



Universidade do Estado do Rio de Janeiro
Centro Biomédico
Instituto de Medicina Social

Simone de Souza Cardoso

**Modelos de efeitos mistos para previsão da poluição do ar:
Avaliação global da mortalidade atribuível à poluição do ar nas regiões
metropolitanas do Brasil**

Rio de Janeiro

2018

Simone de Souza Cardoso

**Modelos de efeitos mistos para previsão da poluição do ar:
Avaliação global da mortalidade atribuível à poluição do ar nas regiões metropolitanas
do Brasil**

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Saúde Coletiva, da Universidade do Estado do Rio de Janeiro.

Orientador: Prof. Dr. Antônio Carlos Monteiro Ponce de Leon

Rio de Janeiro

2018

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/CB/C

C268 Cardoso, Simone de Souza
Modelos de efeitos mistos para previsão da poluição do ar :
avaliação global da mortalidade atribuível à poluição do ar nas
regiões metropolitanas do Brasil / Simone de Souza Cardoso –
2018.
194 f.

Orientador: Antônio Carlos Monteiro Ponce de Leon.

Tese (doutorado) – Universidade do Estado do Rio de Janeiro,
Instituto de Medicina Social.

1. Ar - Poluição – Teses. 2. Doenças cardiovasculares –
Teses. 3. Doenças respiratórias – Teses. 4. Mortalidade – Teses.
I. Ponce de Leon, Antônio Carlos Monteiro. II. Universidade do
Estado do Rio de Janeiro. Instituto de Medicina Social. III.
Título.

CDU 614.71:314.14

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta tese,
desde que citada a fonte.

Assinatura

Data

Simone de Souza Cardoso

**Modelos de efeitos mistos para previsão da poluição do ar:
Avaliação global da mortalidade atribuível à poluição do ar nas regiões metropolitanas
do Brasil**

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Saúde Coletiva, da Universidade do Estado do Rio de Janeiro.

Aprovada em 05 de julho de 2018.

Orientador: Prof. Dr. Antônio Carlos Monteiro Ponce de Leon

Instituto de Medicina Social – UERJ

Banca Examinadora:

Prof. Dr. Antônio Carlos Monteiro Ponce de Leon

Instituto de Medicina Social - UERJ

Prof. Dr. Guilherme Loureiro Werneck

Instituto de Medicina Social - UERJ

Prof. Dr. Antônio Guilherme Fonseca Pacheco

Fundação Oswaldo Cruz

Prof. Dr. Cleber Nascimento do Carmo

Fundação Oswaldo Cruz

Prof.^a Dra. Maria Tereza Serrano Barbosa

Universidade Federal do Estado do Rio de Janeiro

Rio de Janeiro

2018

DEDICATÓRIA

Dedico esse trabalho ao meu pai Valter Francisco Cardoso (*in memoriam*) com todo o meu amor e gratidão por tudo que fez por mim e para minha formação.

AGRADECIMENTOS

A Deus pela força, pela perseverança dada na busca do cumprimento da meta e pela inspiração, dada através do Espírito Santo, para desenvolver esta Tese; por deixar permanecer ao meu lado pessoas tão especiais como a minha família e meus amigos, que me deram força e estímulo durante esse Doutorado; a Jesus Cristo e Nossa Senhora de Fátima, que estão sempre presentes em minha vida.

À minha mãe, Lúcia de Souza Cardoso, e ao meu irmão, Fábio de Souza Cardoso, por todo amor dedicado e por estarem sempre me apoiando e ajudando em todos os momentos da minha vida.

Ao meu orientador, Professor Antônio Ponce de Leon, grande professor e de extrema competência, que acreditou na minha capacidade em desenvolver o tema da tese. Obrigada pelos seus ensinamentos, eu os levarei comigo para sempre.

Aos professores Luciana Tricai Cavalini e Alexandre Marinho, que foram fonte de inspiração para o meu caminho acadêmico e por acreditarem no meu potencial.

À minha colega, Vivian Vicente de Almeida, pelo incentivo aos estudos e pesquisas.

A todos os meus colegas da turma de Epidemiologia do Mestrado e Doutorado 2014. Agradeço a vocês pela demonstração de união e da ajuda mútua.

Ao colega, Gustavo dos Santos Souza, agradeço pelo carinho, amizade e colaboração nas informações dos dados deste trabalho.

À minha amiga, Adma do Nascimento Lima, por ter me ajudado muito neste doutorado e pelo companheirismo e amizade.

A todos os professores do IMS, os quais tive a oportunidade de receber seus conhecimentos durante as disciplinas que cursei, pelos seminários e palestras que eu assisti.

Ao Instituto de Pesquisa Econômica Aplicada – IPEA, lugar que me deu aspiração à pesquisa.

Aos institutos, secretarias e fundações do meio ambiente que forneceram dados e informações.

E àqueles que de alguma forma colaboraram para a realização desta tese.

Os erros causados por dados inadequados são muito menores do que aqueles devido à falta total de dados.

Charles Babbage

RESUMO

CARDOSO, Simone de Souza. *Modelos de efeitos mistos para previsão da poluição do ar: Avaliação global da mortalidade atribuível à poluição do ar nas regiões metropolitanas do Brasil*. 2018. 194 f. Tese (Doutorado em Saúde Coletiva) – Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. Rio de Janeiro. 2018.

As maiores fontes de poluentes atmosféricos estão na queima de combustíveis oriundas de fontes estacionárias (usinas e energia elétrica) e móveis (frota de veículos automotores). Grande parte dessas fontes está concentrada nas grandes cidades, onde são esperados impactos significativos na saúde, sobretudo no que se refere às doenças respiratórias e cardiovasculares. A exposição aos poluentes atmosféricos, especificamente ao material particulado, está associada a uma série de efeitos sobre a saúde, mas os efeitos sobre a mortalidade são indiscutivelmente os mais importantes e são mais favoráveis à avaliação global. Quantificar a magnitude do impacto da poluição do ar na saúde nas grandes cidades apresenta desafios consideráveis devido à disponibilidade limitada de informações. Esta tese apresenta dois estudos. O primeiro aborda aspectos metodológicos sobre previsão de concentração de material particulado em áreas urbanas onde não há redes de monitoramento da qualidade do ar. O método utilizado foi o modelo linear de efeitos mistos para previsão das médias anuais de material particulado inalável com diâmetro $\leq 10\mu\text{m}$ (PM₁₀) no período de 2001 a 2014 nas áreas metropolitanas do Brasil. A metodologia baseou-se no modelo proposto por Cohen *et al.* (2004), desenvolvido pelo Banco Mundial, para estimar as concentrações de PM₁₀ em função de características econômicas, meteorológicas, demográficas e outras das regiões metropolitanas. O resultado mostrou que apenas umidade relativa do ar, produto interno bruto per capita do setor serviços, altitude média e focos de queimadas foram associados aos níveis de concentração de PM₁₀. O objetivo do segundo estudo foi estimar a carga de mortalidade no conjunto das regiões metropolitanas, a partir do modelo de efeitos mistos aplicado a dois grupos de doenças e faixas etárias específicas: doenças cardiovasculares em adultos 30 anos ou mais e doenças respiratórias em crianças menores de um ano e de um a cinco anos. Foram estimados o número de mortes atribuíveis e os anos de vida perdidos (*Years of Life Lost* - YLL) para adultos usando estimativas de risco de um estudo de coorte da Associação Americana de Câncer. E com relação às faixas etárias de um ano e de um a cinco anos, as estimativas de risco consideradas foram extraídas do estudo meta-analítico do projeto - Estudo de Saúde e Poluição Atmosférica na América Latina.

Palavras-chave: Modelos de efeitos mistos. Previsão. Material Particulado. Regiões Metropolitanas. Carga Global de Doenças. Mortalidade. Doenças Cardiovasculares. Doenças Respiratórias

ABSTRACT

CARDOSO, Simone de Souza. *Mixed effects models for predicting air pollution: Global assessment of mortality attributable to air pollution in metropolitan regions of Brazil*. 2018. 194 f. Tese (Doutorado em Saúde Coletiva) – Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. Rio de Janeiro. 2018.

The largest sources of the air pollutants are the burning of fuels from stationary sources (power plants and electric power) and mobile sources (fleet of motor vehicles). Most of these sources is concentrated in the large cities, where significant health impacts are expected, principally in relation to respiratory and cardiovascular diseases. Exposure to air pollutants, especially particulate matter, is associated with several health effects, but the effects on the mortality are undoubtedly the most important and are more favorable overall assessment. Quantifying the magnitude of the impact air pollution on healthy in large cities presents considerable challenges due to limited availability of information. This thesis presents two studies. The first deals with methodological aspects on the prediction of concentration of particulate matter in urban areas where there aren't air quality monitoring networks. The method used was the Linear Mixed-Effects Models for prediction of annual averages of inhalable particulate matter with diameter $\leq 10\mu\text{m}$ (PM₁₀) in the period from 2001 to 2014 in the metropolitan areas of Brazil. The methodology was based on the model proposed by Cohen *et al.* (2004), developed by the World Bank, to estimate PM₁₀ concentrations in terms of economic, meteorological, demographic and other characteristics of metropolitan regions. The results showed that only the relative air humidity, GDP per capita of the services sector, average altitude and fires were associated with PM₁₀ concentration levels. The objective of the second study was to estimate the mortality burden in the metropolitan regions as whole, using mixed-effects models applied to two specific disease groups and age groups: Cardiovascular diseases in adults with 30 years or older and respiratory diseases in children under one year and from one to five years. We estimated the number of attributable deaths and *Years of Life Lost* (YLL) for adults using risk estimates from an American Cancer Society cohort study. And regarding the one-year and one-to-five-year age groups, the risk estimates considered were taken from meta-analytical study of the project - Multicity Study of Air Pollution and Mortality in Latin America (the ESCALA Study).

