,5
Temperatura mínima (°C)	20,9	2,5	13,1	19,1	21,1	22,8	26,0
Temperatura média (°C)	24,4	2,4	17,0	22,6	24,4	26,3	30,8
Temperatura máxima (°C)	29,3	3,3	19,4	27,2	29,4	31,6	39,7
Umidade relativa do ar (%)	77,4	6,0	61,6	73,2	77,2	81,1	97,3
Atendimentos	27,1	18,1	1,0	13,0	24,0	37,0	121,0

Houve indicação de correlação entre os poluentes atmosféricos, considerada moderada e fraca (Tabela 2).

As três primeiras componentes explicaram 83,2% da variabilidade total das variáveis. Utilizou-se a proporção da variância acumulada como critério de escolha das componentes no MAG. Resultados similares na modelagem foram encontrados quando utilizadas as quatro primeiras componentes. Decidiu-se utilizar as três primeiras pelo critério de parcimônia como covariáveis e foram destacadas em negrito (Tabela 3).

Os maiores coeficientes (autovetores) das componentes 1, 2 e 3 foram dos poluentes CO, O₃ e SO₂, respectivamente. Foi sugerida divisão de *clusters* em cada componente que agrupasse, e.g., poluentes com cargas fatoriais > 0,5. As sugestões foram indicadas por (*) para cada componente principal (Tabela 3).

Foi ajustado um modelo sazonal autorregressivo de média móvel (SARMA) (1,0,0)(1,0,0)₇ para os resíduos do MAG-ACP, o que resultou no modelo final MAG-ACP-SAR. Isso eliminou a autocorrelação dos dados.

Estudo comparativo da qualidade dos ajustes dos dois modelos propostos foi feito por meio do \overline{EQM} , definido da forma:

$$\overline{EQM} = \sum_{i=1}^n \frac{e_i^2}{N},$$

em que $e_i = Y_i - \bar{Y}_i$, com \bar{Y}_i sendo o valor predito de Y_i , o número de atendimentos hospitalares. Os resultados para o MAG foi aproximadamente 35,0% superior ao obtido pelo MAG-ACP.

Os valores do RR para cada poluente e modelo foram calculados para a comparação do desempenho dos modelos ajustados MAG e MAG-ACP-SAR nas variáveis. Os resultados foram expressos por incremento da variação interquartilica, uma vez que a análise do RR foi realizada para poluentes de diferentes escalas (Tabela 4). Os resultados do RR estimado foram significativos para todos os modelos. Os valores mais expressivos do RR estimado foram obtidos principalmente por meio do modelo proposto MAG-ACP.

Tabela 2. Correlação entre os poluentes, variáveis meteorológicas e atendimentos. Grande Vitória, ES, janeiro de 2005 a dezembro de 2010.

Variável	PM ₁₀	SO ₂	NO ₂	CO	O ₃	T(máx)	T(mín)	UR	Atendimento
PM ₁₀	1,00								
SO ₂	0,31	1,00							
NO ₂	0,34	0,04	1,00						
CO	0,35	0,22	0,61	1,00					
O ₃	0,04	0,08	0,04	0,40	1,00				
T(máx)	0,20	0,44	0,43	0,06	0,23	1,00			
T(mín)	0,10	0,16	0,48	0,10	0,16	0,62	1,00		
UR	0,28	0,29	0,23	0,26	0,22	0,44	0,03	1,00	
Atendimentos	0,05	0,33	0,09	0,09	0,08	0,15	0,19	0,14	1,00

Tabela 3. Resultado das cargas fatoriais e estatísticas da aplicação de análise de componentes principais (CP). Grande Vitória, ES, janeiro de 2005 a dezembro de 2010.

Variável	CP 1	CP 2	CP 3	CP 4	CP 5
Desvio padrão	1,4315	10,431	10,115	0,7741	0,4904
Proporção da variância	0,4098	0,2176	0,2046	0,1198	0,0481
Proporção da variância acumulada	0,4098	0,6274	0,832	0,9519	1,0000
CO	-0,6074*	-0,1999	-0,2311	-0,2146	-0,7012
NO ₂	-0,5058*	0,3316	-0,0486	-0,2599	-0,5810
O ₃	0,2523	0,8615*	-0,0363	-0,1995	-0,3911
PM ₁₀	-0,4680	0,3213	0,2784	0,7746	-0,0151
SO ₂	-0,3041	0,068	0,7992*	-0,4966	0,1327

Os valores em negrito referem-se às componentes utilizadas pelo critério de parcimônia como covariáveis.

