



**MODELAGEM ESTATÍSTICA DA RELAÇÃO ENTRE POLUIÇÃO  
ATMOSFÉRICA E MORTALIDADE POR DOENÇAS RESPIRATÓRIAS**

**NILSON DE OLIVEIRA LEITE**

**2005**

**NILSON DE OLIVEIRA LEITE**

**MODELAGEM ESTATÍSTICA DA RELAÇÃO ENTRE POLUIÇÃO  
ATMOSFÉRICA E MORTALIDADE POR DOENÇAS RESPIRATÓRIAS**

Dissertação apresentada à Universidade Federal de Lavras como exigência do Programa de Pós-graduação em Agronomia, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Orientador

Prof. Dr. Augusto Ramalho de Moraes

LAVRAS

MINAS GERAIS – BRASIL

2005

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da  
Biblioteca Central da UFLA**

**Leite, Nilson de Oliveira**

Modelagem estatística da relação entre poluição atmosférica e mortalidade por  
doenças respiratórias / Nilson de Oliveira Leite. -- Lavras : UFLA, 2005.

64 p. : il.

Orientador: Augusto Ramalho de Moraes.

Dissertação (Mestrado) - UFLA.

Bibliografia.

1. Poluição do ar. 2. Doenças respiratórias. 3. Mortalidade. 4. Modelagem  
estatística. 5. Série temporal. 6. Autocorrelação. I. Universidade Federal de  
Lavras. II. Título.

CDD-519.537

**NILSON DE OLIVEIRA LEITE**

**MODELAGEM ESTATÍSTICA DA RELAÇÃO ENTRE POLUIÇÃO  
ATMOSFÉRICA E MORTALIDADE POR DOENÇAS RESPIRATÓRIAS**

Dissertação apresentada à Universidade Federal de Lavras como exigência o Programa de Pós-graduação em Agronomia, área de concentração em Estatística e Experimentação Agropecuária para a obtenção do título de "Mestre".

APROVADA em 10 de janeiro de 2005.

Prof. <sup>a</sup> . Dr. <sup>a</sup> . Eliane Scheid Gazire	PUC/BH
Prof. Dr. Júlio Silvio de Souza Bueno Filho	UFLA
Prof. <sup>a</sup> . Dr. <sup>a</sup> . Thelma Sáfadi	UFLA

  
Prof. Dr. Augusto Ramalho de Moraes

UFLA

(Orientador)

LAVRAS  
MINAS GERAIS – BRASIL

Ao meu querido pai, Eufrônio (“in memorian”)  
e minha querida mãe, Francisca, pela sabedoria e ensinamentos.

### **OFEREÇO**

Aos meus irmãos, Gilson (“in memorian”), Gilda e Raimundo Antônio, pelo  
apoio e amizade.

Aos sobrinhos Gilvaney, Gilney, Letícia e Luiz Otávio, pela amizade, ajuda e  
companheirismo.

À minha esposa, Maria José, minhas filhas, Ana Flávia e Raquel, pelo amor,  
incentivo e compreensão nas horas difíceis.

### **DEDICO**

## **AGRADECIMENTOS**

A Deus, pela realização desta etapa

À Universidade Federal de Lavras, pela oportunidade de realização do curso.

Aos orientadores, Augusto Ramalho de Moraes e Thelma Sáfydi, pela orientação, ensinamentos e amizade.

Aos professores do curso de Pós-Graduação do DEX, Júlio Sílvio, Paulo César, Joel Augusto, Lucas Chaves, Eduardo, Daniel, Thelma e Maria do Carmo, pela atenção, ensinamentos transmitidos durante o curso.

Às professoras Maria José da Rocha Ribeiro e Eliane Scheid Gazire, pela amizade e incentivo durante a realização do curso.

Aos amigos Antônio Assumpção, Nicinha Prado, Marcelo e Marcos Prado, pelo carinho, apoio e amizade.

Ao grande amigo e Professor Ladival Ignácio Pereira, que esteve presente em todos os momentos decisivos da minha vida.

Ao grande amigo Marcelo Ângelo Cirillo, pela amizade, orientação e auxílio no desenvolvimento deste trabalho.

Aos amigos que encontrei, Washington, José Waldemar, Janser, Fabyano e Denismar, pela atenção e apoio.

A Elisete Gomides, da Fundação Estadual do Meio Ambiente, pelo fornecimento dos dados.

Aos amigos Ester Ismênia, Edel de Oliveira, Fernanda, Maria Helena, Tia Helena, Tio Miguel, Fátima, Ligia, Alice, Jerry, Helenice e Amanda, pelo incentivo.

Aos médicos Claudete Araújo e Arturo Pacheco, pela atenção e ajuda na seleção dos dados.

Aos amigos de jornada, Taciana, Roberta, Imaculada, Gisele, Antônio, Regilson, Deive, Eric, Rômulo, Luciane, Mônica e Devanil.

## SUMÁRIO

LISTA DE TABELAS.....	i
LISTA DE FIGURAS.....	iii
RESUMO.....	v
ABSTRACT.....	vi
1 INTRODUÇÃO.....	1
2 REFERENCIAL TEÓRICO.....	3
2.1 Caracterização da Região Metropolitana de Belo Horizonte (RMBH).....	3
2.2 Principais fontes de poluição.....	4
2.3 Aspectos metodológicos.....	6
2.4 Modelos lineares.....	7
2.5 Modelagem dos erros por meio de séries temporais.....	9
2.5.1 Modelo auto-regressivo AR (p).....	10
2.5.2 Modelo de média móvel MA (q).....	11
2.6 Modelos lineares e diferentes estrutura de covariâncias.....	12
2.7 Função de autocorrelação.....	14
2.8 Teste de ruído branco.....	15
2.9 Critério de seleção do melhor modelo.....	16
2.10 Estimação dos parâmetros.....	17
3 MATERIAL E MÉTODOS.....	21
3.1 Material.....	21
3.1.1 Dados de mortalidade.....	21
3.1.2 Dados de poluição atmosférica e variáveis meteorológicas.....	22
3.2 Métodos.....	23
3.2.1 Análise exploratória dos dados.....	23
3.2.2 Modelagem estatística .....	24

4 RESULTADOS E DISCUSSÃO.....	28
4.1 Resultados da análise exploratória dos dados.....	28
4.2 Resultados da análise inferencial dos dados.....	37
4.2.1 Modelagem estatística.....	37
4.2.2 Modelo com inclusão dos poluentes.....	41
5 CONCLUSÃO.....	55
REFERENCIAL BIBLIOGRÁFICO.....	56
ANEXOS.....	59



## LISTAS DE TABELAS

TABELA 1 Valores máximos permitidos dos principais poluentes, considerados como padrão no Brasil e no Estados Unidos (EPA).....	5
TABELA 2 Modelos utilizados para a comparação do efeito dos poluentes SO <sub>2</sub> (dióxido de enxofre) e PM-10 (material particulado) nas estações de Betim e Contagem.....	27
TABELA 3 Valores do número de total de óbitos por causas respiratórias, outras causas e total e respectivas médias diárias, medianas diárias e desvio padrão do número diário de óbitos, para os anos de 2001 e 2002, na RMBH.....	28
TABELA 4 Valores médios, medianas diárias, mínimos, máximos, desvio padrão e número de ultrapassagens (NUP) da concentração do poluente PM-10 ( $\mu\text{g} \cdot \text{m}^{-3}$ ), nas estações de Betim e Contagem, anos de 2001 a 2002.....	32
TABELA 5 Valores da concentração médias diárias, medianas diárias, mínimos, máximos, desvio padrão e número de ultrapassagens (NUP), para o dióxido de enxofre SO <sub>2</sub> ( $\mu\text{g} \cdot \text{m}^{-3}$ ), nas estações de Betim e Contagem, nos anos de 2001 a 2002.....	34
TABELA 6 Valores médios, mínimos, máximos e desvio padrão de temperatura média diária (em °C) e umidade relativa do ar (em %), na RMBH, para os anos de 2001 a 2002.....	35
TABELA 7 Estimativas dos coeficientes do modelo estatístico com erros ARMA (0,3) e respectivos erros padrões e probabilidade de serem significativas (p-valor), estatística de Box & Pierce e critério de informação de Akaike (AIC).....	39
TABELA 8 Valores estimados para o critério de Akaike em função de diferentes modelos ARMA (p,q) pesquisados.....	41
TABELA 9 Estimativas dos coeficientes do modelo completo e respectivos erros padrões, p-valor, estimativa da estatística de Box & Pierce e do critério de informação de Akaike (AIC).....	43

TABELA 10 Estimativas dos parâmetros dos diversos modelos, respectivos erros padrões, estatística do teste de Box & Pierce e critério de informação de Akaike (AIC).....47

TABELA 11 Estimativas dos parâmetros do modelo AR (p), MA (q) e ARMA (p,q), erros padrões (ep), p-valores da estatística de Box & Pierce e critério de Akaike (AIC), quando são incluídas separadamente no modelo (18) as variáveis dióxido de enxofre e material particulado da Estação de Betim.....50

TABELA 12 Estimativas dos parâmetros do modelo AR (p), MA (q) e ARMA (p,q), erros padrões (ep), p-valores da estatística de Box & Pierce e critério de Akaike (AIC), quando são incluídas separadamente no modelo (18) as variáveis dióxido de enxofre e material particulado da Estação de Contagem.....51

## LISTA DE FIGURAS

- FIGURA 1 Representação gráfica na forma de box-plot do número diário de óbitos de pessoas acima de 60anos, considerando causas respiratórias, outras causas e total, na RMBH, 2001 – 2002.....29
- FIGURA 2 Representação gráfica do número de óbitos diário de pessoas com idade acima de 60 anos, considerando causas respiratórias, outras causas e todas as causas, na RMBH, 2001 - 2002.....30
- FIGURA 3 Representação gráfica na forma de histograma do número diária de óbitos para as pessoas acima de 60 anos, RMBH, no período de jan/2001 a dez/2002.....31
- FIGURA 4 Representação gráfica da concentração diária do poluente material particulado (PM-10), registrados nas estações de Betim e Contagem, da RMBH, no período de jan/2001 a dez/2002.....33
- FIGURA 5 Representação gráfica da concentração diária do poluente dióxido de enxofre  $SO_2$ , ( $\mu g.m^{-3}$ ), para as estações de Betim e Contagem, da RMBH, no período de jan de 2001 a dez de 2002.....35
- FIGURA 6 Representação gráfica da temperatura média diária e umidade relativa do ar (%), na RMBH, no período de jan/2001 a dez/2002,.....36
- FIGURA 7 Representação gráfica da função de autocorrelação e diagrama de dispersão dos resíduos para o modelo ARMA (0,3) para as variáveis temperatura e umidade relativa do ar.....40
- FIGURA 8 Representação gráfica da função de autocorrelação e diagrama de dispersão dos resíduos para o modelo completo ARMA (2,2).....44
- FIGURA 9 Representação gráfica da função de autocorrelação dos resíduos para avaliação do efeito do dióxido de enxofre ( $SO_2$ ) da estação de Betim para o modelo ARMA (2,2).....48
- FIGURA 10 Representação gráfica da função de autocorrelação dos resíduos para avaliação do efeito do material particulado (PM-10) da estação de Betim para o modelo ARMA (2,2).....48

FIGURA 11 Representação gráfica da função de autocorrelação dos resíduos para avaliação do efeito do dióxido de enxofre (SO<sub>2</sub>) da estação de Contagem para o modelo ARMA (2,2).....48

FIGURA 12 Representação gráfica da função de autocorrelação dos resíduos para avaliação do efeito do material particulado (PM-10) da estação de Contagem para o modelo ARMA (2,2).....48

FIGURA 13 Diagrama de dispersão dos resíduos para avaliação do efeito do dióxido de enxofre (SO<sub>2</sub>) na estação Betim para o modelo ARMA (2,2).....49

FIGURA 14 Diagrama de dispersão dos resíduos para avaliação do efeito do material particulado (PM-10) na da estação de Betim para o modelo ARMA (2,2).....49

FIGURA 15 Diagrama de dispersão dos resíduos para avaliação do efeito do dióxido de enxofre (SO<sub>2</sub>) na estação Contagem para o modelo ARMA (2,2).....49

FIGURA 16 Diagrama de dispersão dos resíduos para avaliação do efeito do material particulado (PM-10) na da estação de Contagem.....49

FIGURA 17 Representação gráfica da função de autocorrelação dos resíduos para dióxido de enxofre (SO<sub>2</sub>) utilizando o modelo AR (1) e para o material particulado (PM-10) utilizando o modelo ARMA (2,2) estação de Betim.....52

FIGURA 18 Representação gráfica da função de autocorrelação dos resíduos para dióxido de enxofre (SO<sub>2</sub>) utilizando o modelo AR (1) e para o material particulado (PM-10) utilizando o modelo ARMA (2,2) estação de Contagem.....52

FIGURA 19 Diagrama de dispersão dos resíduos para dióxido de enxofre (SO<sub>2</sub>) utilizando o modelo AR (1) e para o material particulado (PM-10) utilizando o modelo ARMA (2,2) estação de Betim.....53

FIGURA 20 Diagrama de dispersão dos resíduos para dióxido de enxofre (SO<sub>2</sub>) utilizando o modelo AR (1) e para o material particulado (PM-10) utilizando o modelo ARMA (2,2) estação de Contagem.....53

## RESUMO

**LEITE, Nilson de Oliveira. Modelagem estatística da relação entre poluição atmosférica e mortalidade por doenças respiratórias. 2005. 64 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, MG.\***

A relação entre mortalidade (número diário de óbito) por causas respiratórias, em pessoas com idade superior a 60 anos, e índice de poluição e ambientais, na Região Metropolitana de Belo Horizonte (RMBH), foi avaliada por meio de vários modelos estatísticos. Procurou-se verificar se as variáveis ambientais mês, ano, temperatura, umidade relativa do ar e os poluentes dióxido de enxofre e material particulado apresentam associação com o número de óbitos, utilizando modelos de regressão, tendo os parâmetros sido estimados pelo método dos quadrados mínimos generalizados. Foram considerados, ainda, estrutura de autocorrelação dos resíduos por meio dos modelos autoregressivo e de médias móveis. As análises foram realizadas utilizando-se de rotina implementada no programa R. A qualidade do ajuste dos modelos foi verificadas pelo teste de Box & Pierce e a seleção do melhor modelo pelo critério de Akaike e diagrama de dispersão dos resíduos. Verificou-se que o número de óbitos em pessoas com idade acima de 60 por doenças do aparelho respiratório é explicado de modo satisfatório por modelo que inclui as variáveis mes, ano, temperatura, umidade relativa do ar, material particulado (PM-10) medido na Estação de Contagem, material particulado (PM-10) e dióxido de enxofre (SO<sub>2</sub>) medido na Estação de Betim. Observaram-se efeitos significativos e positivos de umidade relativa do ar. O número de óbitos relaciona-se negativamente com os meses de novembro e dezembro, caracterizados como de temperaturas e umidades relativa do ar elevadas e positivamente com o mês de julho, caracterizado como de baixa temperatura.