Keywords: Mixed Effects Models. Forecasting. Particulate Matter. Metropolitan Regions. Global Burden of Disease. Mortality. Cardiovascular Diseases. Respiratory Tract Diseases

LISTA DE FIGURAS

| | | |
|-------------|--|-----|
| Figura 1 - | Esquema de análise do modelo linear de efeitos mistos para previsão..... | 57 |
| Figura 2 - | Médias anuais de PM ₁₀ (µg/m ³) em nove regiões metropolitanas do Brasil durante o período 2001 – 2014..... | 80 |
| Figura 3 - | Previsões das médias anuais de PM ₁₀ do modelo de efeitos mistos nas regiões metropolitanas do Brasil em 2014..... | 89 |
| Figura 4 - | Consumo anual per capita de etanol (litros/hab.) nas regiões metropolitanas, 2014..... | 128 |
| Figura 5 - | Consumo anual per capita de gasolina (litros/hab.) nas regiões metropolitanas, 2014..... | 128 |
| Figura 6 - | Consumo anual per capita de óleo diesel (litros/hab.) nas regiões metropolitanas, 2014..... | 129 |
| Figura 7 - | Média anual de nebulosidade (decimais) nas regiões metropolitanas, 2014..... | 129 |
| Figura 8 - | Média anual da pressão atmosférica (hPa) nas regiões metropolitanas, 2014..... | 130 |
| Figura 9 - | Média anual da temperatura (Celsius) nas regiões metropolitanas, 2014..... | 130 |
| Figura 10 - | Média anual de umidade relativa (%) nas regiões metropolitanas, 2014..... | 131 |
| Figura 11 - | Média anual da velocidade do vento (m/s) nas regiões metropolitanas, 2014..... | 131 |
| Figura 12 - | Total anual do índice pluviométrico (mm) nas regiões metropolitanas, 2014..... | 132 |
| Figura 13 - | Altitude média (em metros) das regiões metropolitanas, 2014..... | 132 |
| Figura 14 - | PIB per capita (em R\$ 1.000) nas regiões metropolitanas, 2014..... | 133 |
| Figura 15 - | PIB per capita do setor serviços (em R\$ 1.000) nas regiões metropolitanas, 2014..... | 133 |
| Figura 16 - | PIB per capita do setor agropecuário (em R\$ 1.000) nas regiões metropolitanas, 2014..... | 134 |

| | |
|--|-----|
| Figura 17 - PIB per capita do setor indústria (em R\$ 1.000) nas regiões metropolitanas, 2014..... | 134 |
| Figura 18 - Total de decolagens (aviões) nas regiões metropolitanas, 2014..... | 135 |
| Figura 19 - Total de frotas de veículos nas regiões metropolitanas, 2014..... | 135 |
| Figura 20 - Taxa de crescimento anual de veículos novos (%) nas regiões metropolitanas, 2014..... | 136 |
| Figura 21 - Total de focos de queimadas nas regiões metropolitanas, 2014..... | 136 |
| Figura 22 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014 na região metropolitana de São Paulo – SP..... | 137 |
| Figura 23 - Estações de monitoramento da região metropolitana de São Paulo - SP com menos de 40% de dados faltantes de PM ₁₀ durante o período de 2001 a 2014..... | 137 |
| Figura 24 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014 na região metropolitana do Rio de Janeiro – RJ..... | 138 |
| Figura 25 - Estações de monitoramento da região metropolitana do Rio de Janeiro - RJ com menos de 40% de dados faltantes de PM ₁₀ durante o período de 2001 a 2014..... | 138 |
| Figura 26 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014 na Baixada Santista – SP..... | 139 |
| Figura 27 - Estações de monitoramento da Baixada Santista - SP com menos de 40% de dados faltantes de PM ₁₀ durante o período de 2001 a 2014..... | 139 |
| Figura 28 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014 em Belo Horizonte – MG..... | 140 |
| Figura 29 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014 em Campinas – SP..... | 140 |
| Figura 30 - Estações de monitoramento de Campinas - SP com menos de 40% de dados faltantes de PM ₁₀ durante o período de 2001 a 2014..... | 141 |
| Figura 31 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014 em Curitiba – PR..... | 141 |
| Figura 32 - Estações de monitoramento de Curitiba - PR com menos de 50% de dados faltantes de PM ₁₀ durante o período de 2001 a 2014..... | 142 |

| | |
|--|-----|
| Figura 33 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014, Grande Vitória – ES..... | 142 |
| Figura 34 - Estações de monitoramento da Grande Vitória - ES com menos de 40% de dados faltantes de PM ₁₀ durante o período de 2001 a 2014..... | 143 |
| Figura 35 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014 na Região metropolitana de Porto Alegre – RS..... | 143 |
| Figura 36 - Estações de monitoramento de PM ₁₀ com informações disponíveis durante o período de 2001 a 2014 na Região metropolitana de Salvador – BA..... | 144 |
| Figura 37 - Estações de monitoramento da Região metropolitana de Salvador - BA com menos de 40% de dados faltantes de PM ₁₀ durante o período de 2001 a 2014..... | 144 |
| Figura 38 - Evolução das médias anuais de PM ₁₀ nas regiões metropolitanas durante o período 2001 a 2014..... | 171 |
| Figura 39 - Probabilidade dos resíduos padronizados (A), função de autocorrelação dos resíduos (B), função de autocorrelação parcial dos resíduos (C) e gráfico dos resíduos padronizados versus valores ajustados às médias anuais de PM ₁₀ (*Y=log(Y)) (D) e box plot dos resíduos padronizados do modelo 1.3 (E)..... | 179 |
| Figura 40 - Probabilidade dos resíduos padronizados (A), função de autocorrelação dos resíduos (B), função de autocorrelação parcial dos resíduos (C), gráfico dos resíduos padronizados versus valores ajustados às médias anuais de PM ₁₀ (*Y=log(Y)) (D) e box plot dos resíduos padronizados (E) do Modelo 1.4.4..... | 182 |

LISTA DE GRÁFICOS

| | | |
|--------------|--|-----|
| Gráfico 1 - | Histograma das médias anuais de PM ₁₀ em áreas metropolitanas..... | 82 |
| Gráfico 2 - | Histograma das médias anuais de PM ₁₀ , após as imputações, em áreas metropolitanas durante o período 2001 a 2014..... | 85 |
| Gráfico 3 - | Comparação das curvas das funções C-R linear e log-linear para o RR às exposições do PM _{2,5} sobre a mortalidade por doenças cardiovasculares..... | 105 |
| Gráfico 4 - | Evolução do consumo anual de gasolina per capita (em R\$ 1.000 l/ano) - Regiões metropolitanas, 2001 a 2014..... | 145 |
| Gráfico 5 - | Evolução do consumo anual de óleo diesel per capita (em R\$ 1.000 l/ano) -Regiões metropolitanas, 2001 a 2014..... | 146 |
| Gráfico 6 - | Evolução do consumo anual de etanol per capita (em R\$ 1.000 l/ano) - Regiões metropolitanas - 2001 a 2014..... | 147 |
| Gráfico 7 - | Evolução da média anual de nebulosidade (decimais) - Regiões metropolitanas, 2001 a 2014..... | 148 |
| Gráfico 8 - | Evolução da média anual da pressão barométrica (hPa) - Regiões metropolitanas, 2001 a 2014..... | 149 |
| Gráfico 9 - | Evolução da média anual da temperatura média (Celsius) - Regiões metropolitanas, 2001 a 2014..... | 150 |
| Gráfico 10 - | Evolução da média anual da umidade relativa do ar (%) - Regiões metropolitanas, 2001 a 2014..... | 151 |
| Gráfico 11 - | Evolução da média anual da velocidade do vento (m/s) - Regiões metropolitanas, 2001 a 2014..... | 152 |
| Gráfico 12 - | Evolução da média anual do índice pluviométrico (mm) - Regiões metropolitanas, 2001 a 2014..... | 153 |
| Gráfico 13 - | Evolução anual da densidade demográfica (Hab./Km ²) - Regiões metropolitanas, 2001 a 2014..... | 153 |
| Gráfico 14 - | Evolução anual do PIB per capita (Total em R \$ 100.000) - Regiões metropolitanas, 2001 a 2014..... | 154 |
| Gráfico 15 - | Evolução anual do PIB setor serviços (Total em R\$ 100.000) - Regiões metropolitanas, 2001 a 2014..... | 154 |

| | |
|---|-----|
| Gráfico 16 - Evolução anual do PIB setor agropecuária (total em R\$ 100.000) - Regiões metropolitanas, 2001 a 2014..... | 155 |
| Gráfico 17 - Evolução anual do PIB setor indústria (total em R\$ 100.000) - Regiões metropolitanas, 2001 a 2014..... | 155 |
| Gráfico 18 - Evolução anual do número de decolagens (aviões) - Regiões metropolitanas, 2001 a 2014..... | 156 |
| Gráfico 19 - Evolução anual de frotas de veículos por km2 - Regiões metropolitanas, 2001 a 2014..... | 157 |
| Gráfico 20 - Evolução da taxa de crescimento anual de veículos novos - Regiões metropolitanas, 2001 a 2014..... | 157 |
| Gráfico 21 - Evolução do total anual do número de focos de queimadas - Regiões metropolitanas, 2001 a 2014..... | 158 |

LISTA DE QUADROS

| | |
|--|-----|
| Quadro 1 - Classes das matrizes de variâncias-covariâncias dos efeitos aleatórios da biblioteca nlme..... | 49 |
| Quadro 2 - Órgãos Estaduais de Meio Ambiente – OEMA..... | 125 |
| Quadro 3 - Critérios de representatividade temporal dos dados utilizados pela CETESB, SMAC, FEAM e IEMA..... | 126 |
| Quadro 4 - Critérios de representatividade temporal dos dados utilizados pela IAP..... | 126 |
| Quadro 5 - Critérios de representatividade de dados utilizados pela FEPAM..... | 127 |
| Quadro 6 - Critérios de representatividade de dados utilizados pela INEA..... | 127 |

LISTA DE TABELAS

| | | |
|-------------|---|-----|
| Tabela 1 - | Padrões das funções variância..... | 50 |
| Tabela 2 - | Estimativas do RR de mortalidade, das inclinações da função C-R e os tipos de estudo recomendados para o cálculo da carga de mortalidade..... | 65 |
| Tabela 3 - | Municípios e população urbana com redes de monitoramento de PM ₁₀ , por região metropolitana, período 2001 a 2014..... | 79 |
| Tabela 4 - | Estatística descritiva das médias anuais de PM ₁₀ (µg/m ³) de acordo com as regiões metropolitanas durante o período de 2001 a 2014..... | 81 |
| Tabela 5 - | Estatística descritiva das características das 36 regiões metropolitanas, 2001 – 2014..... | 83 |
| Tabela 6 - | Estimativas dos parâmetros para os efeitos fixos e aleatórios, erro-padrão, estatística do teste-t e p-valores: modelo AR(1) com estrutura de variância identidade..... | 86 |
| Tabela 7 - | Médias anuais de PM ₁₀ das estações de monitoramento da qualidade do ar e as previsões do modelo de efeitos mistos nas regiões metropolitanas do Brasil..... | 87 |
| Tabela 8 - | Médias anuais dos níveis de concentração de PM ₁₀ das regiões metropolitanas, estimadas pelo modelo de efeitos mistos no período de 2001 – 2014..... | 88 |
| Tabela 9 - | Estimativas do RR de mortalidade, das inclinações da função C-R e os tipos de estudo recomendados para o cálculo da carga de mortalidade e dos óbitos atribuíveis..... | 101 |
| Tabela 10 - | Estimativa da média anual de PM ₁₀ proveniente do modelo de efeitos mistos, média anual de PM _{2,5} e total de óbitos por doenças cardiovasculares e respiratórias em áreas metropolitanas do Brasil, 2014..... | 103 |
| Tabela 11 - | Risco relativo e fração atribuível da mortalidade global associados ao PM _{2,5} para doenças cardiovasculares e PM ₁₀ para doenças respiratórias (IC 95%) | 104 |