Os RR estimados para PM₁₀ aumentaram aproximadamente de 2,0% ($\bar{R}\bar{R}$) para 3,0% ($\bar{R}\bar{R}^*$). Aumentos expressivos nos valores estimados do RR foram observados para o poluente CO. Nesse caso, $\bar{R}\bar{R} = 1,020$, $\bar{R}\bar{R}^* = 1,048$. Portanto, o modelo proposto MAG-ACP-SAR apresentou resultados mais expressivos no aumento esperado do número de atendimentos por causas respiratórias, comparado ao MAG usual.

DISCUSSÃO

O presente artigo propôs o uso de duas técnicas estatísticas a fim de aperfeiçoar o desempenho do modelo de associação entre os poluentes atmosféricos e os atendimentos hospitalares por doenças respiratórias. Verificou-se que as estimativas de RR foram melhores em níveis gerais de análise, sobretudo quando comparadas com as modelagens usualmente utilizadas na literatura. A técnica de ACP eliminou a correlação entre os poluentes estudados.

O modelo proposto por este estudo é denotado por MAG-ACP e utiliza as componentes principais dos dados originais como covariáveis no modelo MAG. Como essas componentes possuem autocorrelação, essa propriedade é transferida para os resíduos do modelo ajustado MAG-ACP. Os resíduos desse modelo foram ajustados por meio do modelo sazonal autorregressivo integrado de média móvel (SARIMA) $(1,0,0)(1,0,0)_7$. O modelo final é definido como MAG-ACP-SAR, com a utilização do modelo SAR, caso particular do modelo SARIMA.

A qualidade do ajuste dos modelos supracitados foi calculada por meio do erro quadrático médio (EQM) estimado. Os resultados indicam que o EQM do modelo MAG usual foi 35,0% superior ao do modelo MAG-ACP-SAR

proposto, ou seja, o modelo proposto apresenta melhores resultados que o usualmente utilizado na literatura.

Os níveis de concentração dos poluentes estudados não ultrapassaram o padrão primário de qualidade do ar recomendados pelo Conselho Nacional de Meio Ambiente (CONAMA),^c nem os limites estabelecidos pela diretriz da Organização Mundial da Saúde (OMS).^d Porém, outros estudos mostram que os poluentes, mesmo em níveis de exposição abaixo dos padrões aceitáveis, causam efeitos deletérios para a saúde humana.^a

Os efeitos deletérios provocados pela exposição às concentrações dos poluentes na saúde da população infantil da Região da Grande Vitória foram obtidos por meio da estimativa do Risco Relativo nos modelos de regressão propostos MAG, MAG-ACP-SAR.

Com base nos estudos teóricos e empíricos apresentados por Zamprogno,^a a técnica de ACP pode ser aplicada sem causar interpretações e testes espúrios quando o processo possui uma estrutura de autocorrelação fraca.

As análises descritivas e gráficas motivaram o uso da técnica de ACP nos dados de poluentes atmosféricos, mesmo com a indicação de que os poluentes possuem fraca estrutura de correlação e autocorrelação.

Aumento de 10,5 $\mu\text{g}/\text{m}^3$ (intervalo interquartilico) nos níveis de material particulado (PM₁₀) acarretou num $\bar{R}\bar{R}^*$ para 1,029 (IC95% 0,991;1,09) no modelo MAG-ACP-SAR. Interpretações similares podem ser observadas para os outros poluentes e no MAG usual. Os resultados encontrados neste estudo por meio dos MAG e MAG-ACP são coerentes com os estudos realizados por Roberts & Martins⁷ (2006). Os autores consideraram a relação entre a morbidade e as concentrações de poluentes atmosféricos para dados registrados na

^c Ministério do Meio Ambiente. Conselho Nacional do Meio Ambiente. Resolução CONAMA n° 003, de 28 de junho de 1990. Dispõe sobre o estabelecimento de padrões nacionais de qualidade do ar determinando as concentrações de poluentes atmosféricos. *Diário Oficial Uniao*. 22 ago 1990 [2013 Out 21] Seção1:15937-9. Disponível em: <http://www.mma.gov.br/port/conama/legiabre.cfm?codlegi=100>

^d World Healthy Organization. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Genebra; 2006.

Tabela 4. Risco Relativo (RR) e intervalo de confiança de 95% nos atendimentos por doenças respiratórias em < 6 anos para variação interquartilica dos poluentes PM₁₀, SO₂, NO₂, O₃ e CO na Região da Grande Vitória. Grande Vitória, ES, janeiro de 2005 a dezembro de 2010.