---

\* Comitê Orientador: Prof. Dr. Augusto Ramalho de Moraes - UFLA  
(Co - orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Thelma Sáfyadi - UFLA)

## **Abstract**

**LEITE, Nilson de Oliveira. Statistical modeling of the relationship between atmospheric pollution and mortality from respiratory diseases. 2005. 64 p. Dissertation (Master's Degree in Statistics and Agricultural Experimentation) Federal University of Lavras, Lavras, MG. \***

The relationship between mortality (daily number of deaths) due to respiratory causes, in people over the age of 60, and the pollution index in the Metropolitan Area of Belo Horizonte (MABH) was evaluated through several statistical models. Using regression models, the parameters being estimated by the generalized least squares method, an attempt to verify if the environmental variables of month, year, temperature, relative humidity and the pollutants sulfur dioxide and particulate material present an association with the number of deaths. Also considered was the auto-correlation structure of the residues through the auto-regressive and moving averages models. The analysis was done using the routine implemented in the program R. The quality of the model adjustment was verified through the Box & Pierce test and the selection of the best model by the Akaike criterion and by the residue dispersion diagram. It was verified that the number of deaths in people over 60 from respiratory system diseases is explained in a satisfactory way by a model that includes the variables of months, years, meteorological variables (temperature and relative humidity) and particulate material (PM-10) measured at the Contagem Station and particulate material (PM-10) and sulfur dioxide (SO<sub>2</sub>) measured at the Betim Station. Significant and positive effects of the relative humidity of the air was observed. The number of deaths relate negatively to the months of November and December, characterized by elevated temperatures and relative humidities and positively with the month of July, characterized by low temperatures.

---

\* Guidance committee: Prof. Dr. Augusto Ramalho de Moraes - UFLA  
(Co - Advisor: Profa. Dra. Thelma Sáfyadi - UFLA)

# 1 INTRODUÇÃO

Em algumas áreas da pesquisa científica, pode-se encontrar as situações especiais de estudo, como é o caso de relações entre variáveis, tais como a taxa de mortalidade e índices de poluição e meteorológica.

A modelagem dessas relações nem sempre é fácil, pois essas variáveis possuem características bem distintas, umas podem ser ajustada por uma distribuição normal, outras não, como é o caso do número de óbitos.

A modelagem da associação entre poluição atmosférica e mortalidade em centros urbanos requer a utilização de métodos de ajustes apropriados e modelos adequados para a estimação de parâmetros.

Os modelos estatísticos constituem ferramentas extremamente úteis para resumir e interpretar dados que podem facilitar a avaliação da forma e da intensidade de associações de interesse em estudos epidemiológicos.

Como o número de óbitos diários em uma região e a concentração de poluentes, em geral, não seguem uma distribuição normal, então, as técnicas estatísticas usuais baseadas em um modelo de Gauss-Markov não são as mais adequadas para o estudo de relações entre elas.

Assim, surgem como alternativa aos modelos lineares clássicos a utilização de transformação de dados, de modelos lineares generalizados (McCullagh & Nelder, 1989) e de modelos aditivos generalizados (Hastie & Tibshirani, 1990), que são adequados para a análise de dados cuja distribuição pertence à família exponencial.

Quando se avaliam dados de contagem que, em geral, seguem uma distribuição de Poisson e, se esses dados são obtidos ao acaso do tempo, convém considerar no modelo uma estrutura de autocorrelação nos resíduos (Fokianos & Kedem, 2004).

A falta de informação no que se refere à relação entre poluição atmosférica e doenças respiratórias no município de Belo Horizonte motivaram a verificação dessa associação.

Em virtude do que foi mencionado, este trabalho teve por objetivo inferir o comportamento da mortalidade em pessoas com idade acima de 60 anos, em função de variáveis climáticas e de poluentes. De modo específico procurou-se avaliar se as variáveis ano, mês, temperatura e umidade relativa do ar e os poluentes: dióxido de enxofre e material particulado afetam o número de óbitos por doenças respiratórias, por meio de modelos lineares, considerando diversas estruturas de autocorrelação.



## **2 REFERENCIAL TEÓRICO**

### **2.1 Caracterização da Região Metropolitana de Belo Horizonte (RMBH)**

A Região Metropolitana de Belo Horizonte, aqui denominada por RMBH (Figura-1A), é constituída por 34 municípios com uma população estimada em 4.357.942, entre homens e mulheres e ocupando uma área de aproximadamente uma área de 9.459,10 (Tabelas 1B e 2B).

A região possui uma frota de 668.662 veículos de diversos tipos, segundo a Tabela-3B.

As principais atividades econômicas da RMBH, podem ser agrupadas em quatro setores:

- a) Produção agrícola: mandioca, banana, laranja, milho, feijão, arroz, hortifrutigranjeiros e cana-de-açúcar;
- b) pecuária: leite, gado de corte e suinocultura;
- c) mineração: extração de minério de ferro, areia, argila, pedra de calçamento, cimento, cal, brita, cascalho e pedra cristal;
- d) indústria: siderurgia, cerâmica, peças de veículos automotores, metalurgia, mecânica, eletro-mecânica, fundição, produção de ferro gusa e indústria de transformação.

Como vias principais de escoamento dos bens de consumo e de transformação, a RMBH, tem as rodovias estaduais MG-020, MG-060, MG-50, MG-424 e as federais BR-040, BR-262 e BR-381.

## 2.2 Principais fontes de poluição

As áreas de proteção ambiental e, principalmente, aquelas do cerrado, estão passando por temporadas de incêndios, provocadas pelo uso das queimadas com o intuito de manejo (errôneo) do solo para as plantações e produção de carvão, o qual, por sua vez, alimenta as indústrias de transformação e siderurgias que estão localizadas na RMBH.

A poluição atmosférica é constituída basicamente de duas fontes poluidoras: fixa e naturais. Fontes fixas são aquelas que acontecem com forte influência da ação do homem, como fábricas, veículos automotores, queimadas e atividades comerciais. As fontes naturais, que também produzem poluição atmosférica, são os ventos e a baixa umidade relativa do ar, os quais provocam a combustão espontânea, erupção vulcânica, descargas elétricas e os processos de decomposição espontânea.

Os motores de combustão interna produzem o monóxido de carbono, gás inodoro e altamente prejudicial à saúde, hidrocarbonetos e óxido de nitrogênio que, sob a influência da luz do sol, formam o ozônio, um agente irritante e poluente. Muitas vezes, o monóxido de carbono é produzido em excesso, devido aos gases emitidos pelos veículos automotores, como, por exemplo as motocicletas que, devido ao baixo custo, têm aumentado em número, faltando, contudo, de fiscalização e regulamentação adequadas.

Outras fontes de poluição são as fábricas de cimento, as recauchutagens de pneus, produção de ferro gusa e outros processos industriais que emitem dióxido de enxofre e óxido de nitrogênio que, combinados com a umidade relativa do ar, criam a chuva ácida.

Os atos regulamentares de poluição no Brasil vêm sendo adotados desde 1976. A resolução do Conselho Nacional do Meio Ambiente (CONAMA) nº 3,

de 1990, estabelece padrão de qualidade do ar dos principais agentes causadores de poluição atmosférica. Os principais índices estão apresentados na Tabela 1

TABELA 1 Valores máximos permitidos dos principais poluentes, considerados como padrão no Brasil e no Estados Unidos (EPA).

Poluentes	Brasil ( $\mu\text{g.m}^{-3}$ )		EPA (EUA) ( $\mu\text{g.m}^{-3}$ )	
	Primário	Secundário	Primário	Secundário
Dióxido de enxofre ( $\text{SO}_2$ )	365	80	365	80
Monóxido de carbono (CO)	35 ppm	9 ppm	35 ppm	9 ppm
Material particulado (PM-10)	150	50	150	50
Dióxido de nitrogênio ( $\text{NO}_2$ )	320	100	100	-
Ozônio ( $\text{O}_3$ )	160	-	0,12 ppm	0,08 ppm

Fonte: Resolução CONAMA nº 3 de 28/06/90 e EPA

Estes valores são classificados em padrões primários: que são as concentrações de poluentes atmosféricos que, se ultrapassadas, poderão afetar a saúde da população, podendo ser entendidas como os níveis máximos toleráveis e os padrões secundários são as concentrações de poluentes atmosféricos abaixo dos quais se prevê o mínimo efeito adverso sobre o bem-estar da população, assim como o mínimo dano à fauna, à flora e ao meio ambiente em geral.

Os poluentes que causam danos à saúde, de acordo com Zurita & Tolfo (2000) são:

- monóxido de carbono: quando ultrapassados os padrões primários e exposição prolongada pode provocar dores de cabeça, afetar a visão e provocar efeitos no coração;
- dióxido de enxofre: exposições prolongadas provocam intoxicação, inflamação nas mucosas e problemas respiratórios;

- material particulado: tal poluente vem sendo associado ao aumento de mortes totais em idosos e crianças, por doenças cardiovasculares e respiratórias.

Com um fluxo intenso de veículos automotores, alinhado a indústrias mecânicas, mineração e de transformação e siderurgias instaladas, a Região Metropolitana de Belo Horizonte, é um lugar apropriado para o desenvolvimento de estudos sobre os efeitos da poluição atmosférica sobre a saúde de sua população.

### **2.3 Aspectos metodológicos**

Neste tipo de estudo, o interesse é avaliar a associação entre poluição atmosférica e mortalidade ou internação (Dockery & Pope III, 1994), mas, ambas podem ser influenciados por fatores meteorológicos. Em geral, a mortalidade é maior em baixas temperaturas, principalmente em idosos. O mesmo ocorre com os níveis de concentração dos poluentes já que, em baixa umidade relativa do ar e temperatura, desfavorecem a dispersão das partículas.

Assim, as variáveis meteorológicas devem, portanto, ser consideradas também como variáveis explicativas em vários modelos sob investigação.

A poluição atmosférica e as variáveis meteorológicas, em geral, apresentam tanto tendência quanto sazonalidade, principalmente quando o período é muito grande (vários anos ou meses). Para controlar esses fatores, diversas alternativas podem ser utilizadas. Uma abordagem é a utilização de variáveis indicadoras para os meses do ano e ano (Gouveia & Fletcher, 2000). Diversas metodologias são utilizadas como modelos lineares generalizados por Schawrtz & Dockery, (1992) e Schawrtz, (1999) e modelos aditivos generalizados por Dominici & Zeger, (2000) e Lima (2001).

Observações realizadas em dias consecutivos na mesma população podem estar correlacionadas, o que contraria as suposições dos modelos de regressão (Hoffman & Vieira, 1998).

Alternativas para lidar com erros correlacionados, incluem o uso de modelos auto-regressivos, médias móveis ou auto-regressivo de médias móveis (Zeger, 1998; Fokianos & Kedem, 2004; Morettin & Toloi, 2004).

## 2.4 Modelos lineares

Os modelos lineares são usados na análise de conjunto de dados, cujos erros experimentais seguem uma distribuição normal, e foram discutidos em muitos textos de estatística, como em Graybill (1961) e Searle (1971), entre outros.

Esses autores consideram que um modelo linear pode ser definido por:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (1)$$

no qual :

$\mathbf{Y}$  é o vetor com valores observados, de dimensões  $n \times 1$ ;

$\mathbf{X}$  é matriz do delineamento, de dimensões  $n \times p$ ;

$\boldsymbol{\beta}$  é o vetor de parâmetros desconhecidos, dimensões  $p \times 1$ ;

$\boldsymbol{\epsilon}$  é o vetor de erros aleatórios, de dimensões  $n \times 1$ ,

Esse modelo pode assumir várias caracterizações de acordo com o planejamento realizado e descrito na matriz  $\mathbf{X}$ , cuja finalidade é modelar o vetor  $\mathbf{Y}$ , utilizando-se os parâmetros de  $\boldsymbol{\beta}$ . Em geral, considera-se que o vetor  $\boldsymbol{\epsilon}$  é constituído de variáveis aleatórias independentes e distribuídas com média  $\mathbf{0}$  e

variância  $\sigma^2$ . No caso de distribuição normal, tem-se  $\epsilon \sim N(0, \sigma^2 I)$ , o que implica em  $Y \sim N(X\beta, \sigma^2 I)$ .

Conhecendo-se a distribuição de probabilidade do erro aleatório, pode-se fazer uso do método da máxima verossimilhança (McCulloch & Searle, 2000) para se obter os estimadores dos parâmetros. Assim, a função log-verossimilhança é dada por:

$$L(\beta, \sigma^2 / y) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}N}} \exp\left[-\frac{1}{2}(y - X\beta)^T \frac{1}{\sigma^2}(y - X\beta)\right] \quad (2)$$

As estimativas de máxima verossimilhança dos parâmetros desconhecidos podem ser obtidas resolvendo-se o sistema de equações normais:

$$X^T X \beta = X^T Y$$

Para a obtenção de  $\beta$ , têm-se dois casos:

a) Quando a matriz  $X$  é de posto coluna completo,  $X^T X$  é não singular, e a solução única é dada por:  $\hat{\beta} = (X^T X)^{-1} X^T Y$  e a matriz de variância e covariância é  $V(\hat{\beta}) = (X^T X)^{-1} \sigma^2$

b) A matriz  $X$  é de posto incompleto,  $X^T X$  é singular e o sistema de equações normais é sempre consistente, possuindo várias soluções. Entre as alternativas possíveis, pode-se utilizar as inversas generalizadas e, nesses casos, as soluções podem ser obtidas por:  $\beta^o = (X^T X)^o X^T Y$

sendo  $\beta^o$  uma das soluções e  $(X^T X)^o$  uma matriz inversa generalizada de  $X^T X$ .

Nesse caso, não faz sentido obter  $V(\beta^o) = (X^T X)^o \sigma^2$ , pois ela não é única, variando de acordo com a matriz inversa generalizada.

## 2.5 Modelagem de erros por meio de séries temporais

As bases teóricas do estudo de séries temporais foram elaboradas a partir de Wold (1938), citado por Harvey (1994), que relacionou os modelos auto-regressivos AR (p) e de médias móveis MA (q) à teoria matemática em diferentes campos do conhecimento, com aplicações em ciências econômicas, ciências agrárias e sociais e previsões.

Essa metodologia teve maior desenvolvimento a partir do trabalho de Box & Jenkins (1976), que apresentaram, em um único texto, os conceitos básicos e os principais modelos. Maiores detalhes e outras informações podem ser encontrados em Pankratz (1983), Wei (1990), Fuller (1996) e Morettin & Tolo (2004).

Box & Jenkins (1976) perceberam a relação entre média móvel e o modelo auto-regressivo e propuseram a metodologia dos modelos auto-regressivos de médias móveis, que denominaram modelo ARMA (p,q).

Uma das suposições de um processo estacionário de uma série temporal é que ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, repercutindo de alguma maneira em um equilíbrio estável.

Uma justificativa para tal modelo está no fato de que, muitas vezes, as características dos dados de um determinado processo não permitem uma descrição satisfatória do processo puramente de média móvel ou puramente auto-regressivo. Isso implica que as estimativas dos parâmetros ou seleção do melhor modelo que represente os dados, requer em a inclusão de ambos os processos no modelo, o que pode resultar em um modelo com menor número de parâmetros do que seria necessário para descrever satisfatoriamente o fenômeno que está sendo pesquisado.

Um modelo misto auto-regressivo e de média móvel (ARMA) pode ser caracterizado, segundo Morettin & Toloí (2004), como aquele que considera as seguintes características:

- i)  $X_t$  é uma seqüência de variáveis aleatórias indexadas no tempo  $t$ ;
- ii)  $e_t$  é um ruído branco com as seguintes propriedades:

$$E(e_t) = 0, E(e_t^2) = \sigma^2 (\text{constante}) \text{ e } \text{cov}(e_t, e_r) = 0, t \neq r.$$

O operador de defasagem (ou de retardo),  $B$  é definido por:  $BX_t = X_{t-1}$ , no qual  $B^n X_t = X_{t-n}$ , e a primeira diferença da seqüência é definida por:  $\Delta X_t = X_t - X_{t-1}, t = 1, 2, 3, \dots, n$ ,

Assim,  $\Delta X_t$  pode ser escrito por  $\Delta X_t = X_t - BX_t = (1 - B)X_t$ ,

### 2.5.1 Modelo auto-regressivo AR (p)

Quando a série observada  $X_t$  pode ser gerada por uma média ponderada de somente as “p” primeiras observações próximas passadas da variável, acrescidas de perturbações aleatórias ( $e_t$ ) do período presente, tem-se o processo auto regressivo de ordem p representado por AR (p), que é um caso especial do modelo ARMA (p,q), para q = 0.