| | | |
|-------------|---|-----|
| Tabela 12 - | Resumo dos óbitos e da carga de mortalidade atribuíveis à poluição do ar urbana - Brasil, 2014..... | 105 |
| Tabela 13 - | Análise de sensibilidade da mortalidade cardiovascular para adultos ≥ 30 anos relacionada à exposição a longo prazo de $PM_{2,5}$, regiões metropolitanas - Brasil, 2014..... | 107 |
| Tabela 14 - | Correlação de Spearman dos parâmetros meteorológico entre CETESB e INMET na Região metropolitana de São Paulo, 2001 -2014..... | 159 |
| Tabela 15 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana da Baixada Santista..... | 160 |
| Tabela 16 - | Correlações de Spearman entre as estações para as médias anuais de PM_{10} na Baixada Santista..... | 160 |
| Tabela 17 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana de Belo Horizonte..... | 161 |
| Tabela 18 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana de Campinas..... | 161 |
| Tabela 19 - | Correlações de Spearman entre as estações para as médias anuais de PM_{10} na região metropolitana de Campinas..... | 161 |
| Tabela 20 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana de Curitiba..... | 162 |
| Tabela 21 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana do Espírito Santo..... | 162 |
| Tabela 22 - | Correlações de Spearman entre as estações para as médias anuais de PM_{10} na região metropolitana do Espírito Santo..... | 162 |
| Tabela 23 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana de Porto Alegre..... | 163 |

| | | |
|-------------|--|-----|
| Tabela 24 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana do Rio de Janeiro..... | 163 |
| Tabela 25 - | Correlações de Spearman entre as estações para as médias anuais de PM_{10} na região metropolitana do Rio de Janeiro..... | 165 |
| Tabela 26 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana de Salvador..... | 166 |
| Tabela 27 - | Correlações de Spearman entre as estações para as médias anuais de PM_{10} na região metropolitana de Salvador..... | 166 |
| Tabela 28 - | Percentual de missing nos dados de médias anuais de PM_{10} das estações de monitoramento da poluição do ar na região metropolitana de São Paulo..... | 167 |
| Tabela 29 - | Correlações de Spearman entre as estações para as médias anuais de PM_{10} na região metropolitana de São Paulo..... | 168 |
| Tabela 30 - | Estatística descritiva das médias anuais de PM_{10} , após as imputações de dados, de acordo com as regiões metropolitanas durante 2001 a 2014..... | 169 |
| Tabela 31 - | Estatísticas descritivas das médias anuais de PM_{10} de acordo com as regiões metropolitanas durante 2001 – 2014..... | 170 |
| Tabela 32 - | Modelos considerados na seleção dos efeitos fixos dos termos temporais para a análise das médias anuais de PM_{10} | 172 |
| Tabela 33 - | Modelos considerados na seleção dos efeitos aleatórios para a análise das médias anuais de PM_{10} | 173 |
| Tabela 34 - | Modelos considerados na modelagem dos efeitos aleatórios para a análise das médias anuais de PM_{10} | 173 |
| Tabela 35 - | Estimativas dos parâmetros para os efeitos fixos, erro-padrão, estatística do teste-t e p-valores do modelo 1.1..... | 173 |
| Tabela 36 - | VIF das variáveis independentes..... | 175 |
| Tabela 37 - | Estimativas dos parâmetros para os efeitos fixos, erro-padrão, estatística do teste-t e p-valores do modelo 1.2..... | 176 |
| Tabela 38 - | Estimativas dos parâmetros para os efeitos fixos e aleatórios, erro-padrão, estatística do teste-t e p-valores, modelo 1.3..... | 177 |

| | | |
|-------------|---|-----|
| Tabela 39 - | Modelos considerados para o teste de estruturas de correlação dos erros aleatórios..... | 180 |
| Tabela 40 - | Resumo dos critérios utilizados na modelagem da estrutura de variância-covariância dos erros aleatórios, assumindo diferentes funções variância..... | 181 |
| Tabela 41 - | Estimativas dos parâmetros para os efeitos fixos e aleatórios, erro-padrão, estatística do teste-t e p-valores: modelo AR(1) com estrutura da função variância varIdent..... | 184 |
| Tabela 42 - | Médias anuais dos níveis de concentração de PM ₁₀ das regiões metropolitanas, estimadas pelo modelo de efeitos mistos no período de 2001 – 2014..... | 185 |
| Tabela 43 - | Previsão das médias anuais de PM ₁₀ do modelo de efeitos mistos, estimativas das médias anuais de PM _{2,5} e óbitos por doenças cardiovasculares e respiratórias em áreas metropolitanas do Brasil, 2014..... | 186 |
| Tabela 44 - | Risco relativo da mortalidade associada com PM _{2,5} para doenças cardiovasculares e PM ₁₀ para doenças respiratórias de acordo com o cenário base (IC95%)..... | 187 |
| Tabela 45 - | Fração atribuível da mortalidade associada com PM _{2,5} para doenças cardiovasculares e PM ₁₀ para doenças respiratórias de acordo com o cenário base (IC 95%)..... | 188 |
| Tabela 46 - | Óbitos atribuíveis associados com PM _{2,5} para doenças cardiovasculares e PM ₁₀ para doenças respiratórias de acordo com o cenário base (IC 95%)..... | 189 |
| Tabela 47 - | RM São Paulo: Características das estações de monitoramento PM ₁₀ | 190 |
| Tabela 48 - | RM Baixada Santista: Características das estações de monitoramento PM ₁₀ | 191 |
| Tabela 49 - | RM Campinas: Características das estações de monitoramento PM ₁₀ | 191 |
| Tabela 50 - | RM Rio de Janeiro: Características das estações de monitoramento PM ₁₀ | 191 |

| | | |
|-------------|--|-----|
| Tabela 51 - | RM Salvador: Características das estações de monitoramento PM ₁₀ | 193 |
| Tabela 52 - | RM Belo Horizonte: Características das estações de monitoramento PM ₁₀ | 193 |
| Tabela 53 - | RM Curitiba: Características das estações de monitoramento PM ₁₀ | 193 |
| Tabela 54 - | RM Porto Alegre: Características das estações de monitoramento PM ₁₀ | 194 |
| Tabela 55 - | RM Grande Vitória: Características das estações de monitoramento PM ₁₀ | 194 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|----------|---|
| ACS | <i>American Cancer Society</i> |
| AIC | <i>Akaike Information Criterion</i> |
| ANAC | Agência Nacional de Aviação Civil |
| ANP | Agência Nacional do Petróleo |
| ARIMA | <i>Autoregressive Integrated Moving Average</i> |
| BIC | <i>Bayesian Information Criterion</i> |
| BLUP | <i>Best Linear Unbiased Predictors</i> |
| CETESB | Companhia Ambiental do Estado de São Paulo |
| CID-10 | Décima Revisão da Classificação Internacional de Doenças e Problemas Relacionados à Saúde |
| CONAMA | Conselho Nacional do Meio Ambiente |
| C-R | Concentração-Resposta |
| DP | Desvio-padrão |
| DALY | <i>Disability-Adjusted Life-Years</i> |
| DENATRAN | Departamento Nacional de Trânsito |
| EM | <i>Expectation – Maximization</i> |
| EPA | <i>Environmental Protection Agency</i> |
| ESCALA | <i>Estudio de Salud y Contaminación del Aire en Latino américa</i> |
| ESCAPE | <i>European Study of Cohorts for Air Pollution Effects</i> |
| FA | Fração atribuível |
| FAC | Função de Autocorrelação |
| FACP | Função de Autocorrelação Parcial |
| FEAM | Fundação Estadual de Meio Ambiente |
| FEPAM | Fundação Estadual de Proteção Ambiental |
| GBD | <i>Global Burden of Disease</i> |
| GMAPS | <i>Global Model of Ambient Particulates Ambientais</i> |
| IAP | Instituto Ambiental do Paraná |
| IBGE | Instituto Brasileiro de Geografia e Estatística |
| IC | Intervalo de confiança |

| | |
|-------------------|--|
| IEMA | Instituto Estadual de Meio Ambiente e Recursos Hídricos |
| INEA | Instituto Estadual do Ambiente |
| INEMA | Instituto de Meio Ambiente e Recursos Hídricos |
| INMET | Instituto Nacional de Meteorologia |
| INPE | Instituto Nacional de Pesquisas Espaciais |
| IRA | Infecções Respiratórias Agudas |
| LUR | <i>Land use Regression models</i> |
| MAE | <i>Mean Absolute Error</i> |
| ML | <i>Maximum Likelihood</i> |
| MP | Material Particulado |
| MQG | Mínimos Quadrados Generalizados |
| OEMA | Órgãos Estaduais de Meio Ambiente |
| OLS | <i>Ordinary Least Squares</i> |
| OMS | Organização Mundial da Saúde |
| PIB | Produto Interno Bruto |
| PM ₁₀ | Material Particulado inalável com diâmetro $\leq 10\mu\text{m}$, corresponde à fração inalável das Partículas Totais em Suspensão |
| PM _{2,5} | Material Particulado fino com diâmetro $\leq 2,5\mu\text{m}$ |
| PTS | Partículas Totais em Suspensão |
| REML | <i>Restricted Maximum Likelihood</i> |
| RMSE | <i>Root Mean Square Error</i> |
| RR | Risco Relativo |
| SIM | Sistema de Informações de Mortalidade |
| SMAC | Secretaria Municipal de Meio Ambiente |
| TM5 | Modelos de Transporte Químicos Global |
| TRV | Teste da Razão de Verossimilhança |
| VIF | Variance Inflation Factors |
| YLL | <i>Years of Life Lost</i> |

LISTA DE SÍMBOLOS

| | |
|--------------------------|--|
| $\mu\text{g}/\text{m}^3$ | Micrograma por metro cúbico, concentração gravimétrica do poluente no ar |
| μm | Micrômetro |
| hPa | Hectopascal, unidade de medida da pressão exercida pela atmosfera sobre a superfície (pressão atmosférica) |