	\bar{RR}	IC95%	\bar{RR}^*	IC95%
PM ₁₀	1,02	1,010;1,039	1,03	1,001;1,090
SO ₂	1,04	1,010;1,080	0,98	0,972;1,001
CO	1,02	1,010;1,030	1,05	1,002;1,071
NO ₂	1	0,990;1,020	1,03	1,010;1,040
O ₃	0,98	0,972;1,001	1,08	1,003;1,093

PM₁₀: material particulado; SO₂: dióxido de enxofre; O₃: ozônio; CO: monóxido de carbono; NO₂: dióxido de nitrogênio

Coreia. O artigo comprova que a utilização da técnica de ACP melhora as estimativas finais do risco relativo.

Os resultados do presente estudo apontam para relação significativa entre os níveis de concentração dos poluentes e o número de atendimentos hospitalares em crianças < 6 anos, mesmo em um ambiente com níveis abaixo dos padrões recomendados pelo CONAMA^c e OMS.^d As componentes principais obtidas por meio da matriz de variância/covariância aplicada a processos indexados no tempo apresentam as propriedades de correlação temporal. Propõe-se a utilização de filtros paramétricos nas variáveis originais para remover a correlação temporal. O método de filtragem, por meio do modelo vetorial autoregressivo (VAR), é sugerido como procedimento alternativo para transformar os dados atmosféricos num processo ruído branco.

Recomenda-se a utilização da técnica de ACP na análise de dados de frequência. Outros grupos suscetíveis da população devem ser estudados, bem como outras doenças, por exemplo, as cardiovasculares. Podem-se utilizar outras técnicas como a do *bootstrap* para estimar os intervalos de confiança com mais precisão e a modelagem *generalized linear autoregressive moving average* (GLARMA), para resolver o problema da autocorrelação serial dos dados.

AGRADECIMENTOS

Ao Instituto Estadual de Meio Ambiente (IEMA) e ao Hospital Infantil Nossa Senhora da Glória (HINSG), pelo fornecimento dos dados ambientais e hospitalares, respectivamente.

REFERÊNCIAS

1. Chen R, Chu C, Tan J, Cao J, Song W, Xu X, et al. Ambient air pollution and hospital admission in Shanghai, China. *J Hazard Mater*. 2010;181(1-3):234-40.
2. Friedman J. Multivariate adaptive regression splines. *Ann Stat*. 1991;19(1):1-67.
3. Gouveia N, Bremner SA, Novaes HM. Association between ambient air pollution and birth weight in São Paulo, Brazil. *J Epidemiol Community Health*. 2004;58(1):11-7. DOI:10.1136/jech.58.1.11
4. Johnson RA, Wichern DW. Applied multivariate statistical analysis. 6th ed. New Jersey: Prentice Hall; 2007.
5. Kedem B, Fokianos K. *Regression models for time series analysis*. 2th ed. Hoboken: Wiley; 2002. DOI:10.1002/0471266981
6. Martins L, Latorre MRDO, Saldiva PHN, Braga ALF. Relação entre poluição atmosférica e atendimentos por infecções de vias aéreas superiores no município de São Paulo: avaliação do rodízio de veículos. *Rev Bras Epidemiol*. 2001;4(3):220-9. DOI:10.1590/S1415-790X2001000300008
7. Roberts S, Martin M. Using Supervised Principal Components Analysis to Assess Multiple Pollutant Effects. *Environ Health Perspect*. 2006;114(12):1977-82.
8. Freitas C, Bremner SA, Gouveia N, Pereira LAA, Saldiva PHN. Internações e óbitos e sua relação com a poluição atmosférica em São Paulo, 1993 a 1997. *Rev Saude Publica*. 2004;38(6):751-7. DOI:10.1590/S0034-89102004000600001
9. Schwartz J. Harvesting and long term exposure effects in the relationship between air pollution and mortality. *Am J Epidemiol*. 2000;151(5):440-8.
10. Wahba G. Splines in nonparametric regression. *Encyclopedia of Environmetrics*. 2000.
11. Wang Y, Pham H. Analyzing the effects of air pollution and mortality by generalized additive models with robust principal components. *Int J Syst Assur Eng Manag*. 2011;2(3):253-9. DOI:10.1007/s13198-011-0072-7
12. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159(7):702-6. DOI:10.1093/aje/kwh09

Artigo baseado na dissertação de mestrado de Souza JB, intitulada: "Análise de Componentes Principais e a Modelagem Linear Generalizada: uma associação entre o número de atendimentos hospitalares por causas respiratórias e a qualidade do ar, na Região da Grande Vitória, ES", apresentada ao Programa de Pós-graduação em Engenharia Ambiental, Universidade Federal do Espírito Santo, em 2013. Os autores declaram não haver conflito de interesses.