De forma geral um modelo AR (p), pode ser escrito na forma

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t \quad (3)$$

no qual  $\phi_1, \phi_2, \dots, \phi_p$  são parâmetros a serem estimados e ( $e_t$ ) é o resíduo ou desvio associado à t-ésima observação.

O processo de auto-regressivo pode também ser escrito da seguinte maneira:





$\phi(B)X_t = e_t$ , em que  $\phi(B)$  é o polinômio auto regressivo, dado por:

$$\phi(B) = 1 - BX_t - B^2X_t - \dots - B^pX_t. \quad (4)$$

### 2.5.2 Modelo de média móvel MA (q)

No caso em que a série observada  $X_t$  é gerada por uma média ponderada do presente e “q” valores passados de um processo de ruído branco, o processo de média móvel, MA (q), de ordem q, é um caso especial do modelo ARMA (p,q), com  $p = 0$ .

De maneira geral, um modelo MA (q) pode ser escrito na forma:

$X_t = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$ , em que  $\theta_1, \theta_2, \dots, \theta_q$  representa um conjunto finito de parâmetros a serem estimados e,  $(e_t)$  é o resíduo associado à t-ésima observação.

Pode-se reescrever o processo de média móvel da seguinte maneira:

$X_t = \theta_q(B)e_t$ , no qual  $\theta_q(B)$  é o polinômio de média móvel, dado por:

$$\theta(B) = (1 - Be_t - B^2e_t - \dots - B^qe_t) \quad (5)$$

Um modelo ARMA (p,q) é, portanto, definido por  $\phi_p(B)X_t = \theta_q(B)e_t$

## 2.6 Modelos lineares e diferentes estruturas de covariâncias

Avaliar a relação entre o efeito da poluição sobre a saúde humana não é uma tarefa fácil, pois os níveis de poluição comumente encontrados nos grandes centros urbanos são menores do que o estabelecido pelo Conselho Nacional do Meio Ambiente (CONAMA) e, portanto, apresentam maiores dificuldades de percepção e medição.

Alem disso, fatores como variações climáticas e problemas técnicos de aparelhos apresentam maiores dificuldades nos sistema de medição e, conseqüentemente, um esforço analítico e metodológico é indispensável.

Quando os dados são coletados diariamente, possivelmente os erros tendem a ser correlacionados, isto porque os valores obtidos em dias consecutivos têm uma associação maior do que aqueles obtidos em dias mais distantes. Tal correlação pode ser modelada por meio de uma estrutura de covariâncias. Nesses casos, a idéia básica é estimar os parâmetros do modelo, associando diferentes estruturas de correlação nos resíduos.

Dessa forma, têm-se os modelos auto-regressivos. Os mais comuns são os auto-regressivos AR (p), médias móveis MA (q) e modelos auto-regressivos e de médias móveis ARMA (p,q) (Morettin & Toloí, 2004).

Para a obtenção das estimativas dos parâmetros provenientes desses modelos são utilizados os métodos:

i) Método dos quadrados mínimos ordinários, utilizado quando são parte de uma seqüência de variáveis aleatórias possui erros homocedásticos e não correlacionados, ou seja,  $\varepsilon \sim N(0, \sigma^2)$  e, ainda, satisfazendo às pressuposições do modelo linear clássico de aditividade dos efeitos, normalidade, variância constante e independência dos erros.

ii) Método dos quadrados mínimos ponderados, utilizado quando uma seqüência de variáveis aleatórias possui heterogeneidade de variâncias, sendo mais eficiente por fornecer estimadores não tendenciosos (Hoffman & Vieira, 1998; Draper & Smith, 1998).

Nesse caso, o erro têm distribuição  $\varepsilon \sim N(0, V\sigma^2)$ , sendo V uma matriz diagonal, positiva definida, que pondera a variância  $\sigma^2$ .

Assim, a minimização da soma de quadrados dos erros  $\varepsilon'\varepsilon$  fornece o sistema de equações normais (SEN) dado por:  $X'V^{-1}X\hat{\beta} = X'V^{-1}Y$  e que tem como solução  $\hat{\beta} = (X'V^{-1}X)^{-1} X'V^{-1}Y$ .

iii) Quadrados mínimos generalizados, utilizados, segundo Hoffman & Vieira (1998), quando a estrutura de erros considera a existência de dependência entre eles, ou seja, os erros podem estar correlacionados. Assim,  $\varepsilon_i \sim N(0, W\sigma^2)$ , sendo W uma matriz positiva definida e simétrica, contendo as variâncias e covariâncias.

De acordo com Souza (1998) e Fox (2002), o método de quadrados mínimos generalizados é uma abordagem para o estudo de modelos, com estrutura de erros associados aos resíduos dos modelos auto-regressivos AR (p), médias móveis MA (q) e modelos auto-regressivos e de médias móveis ARMA (p,q).

## 2.7 Função de autocorrelação

No estudo de série de valores medidos seqüencialmente no tempo, uma característica dos resíduos é a existência de uma relação entre o erro da observação relativa a um período estar correlacionado com as observações anteriores, justificando assim o termo auto-regressivo.

Segundo Morettin & Toloí (2004), o coeficiente de autocovariância com defasagem  $k$ , sendo  $k$  o número de intervalos de tempo defasados na série de valores  $X_t$ , representa a covariância entre  $X_t$  e  $X_{t+k}$ , ou seja:

$$\gamma_k = \text{cov}(X_t, X_{t+k}) = E[(X_t - \mu_t)(X_{t+k} - \mu_{t+k})] \quad (6)$$

é a autocovariância entre dois pontos quaisquer ou o valor esperado do produto do desvio de cada ponto em relação à média do processo.

Outra medida importante é a autocorrelação, que é obtida quando divide-se a autocovariância pela variância, ou seja,  $\rho_k = \gamma_k / \gamma_0$ , com  $\gamma_k$  sendo a autocovariância,  $\gamma_0$ , a variância e  $\rho_k$  o coeficiente de autocorrelação, que mostra a relação linear que existe entre os valores da variável espaçados em uma distância de  $k$  defasagens.

A função que relaciona cada distância de  $k$  com seu correspondente coeficiente  $\rho_k$  é definida como função de autocorrelação.

A seqüência  $\rho_k$  com  $k = 0, 1, 2, 3, \dots$  é uma indicação do intervalo para o qual um valor do processo está correlacionado com seus valores, defasados  $k$  períodos de tempo. A função de autocorrelação da série, portanto, torna-se importante à medida que proporciona graus de intensidade da correlação existente entre observações vizinhas da série.

Um processo estacionário é caracterizado pelo fato das observações oscilar em torno de uma média e possuir as seguintes propriedades (Morettin & Tolo, 2004):

i)  $\gamma_k > 0$ ;

ii)  $\gamma_{-k} = \gamma_k$ ;

iii)  $\gamma_0 \geq |\gamma_k| \quad \forall k$

Nota-se que, de acordo com a propriedade (ii) a função de autocorrelação é simétrica em relação ao gráfico da mesma, que se chama de correlograma. Basta incluir a parte não negativa, que pode ser obtida colocando-se em um gráfico os valores de  $\rho_k$ , em função dos níveis de defasagem (lags).

## 2.8 Teste de ruído branco

Na literatura são propostos vários testes que geralmente são aplicados aos resíduos. Particularmente nos modelos de séries temporais, utiliza-se o teste de Box & Pierce (1970). Convém salientar que a verificação do ruído branco pode ser feita graficamente por meio da função de autocorrelação.

A estatística do teste, segundo Box & Pierce (1970), é dada por:

$$Q_c(K) = N \sum_{j=1}^K \hat{\rho}_j^2 \quad (7)$$

com uma distribuição aproximada de qui-quadrado  $\chi^2$ , com k-p-q número de graus de liberdade, em que k representa o numero de lags que, em geral, está situado entre 10 a 15; p é a ordem do modelo auto-regressivo; q é a ordem do

modelo de médias móveis,  $N$  o número de observações da série temporal e  $\hat{\rho}_j$  estimativa de autocorrelação.

Uma modificação desse teste, proposta por Ljung & Box (1978), considera a estatística.

$$Q_c(K) = N(N+2) \sum_{j=1}^K \frac{\hat{\rho}_j^2}{N-j} \quad (8)$$

A regra de decisão do teste é dada de tal forma que quando o valor de  $Q_c(K)$  for menor do que o valor crítico da distribuição qui-quadrado  $\chi^2$ , com  $K-p-q$  número de graus de liberdade a um nível de significância  $\alpha$ , tem-se a indicação de que o modelo em pesquisa está bem ajustado aos dados, ou seja, o resíduo não é correlacionado, sendo independente e identicamente distribuído. Logo, é ruído branco.

## 2.9 Critério de seleção do melhor modelo

Um critério para a seleção de modelos é o critério de informação de Akaike, denotado de AIC (Akaike, 1974). O método é aplicado a modelos cujos parâmetros são estimados pelo do método da máxima verossimilhança e a fórmula para modelo ARMA (p,q), de acordo com Hannan (1980) e Morettin & Tolo (2004):

$$AIC_{p,q} = N \ln(\hat{\sigma}^2) + 2(p+q+2), \quad (9)$$

em que:

$\hat{\sigma}^2$  é a variância do erro estimado pelo método da máxima verossimilhança;

$p$  é a ordem do modelo auto-regressivo;

$q$  é a ordem do modelo de média móvel;

$N$  é o número de observações.

Este método consiste na comparação de modelos passo-a-passo, nos quais a cada passo ocorre a inclusão ou exclusão de uma variável no modelo, buscando minimizar a estatística AIC. O modelo que apresentar menor valor para o AIC será escolhido como sendo o melhor.

## **2.10 Estimação dos parâmetros**

Após a etapa de identificação da ordem do modelo que melhor representa o fenômeno em estudo, torna-se necessária a obtenção de estimativas dos parâmetros do modelo identificado.

Na estimação dos parâmetros de um modelo, Mood et al. (1974) consideram que podem ser empregados os métodos de máxima verossimilhança, de mínimos quadrados ou o método dos momentos. Desses métodos, comentam que o mais comumente utilizado é o de máxima verossimilhança.

No contexto de modelos auto-regressivos, entre os métodos utilizados para a estimação dos parâmetros, o mais citado por Morettin & Tolo (2004) é o método da máxima verossimilhança, cujo procedimento basicamente consiste em resolver as equações de Yule & Walker, as quais, matricialmente, são representadas conforme a expressão:

$$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \cdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdots \\ \rho_p \end{bmatrix} \quad (10)$$

nas quais  $\phi_i$  é o  $i$ -ésimo coeficiente do modelo auto regressivo e  $\rho_i$  é o  $i$ -ésimo coeficiente de autocorrelação, com  $i = 1, 2, 3, \dots, p$ .

Convém salientar que a resolução dessas equações é feita mediante a um procedimento recursivo para ordens sucessivas  $p = 1, 2, 3, \dots$ . Este procedimento é conhecido como algoritmo de Durbin & Levinson (Morettin & Toloí, 2004).

Os valores iniciais a serem utilizados na execução deste procedimento provêm de estimativas preliminares dos parâmetros do modelo identificado. Estas estimativas são obtidas por meio das autocorrelações amostrais de uma série.

O processo para a obtenção das estimativas, de acordo com Morettin & Toloí (2004), é o seguinte:

i) Para o processo AR ( $p$ ), resolvem-se as equações de Yule & Walker (10), em geral, considerando-se os parâmetros como  $\phi_1, \phi_2, \dots, \phi_p$ , que podem ser obtidos de (10), ou seja, se  $\rho_1, \rho_2, \dots, \rho_p$  são conhecidos e, em geral, os são, pois podem ser estimadas por meio da função de autocorrelação amostral, então:

$$\rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \cdots + \phi_p \rho_0 \quad (11)$$

e

$$\sigma_a^2 = \gamma_0 (1 - \phi_1 \rho_1 - \cdots - \phi_p \rho_p) \quad (12)$$

é a variância do erro;



ii) Para o processo MA (q), utiliza-se a equação em incógnitas,  $\theta_1, \theta_2, \dots, \theta_q$  e uma estimativa para a variancia  $\hat{\sigma}_a^2$  é obtida de  $\gamma_0 = \sigma_a^2 (1 + \theta_1^2 + \dots + \theta_q^2)$ ;

iii) Para os processos ARMA (p,q), obtêm-se estimativas iniciais para  $\phi_1, \phi_2, \dots, \phi_p$ , resolvendo-se

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \dots + \phi_p \rho_{j-p} \quad (13)$$

para  $j = q + 1, \dots, q + p$ .

Mood et al. (1974) consideram que, para amostras grandes, os estimadores da máxima verossimilhança possuem as seguintes propriedades assintóticas:

- eficiência, com menor variância que qualquer outro estimador;
- consistência;
- no caso de haver um outro estimador suficiente, ele será produzido pelo estimador de máxima verossimilhança;
- ser normalmente distribuído, com facilidade para a obtenção dos parâmetros média e variância.

Segundo Souza (1998), com o modelo de regressão múltipla:

$$Y_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ji} + \varepsilon_i, \quad (14)$$

pode-se ajustar para  $\varepsilon_i$  :

- um modelo auto-regressivo de ordem p AR(p), por meio de:

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \dots + \phi_p \varepsilon_{t-p} + e_t, \quad (15)$$

no qual  $e_t \sim N(0, \sigma_e^2)$

• um modelo de média móvel de ordem q MA(q), por meio de:

$$\varepsilon_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} + e_t, \quad (16)$$

no qual  $e_t \sim N(0, \sigma_e^2)$

Outro modelo possível é o modelo misto auto-regressivo e de médias móveis ARMA (p,q), por meio de:

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \dots + \phi_p \varepsilon_{t-p} + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} + e_t, \quad (17)$$

no qual  $e_t \sim N(0, \sigma_e^2)$

O processo de estimação ARMA(p,q) segundo Morettin & Tolo (2004), consiste em achar um vetor de parâmetros auto-regressivos  $\phi_p = (\phi_1, \phi_2, \phi_3, \dots, \phi_p)$  e um vetor de médias móveis  $\theta_q = (\theta_1, \theta_2, \theta_3, \dots, \theta_q)$ , tais que minimizem a soma de quadrados de erros, sendo esses, obtidos pela diferença entre os valores observados e os esperados. Os valores esperados são determinados utilizando-se as estimativas dos parâmetros  $\hat{\phi}_p = (\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \dots, \hat{\phi}_p)$  e  $\hat{\theta}_q = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \dots, \hat{\theta}_q)$ . Logo, deve-se achar um valor para  $S(\phi, \theta) = \sum \varepsilon_t^2$ , que seja mínimo.

## 3 MATERIAL E MÉTODOS

### 3.1 Material

#### 3.1.1 Dados de mortalidade

O Sistema de Informação sobre Mortalidade da Diretoria de Processamento e Monitoramento de Dados Epidemiológicos da Secretaria de Estado da Saúde (SES/SE/DPMDE/SIM/MG), é um órgão governamental com a finalidade de gerar registros e estatísticas confiáveis de mortalidade no Estado.