SUMÁRIO

| | | |
|---------|---|----|
| | INTRODUÇÃO | 27 |
| 1 | REVISÃO DE LITERATURA | 32 |
| 1.1 | Estudos de modelos de efeitos mistos para previsão da poluição do ar | 32 |
| 1.2 | Outras abordagens para previsão da poluição do ar | 34 |
| 1.2.1 | <u>Análise espacial integrada aos estudos de previsão da poluição do ar</u> | 34 |
| 2 | JUSTIFICATIVAS | 37 |
| 2.1 | O uso de modelos de efeitos mistos para prever a poluição do ar nas regiões metropolitanas do Brasil | 37 |
| 2.2 | Avaliação do impacto da poluição do ar na mortalidade em regiões metropolitanas do Brasil | 37 |
| 3 | OBJETIVOS | 38 |
| 3.1 | Objetivos gerais | 38 |
| 3.2 | Objetivos específicos | 38 |
| 4 | MÉTODOS | 39 |
| 4.1 | Modelo linear de efeitos mistos para dados longitudinais | 39 |
| 4.1.1 | <u>O modelo</u> | 39 |
| 4.1.2 | <u>Inferência para os efeitos fixos</u> | 42 |
| 4.1.2.1 | Método da Máxima Verossimilhança..... | 42 |
| 4.1.2.2 | Método da Máxima Verossimilhança Restrita..... | 43 |
| 4.1.3 | <u>Teste da Razão de Verossimilhanças</u> | 44 |
| 4.1.4 | <u>Testes T e F condicionais</u> | 45 |
| 4.1.5 | <u>Intervalos de confiança</u> | 46 |
| 4.1.6 | <u>Modelagem das estruturas residuais das matrizes de variância-covariância</u> | 47 |
| 4.1.7 | <u>Adequação do Modelo</u> | 49 |
| 4.1.7.1 | Violação da hipótese de normalidade..... | 49 |
| 4.1.7.2 | Modelagem da heterocedasticidade..... | 50 |
| 4.1.7.3 | Modelagem da dependência serial..... | 51 |
| 4.1.8 | <u>Validações - Performance do Modelo</u> | 54 |
| 4.1.8.1 | Coeficiente de determinação..... | 54 |
| 4.1.8.2 | Estatísticas de performance e "Data Splitting"..... | 55 |

| | | |
|---------|---|-----|
| 4.1.9 | <u>Valores ajustados e previsões</u> | 56 |
| 4.2 | Análise Estatística | 58 |
| 4.2.1 | <u>Fontes de dados</u> | 58 |
| 4.2.2 | <u>Estimação das médias anuais de PM₁₀ nas regiões metropolitanas do Brasil</u> | 60 |
| 4.2.3 | <u>Estimação da carga de mortalidade e dos óbitos atribuíveis ao material particulado</u> | 63 |
| 4.2.3.1 | Estudo da Associação Americana de Câncer..... | 63 |
| 4.2.3.2 | Caso base..... | 63 |
| 4.2.4 | <u>Análise de sensibilidade</u> | 66 |
| 4.2.5 | <u>Aspectos computacionais</u> | 67 |
| 5 | RESULTADOS | 68 |
| 5.1 | O uso de modelos de efeitos mistos para prever a poluição do ar nas regiões metropolitanas do Brasil (manuscrito 1) | 68 |
| 5.2 | Avaliação do impacto da poluição do ar na mortalidade em regiões metropolitanas do Brasil (manuscrito 2) | 94 |
| | CONCLUSÃO | 115 |
| | REFERÊNCIAS | 116 |
| | APÊNDICE A - Estruturação das redes de monitoramento de PM ₁₀ nas regiões metropolitanas do Brasil..... | 124 |
| | APÊNDICE B - Critérios de validação dos dados das redes de monitoramento..... | 126 |
| | APÊNDICE C - Mapas dos fatores contextuais das regiões metropolitanas do Brasil – 2014..... | 128 |
| | APÊNDICE D - Mapas ilustrativos da distribuição das estações de monitoramento de PM ₁₀ nas regiões metropolitanas do Brasil..... | 137 |
| | APÊNDICE E - Gráficos ilustrativos da evolução dos fatores contextuais nas regiões metropolitanas brasileiras durante o período 2001 a 2014..... | 145 |
| | APÊNDICE F - Correlações entre fontes de dados dos parâmetros meteorológicos..... | 159 |
| | APÊNDICE G - Sumário de dados faltantes, correlações entre as estações das médias anuais de PM ₁₀ e estatística descritiva após a imputação de dados..... | 160 |

| | |
|--|-----|
| APÊNDICE H - Modelagem e estimativas dos níveis de concentração de PM ₁₀ das regiões metropolitanas do Brasil no período de 2001 a 2014..... | 170 |
| APÊNDICE I - Tabelas ilustrativas de avaliação da carga de mortalidade nas regiões metropolitanas brasileiras, 2014..... | 186 |
| APÊNDICE J - Características das redes de monitoramento das regiões metropolitanas do Brasil..... | 190 |

APRESENTAÇÃO

Esta tese de doutorado é uma atualização e extensão dos trabalhos De Leon *et al.* (2012) e Jacobson *et al.* (2012) em que o Modelo Brasileiro de Partículas Ambientais (BMAPS) foi aplicado para prever a média anual de PM₁₀ em todas as regiões metropolitanas do Brasil com a finalidade de avaliar o impacto dos níveis de material particulado na saúde. Os temas desenvolvidos no presente trabalho também abordam questões metodológicas relacionadas com previsões e medidas de impactos da poluição do ar na saúde restritas às áreas urbanas. Ainda em De Leon *et al.* (2012) e Jacobson *et al.* (2012), as previsões da poluição do ar em áreas urbanas foram geradas para o período 1998 a 2008 pelo modelo BMAPS composto por efeitos fixos tais como tendência do tempo, umidade anual, temperatura e Produto Interno Bruto (PIB) para Serviços. E por efeitos aleatórios somente com tendência do tempo. A proposta que aqui é apresentada faz além da atualização, gerando previsões para regiões metropolitanas durante o período de 2001 a 2014, faz também uma expansão do modelo com um novo conjunto de variáveis explicativas.

A tese está composta de duas partes que se complementam. O artigo 1 tem caráter metodológico e objetiva contribuir com ferramentas para o desenvolvimento de estudos sobre impactos da poluição do ar na saúde, enfatizando estudos epidemiológicos de séries temporais. Apesar de orientados para os estudos dos efeitos da poluição do ar na saúde, os aspectos metodológicos abordados e desenvolvidos nesta tese podem ser aplicados não só na área da Epidemiologia, mas também em outras áreas como ciências econômicas, ambientais e geográficas; O artigo 2 aborda um tema investigativo sobre os efeitos da poluição do ar na saúde em áreas metropolitanas, com o objetivo de contribuir para Saúde Pública na conscientização sobre o impacto que a poluição do ar pode causar na mortalidade em áreas urbanas.

O capítulo sobre a revisão de literatura focaliza o tema de poluição do ar e seus impactos na saúde com ênfase nos métodos epidemiológicos relacionados ao método abordado na tese. Ainda outras abordagens metodológicas para o apoio nos cálculos de impactos também foram citadas. Já na seção de introdução de cada artigo, o leitor poderá identificar os elementos da revisão bibliográfica destacando o tema abordado no respectivo artigo. Poderá observar também que algumas referências bibliográficas são comuns em ambos os artigos.

Esta tese está organizada da seguinte forma: introdução, onde foi feita uma breve contextualização sobre o assunto da tese; o capítulo 1, apresenta a revisão de literatura; capítulo 2, apresenta as justificativas para a realização de cada estudo; no capítulo 3 estão descritos os

objetivos gerais e os objetivos específicos de cada artigo; capítulo 4, destinado à abordagem do método da tese; capítulo 5, apresenta como resultado dois artigos: o primeiro artigo, contido na seção 5.1, refere-se a um estudo metodológico. E o segundo, contido na seção 5.2, refere-se a um estudo de aspecto epidemiológico. Ao final de cada artigo estão listadas as referências bibliográficas. Nas páginas finais da tese uma breve conclusão foi redigida. Estão listadas também todas as referências bibliográficas utilizadas e os apêndices, onde estão incluídos tabelas, gráficos e mapas não publicáveis dos artigos.

INTRODUÇÃO

O Material Particulado (MP) é uma mistura de partículas líquidas e sólidas em suspensão na atmosfera conhecidas como aerossóis, sendo mais impactante em áreas urbanas. Pode ser classificado como um poluente primário (poluentes emitidos para a atmosfera) ou um poluente secundário (formado na própria atmosfera). Quanto ao seu tamanho, com base no diâmetro aerodinâmico da partícula, suas medidas variam em três principais frações: partículas grossas (2,5-10 μm), partículas finas ($\leq 2,5\mu\text{m}$) - PM_{2,5}- e partículas ultrafinas ($\leq 0,1\mu\text{m}$). O material particulado inalável tem o diâmetro $\leq 10 \mu\text{m}$ (PM₁₀), em sua composição, inclui partículas finas e grossas e são tipicamente medidos na atmosfera para estudos dos efeitos em saúde (WHO, 2006). Para políticas de saúde pública, em termos de estabelecer diretrizes ou padrões para níveis aceitáveis de MP, concentraram-se principalmente em indicadores de partículas finas (PM_{2,5}), partículas inaláveis (PM₁₀) e partículas grossas (PM_{10-2,5}). As ações regulatórias vinda da Organização Mundial da Saúde (OMS), *Environmental Protection Agency* (EPA) e do Conselho Nacional do Meio Ambiente (CONAMA), órgão que estabelece políticas para o meio ambiente no Brasil, definem padrões de qualidade do ar referentes a esses poluentes. Os limites das médias anuais definidos pela OMS para PM₁₀ é de 20 $\mu\text{g}/\text{m}^3$ e para PM_{2,5} de 10 $\mu\text{g}/\text{m}^3$. Pela EPA é de 50 $\mu\text{g}/\text{m}^3$ para PM₁₀ e de 15 $\mu\text{g}/\text{m}^3$ para PM_{2,5}. E pelo CONAMA a média anual é de 50 $\mu\text{g}/\text{m}^3$ para PM₁₀. No Brasil, não há uma definição nos limites da média anual para PM_{2,5}, justamente este, o poluente de partículas finas que tem uma maior importância para a saúde, por ser mais tóxicos, alcança regiões mais profundas dos pulmões, permanece suspenso por longos períodos de tempo e é facilmente transportados para longas distâncias (Pope e Dockery, 2006).

É comprovado que elevados níveis de concentração de MP e outros poluentes atmosféricos causam aumento na mortalidade, nos atendimentos ambulatoriais e em admissões hospitalares. Hipótese esta que foi testada nos EUA e na Europa em meados do século XX, quando houve eventos maiores de poluição do ar seguidos por aumentos significativos na mortalidade, principalmente em pacientes com doença cardiorrespiratória existente (Firket, 1936; Logan, 1953; Schwartz e Marcus, 1990; Schwartz, 1994). Contudo, estudos mais recentes demonstraram que os efeitos da poluição do ar sobre a saúde estão presentes mesmo em níveis mais baixos, principalmente no que se refere às exposições a poluentes como o MP (Ostro, 1984; Pope *et al.*, 1992; Dockery e Pope, 1994; Schwartz, 1995; Brunekreef, 1997; Mage *et al.*, 1999; Brunekreef e Holgate, 2002), dando indícios de que não existe um limiar de efeito zero

para as partículas de poluição e que os riscos para a saúde podem estar presentes em qualquer nível de exposição. Desse modo, grande parte das pesquisas epidemiológicas têm se concentrado em doenças respiratórias e cardiovasculares encontrando evidências consistentes à premissa de que o MP é um fator de risco importante para essas doenças (Dockery e Pope, 1994; Schwartz, 1994; Lin *et al.*, 1999; Pope e Dockery, 1999; Dominici *et al.*, 2000; Samet *et al.*, 2000; Katsouyanni *et al.*, 2001; Dominici *et al.*, 2002; Dominici *et al.*, 2004; Wong *et al.*, 2008; Romieu *et al.*, 2012). Essas descobertas geraram preocupações em relação aos efeitos da poluição atmosférica na saúde pública. Apesar da maior parte das recentes investigações em estudos de séries temporais epidemiológicas tenha incidido em análises sobre os efeitos de exposições agudas da poluição do ar¹, a avaliação sobre o impacto na mortalidade é limitada porque os estudos sobre os efeitos agudos na mortalidade diária não permitem, por si só, estimar o tempo de vida perdido, mas apenas o número diário de mortes antecipadas (Mcmichael *et al.*, 1998; Kunzli *et al.*, 2001), desta forma esses efeitos fornecem poucas informações sobre o quanto a vida é reduzida, uma vez que, a poluição do ar afeta também as taxas de mortalidade a longo prazo no modo de indução de doenças crônicas (Abbey *et al.*, 1999; Pope e Dockery, 1999; Zeger *et al.*, 1999; Schwartz, 2000; 2001; Zanobetti *et al.*, 2002; Dominici *et al.*, 2003)². Essa é uma questão muito importante para a saúde pública e exige, portanto, o conhecimento da carga global de doenças – “*Global burden of disease* (GBD) e de mortes prematuras evitáveis para o controle de poluentes atmosféricos específicos (Kunzli *et al.*, 2000). Além disso, vale lembrar que, embora os efeitos da poluição atmosférica sobre as taxas de mortalidade nos estudos de séries temporais sejam pequenas (Ostro, 1996; Kunzli *et al.*, 2000), a carga de doença atribuível à poluição do ar pode ser substancial, quando uma grande população exposta é considerada.

Apesar da importância do conhecimento da carga global de doença e de mortes atribuíveis, quantificar a magnitude dos impactos da poluição do ar na saúde apresenta problemas consideráveis, como a não representatividade das medições dos poluentes, variações das exposições de poluentes, principalmente, em áreas urbanas (Sarnat *et al.*, 2010) e/ou variações das exposições individuais (Navidi e Lurmann, 1994; Zeger *et al.*, 2000). Além disso,

¹ Os estudos sobre os efeitos das exposições agudas da poluição do ar são estudos sobre variações temporais de curto prazo da poluição associadas a variações (também de curto prazo) das medidas de saúde. Esses efeitos tipicamente avaliam relações temporais de 1 a 5 dias (Pope e Dockery, 1999).

² Os estudos de exposição crônica usam diferenças espaciais da poluição como fonte de variabilidade de exposição. Comparam vários resultados de saúde em comunidades ou bairros com diferentes níveis de poluição. Esses estudos são principalmente transversais e usam dados de poluição de longo prazo (um ano ou mais). Esses estudos são frequentemente interpretados como avaliações de doenças crônicas e/ou efeitos cumulativos da exposição (Pope e Dockery, 1999).

a disponibilidade limitada de monitores de medição dos poluentes atmosféricos torna-se, também, um problema pois impossibilita obter medidas globais de impacto.

É de conhecimento na literatura que impactos significativos do MP na saúde são esperados nos grandes centros urbanos (Brunekreef, 1997; Kunzli *et al.*, 2000) já que neles estão concentradas as maiores fontes de emissão da poluição do ar como queima de combustíveis, a partir de fontes estacionárias, como usinas e energia elétrica; e fontes móveis como, por exemplo, frota de veículos automotores (carros, caminhões e ônibus) principais responsáveis pelo aumento nos níveis de poluição (Holman, 1999). Fontes como poeira rodoviária, queima de biomassa e processos de fabricação também contribuem para o aumento da poluição atmosférica. Sendo assim, em muitos países, grande parte dos sistemas de monitoramento da poluição do ar estão localizados nesses grandes centros urbanos e já fazem parte de programas regulatórios destinados a proteger a saúde pública e o meio ambiente (Grant *et al.*, 1999). Os sistemas mais extensos são encontrados nos Estados Unidos e nos países da Europa Ocidental, onde estão vigentes desde meados da década de 1970. Porém países de baixa e média renda, localizados na África, na Ásia e na América Latina, incluindo o Brasil, não são tão aprimorados quanto os países de alta renda porque o acesso aos dados coletados por esses sistemas e a padronização internacional dos métodos de monitoramento são limitados e dessa forma não fornecem dados necessários ou fornecem de maneira precária dados para a avaliação do impacto da poluição do ar na saúde, que são muitas vezes solicitados nas pesquisas epidemiológicas (Krzyzanowsk e Schwela, 1999). Além disso, em muitas regiões, desses países, não há nem sistema de monitoramento da poluição do ar, e estas devem ser estimadas utilizando metodologias para previsão.

De acordo com o relatório de 2014 do Instituto de Energia e Meio Ambiente (IEMA, 2014), a estruturação de redes de monitoramento do ar, no Brasil, ainda é muito restrita e tímida, seja na cobertura territorial, no número de parâmetros monitorados, na representatividade das medições e no histórico amostral. Isso dá indícios de que esse tipo de atividade ainda não é tão prioritário na gestão ambiental brasileira. Reforçado, em tese, pelas dificuldades gerenciais, ao baixo número de técnicos envolvidos e à falta de recursos orçamentárias por parte dos próprios órgãos estaduais de meio ambiente (OEMA) para implantação, manutenção e expansão das redes. As dificuldades encontradas, muitas vezes acarretam na descontinuidade e na baixa cobertura do monitoramento no país, contribuindo, portanto, para a ocorrência de lacunas nos dados de poluição. Apesar do reconhecimento das fragilidades, alguns caminhos foram tomados por parte desses órgãos para o enfrentamento desses obstáculos como, por exemplo, arranjos

gerenciais de redes, diversificação das fontes de recursos e outras medidas destinadas a redução de custos (IEMA, 2014).

Assim sendo, diante de tais problemas, nos estudos de séries temporais epidemiológicos, uma forma de contornar o problema de dados faltantes das exposições à poluição do ar de longo prazo é a aplicação de modelos estatísticos. A obtenção de estimativas confiáveis nesses modelos, baseia-se na completa captura de estruturas temporais (por exemplo, tendências temporais de curto e longo prazo) e de variações regionais ou locais. Porém, a utilização de tais modelos é desafiadora porque os dados de medição são limitados nos locais e no tempo, e existem associações complexas e não-lineares entre os preditores (como variáveis meteorológicas e de tráfego) e concentrações de poluentes.

Neste contexto, a metodologia dos modelos de efeitos mistos é capaz de estimar, de forma confiável, médias anuais das concentrações de MP em áreas urbanas sem dados de monitoramento disponíveis, o que possibilita cálculos de medidas globais de impacto em áreas urbanas. Dessa forma, mais especificamente, esses modelos não somente podem ser utilizados para abordar questões de associação, quantificando fontes de variabilidade, como também para fazer previsões da poluição do ar. Fornecem uma ferramenta flexível e poderosa para a análise de dados agrupados, permitindo a utilização de diferentes estruturas temporais, assim como a captura das variações locais. Suas estimativas estão associadas às incertezas levando em consideração as características contextuais das áreas urbanas onde nem todos os locais de monitoramento medem o mesmo conjunto de poluentes. O método produz uma distribuição preditiva que engloba vários locais em diferentes pontos do tempo, utilizando dados disponíveis e permitindo uma previsão simultânea seccional e temporal. Pontos importantes num bom modelo de regressão para previsão devem ser considerados como a apresentação de um bom ajuste (R^2 alto), plausibilidade e estabilidade nos seus coeficientes estimados.

Formalizada por Cohen *et al.* (2004) e desenvolvida pelo banco mundial, a metodologia de previsão denominada como “*Global model of ambient particulates*” (GMAPS) é um modelo linear de efeitos mistos (econométrico) para dados longitudinais que tem o propósito de estimar níveis atuais de concentrações de MP em cidades do mundo com populações maiores que 100.000 habitantes. e incorpora os principais determinantes das variações de concentrações do poluente no ambiente externo (“*outdoors*”) - variáveis de consumo de energia, socioeconômicas, ambientais, demográficas e outras variáveis locais. O modelo gera estimativas com base em medições disponíveis de MP em área já monitorada. Não sendo um modelo causal, apenas prediz níveis de concentrações de poluição em áreas não monitoradas com intuito de apoiar cálculos, em termos globais, para avaliação do impacto dos níveis de

concentração da poluição do ar na saúde (Cohen *et al.*, 2004; Pandey *et al.*, 2006). O modelo GMAPS foi utilizado como metodologia para estimar MP em cidades > 100.000 hab. para o GBD 2000.

1 REVISÃO DE LITERATURA

1.1 Estudos de modelos de efeitos mistos para previsão da poluição do ar

Para o desenvolvimento do estudo com enfoque no modelo GMAPS para avaliação do impacto da poluição do ar na saúde, foram realizadas buscas da literatura na seguinte web site: Pubmed – www.pubmed.com. Os descritores de assunto utilizados na base de dados foram cruzados da seguinte forma:

- Pesquisa 1: air pollution AND Global Burden of Disease AND mortality
- Pesquisa 2: Global Burden of Disease AND Particulate Matter

Foram utilizados como critérios de inclusão os artigos publicados em qualquer período do tempo, de qualquer idioma, bem como a inclusão de todos os artigos publicados que abordassem em todos os campos (título, resumo, MeSH e outros) os descritores.

Apenas três estudos encontrados (Cohen *et al.*, 2004; Norman *et al.*, 2007; Brauer *et al.*, 2012) abordaram modelos de efeitos mistos ou GMAPS para estimar as concentrações de MP para avaliação do impacto da poluição do ar na saúde.