Este órgão mantém arquivos que são atualizados periodicamente contendo informações existentes na Declaração de Óbitos. O procedimento adotado após um falecimento exige que a declaração de óbito seja preenchida criteriosamente e assinada por um médico, contendo entre outras informações, o código do município de ocorrência e o tipo de óbito.

Foram obtidos da DPMDE, os dados referentes a todos os óbitos ocorridos nos anos de 2001 a 2002, no estado de Minas Gerais. Desses, foram selecionados somente os óbitos da Região Metropolitana de Belo Horizonte (RMBH), de acordo com o código do município de ocorrência.

Foram excluídos, para este estudo, todos os óbitos por causas não naturais, como aquelas por acidentes de qualquer natureza ou homicídios. Demais foram classificados segundo a faixa etária de interesse e a causa básica de morte. A faixa etária de interesse, neste estudo, foi acima de 60 anos de idade.

Para a realização deste trabalho, foi considerado como variável resposta ou de interesse o número de óbitos em pessoas com idade acima de 60 anos, cuja causa básica, foram as doenças do aparelho respiratório, de acordo com o Código Internacional de Doenças nº 10 (CID-10) (Edusp, 1993).



### 3.1.2 Dados de poluição atmosférica e variáveis meteorológicas

O monitoramento da qualidade do ar na RMBH é realizado com equipamentos automáticos, denominados fonte fixa e é coordenado pela Divisão de Qualidade do Ar da Fundação Estadual do Meio Ambiente (FEAM).

As estações são constituídas por cabinas climatizadas, equipadas com analisadores e sensores que executam a amostragem do ar atmosférico e fornecem as concentrações dos poluentes.

Os dados referem-se ao período de monitoramento de 2001 a 2002, em três estações, que são:

- Estação de Belo Horizonte: localizada na Praça Rui Barbosa em terreno da Prefeitura Municipal de Belo Horizonte (PMBH). O local situado próximo à linha de transporte ferroviário e metrô, é caracterizado por fluxo intenso de automóveis e ônibus urbanos;

- Estação de Contagem: está localizada na avenida Babita Camargos, em área externa do Departamento Nacional de Estradas de Rodagem (DNER), em local próximo à Praça dos Trabalhadores e da Companhia Energética de Minas Gerais (CEMIG). A região é circundada por diversas atividades industriais, o que gera poluição de origem industrial, concentra grande quantidade de veículos leves e pesados e há, ainda, na região, um fluxo intenso de pessoas;

- Estação de Betim foi inicialmente instalada na área externa da Escola Municipal Valério Ferreira Palhares, no Bairro Jardim das Alterosas, em área de Administração da Prefeitura Municipal de Betim. A região é residencial, apresenta fluxo moderado de veículos e fica próxima a várias unidades industriais.

As três estações são responsáveis pela medida das concentrações dos seguintes poluentes: material particulado, dióxido de enxofre, monóxido de

carbono e dos seguintes parâmetros meteorológicos: direção e velocidade dos ventos, temperatura e umidade relativa do ar.

Neste estudo considerou-se somente a medida de concentração diária do poluente material particulado (PM-10) e do dióxido de enxofre (SO<sub>2</sub>) das estações de Betim e Contagem, uma vez que os dados são mais representativos. Os parâmetros meteorológicos temperatura média diária (°C) e umidade relativa do ar diária (%), também foram computados. Foram desconsiderados os dados da Estação de Belo Horizonte, pois estavam faltantes.

O período de estudo considerado apresentou somente 585 observações, em função da falta de medições eventuais.

## **3.2 Métodos**

### **3.2.1 Análise exploratória dos dados**

Com a finalidade de mostrar a localização específica da Região Metropolitana de Belo Horizonte em Minas Gerais, apresenta-se, na Figura 1A o mapa da RMBH. A população, a área e a densidade demográfica, constando na Tabela 1B; a população existente na Tabela 2B e o número de veículos por categoria em circulação da RMBH na Tabela 3B.

Foram calculados e utilizados como medidas descritivas (média, mediana, desvio padrão, valores mínimo e máximo) os números diários de óbitos por doenças do aparelho respiratório em pessoas com idade acima de 60 anos no período de janeiro de 2001 até dezembro de 2002, temperatura média diária (° C), umidade relativa diária do ar (%) e os poluentes atmosféricos material particulado (PM-10) e dióxido de enxofre (SO<sub>2</sub>)

A partir dos dados coletados sobre número de óbitos, temperatura, umidade, material particulado e dióxido de enxofre, foram construídos gráficos de linha, diagrama box-plots e histogramas, para representar o comportamento geral dos dados, fornecendo idéia sobre posição, dispersão, simetria e dados periféricos – *outliers*.

A importância da análise exploratória de dados é caracterizar a distribuição da variável resposta ou de interesse, o número de óbitos em pessoas acima de 60 anos e utilizar esta informação na modelagem estatística dos dados.

Para a realização dessas análises exploratórias foi utilizado o programa computacional STATISTICA.

### **3.2.2 Modelagem estatística**

Para a modelagem estatística dos dados de mortalidade, foi tomada como premissa que os valores de mortalidade em um dado dia podem estar relacionados com concentrações de poluentes, com temperaturas e com umidade ocorridos em algum tempo (ou dia) anterior. Desse modo, utilizaram-se essas variáveis defasadas no tempo, com variações de 2 a 7 dias anteriores ao registro da ocorrência do óbito.

Para a inclusão da variável no modelo, foi selecionado aquele conjunto de valores que apresentou maior associação com o número de óbitos em pessoas com idade acima de 60 anos por doenças do aparelho respiratório.

Os procedimentos de análise e ajuste dos modelos de regressão na tentativa de explicar e modelar a associação entre as variáveis ambientais e a mortalidade por causas respiratórias em pessoas com idade acima de 60 anos, foram baseados em Schwartz & Dockery (1992), Dockery & Pope III (1994), Schwartz (1999) e Gouveia & Fletcher (2000).

Para o estudo da relação entre o índice de mortalidade (número diário de óbito) e os índices ambientais (temperatura e umidade relativa do ar) e os de poluição (material particulado PM-10 e dióxido de enxofre SO<sub>2</sub>) foram utilizados vários modelos estatísticos. Esses modelos têm por finalidade a descrição ou estudo da variável número de óbitos por causas respiratórias, em pessoas com idade superior a 60 anos.

Foram consideradas como variáveis independentes ou explicativas:

- onze meses do ano, variáveis indicadoras, utilizado com a finalidade de possibilitar o controle da autocorrelação parcial dos resíduos ao longo do tempo (jan, fev, mar, mai, jun, jul, ago, set, out, nov, dez) e abril será o mês de referência;
- duas variáveis meteorológicas (temperatura média diária (°C) e umidade relativa do ar (%));
- a variável ano (2001 e 2002);
- concentração dos poluentes: dióxido de enxofre - SO<sub>2</sub> (µg.m<sup>-3</sup>) material particulado - PM-10 (µg.m<sup>-3</sup>).

A construção e a avaliação dos modelos estatísticos foram realizadas considerando-se os seguintes passos:

- 1) inclusão de variáveis independentes (os meses do ano, o ano, a temperatura média diária e umidade relativa do ar média diária) considerando aqueles cuja defasagem apresentou maiores estimativas do coeficiente de correlação de Pearson;
- 2) inclusão das concentrações de poluentes, separadamente e em conjunto, como variáveis independentes.

Assim, o modelo adotado inicialmente apresentou a seguinte forma:

$$Y_i = \beta_0 + \beta_1 jan + \dots + \beta_{11} dez + \beta_{12} ano + \beta_{13} temp + \beta_{14} umid + \varepsilon_i \quad (18)$$

no qual:

$Y_i$ : valores observados do número de óbitos;

$\beta_0, \beta_1, \dots, \beta_{14}$ : parâmetros a serem estimados;

jan: variável indicadora, que recebe o valor 1 (um) se o número de óbitos ocorre em janeiro e 0 (zero) nos demais casos; de modo similar para as variáveis independentes fev, mar, mai, jun, jul, ago, set, out, nov e dez;

temp: valores da temperatura média diária, associados ao  $i$ -ésimo número de óbitos;

umid: umidade relativa do ar média diária;

$\varepsilon_i$ : erros aleatórios associados ao  $i$ -ésimo valor observado  $Y_i$  e, tais que,

$\varepsilon_i \sim ARMA(p, q)$ .

A partir do modelo (18), sendo este denominado de modelo principal por meio da inclusão e ou exclusão de covariáveis, foram obtidos novos modelos, de tal forma que, ao se comparar os modelos reduzidos com o principal, avaliou-se a contribuição da variável ausente no modelo reduzido, na explicação do número de óbitos por doenças do aparelho respiratório.

A Tabela 2 lista os modelos utilizados para a comparação do efeito da poluição atmosférica. O **modelo 0** é denominado modelo completo; o **modelo 1** para avaliar o efeito do SO<sub>2</sub> (dióxido de enxofre) e o **modelo 2** o efeito do PM-10 (material particulado) da Estação de Medição de Betim; o **modelo 3**, avalia o efeito do SO<sub>2</sub> (dióxido de enxofre) e o **modelo 4** o PM-10 (material particulado) da Estação de Medição de Contagem.



Para cada modelo foi utilizada medida de diagnósticos para verificar a qualidade de ajuste, por meio do gráfico da função de autocorrelação, diagrama de dispersão dos resíduos e da estatística de Ljung & Box (1978), conforme descrito nos itens 2.7 e 2.8.

Para a escolha do “melhor” modelo foi utilizado o critério de informação de Akaike, descrito no item 2.9.

O recurso computacional utilizado foi o software R, cujos comandos estão listados no Anexo C, e foram baseados em Venables & Ripley (2002).

**TABELA 2** Modelos utilizados para a comparação do efeito dos poluentes SO<sub>2</sub> (dióxido de enxofre) e PM-10 (material particulado) nas estações de Betim e Contagem.

Modelos	Componentes dos modelos
Modelo 0	$\beta_0 + \beta_1 jan + \dots + \beta_{11} dez + \beta_{12} temp + \beta_{13} umid + \beta_{14} ano + \beta_{15} pmbetim + \beta_{16} so_2 betim + \beta_{17} pmcont + \beta_{18} so_2 cont + \varepsilon_t$
Modelo 1	$\beta_0 + \beta_1 jan + \dots + \beta_{11} dez + \beta_{12} temp + \beta_{13} umid + \beta_{14} ano + \beta_{15} pmbetim + \beta_{16} pmcont + \beta_{17} so_2 cont + \varepsilon_t$
Modelo 2	$\beta_0 + \beta_1 jan + \dots + \beta_{11} dez + \beta_{12} temp + \beta_{13} umid + \beta_{14} ano + \beta_{15} so_2 betim + \beta_{16} pmcont + \beta_{17} so_2 cont + \varepsilon_t$
Modelo 3	$\beta_0 + \beta_1 jan + \dots + \beta_{11} dez + \beta_{12} temp + \beta_{13} umid + \beta_{14} ano + \beta_{15} pmbetim + \beta_{16} so_2 betim + \beta_{17} pmcont + \varepsilon_t$
Modelo 4	$\beta_0 + \beta_1 jan + \dots + \beta_{11} dez + \beta_{12} temp + \beta_{13} umid + \beta_{14} ano + \beta_{15} pmbetim + \beta_{16} so_2 betim + \beta_{17} so_2 cont + \varepsilon_t$

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Análise exploratória dos dados

Na Tabela 3 encontra-se o número total de óbitos na RMBH no período de 2001 e 2002, totalizando 15.676 óbitos, dos aproximadamente 18,99 % são relativos a doenças do aparelho respiratório como causa básica.

A média relativa a essas doenças no ano de 2001 foi de 4 óbitos diários. No ano seguinte, notou-se um aumento de, em média, 3 óbitos diários. Entretanto, observando-se a variabilidade do número de óbitos desses anos, notou-se uma homogeneidade, pois os valores dos desvios padrões apresentaram pouca oscilação.

**TABELA 3** Valores do número total de óbitos por causas respiratórias, outras causas e total e respectivas médias diária, medianas diária e desvio padrão do número diário de óbitos, para os anos de 2001 e 2002, na RMBH.

Variável	Causas respiratórias		Outras causas		Total	
	2001	2002	2001	2002	2001	2002
Total	1.136	1.841	4.503	8.199	5.639	10.040
Média diária	4	7	15	30	19	37
Mediana diária	3	6	15	29	18	36
Desvio padrão	2	3	4	6	4	6

Fonte: SES/SE/DPMD/DES/SIM/MG

Conforme os resultados encontrados na Figura 1, pode-se observar que o número de óbitos relativo às causas respiratórias de fato apresentam-se homogêneas quanto à variabilidade, o mesmo ocorrendo nas mortes ocasionadas por outras causas. Um outro importante resultado, foi que, em ambas as situações, notaram-se observações discrepantes o que, possivelmente, poderá dificultar o ajuste de modelos.

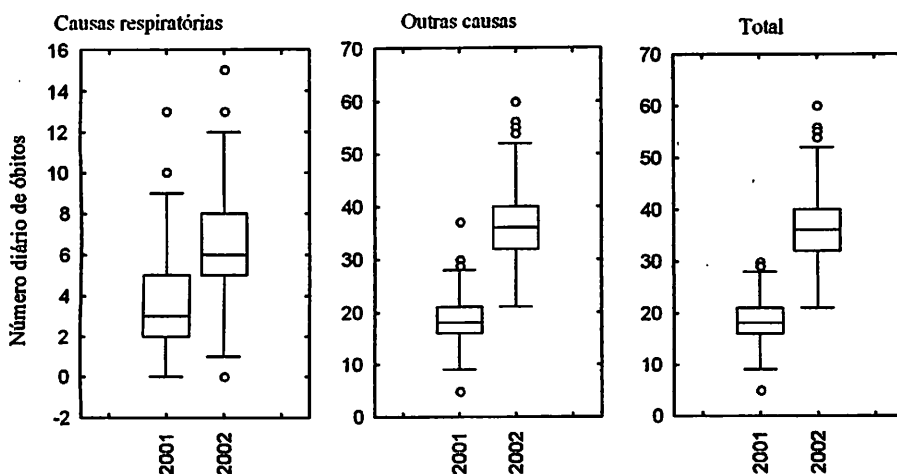


FIGURA 1 Representação gráfica na forma de box-plot do número diário de óbitos de pessoas acima de 60 anos considerando causas respiratórias, outras causas e total, na RMBH. 2001-2002.

Analisando-se a Figura 2, observa-se que os picos de mortalidade ocorrem nas observações de número 180, 320 e 490, com valores próximos a 15 óbitos diários, e que correspondem aos dias (ou meses) de inverno. Nota-se também, que após a 300ª observação ocorreu um aumento no número diário de óbitos, mostrando que a variável ano é relevante, afetando a mortalidade; no entanto, deve-se buscar as causas desse comportamento.

É importante ressaltar que, pelo fato das observações estarem em função do tempo, torna-se necessário incorporar uma estrutura de autocorrelação nos resíduos.

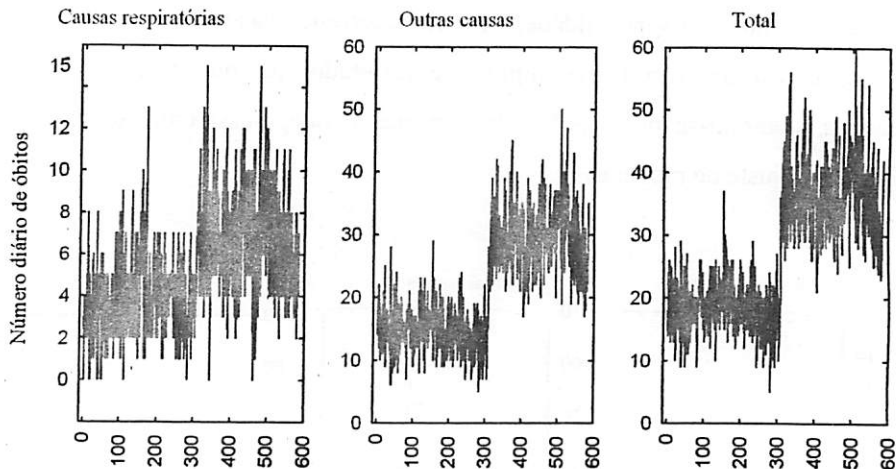


FIGURA 2 Representação gráfica do número de óbitos diários de pessoas com idade acima de 60 anos, considerando causas respiratórias, outras causas e todas as causas, RMBH, 2001-2002.

Os histogramas relativos à descrição do número de óbitos apresentam certo grau de assimetria (Figura 3) à direita, com comportamento relativamente semelhante entre as diferentes causas. Observa-se que o número de óbitos aproxima-se mais de uma distribuição de Poisson do que de uma normal. No entanto, para as causas respiratórias notam-se dois picos nos valores 3 e 5, o que pode ser devido à variação de ano, como salientado na Figura 2.

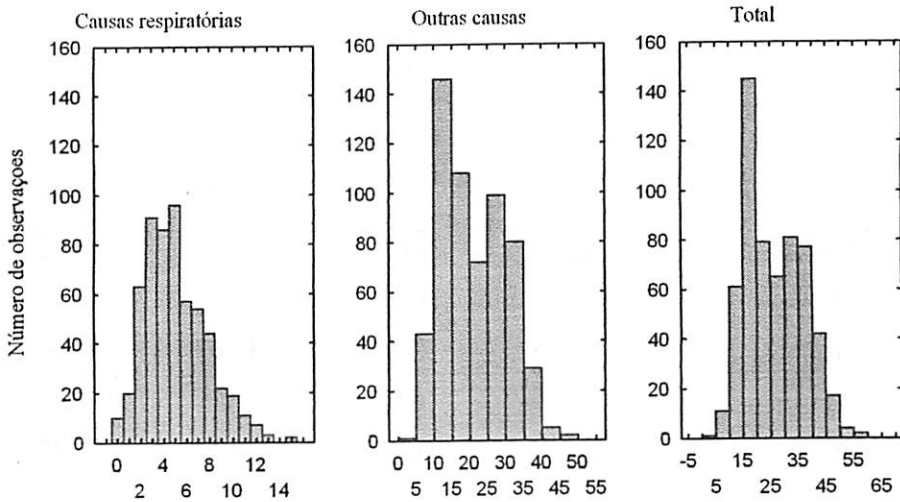


FIGURA 3 Representação gráfica na forma de histograma do número diário de óbitos para as pessoas acima de 60 anos, RMBH, no período de jan/2001 a dez/2002.

Os resultados referentes à concentração do poluente material particulado, PM-10, das Estações de Betim e Contagem, variaram de 0 a  $124,67 \mu\text{g} \cdot \text{m}^{-3}$ , abaixo do padrão  $150 \mu\text{g} \cdot \text{m}^{-3}$  exigido pelo CONAMA. Os valores médios diários da Estação de Betim foram maiores do que os da Estação de Contagem, (Tabela 4) e os valores medianos variaram de  $35,13$  a  $43,97 \mu\text{g} \cdot \text{m}^{-3}$ .

Observa-se que houve um aumento nas concentrações do poluente material particulado PM-10 de  $43,27 \mu\text{g} \cdot \text{m}^{-3}$  referente ao ano de 2001, para  $49,39 \mu\text{g} \cdot \text{m}^{-3}$  obtido em 2002, medidos na Estação de Betim. Já na Estação de Contagem, as concentrações de PM-10 variaram de  $36,45 \mu\text{g} \cdot \text{m}^{-3}$  em 2001, para  $45,97 \mu\text{g} \cdot \text{m}^{-3}$  em 2002; indicando que, no ano de 2001, ocorreu uma concentração menor nos níveis desse poluente.

TABELA 4 Valores médios diários, medianas diárias, mínimos, máximos, desvio padrão e número de ultrapassagens ao padrão (NUP) da concentração do poluente PM-10 ( $\mu\text{g} \cdot \text{m}^{-3}$ ), nas estações de Betim e Contagem, nos anos de 2001 a 2002.

Variável	Estação de Betim		Estação de Contagem	
	2001	2002	2001	2002
Média diária	43,27	49,39	36,45	45,97
Mediana diária	35,60	43,72	35,13	43,97
Desvio padrão	23,75	21,04	15,02	15,96
Mínimo	0	6,57	0	0
Máximo	124,67	112,17	85,23	96,47
NUP	0	0	0	0

Fonte: Federação Estadual de Meio Ambiente-FEAM

A representação gráfica de concentração diária do poluente PM-10 (Figura 4) apresentou comportamento bem distinto entre as duas estações. Na estação de Betim, tem-se um comportamento bem definido com aumento do poluente em certas épocas do ano e diminuição em outras; na estação de Contagem não foi evidenciado este fato.

Notam-se picos de altas concentrações do material particulado (PM-10), principalmente nas épocas próximas dos períodos secos, que ocorrem no inverno.

Para a estação de Betim, os gráficos sugerem que esses períodos de maiores ocorrências do material particulado são em maio e junho, e, para Contagem, em abril e maio.

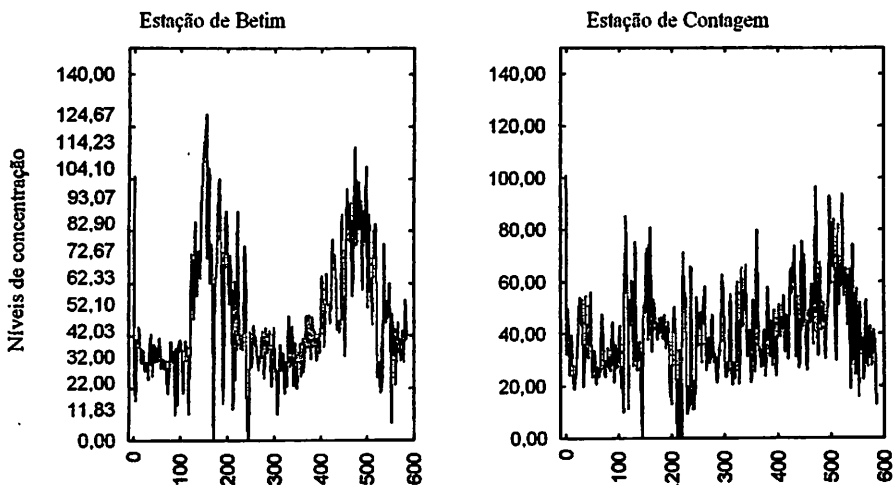
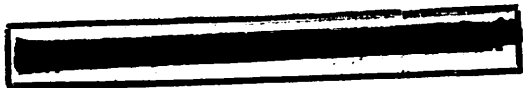


FIGURA 4 Representação gráfica da concentração diária do poluente material particulado PM-10, registrados nas estações de Betim e Contagem, da RMBH, no período de jan-2001 a dez-2002.

Os resultados encontrados na Tabela 5 referem-se às características descritivas para os valores da concentração do poluente dióxido de enxofre.

De acordo com esses resultados, pode-se afirmar que o dióxido de enxofre ( $\text{SO}_2$ ) das estações de medições Betim e Contagem variaram de 0 a  $11,57\mu\text{g.m}^{-3}$ , bem abaixo do padrão CONAMA ( $365\mu\text{g.m}^{-3}$ ). Os valores médios diários da estação de Contagem foram maiores do que os da estação de Betim.

Observa-se que houve uma diminuição nas concentrações do poluente dióxido de enxofre de  $0,96\mu\text{g.m}^{-3}$  no ano de 2001, para  $0,56\mu\text{g.m}^{-3}$  em 2002, medidos na estação de Betim. Já na estação de Contagem, as concentrações médias foram de  $0,84\mu\text{g.m}^{-3}$ , em 2002. Em geral, verificou-se que, no ano de 2002, ocorreram menores concentrações nos níveis desse poluente.

TABELA 5 Valores da concentração média diária, medianas diárias, mínimos, máximos, desvio padrão e número de ultrapassagens (NUP), para o dióxido de enxofre  $\text{SO}_2$  ( $\mu\text{g} \cdot \text{m}^{-3}$ ), nas estações de Betim e Contagem, nos anos de 2001 a 2002.

Variável	Estação de Betim		Estação de Contagem	
	2001	2002	2001	2002
Média	0,96	0,56	0,88	0,84
Mediana	0,20	0,20	0,37	0,62
Desvio padrão	1,48	0,84	1,45	0,77
Mínimo	0	0	0	0
Máximo	4,73	4,33	11,57	3,77
NUP	0	0	0	0

Fonte: Federação Estadual de Meio Ambiente-FEAM

De acordo com a representação gráfica (Figura 5), nota-se um comportamento bem distinto das concentrações de dióxido de enxofre nas estações de Betim e Contagem. Não observou-se comportamento bem distinto, como diminuição do poluente em certas épocas do ano e aumento em outras. Os picos de concentrações ocorreram no início de 2001 e próximo ao inverno de 2002, na estação de Betim. Já na estação de Contagem, as maiores concentrações ocorreram em abril de 2001 e início de 2002.



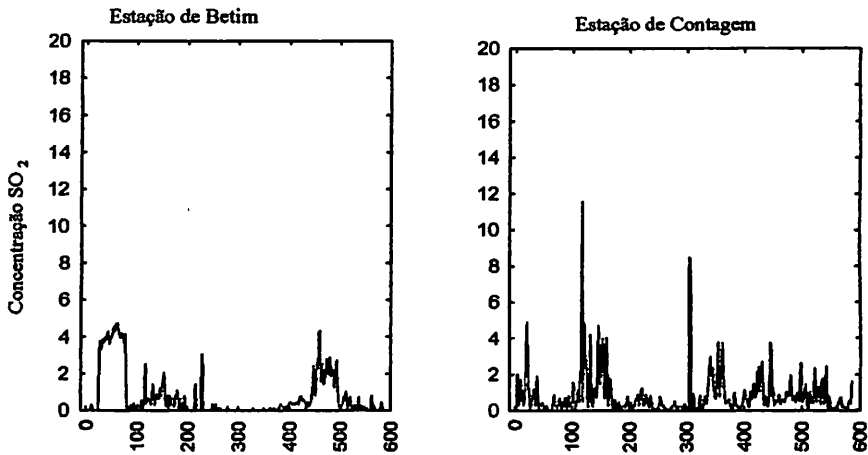


FIGURA 5 Representação gráfica da concentração diária do poluente dióxido de enxofre SO<sub>2</sub> ( $\mu\text{g}\cdot\text{m}^{-3}$ ), para as estações de Betim e Contagem, da RMBH, no período de janeiro de 2001 a dezembro de 2002.

A Tabela 6 mostra que os valores médios de temperatura e umidade relativa do ar do ano de 2001 são bastante próximos aos do ano de 2002, com valor médio da temperatura de 23,8 °C e de 65,4 % de umidade relativa do ar.

TABELA 6 Valores médios, mínimos, máximos e desvio padrão de temperatura média diária (em °C) e umidade relativa do ar (em %), na RMBH para os anos de 2001 a 2002.

Variável	Temperatura °C		Umidade relativa do ar %	
	2001	2002	2001	2002
Média	23,71	23,82	64,06	66,80
Mínimo	16,70	17,30	34,70	45,30
Máximo	30,80	28,20	96,40	92,30
Desvio padrão	2,91	2,43	10,32	8,80

Fonte: Federação Estadual de Meio Ambiente-FEAM

O comportamento cíclico da temperatura (Figura 6) sugere que as épocas quentes ocorrem no verão, com temperaturas chegando aos 30°C e as menores temperaturas ocorrendo no inverno, com valores próximos aos 17°C.

A umidade relativa do ar (%) apresentou comportamento um pouco semelhante ao da temperatura, com maiores valores na época de verão e menores no inverno. Esse comportamento já esperado, pois no verão ocorre a maior parte das chuvas na região.

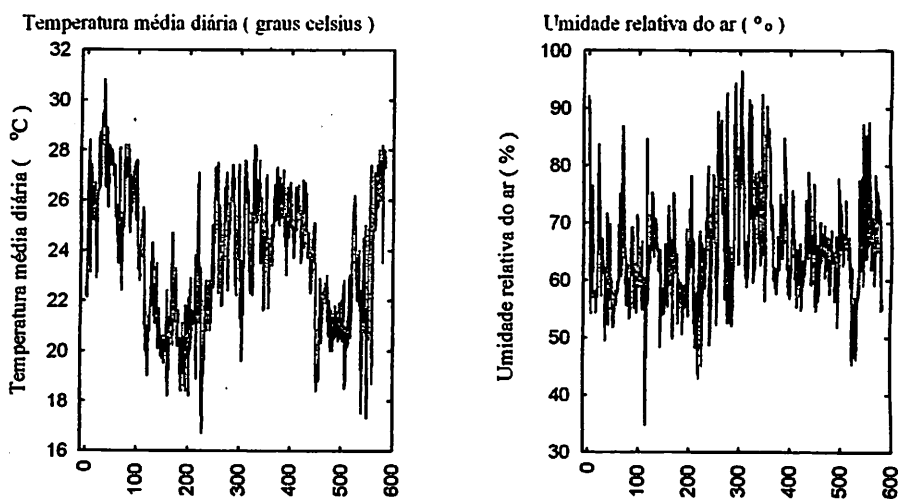


FIGURA 6 Representação gráfica da temperatura média diária (°C) e umidade relativa do ar (%), na RMBH, no período de jan/2001 a dez/2002.

## **4.2 Resultados da análise inferencial dos dados**

São apresentados e discutidos, neste tópico, alguns aspectos que devem ser considerados na análise com dados de contagem, bem como a relação entre as variáveis meteorológicas, defasagem entre a causa (poluição atmosférica) e o efeito (óbito em pessoas por doenças do aparelho respiratório), e a escolha do melhor modelo que explique essa relação.

Zeger (1988) e Schwartz (1999) observaram que, no estudo dessa relação, caso exista, deve-se considerar que as causas não se dão no mesmo dia em que ocorre o óbito ou internação, ou seja, o número de óbitos ou internações ocorridos num determinado dia pode ser consequência das condições meteorológicas e de poluição atmosférica de vários dias anteriores. Assim, neste estudo verificou-se que a melhor associação entre temperatura e umidade relativa do ar ocorreu dois dias anteriores ao óbito, e para o material particulado, três dias, mantendo-se fixo o poluente dióxido de enxofre.

Assim, torna-se propício incorporar no processo de modelagem uma estrutura de correlação nos resíduos. Dessa forma, avaliaram-se diversos modelos para as duas estações de medições dos poluentes e das variáveis meteorológicas de Betim e Contagem.

### **4.2.1 Modelagem estatística**

As estimativas dos parâmetros do modelo, definido em (18), estão apresentadas na Tabela 7. Observa-se que poucas variáveis tiveram contribuição significativas. Entre esses, o efeito de ano positivo indica que está ocorrendo aumento médio de 2,8549 óbitos por ano, o mesmo ocorrendo no mês de julho, com aumento de 1,1333 óbitos. No mês de julho, esse resultado já era esperado

pois é um mês em que ocorre a menor temperatura, o clima está mais seco e esses fatores contribuem para aumentar as doenças do aparelho respiratório, principalmente em idosos e crianças.