O estudo de Cohen *et al.* (2004) obteve estimativas comparáveis de exposição à poluição do ar para 14 grupos de países definidos pela OMS de acordo com os estratos de mortalidade, no período de 1985 a 1999. Utilizou o método GMAPS para estimar as concentrações de PM₁₀ para o ano de 1999. Especificamente, as variáveis de consumo de energia, socioeconômicas, ambientais, demográficas, e medições de PM₁₀, disponíveis em 304 cidades, foram usadas para prever os níveis de PM₁₀ em todas as 3.211 cidades em todo o mundo com população superior a 100.000 habitantes. O modelo explicou 88% da variabilidade indicando um bom ajuste. A correlação total entre os dados observados e previstos para as observações de PM₁₀ e de partículas totais em suspensão (PTS) foi de cerca de 90% ($r=0,9$). As estimativas de PM₁₀ foram convertidas em estimativas de MP fino, com diâmetro $\leq 2,5 \mu\text{m}$ (PM_{2,5}), utilizando informações disponíveis sobre a variação geográfica na proporção PM_{2,5}/PM₁₀. As médias anuais regionais de PM_{2,5} e PM₁₀ foram ponderadas pela população. As estimativas para o GBD 2000 foram baseadas nas contribuições de três desfechos de saúde: mortalidade por causas cardiovasculares em adultos, câncer de pulmão e infecções respiratórias agudas em crianças de 0 a 5 anos de idade. Os números atribuíveis às mortes e anos de vida perdidos – “*Years of Life Lost*” (YLL) – em adultos e crianças (<5 anos) foram estimados utilizando as estimativas de risco de um grande estudo de coorte dos EUA feito em adultos, o “*American Cancer Society*” (ACS) (Pope

et al., 2002), e de uma metanálise de cinco estudos de séries temporais sobre mortalidade em crianças. Foi considerado como caso base: estimativas de risco da função concentração-resposta (C-R) linear e média anual contrafactual de concentração de $PM_{2,5}$ e PM_{10} de 7,5 e 15 $\mu\text{g}/\text{m}^3$, respectivamente. Neste estudo, a poluição do ar foi estimada em áreas urbanas de todo mundo em termos de MP para o ano 2000. Cerca de 3% da mortalidade atribuível ao MP foi por doenças cardiovasculares em adultos, 5% por câncer de traqueia, brônquios e pulmão, e 1% por infecções respiratórias agudas em crianças de 0-4 anos. Isso equivaleu a cerca de 800.000 mortes (1,4% do total global) e 6,4 milhões de YLL (0,7% do total global). Os resultados indicaram que o impacto da poluição do ar sobre a carga de doença nas cidades do mundo é grande, com provável subestimação devido às incertezas. Houve também uma variação considerável nas estimativas entre os 14 grupos, tendo maior impacto nas cidades mais poluídas e crescimento rápido nos países em desenvolvimento.

Outro estudo, no qual se propôs utilizar estimativas de PM_{10} geradas para a África do Sul pelo método GMAPS, foi o de Norman *et al.* (2007) para quantificar o impacto na mortalidade por doenças cardiovasculares e câncer em adultos ≥ 30 anos e por doenças respiratórias em crianças < 5 anos atribuídas à poluição do ar na África do Sul em 2000. Entretanto segundo os autores, as estimativas se demonstraram subestimadas quando comparadas as estimativas vindas de dados de monitores locais, sendo estas, portanto, escolhidas para avaliação do impacto da poluição do ar na saúde. As estimativas de PM_{10} foram convertidas para estimativas de $PM_{2,5}$, utilizando informações disponíveis sobre variação geográfica. As médias anuais de concentração de ambos os poluentes foram ponderados de acordo com o tamanho da população. Para estimar o risco relativo (RR) de mortalidade por doenças cardiovasculares e câncer de pulmão em adultos com 30 anos ou mais, foram utilizados os resultados do estudo da ACS. Os autores consideraram como caso base: as estimativas de risco da função C-R linear do estudo da ACS e média anual contrafactual (ou mínimo teórico) de concentração de $PM_{2,5}$ e PM_{10} de 7,5 e 15 $\mu\text{g}/\text{m}^3$, respectivamente. Para estimar a relação entre a exposição à MP e a mortalidade por infecções respiratórias agudas (IRA) entre crianças de 0 a 4 anos, foi tida como base a estimativa resumida de cinco estudos de séries temporais do estudo Cohen *et al.* (2004). Esses estudos foram resumidos através de uma média ponderada das estimativas dos estudos individualmente. Estimou-se que MP impactaria 3,7% na mortalidade por doenças cardiovasculares e 5,1% na mortalidade por câncer de traqueia, brônquios e pulmão em adultos ≥ 30 anos, e 1,1% na mortalidade por doenças respiratórias em crianças < 5 anos. Isso equivaleu a 4.637 mortes ou 0,9% (IC95%:0,3-1,5%) sobre todos os

óbitos e cerca de 42.000 YLL, ou 0,4% (IC 95%:0,1 – 0,7%) sobre todos os YLL na África do Sul em 2000.

O estudo de Brauer *et al.* (2012) abordou a metodologia GMAPS para previsão de PM_{2,5}. Porém, foram identificadas limitações significativas da metodologia no estudo, como, por exemplo, a não aplicabilidade do modelo em áreas rurais e ao poluente ozônio. Sendo assim, foram utilizados outros métodos de previsão, como satélites e modelos de transporte químicos global (TM5) para gerar estimativas globais de concentrações médias de PM_{2,5} e ozônio para 1990 e 2005. Como resultado, constataram que em 2005, 89% da população mundial vivia em áreas onde a diretriz da qualidade do ar da OMS de 10 µg/m³ para PM_{2,5} (média anual) foi excedida. Globalmente, 32% da população vivia em áreas que excediam da meta provisória da fase 1 da OMS de 35 µg/m³; impulsionado por altas proporções na Ásia Oriental (76%) e Sul (26%). Os maiores níveis sazonais de ozônio foram encontrados na América do Norte e na América Latina, Europa, Sul e Leste da Ásia, e partes da África. Mundialmente, entre 1990 e 2005 o aumento de 6% no PM_{2,5} e uma diminuição de 1% na concentração de ozônio (ponderados pela população) foi evidente.

1.2 Outras abordagens para previsão da poluição do ar

1.2.1 Análise espacial integrada aos estudos de previsão da poluição do ar

Ao contrário dos estudos de séries temporais, um amplo número de estudos de análise espacial é encontrado na literatura sobre previsão dos níveis de concentração da poluição do ar envolvendo métodos como “*Land use regression models*” (LUR), satélites e modelos de transporte químicos e outros para gerar estimativas globais de impacto. São também técnicas que permitiram lidar com os problemas de dados incompletos. A maioria dos estudos de análise espacial sobre avaliações de impacto da exposição à poluição do ar na saúde envolvem comparações entre cidades usando um número limitado de monitores. A disponibilidade de medições de concentrações ambientais de MP varia amplamente em todo o mundo, tornando a estimativa das médias anuais um desafio considerável (Krzyzanowsk e Schwela, 1999).

Os métodos de previsão espacial são utilizados, também, para prever níveis de poluição atmosférica em locais não mensurados, com base nos níveis de concentração observados em

locais já monitorados. Tal abordagem tem sido muito utilizada para doenças crônicas, como câncer com longos períodos de latência, em que a exposição acumulada é relevante.

Modelos espaciais são explorados para avaliações e previsões dos impactos da exposição interurbanas. Esses modelos são classificados sob seis classes de avaliação: modelos baseados em aproximação, interpolação estatística, LUR, modelos de dispersão, modelos integrados e modelos híbridos (Briggs *et al.*, 1997; Mulholland *et al.*, 1998; Lebret *et al.*, 2000; Hoek *et al.*, 2001; Hrubá *et al.*, 2001; Janssen *et al.*, 2001; Hoek *et al.*, 2008); além de satélites e modelos de transporte químicos (Hoff e Christopher, 2009; Anenberg *et al.*, 2010; Emili *et al.*, 2010; Van Donkelaar *et al.*, 2010; Brauer *et al.*, 2012; Chang *et al.*, 2014; Chudnovsky *et al.*, 2014; Kloog *et al.*, 2015; Li *et al.*, 2015; Stieb *et al.*, 2015).

A metodologia LUR visa prever concentrações de poluição em um determinado local. Formalmente, este método usa, como variável dependente, concentrações de poluição medidas como y na localização s e como preditores x 's algumas características urbanas (por exemplo, intensidade de tráfego, densidade populacional, indústria) dentro de áreas próximas à localização s . O método envolve a utilização de modelos de regressão de mínimos quadrados para prever superfícies de poluição com base em dados de monitoramento de poluição e variáveis independentes exógenas existentes (Elliot *et al.*, 2000). Utilizando modelos de regressão para prever concentrações médias anuais, Brauer *et al.* (2003) estimaram exposições médias da poluição atmosférica para Munique, cidade Alemã e Estocolmo, capital da Suécia. Os modelos LUR também foram utilizados para estimar as concentrações de poluição do ar nos endereços domiciliares de participantes nos estudos de saúde envolvidos no “*European Study of Cohorts for Air Pollution Effects*” (ESCAPE) (Eeftens *et al.*, 2012).

Satélites de sensoriamento remoto e TM5 são métodos que permitem estimações de exposições da poluição do ar em populações. Eles simulam componentes de ozônio e aerossóis fornecendo uma completa cobertura global em alto nível de resolução espacial (Brauer *et al.*, 2012).

Brauer *et al.* (2012) utilizaram métodos com base em satélites de sensoriamento remoto e TM5 para prever $PM_{2,5}$ e concentrações de ozônio, ambos medidos em terrenos de 79 países. Verificaram que, em 2005, 89% da população mundial viveu em áreas onde a recomendação da OMS sobre a qualidade do ar foi excedida (média anual de $10 \mu\text{g}/\text{m}^3$ para $PM_{2,5}$). Globalmente, 32% da população viveu em áreas que excediam do objetivo provisório da OMS de $35 \mu\text{g}/\text{m}^3$ de $PM_{2,5}$ (média anual). Os maiores níveis sazonais de ozônio foram encontrados na América do Norte, América Latina, Europa e África. Entre 1990 e 2005, houve um aumento

de 6% de $PM_{2,5}$ e uma diminuição de 1% de ozônio, destacando-se aumentos das concentrações no Leste, Sul e Sudeste da Ásia e diminuições na América do Norte e na Europa.

No GBD 2015, Cohen *et al.* (2017) estimaram as concentrações médias de $PM_{2,5}$ e ozônio para o ano de 2015 com base em satélites e modelos de transporte químico. Utilizaram função C-R para cada causa de morte. Estimaram o RR para doenças cardiovasculares, cerebrovasculares, respiratórias e câncer de pulmão. O $PM_{2,5}$ foi um fator de risco para mortalidade. No total de mortes, a exposição ao $PM_{2,5}$ causou 4,2 milhões de mortes (IC 95% 3,7 a 4,8 milhões) e 103,1 milhões de anos de vidas perdidos ajustado por incapacidade (“*Disability-Adjusted Life-Years*” - DALY) (IC 95% 90,8 a 115,1 milhões), representou 7,6% do total de mortes e 4,2% de DALY. As mortes atribuídas ao $PM_{2,5}$ aumentaram cerca de 3,5 milhões (IC 95% 3 a 4 milhões), em 1990, para 4,2 milhões, em 2015. A exposição ao ozônio causou 254.000 mortes adicionais (IC95% 97.000 a 422.000) e uma perda de 4,1 milhões de DALY (IC95% 1,6 a 6,8 milhões) por doenças Pulmonar Obstrutiva Crônica em 2015.

Evans *et al.* (2013) utilizaram instrumentos de satélite para estimar níveis globais de exposição de $PM_{2,5}$ para 132 países, resumidos em sub-regiões da OMS. Os riscos relativos e as frações atribuíveis da mortalidade foram calculados utilizando funções de C-R. A fração atribuível da mortalidade adulta ao $PM_{2,5}$ foi de 8,0% (IC 95% 5,3-10,5) para doenças cardiovasculares, 12,8% (IC 95% 5,9-18,5) para câncer de pulmão e 9,4% (IC 95% 6,6 -11,8) para doença cardíaca isquêmica.