No entanto, não foi detectada associação significativa entre as variáveis meteorológicas (temperatura e umidade relativa do ar) e mortalidade.

Dos vários modelos avaliados, o auto-regressivo e médias móveis, ARMA (0,3), foi o que apresentou melhor desempenho, pois, de acordo com o teste de Box & Pierce e considerando o número de lags igual a 40, verificou-se uma probabilidade igual a 0,1561. Portanto, há evidências de que os resíduos são ruído branco, isto é, são independentes e normalmente distribuídos com média 0 e variância constante.

TABELA 7 Estimativas dos coeficientes do modelo estatístico com erros ARMA (0,3) e respectivos erros padrões e probabilidade de serem significativas (p-valor), estatística de Box & Pierce e critério de informação de Akaike (AIC).

Variáveis	Estimativas	Erro padrão	p-valor
Intercepto	-6,9371	1,9822	0,0005***
Janeiro	-0,4736	0,4350	0,2768
Fevereiro	-0,4296	0,4461	0,3360
Março	-0,5341	0,4294	0,2141
Maiο	0,4117	0,4854	0,3967
Junho	0,0613	0,5243	0,9070
Julho	1,1333	0,5202	0,0298**
Agosto	0,1260	0,5180	0,8078
Setembro	-0,0061	0,5016	0,9903
Outubro	0,6826	0,8426	0,4182
Novembro	-0,8097	0,4327	0,0618*
Dezembro	-0,4341	0,4804	0,3667
Ano	2,8549	0,2034	0,0000***
Temperatura	0,0515	0,0574	0,3701
Umidade	0,0145	0,0113	0,1977
Box-Pierce	0,1561	-	-
AIC	2582,951	-	-

\* significativo a 10%; \*\* significativo a 5% e \*\*\* significativo 1%

Por meio da análise gráfica da Figura 7, pode-se observar que os erros são não-correlacionados, pois somente 5% das correlações estão fora do intervalo de confiança Morettin & Tolo (2004), caracterizando ruído branco.

O diagrama de dispersão dos resíduos (Figura 7), em função das observações, mostra que, de modo geral, os resíduos não apresentam nenhuma padronização, o que caracteriza um bom ajuste do modelo.

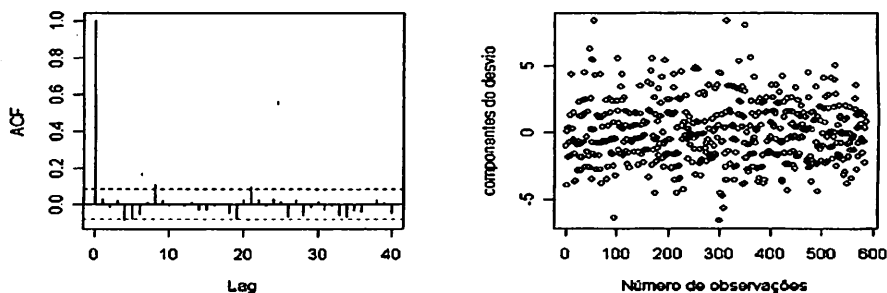


FIGURA 7 Representação gráfica da função de autocorrelação e diagrama de dispersão dos resíduos para o modelo ARMA (0,3) para as variáveis temperatura e umidade relativa do ar.

Portanto, o modelo ARMA (0,3) ajustado, considerando a estrutura de autocorrelação nos resíduos, tem os erros  $\varepsilon_t$  descritos por:

$$\varepsilon_t = e_t + 0,0290e_{t-1} - 0,0298e_{t-2} - 0,0393e_{t-3}, \quad (19)$$

em que  $e_t$  é ruído branco.

#### 4.2.2 Modelos com inclusão dos poluentes

A partir do modelo básico (18) obtido, outros modelos foram ajustados incluindo a concentração de cada um dos quatro poluentes: material particulado da estação de Betim (pmbetim), material particulado de Contagem (pmcont), dióxido de enxofre de Contagem (so<sub>2</sub>cont) e de Betim (so<sub>2</sub>betim).

Diversas estruturas de modelos foram utilizadas para o resíduo, como descritos na Tabela 8. Dentre elas, o de menor critério de informação de Akaike foi o modelo ARMA (2,2), que foi utilizado no modelo completo (modelo 0, da Tabela 2)

TABELA 8 Valores estimados para o critério de Akaike (AIC) em função de diferentes modelos ARMA (p,q) pesquisados.

Estrutura	AIC
ARMA (1,0)	2622,029
ARMA (0,1)	2622,016
ARMA (1,1)	2623,041
<b>ARMA (2,2)</b>	<b>2607,115</b>
ARMA (3,0)	2625,441
ARMA (0,5)	2618,347
ARMA (4,0)	2622,577

O modelo ajustado considerando-se a estrutura de correlação dos resíduos do tipo ARMA (2,2) foi

$$\varepsilon_t = 0,3705\varepsilon_{t-1} + 0,5004\varepsilon_{t-2} + e_t - 0,3976e_{t-1} - 0,6021e_{t-2} \quad (20)$$

no qual  $e_t$  é ruído branco.

A Tabela 9 apresenta as estimativas dos coeficientes do modelo completo e respectivos erros padrões. Com relação aos poluentes, pode-se observar que somente o poluente  $so_2$ betim foi significativo. Para as variáveis meteorológicas, a umidade relativa do ar apresentou certa associação positiva com a mortalidade, o que não era de se esperar, uma vez que para o modelo básico não havia sido não significativo.

Com relação ao efeito de meses, nota-se que julho, novembro e dezembro foram significativos a 5 % de significância, tendo o mês de julho, pertencente ao período de inverno, contribuído para o aumento da mortalidade, enquanto novembro e dezembro, no período de verão, contribuíram para diminuição da mortalidade. O efeito de ano significativo evidencia que a tendência de aumento da mortalidade diária é fator relevante e sugere que este incremento é de 2,6 óbitos.

A estatística de Box & Pierce, que apresentou valor de  $p = 0,1155$  para o modelo com estrutura de quadrado mínimo generalizado, indica que os resíduos do modelo são variáveis aleatórias, constituindo ruído branco.



TABELA 9 Estimativas dos coeficientes do modelo completo e respectivos erros padrões, p-valor, estimativa da estatística de Box & Pierce e do critério de informação de Akaike (AIC).

Parâmetros	Estimativas	Erro padrão	p-valor
Intercepto	-8,0366	2,1054	0,0001***
Janeiro	-0,0511	0,2263	0,8216
Fevereiro	0,1787	0,3010	0,5530
Março	-0,0032	0,2584	0,9902
Maio	0,4474	0,2907	0,1246
Junho	0,2105	0,3895	0,5891
Julho	1,2927	0,3697	0,0006***
Agosto	0,0603	0,3331	0,8564
Setembro	0,1113	0,2921	0,7033
Outubro	0,6162	0,7287	0,3981
Novembro	-0,7924	0,2224	0,0004***
Dezembro	-0,6238	0,2927	0,0335**
Ano	2,6374	0,1846	0,0000***
Temperatura	0,0842	0,0599	0,1605
Umidade	0,0225	0,0126	0,0743*
so <sub>2</sub> betim	-0,1625	0,0845	0,0500**
pmbetim	0,0064	0,0058	0,2674
so <sub>2</sub> cont	0,0182	0,0872	0,8357
pmcont	0,0048	0,0073	0,5098
Box-Pierce	0,1155 (ruído branco)		
AIC	2607,115		

\* significativo a 10%; \*\* significativo a 5% e \*\*\* significativo 1%

Conforme o gráfico de autocorrelação dos resíduos (Figura 8) para o modelo completo, os resíduos podem ser considerados como ruído branco, porque no máximo de 5% das correlações estão fora do intervalo. Nota-se também, no diagrama de dispersão (Figura 8) dos resíduos em função das observações, que, de modo geral, ocorreu uma distribuição aleatória dos desvios, o que caracteriza um bom ajuste do modelo.

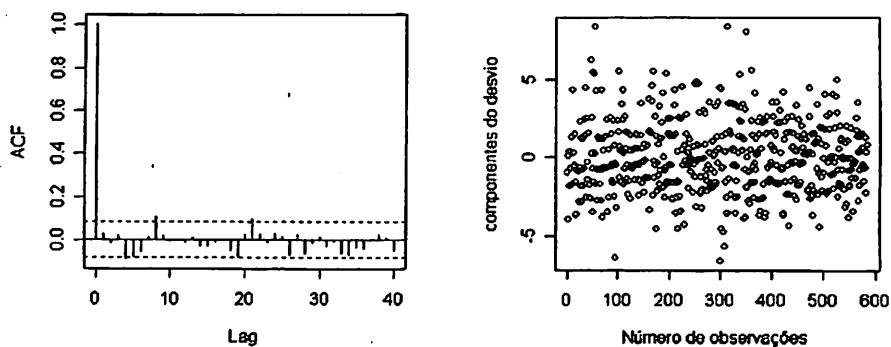


FIGURA 8 Representação gráfica da função de autocorrelação e diagrama de dispersão dos resíduos para o modelo completo ARMA (2,2).

A associação entre a mortalidade e a concentração de PM-10 não é surpreendentes, visto que existem certas informações, principalmente de estudos realizados nos Estados Unidos (USA), apontando nessa direção (Schwartz & Dockery, 1992; Dockery & Pope III, 1994; Dominici & Zeger, 2000). Nesses estudos, e em algumas vezes, no local onde o estudo foi realizado, havia um número maior de estações fixas e móveis para a medição de concentração do poluente PM-10 do que outros. Esse fator pode, assim significar uma maior precisão no sistema de medição do material particulado e, conseqüentemente, favorecer a associação com este poluente.

Estudos epidemiológicos atestam que essa associação ainda não foi totalmente desvendada (Dominici & Zeger, 2000 e Gouveia & Fletcher, 2000),

mas, alguns resultados reforçam cada vez mais a hipótese de causalidade entre poluição atmosférica e mortalidade.

Na Tabela 10 encontram-se as estimativas dos parâmetros dos modelos da Tabela 2 e respectivos erros padrão. Os coeficientes do mês de julho foram significativos em todos os modelos, e coincide com o período de baixas temperaturas, época de inverno, o que favorece o aumento na mortalidade, pois diminui a resistência orgânica, principalmente em idosos. Já os meses de novembro e dezembro tiveram contribuições significativas e negativas, evidenciando uma redução no número de óbitos diário. Esse comportamento já era esperado, pois nesses meses ocorrem épocas de concentrações de chuvas, final da primavera e início de verão.

No entanto, não foram observadas significâncias nas estimativas para temperatura, mas, para a umidade relativa do ar, houve efeito significativo em alguns modelos. Estes resultados sugerem que a temperatura e a umidade relativa do ar não se mostraram altamente associadas à mortalidade.

Nota-se que quando as variáveis (material particulado e dióxido de enxofre) são incluídas nos modelos, apenas o dióxido de enxofre medido na estação de Betim foi significativo, sugerindo que a concentração desse poluente pode estar associada à mortalidade.

Quando foi incluído o material particulado PM-10 das duas estações nos modelos, não se observou efeito significativo de sua presença, indicando que a variabilidade na concentração desse poluente não está diretamente relacionada com a mortalidade. Entretanto, os resultados obtidos estão próximos aos de Schwartz & Dockery (1992), que encontraram estimativas de 0,0381 e erro padrão igual a 0,0082.

Para o dióxido de enxofre, os resultados obtidos não concordam com os resultados de Schwartz & Dockery (1992), que encontraram estimativas de 0,0101 e erro padrão igual a 0,0042.

Não se pode esquecer que os poluentes encontram-se todos juntos na atmosfera e é inadequado atribuir apenas a um deles os efeitos da poluição, mesmo que em baixos níveis, possam acarretar um aumento na mortalidade, principalmente de idosos (Gouveia & Fletcher, 2000).

Este estudo avaliou os efeitos da poluição na mortalidade em pessoas acima de 60 anos por doenças do aparelho respiratório e, quanto aos aspectos metodológicos, pode-se observar que a estatística de Box & Pierce (Tabela 10), para todos os modelos foi significativa, evidenciando que os modelos estão bem ajustados, dado que o resíduo é ruído branco. Isto ocorreu quando os poluentes estão todos juntos ou na falta de um deles no modelo completo.

Nota-se que o modelo sem o poluente dióxido de enxofre ( $\text{SO}_2$ ) na Estação de Contagem (Modelo 3) apresentou menor coeficiente de informação de Akaike (AIC), indicando que este é o melhor modelo quando comparado com os demais, e que o poluente não está aparentemente relacionado com a variável resposta.

Observa-se que, ao retirar do modelo o material particulado (PM-10) da Estação de Contagem (Modelo 4), houve uma alteração em algumas estimativas dos coeficientes, erros padrões e coeficientes de informação de Akaike (AIC), piorando o ajuste do modelo e indicando que este poluente tem contribuição influente.

As funções de autocorrelação para os diferentes modelos foram muito semelhantes (Figuras 9 a 12). Isto sugere que a presença ou ausência de autocorrelação dos resíduos não foi totalmente explicada pelo modelo.

TABELA 10 Estimativas dos parâmetros dos diversos modelos, respectivos erros padrões, estatística de Box & Pierce e critério de informação de Akaike (AIC).

parâmetros	Modelo 0		Modelo 1		Modelo 2		Modelo 3		Modelo 4	
	Estimativas	Erro padrão	Estimativas	Erro padrão	Estimativas	Erro padrão	Estimativas	Erro padrão	Estimativas	Erro padrão
Intercepto	<b>-8,031**</b>	2,105	<b>-8,632**</b>	2,096	<b>-7,516**</b>	2,063	<b>-7,914**</b>	2,010	<b>-7,941**</b>	2,127
Janeiro	-0,051	0,226	-0,171	0,224	-0,114	0,222	-0,048	0,225	-0,048	0,441
Fevereiro	0,178	0,301	-0,236	0,215	0,100	0,296	0,193	0,291	0,074	0,485
Março	-0,003	0,258	-0,167	0,252	-0,060	0,257	-0,004	0,258	-0,106	0,446
Maiο	0,447	0,291	0,311	0,291	0,472	0,294	0,456	0,287	0,589	0,469
Junho	0,210	0,389	-0,047	0,372	0,421	0,342	0,207	0,388	0,267	0,555
Julho	<b>1,292**</b>	0,369	<b>1,231**</b>	0,376	<b>1,428**</b>	0,351	<b>1,276**</b>	0,361	<b>1,251**</b>	0,518
Agosto	0,060	0,333	0,117	0,339	0,075	0,336	0,036	0,311	0,172	0,519
Setembro	0,111	0,292	0,083	0,300	0,083	0,295	0,099	0,286	0,092	0,508
Outubro	0,616	0,728	0,669	0,741	0,445	0,716	0,592	0,719	0,653	0,889
Novembro	<b>-0,792**</b>	0,222	<b>-0,816**</b>	0,229	<b>-0,820**</b>	0,224	<b>-0,803**</b>	0,216	<b>-0,760*</b>	0,443
Dezembro	<b>-0,623**</b>	0,292	<b>-0,627**</b>	0,302	<b>-0,661**</b>	0,294	<b>-0,622**</b>	0,292	-0,417	0,527
Ano	<b>2,637**</b>	0,184	<b>2,751**</b>	0,178	<b>2,652**</b>	0,185	<b>2,629**</b>	0,180	<b>2,705**</b>	0,204
Temperatura	0,084	0,059	0,084	0,060	0,074	0,059	0,081	0,058	0,084	0,060
Umidade	<b>0,022*</b>	0,012	<b>0,026**</b>	0,012	0,020	0,012	<b>0,022*</b>	0,012	0,019	0,012
pmbetim	0,006	0,005	0,005	0,005	-	-	0,006	0,005	0,005	0,006
so <sub>2</sub> betim	<b>-0,162**</b>	0,084	-	-	<b>-0,153*</b>	0,084	<b>-0,166**</b>	0,082	-0,147	0,094
pmcont	0,004	0,007	0,003	0,007	0,006	0,007	0,005	0,006	-	-
so <sub>2</sub> cont	0,018	0,087	0,050	0,086	0,017	0,087	-	-	0,	-
Box-Pierce	0,1151		0,1235		0,1087		0,1156		0,	
AIC	2607,115		2608,839		2606,361		2605,158		26	

\* significativo a 10%; \*\* significativo a 5% e \*\*\* significativo 1%

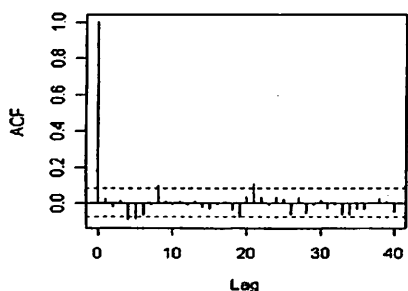


FIGURA 9 Representação gráfica da função de autocorrelação dos resíduos para avaliação do efeito do dióxido de enxofre ( $\text{SO}_2$ ) da estação de Betim para o modelo ARMA (2,2).

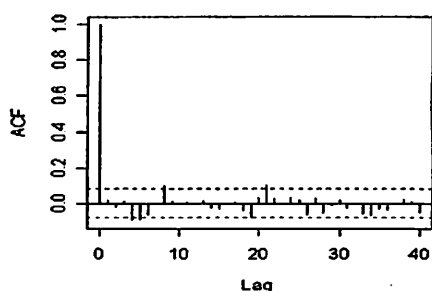


FIGURA 10 Representação gráfica da função de autocorrelação dos resíduos para avaliação do efeito do material particulado (PM-10) da estação de Betim para o modelo ARMA (2,2).

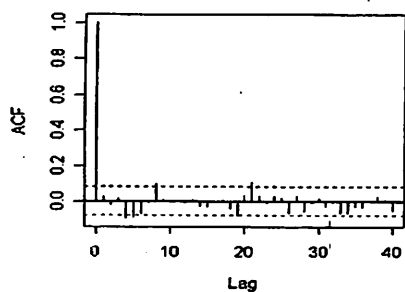


FIGURA 11 Representação gráfica da função de autocorrelação dos resíduos para avaliação do efeito do dióxido de enxofre ( $\text{SO}_2$ ) da estação de Contagem para o modelo ARMA (2,2).

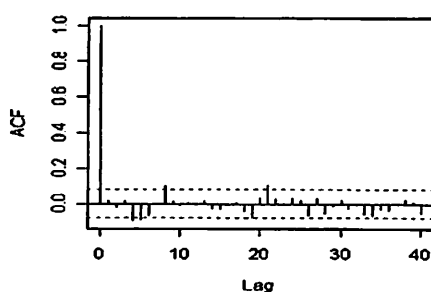


FIGURA 12 Representação gráfica da função de autocorrelação dos resíduos para avaliação do efeito do material particulado (PM-10) da estação de Contagem para o modelo ARMA (2,2).

Os diagramas de dispersão dos resíduos dos modelos em função das observações (Figuras 13 a 16), mostram que os modelos apresentaram bom ajuste, pois as distribuição dos resíduos ocorreu ao longo da média zero. No entanto, pode-se dizer que o efeito da mortalidade não foi totalmente explicado pelos modelos, pois observam-se desvios mais dispersos ou afastados da origem zero.

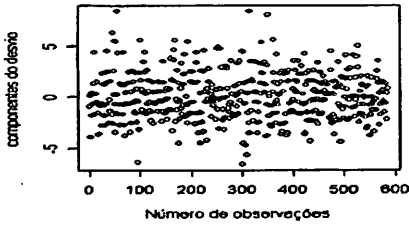


FIGURA 13 Diagrama de dispersão dos resíduos para a avaliação do efeito do dióxido de enxofre ( $SO_2$ ) da estação de Betim para o modelo ARMA (2,2).

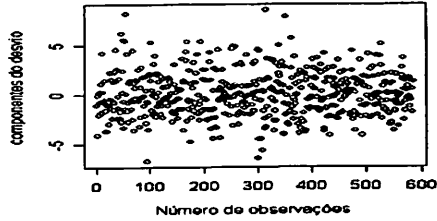


FIGURA 14 Diagrama de dispersão dos resíduos para avaliação do efeito do material particulado (PM-10) da estação de Betim para o modelo ARMA (2,2).

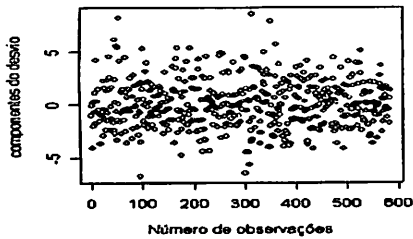


FIGURA 15 Diagrama de dispersão dos resíduos para a avaliação do efeito do dióxido de enxofre ( $SO_2$ ) da estação de Contagem para o modelo ARMA (2,2).

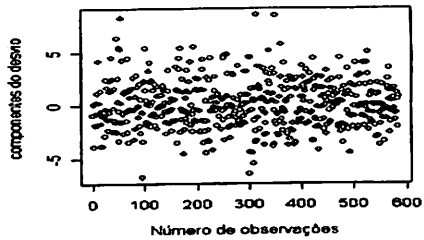


FIGURA 16 Diagrama de dispersão dos resíduos para avaliação do efeito do material particulado (PM-10) da estação de Contagem para o modelo ARMA (2,2).

As Tabelas 11 e 12 apresentam as estimativas dos parâmetros para os modelo AR (p), MA (q) e ARMA (p,q) quando são incluídos no modelo, dado em (18), separadamente, como variáveis explicativas, as concentrações dos poluentes dióxido de enxofre (SO<sub>2</sub>) e material particulado (PM-10) para as Estações de Betim e de Contagem.

Não foi observada nenhuma relação significativa entre os poluentes (material particulado e dióxido de enxofre) e mortalidade em diferentes estruturas de correlação.

Nota-se que, de acordo com a Estatística de Box-Pierce todos os modelos apresentaram p-valor significativo, o que caracteriza que os resíduos são ruído branco.

**TABELA 11** Estimativas dos parâmetros do modelo AR (p), MA (q) e ARMA (p,q), erros padrões (ep), p-valores da estatística de Box & Pierce e critério de informação de Akaike (AIC), quando são incluídas separadamente no modelo (18) as variáveis dióxido de enxofre e material particulado da Estação de Betim.

Estrutura de correlação	Dióxido de enxofre SO <sub>2</sub>				Material particulado PM-10			
	valor	ep	p-valor	AIC	valor	ep	p-valor	AIC
<b>AR (p)</b>								
1	-0,140	0,095	0,123	2617,68	0,004	0,006	0,142	2619,23
2	-0,142	0,093	0,122	2619,34	0,004	0,006	0,141	2620,99
3	-0,141	0,094	0,122	2621,17	0,004	0,006	0,140	2622,75
<b>MA (q)</b>								
2	-0,143	0,093	0,122	2619,23	0,004	0,006	0,141	2620,91
3	-0,141	0,095	0,123	2621,06	0,004	0,006	0,140	2622,63
4	-0,148	0,088	0,122	2618,48	0,005	0,006	0,143	2620,45
<b>ARMA(p,q)</b>								
ARMA(2,2)	-	-	-	-	0,006	0,005	0,125	2605,64



TABELA 12 Estimativas dos parâmetros do modelo AR (p), MA (q) e ARMA (p,q), erros padrões (ep), p-valores da estatística de Box & Pierce e critério de informação de Akaike (AIC), quando são incluídas separadamente no modelo (18) as variáveis dióxido de enxofre e material particulado da Estação de Contagem.

Estrutura de correlação	Dióxido de enxofre SO <sub>2</sub>				Material particulado PM-10			
	valor	ep	p-valor	AIC	valor	ep	p-valor	AIC
<b>AR (p)</b>								
1	<b>0,066</b>	<b>0,087</b>	<b>0,131</b>	<b>2619,29</b>	<b>0,005</b>	<b>0,007</b>	<b>0,131</b>	<b>2619,22</b>
2	<b>0,066</b>	<b>0,087</b>	<b>0,131</b>	<b>2621,07</b>	<b>0,005</b>	<b>0,007</b>	<b>0,130</b>	<b>2620,99</b>
3	<b>0,066</b>	<b>0,087</b>	<b>0,131</b>	<b>2622,80</b>	<b>0,005</b>	<b>0,007</b>	<b>0,129</b>	<b>2622,65</b>
<b>MA (q)</b>								
2	<b>0,066</b>	<b>0,087</b>	<b>0,131</b>	<b>2620,99</b>	<b>0,005</b>	<b>0,007</b>	<b>0,131</b>	<b>2620,91</b>
3	<b>0,065</b>	<b>0,087</b>	<b>0,131</b>	<b>2622,67</b>	<b>0,006</b>	<b>0,007</b>	<b>0,129</b>	<b>2622,51</b>
4	<b>0,069</b>	<b>0,085</b>	<b>0,133</b>	<b>2620,56</b>	<b>0,005</b>	<b>0,006</b>	<b>0,133</b>	<b>2620,70</b>
<b>ARMA(p,q)</b>								
ARMA(2,2)	-	-	-	-	<b>0,005</b>	<b>0,006</b>	<b>0,115</b>	<b>2605,98</b>

Analisando-se as Tabelas 11 e 12 e observa-se que o melhor modelo foi o auto regressivo AR (1), quando o dióxido de enxofre está presente no modelo, apresentando menores valores para o critério de informação de Akaike, de 2.617,68 e 2.619,29, respectivamente, para as estações de Betim e Contagem.

Quando o material particulado foi incluído no modelo, observa-se uma combinação entre os parâmetros auto-regressivo e médias móveis, sendo o modelo ARMA (2,2), aquele que apresentou menor valor para o critério de informação de Akaike, com estimativas de 2.605,64 e 2.605,98, respectivamente, para as estações de Betim e Contagem.

Para o material particulado, observa-se uma combinação entre os parâmetros auto regressivos e médias móveis, ou seja, um modelo do tipo ARMA (2,2), uma vez que apresentou menor valor para o critério de informação de Akaike, sendo 2.605,639 e 2.605,977, respectivamente, para as Estações de Betim e Contagem.

As Figuras 17 e 18 mostram gráficos de autocorrelação de resíduos após a inclusão separadamente dos poluentes pesquisados ao modelo básico. Nota-se que os resíduos podem ser considerados com ruído branco, porque no máximo de 5% das correlações estão fora do intervalo de confiança.

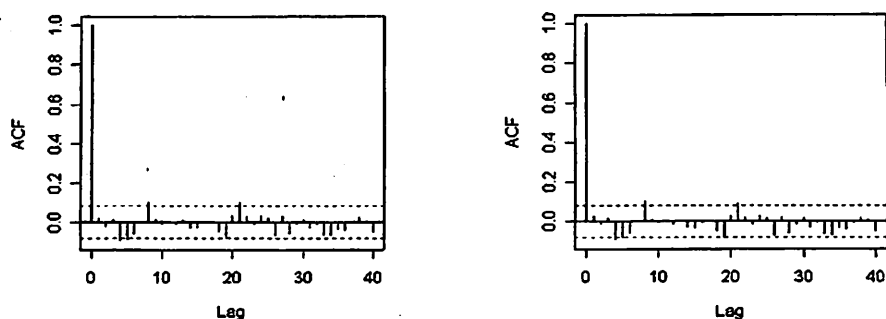


FIGURA 17 Representação gráfica da função de autocorrelação dos resíduos para dióxido de enxofre ( $\text{SO}_2$ ) utilizando o modelo AR (1) e para o material particulado (PM-10) utilizando o modelo ARMA (2,2), estação de Betim.

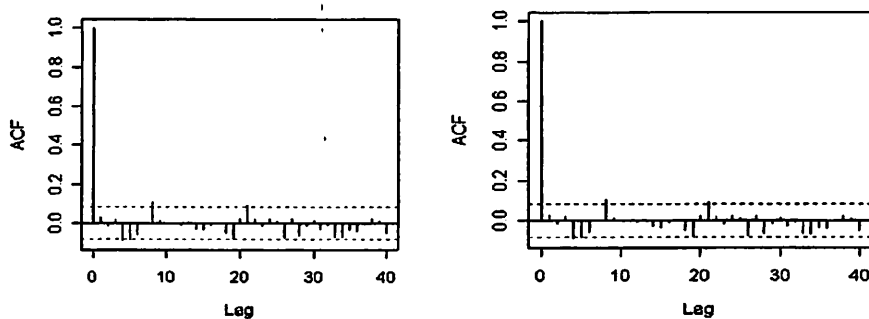


FIGURA 18 Representação gráfica da função de autocorrelação dos resíduos para dióxido de enxofre ( $\text{SO}_2$ ) utilizando o modelo AR (1) e para o material particulado (PM-10) utilizando o modelo ARMA (2,2), estação de Contagem.

As Figuras 19 e 20 apresentam um diagrama de dispersão dos resíduos para os modelos AR (1) considerando o SO<sub>2</sub> presente no modelo e ARMA (2,2) quando o material particulado está presente no modelo. Constata-se que os comportamentos dos resíduos se concentram em torno da média zero, não apresentando nenhuma outra padronização ou tendência.

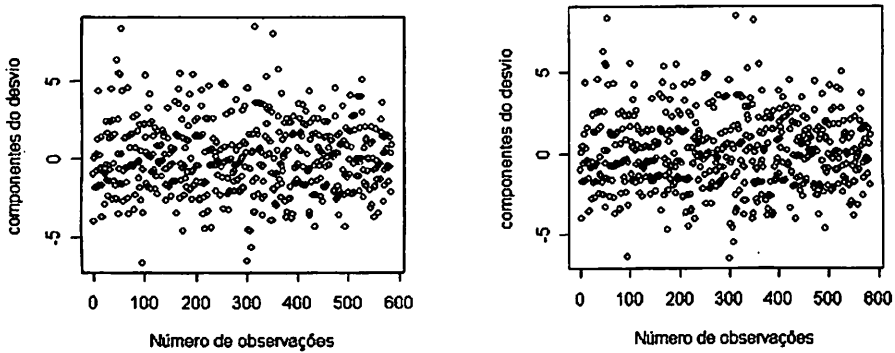


FIGURA 19 Diagrama de dispersão dos resíduos para dióxido de enxofre (SO<sub>2</sub>) utilizando o modelo AR (1) e para o material particulado (PM-10) utilizando o modelo ARMA (2,2), estação de Betim.

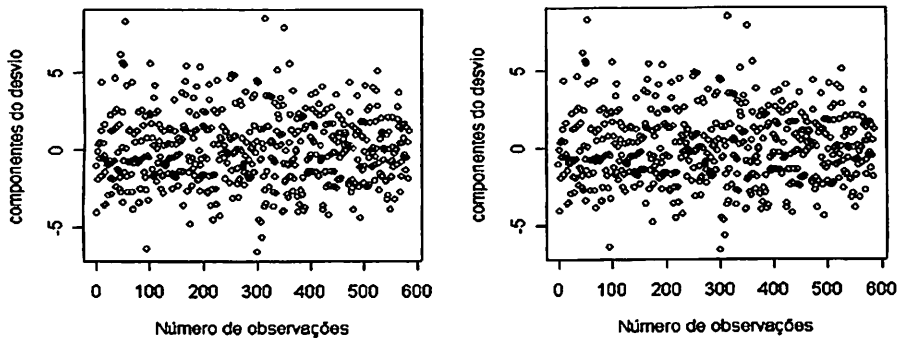


FIGURA 20 Diagrama de dispersão dos resíduos para dióxido de enxofre (SO<sub>2</sub>) utilizando o modelo AR (1) e para o material particulado (PM-10) utilizando o modelo ARMA (2,2), estação de Contagem.