Brauer *et al.* (2016) combinaram estimativas baseadas em satélites, simulações de modelos de transporte químico e medições terrestres de 79 países para produzir estimativas de médias anuais de $PM_{2,5}$ e concentrações de ozônio em intervalos de cinco anos a partir de 1990 a 2010 e o ano 2013. Essas estimativas foram aplicadas para avaliar as concentrações médias para 188 países durante 1990 a 2013. Em 2013, 87% da população mundial viveu em áreas que excederam a $10 \mu g/m^3$ de $PM_{2,5}$ (média anual). Entre 1990 e 2013, o $PM_{2,5}$ aumentou 20,4%, impulsionado pelas tendências no sul da Ásia, no Sudeste Asiático e na China. A diminuição das concentrações médias de $PM_{2,5}$ foi evidente na maioria dos países de alta renda. As concentrações médias de ozônio aumentaram globalmente em 8,9% no período de 1990 a 2013, com aumentos na maioria dos países - com exceção de reduções modestas na América do Norte, partes da Europa e em vários países do Sudeste Asiático.

O GBD 2010 e o relatório de 2016 da OMS sobre a avaliação da poluição do ar e da carga global de doenças para 2012 em cidades do mundo utilizou estimativas de $PM_{2,5}$ e PM_{10} através de satélite em combinação com simulações do TM5 para gerar estimativas globais de concentrações médias de partículas finas e ozônio (Lim *et al.*, 2012; OMS,2016)

2 JUSTIFICATIVAS

2.1 O uso de modelos de efeitos mistos para prever a poluição do ar nas regiões metropolitanas do Brasil

A inexistência de dados é um dos desafios enfrentados em estudos sobre os efeitos da poluição do ar nas grandes cidades. A disponibilidade limitada das informações sobre poluição do ar decorrentes de fatores como falta e defeitos de equipamentos ou erros de leitura, contribui significativamente para a ocorrência de lacunas em uma determinada área e tempo, inviabilizando, portanto, cálculos de medidas globais nos estudos de impacto. Dessa forma, problemas de dados faltantes tem sido extensivamente estudado e vários métodos para previsão de longo prazo das concentrações de poluentes como o MP foram explorados e publicados em análise espacial, porém poucos explorados nos estudos de séries temporais epidemiológicas.

2.2 Avaliação do impacto da poluição do ar na mortalidade em regiões metropolitanas do Brasil

O MP é um fator de risco para doenças respiratórias e cardiovasculares (Dockery e Pope, 1994; Schwartz, 1994; Lin *et al.*, 1999; Pope e Dockery, 1999; Dominici *et al.*, 2000; Samet *et al.*, 2000; Katsouyanni *et al.*, 2001; Dominici *et al.*, 2002; Dominici *et al.*, 2004; Wong *et al.*, 2008; Romieu *et al.*, 2012). E a mensuração do impacto deste fator de risco sobre a carga de doenças nas regiões metropolitanas podem auxiliar nas ações voltadas para a prevenção e controle de poluentes atmosféricos específicos e no debate sobre as novas diretrizes da qualidade do ar (Kunzli *et al.*, 2000).

3 OBJETIVOS

3.1 Objetivos gerais

Propor um modelo de efeitos mistos para previsão das concentrações de PM₁₀ nas regiões metropolitanas do Brasil utilizando-as na avaliação do impacto dos níveis de concentração da poluição do ar na saúde.

3.2 Objetivos específicos

Artigo 1: O uso de modelos de efeitos mistos para prever a poluição do ar nas regiões metropolitanas do Brasil.

Propor um modelo de efeitos mistos como uma metodologia para previsão das concentrações de PM₁₀ para todas as regiões metropolitanas do Brasil que não possuem redes de monitoramento da poluição do ar.

Artigo 2: Avaliação do impacto da poluição do ar na mortalidade em regiões metropolitanas do Brasil.

Estimar e analisar a carga de mortalidade por doenças cardiovasculares em adultos com idade ≥ 30 anos e por doenças respiratórias em crianças menores de um ano e de um a cinco anos atribuída ao MP, utilizando níveis de concentração de PM₁₀ derivados do modelo de efeitos mistos.

4 MÉTODOS

4.1 Modelo linear de efeitos mistos para dados longitudinais

A indisponibilidade de dados de poluição do ar é muito comum nos estudos ambientais, entretanto, esse problema pode ser contornado com a utilização de estimativas de médias anuais de MP com base no modelo linear de efeitos mistos, que incorpora medidas disponíveis de monitores existentes ao longo do tempo. O modelo prevê as concentrações de PM_{10} em áreas urbanas em função de variáveis econômicas, demográficas, geográficas e meteorológicas que afetam o transporte de poluição. Na próxima seção, dentro de um contexto teórico, o modelo linear de efeitos mistos para dados longitudinais será apresentado.

4.1.1 O modelo

Estudos com estruturas de dados em painéis se caracterizam pela sequência temporal de duas ou mais observações de uma mesma característica pertencente a cada indivíduo, bloco, estrato ou área. Aplicado a esse tipo de estrutura de dados nos estudos ecológicos para dados longitudinais, o método dos modelos de efeitos mistos incorpora dados temporais e seccionais. É uma das vantagens desses modelos é a flexibilidade com estrutura de dados desbalanceados.

Os modelos de efeitos mistos descrevem, num determinado período, a relação entre uma variável resposta e uma ou mais variáveis explicativas agrupadas em uma única estrutura hierárquica, em que os níveis mais agregados são blocos ou estratos. Em cada um desses estratos tem-se uma série temporal de observações que correspondem ao menor nível ou ao nível menos agregado. Esses modelos incluem termos de efeitos fixos nos quais procura-se explorar as diferenças entre os estratos (captadas pelo intercepto β_{0j}) condicionadas aos seus próprios fatores contextuais (socioeconômicos, meteorológicos, demográficos, etc.), e termos adicionais de efeitos aleatórios referentes a parte residual do modelo. Portanto, o modelo que será apresentado limita-se a dois níveis de efeitos aleatórios, a generalização para mais níveis é pensada da mesma forma.

Com base em Snijders e Bosker (1999), seja o modelo vazio de dois níveis, no qual o nível mais agregado, o nível 2, referenciado como termo (j) e o menos agregado, o nível 1, como termo (t).

(Modelo 1)

$$Y_{tj} = \beta_0 + u_{0j} + \varepsilon_{tj} \quad j = 1, 2, 3, \dots, M \text{ e } t = 1, 2, 3, \dots, n_j \quad (1)$$

$$\begin{aligned} u_{0j} &\sim N(0, \sigma_{u_0}^2) & \varepsilon_{tj} &\sim N(0, \sigma_{\varepsilon}^2) & \text{Cov}(u_{0j}, \varepsilon_{tj}) &= 0 \\ \text{Var}(Y_{tj}) &= \text{Var}(\beta_0 + u_{0j} + \varepsilon_{tj}) &= \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 & & & (2) \end{aligned}$$

Assim, a variância é particionada em componentes que correspondem cada nível da hierarquia: a variação entre locais $\sigma_{u_0}^2$ e a variação temporal dentro de um determinado grupo ou estrato, σ_{ε}^2 .

Ao condicionamento de uma variável explicativa ou contextual X_1 , o modelo é expandido da seguinte forma:

(Modelo 2)

$$Y_{tj} = \beta_0 + \beta_1 X_{1tj} + u_{0j} + u_{1j} X_{1tj} + \varepsilon_{tj} \quad (3)$$

Fazendo a variável $X_{1tj} = Z_{1tj}$ somente na parte aleatória do modelo temos:

$$Y_{tj} = \beta_0 + \beta_1 X_{1tj} + u_{0j} + u_{1j} Z_{1tj} + \varepsilon_{tj} \quad (4)$$

$$\begin{aligned} u_{0j} &\sim N(0, \sigma_{u_0}^2) & u_{1j} &\sim N(0, \sigma_{u_1}^2) & \text{Cov}(u_{0j}, u_{1j}) &= \sigma_{u_{01}} \\ \text{Cov}(u_{0j}, \varepsilon_{tj}) &= 0 & \text{Cov}(u_{1j}, \varepsilon_{tj}) &= 0 & \varepsilon_{tj} &\sim N(0, \sigma_{\varepsilon}^2) \end{aligned} \quad (5)$$

Em que,

Y_{tj} Variável dependente;

β_0, β_1 são os coeficientes dos efeitos fixos;

X_{1tj} é uma variável independente ou contextual dos efeitos fixos;

Z_{1tj} é uma variável dos efeitos aleatórios;

$u_{0j}, u_{1j}Z_{1tj}$ são termos dos efeitos aleatórios, semelhantes aos erros ε_{tj} . O termo $u_{1j}Z_{1tj}$ pode ser considerado como uma interação entre grupo e a variável Z. Isto implica que os grupos são caracterizados por dois efeitos aleatórios: aleatoriedades de seus interceptos e inclinações;

ε_{tj} é o termo do erro.

O problema da dependência entre as observações ocorre com muita frequência nas pesquisas, principalmente quando a amostra não é tomada aleatoriamente, mas apenas áreas geográficas são utilizadas nos estudos. As observações dentro de uma mesma área ou grupos são mais similares do que observações pertencentes à outras áreas. Dessa forma, em muitos conjuntos de dados agrupados não são verificados os pressupostos da independência entre observações. No caso dos dados longitudinais, as medições realizadas na mesma unidade experimental (isto é, indivíduo, grupo, bloco) tendem a estar relacionadas, ou seja, autocorrelacionadas. O modelo terá, portanto, variância e covariância residuais heterocedásticas e correlacionadas.

$$\begin{aligned} \text{Var}(Y_{tj}) &= \sigma_{u_0}^2 + 2\sigma_{u_{01}}Z_{1tj} + \sigma_{u_1}^2Z_{1tj}^2 + \sigma_{\varepsilon}^2 \\ \text{cov}(Y_{tj}, Y_{t'j}) &= \sigma_{u_0}^2 + (Z_{1tj} + Z_{1t'j})\sigma_{u_{01}} + Z_{1tj}Z_{1t'j}\sigma_{u_1}^2 \end{aligned} \quad (6)$$

Uma medida fundamental para corrigir o problema é a modelagem desta dependência intragrupo, que será apresentada na seção 4.1.7.3.

4.1.2 Inferência para os efeitos fixos

A estimação dos parâmetros da equação (4) se baseia no método da máxima verossimilhança - “*Maximum Likelihood*” (ML) - ou no método da máxima verossimilhança restrita - “*Restricted Maximum Likelihood*” (REML). Os estimadores são consistentes, assintoticamente normais, eficientes, sendo os melhores preditores lineares não tendenciosos dos efeitos fixos – “*Best Linear Unbiased Predictors*” (BLUP).