Neste trabalho, vários refinamentos metodológicos foram introduzidos. Entre eles, destaca-se a utilização de modelos lineares com diferentes estruturas de correlação nos resíduos. Esses fatores contribuíram para aumentar a eficiência da modelagem e ajudar a caracterizar melhor os efeitos da poluição.

Em nenhum modelo ajustado foi detectada associação significativa entre o número de óbitos com a poluição ambiental medida pelo material particulado (PM-10). Isto contraria alguma literatura disponível, como Dockery & Pope III (1994) e Gouveia & Fletcher (2000). Uma justificativa para tal fato, podem ser os baixos níveis de concentração desse poluente na Região Metropolitana de Belo Horizonte, pois os maiores valores não atingiram ao padrão primário. Porém, ressalta-se que, mesmo em pequenos níveis, indivíduos em contato constante podem ser afetados, podendo, nesse caso, o efeito da poluição levar até o óbito.

Neste trabalho, encontraram-se evidências da associação entre o poluente dióxido de enxofre e a mortalidade, o que está de acordo com estudos prévios realizados por Gouveia & Fletcher (2000) na cidade de São Paulo.

Como continuidade do estudo, dada a importância da relação entre poluição ambiental e mortalidade e por se tratar de assunto de relevada importância, vida humana, fica a recomendação de utilizar outras metodologias, como alternativas que permitam uma análise mais acurada do problema proposto.

Caso, em geral, as concentrações dos poluentes PM-10 e SO<sub>2</sub> sejam menores que o padrão da legislação vigente, sugere-se que:

- as indústrias têm procurado manter equipamentos antipoluentes eficientes; tal fato deve ser visto como um incentivo para que continuem a reduzir, ainda mais, as quantidades emitidas atualmente;
- os órgãos públicos têm procurado manter uma fiscalização eficiente, a qual deve ser mantida, para que os índices sejam diminuídos ainda mais.

## 5 CONCLUSÃO

1. O número de óbitos em pessoas com idade acima de 60, causados por doenças do aparelho respiratório, é explicado de modo satisfatório por modelo que inclui as variáveis meses, anos, variáveis meteorológicas (temperatura e umidade relativa do ar) e material particulado - PM-10, medido na Estação de Contagem e material particulado - PM-10 e dióxido de enxofre - SO<sub>2</sub>, medidos na Estação de Betim.
2. O número de óbitos relaciona-se com o poluente dióxido de enxofre (SO<sub>2</sub>), obtido na Estação de Betim.
3. Há associação positiva entre número de óbitos e umidade relativa do ar.
4. O número de óbitos relaciona-se negativamente com os meses de novembro e dezembro, caracterizados como de temperaturas e umidade relativa do ar elevadas.
5. O número de óbitos relaciona-se positivamente com o mês de julho, caracterizado como de baixa temperatura.

## REFERENCIAL BIBLIOGRÁFICO

ANUÁRIO ESTATÍSTICO DE BELO HORIZONTE. Belo Horizonte: Prefeitura Municipal de Belo Horizonte/Secretária Municipal de Planejamento/Departamento de Informações Técnicas, 2001.

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, New York, v. AC-19, n. 6, p. 716-723, 1974.

BOX, G. E. P.; JENKINS, G. M. **Times series analysis, forecasting and control**. San Francisco: Holden-Day, 1976. 575 p.

BOX, G. E. P.; PIERCE, D. A. Distribution of residuals autocorrelations in autoregressive-integrated moving average time series models. **Journal of the American Statistical Association**, Alexandria, v. 65, p. 1509-1529, 1970.

**CID-10. Classificação estatística internacional de doenças e problemas relacionados à saúde**. 10. ed. São Paulo: Editora da Universidade de São Paulo, 1993. Disponível em: <<http://www.datasus.gov.br/cid10/cid10.htm>>. Acesso em: 15 jan. 2005.

DOCKERY, D. W.; POPE III, C. A. Acute respiratory effects of particulate air pollution. **Annual Review Public Health**, Palo Alto, v. 15, p. 107-132, 1994.

DOMINICI, F.; SAMET, J. M.; ZEGER, S. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. **Journal Royal Statistical Society**, v. A, 163, p. 1-22, 2000.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. 3. ed. New York: John Wiley, 1998. 706 p.

FOKIANOS, K.; KEDEM, B. Partial likelihood inference for time series following generalized linear models. **Journal of Time Series Analysis**, Oxford, v. 25, n. 2, p. 173-197, Mar. 2004.

FOX, J. **Time series regression and generalized least squares**. 2002. Disponível em: <[http://cram.r-project.org/dc/comtrib/Fox-comparion/appendix-time\\_series\\_regression.pdf](http://cram.r-project.org/dc/comtrib/Fox-comparion/appendix-time_series_regression.pdf)>. Acesso em: 27 out. 2004.

FULLER, W. A. **Introduction to statistical time series**. 2. ed. New York: John Wiley, 1996. 698 p.

GOUVEIA, N.; FLETCHER, T. Time series analysis of air pollution and mortality: effects by cause, age and socioeconomic status. **Epidemiology Community Health**, London, v. 54, n. 10, p. 750-755, Oct. 2000.

GRAYBILL, F. A. **An introduction to linear statistical models**. New York: McGraw-Hill, 1961. 463 p.

HANNAN, E. J. The estimation of the order of an ARMA process. **Annals of statistics**, Hayward, v. 8, n. 5, p. 1071-1081, 1980.

HASTIE, T. J.; TIBSHIRANI, R. J. **Generalized additive models**. Chapman & Hall, 1990. 335 p.

HARVEY, A. C. **Time series models**. 2. ed. Massachusetts: The MIT press Cambridge, 1994. 308 p.

HOFFMANN, R.; VIEIRA, S. **Análise de regressão: uma introdução à econometria**. 3. ed. São Paulo: HUCITEC, 1998. 379 p.

LIMA, L. P. **Modelos aditivos generalizados: aplicação a um estudo epidemiológico Ambiental**. São Paulo, 2001. Dissertação (Mestrado) - .

Ljung, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. **Biometrika**, London, v. 65, n. 2, p. 297-303, Aug. 1978.

PANKRATZ, A. **Forecasting with univariate Box & Jenkins models, concepts and cases**. New York: John Wiley, 1983. 562 p.

McCULLAGH, P.; NELDER, J. A **Generalized linear models**. 2. ed. London: Chapman & Hall, 1989. 511 p.

McCULLOCH, C. E.; SEARLE, S. R. **Generalized, linear and mixed models**. New York: John Wiley, 2000. 326 p.

MOOD, A. M.; GRAYBILL, F. A; BOES, D. C. **Introduction to the theory of statistics**. 3. ed. McGraw-Hill, 1974. 574 p.

MORETTIN, P. A; TOLOI, C. M. C. **Análise de séries temporais**. São Paulo: Edgard Blucher, 2004. 535 p.

PAULA, G. A. **Modelos de regressão com apoio computacional**. São Paulo: USP. Instituto de Matemática e Estatística, 1999. 226 p.

R : Copyright 2004, The R Foundation for Statistical Computing  
Version 1. 9. 0 (2004-04-12), ISBN 3-900051-00-3. Disponível em:  
<<http://www.economia.unitn.it/micciolo/srs/20041019.doc>>. Acesso em: 15 jan. 2005.

SEARLE, S. R. **Linear models**. New York: John Wiley 1971. 532 p.

SCHWARTZ, J. Air pollution and daily mortality in Birmingham, Alabama. **American Journal of Epidemiology**, Baltimore, v. 137, n. 10, p. 1136-1147, May 1999.

SCHWARTZ, J.; DOCKERY, D. W. Particulate air pollution and daily mortality in Steubenville, Ohio. **American Journal of Epidemiology**, Baltimore, v. 135, n. 1, p. 12-19, Jan. 1992.

SOUZA, G. S. **Introdução aos modelos linear e não linear**. Brasília: Embrapa-SP/Embrapa-SEA, 1998. 489 p.

STATISTICA FOR WINDOWS. Release 5. Copyright. Stat. Soft. 1984-1995.

VENABLES, W. N.; RIPLEY, B. D. **Modern applied statistics with S**. 4. ed. Springer-Verlag, 2002. 495 p.

WEI, W. W. S. **Time series analysis: univariate and multivariate methods**. Red Wood City: Addins-Wesley, 1990. 478 p.

ZEGER, S. L. A regression model for time series of counts. **Biometrika**, London, v. 75, n. 4, p. 621-629, Dec. 1988.

ZURITA, M. L. L.; TOLFO, A. M. (Org.) **A qualidade do ar em Poto Alegre**. Porto Alegre: Secretaria Municipal de Planejamento do Meio Ambiente, 2000. 92 p.



## ANEXOS

ANEXOS	PAGINA
<b>ANEXO A</b>	
FIGURA 1A Mapa da região metropolitana de Belo Horizonte (RMBH).....	60
<b>ANEXO B</b>	
TABELA 1B População (número de habitantes), área total (km <sup>2</sup> ), e densidade demográfica (habitantes, km <sup>2</sup> ) do município d Belo Horiznte, Região Metropolitana de Belo Horizonte, do estado de Minas Gerais e do Brasil, de acordo com censo demográfico de 2000.....	61
TABELA 2B População (número de habitantes) residente total e por sexo do Brasil, de Minas Gerais, da Região Metropolitana de Belo Horizonte e dos municípios que compõem a RMBH, de acordo com censo demográfico de 2000.....	61
TABELA 3B Número de veículos em circulação por categoria, ano de fabricação e tipo, em Belo Horizonte,2000.....	63
<b>ANEXO C</b>	
Rotina do programa R, utilizado para a obtenção das estimativas dos parâmetros dos modelos da análise inferencial dos dados, erro padrão, estatística de Box & Pierce, p-valor, diagrama de autocorrelação e dispersão dos resíduos pelo método dos quadrados mínimos generalizados.	

## ANEXO A



FIGURA 1A Mapa da Região Metropolitana de Belo Horizonte (RMBH)  
 FONTE: Anuário.... 2001

## ANEXO B

TABELA 1B População (número de habitantes), área total (km<sup>2</sup>), e densidade demográfica (habitante km<sup>-2</sup>) do município de Belo Horizonte, da Região Metropolitana de Belo Horizonte, do estado de Minas Gerais e do Brasil, de acordo com censo demográfico de 2000.

Especificação	População	Área (km <sup>2</sup> )	Densidade
BH	2.238.526	330,90	6.764,96
RMBH	4.357.942	9.459,10	460,71
MG	17.891.194	586.552,40	30,50
BR	169.799.170	8.5415.215,30	19,94

Fonte: IBGE. Censo Demográfico 2000.

TABELA 2B População (número de habitantes) residente total e por sexo do Brasil, de Minas Gerais, da Região Metropolitana de Belo Horizonte e dos municípios que compõem a RMBH, de acordo com censo demográfico de 2000.

Especificação	População residente		
	Total	Homens	Mulheres
BR	169.799.170	83.576.015	86.223.155
MG	17.891.494	8.851.587	9.039.907
RMBH	4.357.942	2.106.430	2.251.512
Baldim	8.155	4.071	4.084
Belo Horizonte	2.238.526	1.057.263	1.181.263
Betim	306.675	152.880	153.795
Brumadinho	26.614	13.359	13.255
Caeté	36.299	17.678	18.621
Capim Branco	7.900	3.971	3.929
Confins	4.880	2.392	2.488
Contagem	538.017	263.390	274.627

“...continua...”

“Tabela 2B, cont”

Esmeraldas	47.090	23.895	23.195
Florestal	5.647	2.863	2.784
Ibirité	133.044	66.190	66.854
Igarapé	24.838	12.608	12.230
Itaguara	11.302	5.728	5.574
Itatiaiuçu	8.517	4.337	4.180
Jaboticatubas	13.530	6.936	6.594
Juatuba	16.389	8.233	8.156
Lagoa Santa	37.872	18.767	19.105
Mário Campos	10.535	5.341	5.194
Mateus Leme	24.144	12.020	12.124
Matozinhos	30.164	15.016	15.148
Nova Lima	64.387	31.448	32.939
Nova União	5.427	2.755	2.672
Pedro Leopoldo	53.957	26.539	27.418
Raposos	14.289	6.976	7.313
Ribeirão das Neves	246.846	123.531	123.315
Rio Acima	7.658	3.794	3.864
Rio Manso	4.646	2.440	2.206
Sabará	115.352	56.239	59.113
Santa Luzia	184.903	90.740	94.163
São Joaquim de Bicas	18.152	9.178	8.974
São José da Lapa	15.000	7.519	7.481
Sarzedo	17.274	8.689	8.585
Taquaraçu de Minas	3.491	1.888	1.603
Vespasiano	76.422	37.756	38.666

Fonte: IBGE. Censo Demográfico 2000.

TABELA 3B Número de veículos em circulação por categoria, ano de fabricação e tipo, em Belo Horizonte, 2000.

Detalhamento veículos	Automó- vel	Cami nhão	Cami nhone te	Moto	Ônibus	Outros	Total
Particular	485478	10550	60281	45710	325	12539	614883
Aluguel	6515	11418	7397	3	5352	5582	36267
Oficial	8948	1384	3610	1203	52	422	15619
Experiência	0	1	0	2	0	568	571
Aprendizagem	1203	3	0	72	16	9	1303
Outros	10	2	1	0	0	6	19
<b>Categoria</b>							
2000	21982	349	3046	2980	61	1082	29500
1999	30829	569	3771	4827	185	1296	41476
1998	36378	1015	5913	5179	533	1372	50388
<b>Ano de fabricação</b>							
1997/1996	86744	1811	14461	8061	876	2711	114664
1995/1994	75093	1930	11947	4013	677	2522	96182
1993/1992	40741	828	6194	1699	916	1203	51581
antes 1992	210390	16856	25957	20231	2497	8940	284871
<b>Total</b>	<b>502154</b>	<b>23358</b>	<b>71289</b>	<b>46990</b>	<b>5745</b>	<b>19126</b>	<b>668662</b>

Fonte: DETRAN, 2000

Dados até agosto de 2000.

## ANEXO C

Rotina do programa R, utilizado para a obtenção das estimativas dos parâmetros dos modelos da análise inferencial dos dados, erro padrão, estatística de Box & Pierce, p-valor, diagrama de autocorrelação e dispersão dos resíduos pelo método dos quadrados mínimos generalizados.

```
a<-read.table("A:\\dados.txt",h=T)
attach(a)
a
mes<-a$mes
library(nlme)

m0<-
gls(resp~mes+ano+so2betim+pmbetim+so2cont+pmcont,correlation=corARMA(
p=2,q=2), method="ML")
summary(m0)
anova(m0)
r<-residuals(m0)
acf(r,lag.max=40,xlab="Figura – Representação gráfica de autocorrelação entre
os resíduos",ylab="ACF",main=" ")
plot(r,xlab="Figura -Diagrama de dispersão dos resíduos",ylab="componentes
do desvio",main="")
Box.test(residuals(m0),lag=40)
```