4.1.2.1 Método da Máxima Verossimilhança

As equações apresentadas abaixo referem-se ao modelo de regressão linear clássico, com um único nível de efeitos aleatórios, sendo válidos também para o modelo linear de efeitos mistos.

O modelo linear simples pode ser expresso como:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i \quad i = 1, 2, 3, \dots, n \quad (7)$$

Em que ε_i são independentes e identicamente distribuídos com média 0 e matriz de variância-covariância $\sigma^2 \mathbf{M}_i$, onde $\mathbf{M}_i = \mathbf{I}$, uma matriz identidade.

A abordagem clássica de inferência do modelo (7) baseia-se nos estimadores obtidos a partir da maximização da função de verossimilhança dada por:

$$p(y_i/\beta, \theta, \sigma^2) = (2\pi\sigma^2)^{-\frac{n_i}{2}} \exp\left(\frac{(Y_i - X_i\beta)^T \mathbf{M}_i^{-1} (Y_i - X_i\beta)}{-2\sigma^2}\right) |\mathbf{M}_i|^{-\frac{1}{2}} \quad (8)$$

Nos modelos de efeitos mistos sabemos que a variância apresenta duas fontes de variação, em que

$$Var(Y_{it}) = \sigma_{u_0}^2 + 2\sigma_{u_{01}} Z_{1i} + \sigma_{u_1}^2 Z_{1i}^2 + \sigma_\varepsilon^2 = \mathbf{Z}^T \mathbf{DZ} + \sigma^2 \mathbf{I} \quad (9)$$

Representam, respectivamente, a variação intragrupo e entre grupos.

\mathbf{D} são as componentes da variância. A matriz de variância-covariância é dada por $\sigma^2 \mathbf{M}_i$

em que $\sigma^2 \mathbf{M}_i = \mathbf{Z}^T \mathbf{DZ} + \sigma^2 \mathbf{I} \Rightarrow \mathbf{M}_i = \frac{\mathbf{Z}^T \mathbf{DZ}}{\sigma^2} + \mathbf{I}$.

Dado θ , os estimadores por ML de β e σ^2 são representados por $\hat{\beta}_{ML}$ e $\hat{\sigma}^2_{ML}$, onde

$$\hat{\beta}_{ML}(\theta) = \left(\sum_{i=1}^N X_i^T \mathbf{M}^{-1} X_i \right)^{-1} \left(\sum_{i=1}^N X_i^T \mathbf{M}^{-1} y_i \right) = (\mathbf{X}^T \mathbf{M}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{M}^{-1} \mathbf{Y} \quad (10)$$

$$\hat{\sigma}^2_{ML}(\theta) = \frac{\sum_{i=1}^N (y_i - X_i \hat{\beta}(\theta))^T \mathbf{M}_i^{-1} (y_i - X_i \hat{\beta}(\theta))}{N}$$

De acordo com Goldstein (2011), supondo que os valores das variâncias sejam conhecidos e pudessem construir uma matriz diagonal, o procedimento usual de estimação de Mínimos Quadrados Generalizados (MQG) poderia ser aplicado para obter estimativas dos coeficientes fixos. É um procedimento iterativo que geralmente começa a partir de estimativas "razoáveis" dos parâmetros fixos, inicialmente ajustado por "*Ordinary Least Squares*" OLS (assumindo $\sigma_{u_0}^2 = 0$). A estimativa envolve aplicação de MQG usando a matriz de covariância já estimada pelo método OLS, assumindo a normalidade. Com as estimativas obtidas da primeira iteração, novamente o processo é feito para obter novas estimativas de efeitos fixos. Esse "*looping*" acontece até que o procedimento convirja, isto é, até que as estimativas para todos os parâmetros não mudem mais de um ciclo para o outro.

4.1.2.2 Método da Máxima Verossimilhança Restrita

As estimativas dos parâmetros por ML tendem a subestimar os componentes da variância-covariância, tais como σ^2 , por não levar em conta a variação dos parâmetros fixos, ou seja, não variam perante às alterações de uma variável explicativa e, portanto, ditos invariantes (Pinheiro e Bates, 2000). Em outras palavras, a perda de graus de liberdade proveniente da estimação dos efeitos fixos não é levada em consideração, gerando estimadores viesados. Contudo, o problema pode ser corrigido pelo método REML.

Utilizando o logaritmo natural na equação (8), dá a máxima verossimilhança como:

$$\log L_{ML}(y_i/\beta, \theta, \sigma^2) = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2} \sum_{i=1}^N \log |\mathbf{M}_i^{-1}| - \sum_{i=1}^N \frac{(Y_i - X_i\beta)^T \mathbf{M}_i^{-1} (Y_i - X_i\beta)}{2\sigma^2} \quad (11)$$

O estimador $\hat{\sigma}^2_{REML}(\hat{\theta})$ é dado por:

$$\hat{\sigma}^2_{REML}(\theta) = \frac{\sum_{i=1}^M (y_i - X_i\hat{\beta}(\theta))^T \mathbf{M}_i^{-1} (y_i - X_i\hat{\beta}(\theta))}{N - P} \quad (12)$$

A função da REML é similar à função ML. Para tanto, apenas o termo $\log |\sum_{i=1}^N X_i^T \mathbf{M}_i^{-1} X_i|$ é acrescido na função, conforme é demonstrado abaixo:

$$\log L_{REML}(y_i/\theta) = \log L_{ML}(y_i/\beta, \theta, \sigma^2) - \frac{1}{2} \log \left| \sum_{i=1}^N X_i^T \mathbf{M}_i^{-1} X_i \right|$$

Com relação ao termo $\frac{1}{2} \log |\sum_{i=1}^N X_i^T \mathbf{M}_i^{-1} X_i|$, conclui-se que quaisquer alterações na matriz \mathbf{X} resulta em uma mudança no $\log L_{REML}$. Como consequência, diferentes estruturas dos efeitos fixos que se ajustam utilizando REML não podem ser comparadas com base em suas probabilidades restritas. Em particular, os testes da razão de verossimilhança (TRV) não são válidos sob essas circunstâncias, porém, a estrutura dos efeitos aleatórios do modelo pode ser comparada através desse método (Pinheiro e Bates, 2000).

Vale ressaltar que, a otimização das estimativas, tanto pelo método ML quanto pelo método REML de um modelo linear de efeitos mistos, é geralmente realizada através de iterações EM ou através de iterações de Newton-Raphson.

4.1.3 Teste da Razão de Verossimilhanças

O TRV pode ser usado para modelos ajustados por REML somente se ambos os modelos forem ajustados pelo mesmo método de estimação e se a especificação de efeitos fixos for a mesma (Pinheiro e Bates, 2000). O teste qui-quadrado (χ^2) tende a ser "anticonservador". À medida que os parâmetros dos efeitos fixos são removidos do modelo, aumenta a imprecisão

dos p-valores, isso serve também para os critérios de comparação de modelos de regressão com estruturas não aninhadas, como “*Akaike Information Criterion*” (AIC) e o “*Bayesiana Information Criterion*” (BIC). Por este motivo, alguns autores recomendam a utilização dos testes t e testes F condicionais, para avaliar a significância dos efeitos fixos baseados no método de estimação REML (Pinheiro e Bates, 2000).

A estatística de TRV é dada por:

$$2 \log(L_2/L_1) = 2[\log(L_2) - \log(L_1)] \quad (13)$$

Em que

L_1 é a verossimilhança do modelo restrito e L_2 do modelo completo.

A distribuição assintótica da estatística TRV, sob a hipótese nula de que o modelo restrito é adequado, é uma χ^2 com graus $k_2 - k_1$ de liberdade. E o AIC e o BIC expressos como:

$$\begin{aligned} AIC &= -2l(\hat{\theta}/y) + 2n_{par} \\ BIC &= -2l(\hat{\theta}/y) + 2n_{par} \log(N) \end{aligned} \quad (14)$$

4.1.4 Testes T e F condicionais

Os testes condicionais para a significância de um termo dos efeitos fixos são dados por testes F ou testes t, baseados no método de estimação REML para a variância, $\hat{\sigma}_{REML}^2(\theta)$

Os testes t condicionais avaliam a significância marginal de cada coeficiente dos efeitos fixos, dado que os outros efeitos fixos estão presentes no modelo (Pinheiro e Bates, 2000).

Para testar:

$$H_0: \beta = 0 \quad vs. \quad H_1: \beta \neq 0$$

A estatística do teste t é dada por:

$$T = \frac{\hat{\beta} - \beta}{\hat{\sigma}_{REML} \sqrt{\left[\left(\sum_{i=1}^N X_i^T \mathbf{M}_i^{-1} X_i \right)^{-1} \right]}} \quad (15)$$

Sob a hipótese nula, segue uma distribuição aproximada de t de Student com t-1 graus de liberdade.

Os testes F condicionais testam a significância de termos dos efeitos fixos no modelo, podendo incluir vários coeficientes. Os termos são testados sequencialmente pela ordem com que entram ou saem do modelo. A hipótese nula do teste F é:

$$H_0: L\beta = 0 \quad vs. \quad H_1: L\beta \neq 0$$

Em que L é uma matriz conhecida de dimensão (q× p) com característica q e 0 é o vetor nulo de dimensão q. A estatística de teste do teste F é dada por:

$$F = \frac{\hat{\beta}^T L^T \left[L \left(\sum_{i=1}^N X_i^T \mathbf{M}_i^{-1} X_i \right)^{-1} L^T \right]^{-1} L\beta}{rank(L)} \quad (16)$$

Sob a hipótese nula, a estatística segue uma distribuição aproximada de F de Snedecor com (q, u) graus de liberdade. O número de graus de liberdade do numerador do teste F é dado por rank (L) (Wolfinger, 1993; Verbeke, 1997).

4.1.5 Intervalos de confiança

Os intervalos de confiança aproximados para os efeitos aleatórios e fixos são obtidos utilizando as distribuições aproximadas para as estimativas de ML e REML e os testes t condicionais (Pinheiro e Bates, 2000). Seja (n-1) denominado como graus de liberdade para o

teste t condicional correspondente aos efeitos fixos com base em um $\hat{\beta}$, um intervalo de confiança aproximado de nível $1 - \alpha$ para β é:

$$\hat{\beta} \pm t_{(\tau-1; 1-\frac{\alpha}{2})} \sqrt{\hat{\sigma}^2_{REML} \left[\left(\sum_{i=1}^M X_i^T \mathbf{M}_i^{-1} X_i \right)^{-1} \right]} \quad (17)$$

Onde $t_{(\tau-1; 1-\frac{\alpha}{2})}$ representa o quantil $(1 - \frac{\alpha}{2})$ da distribuição t de *Student* com τ graus de liberdade (Pinheiro e Bates, 2000)

4.1.6 Modelagem das estruturas residuais das matrizes de variância-covariância

Conforme Pinheiro e Bates (2000), muitas vezes há interesse em reduzir as matrizes de variância-covariância \mathbf{D} dos efeitos aleatórios e considerar uma estrutura com menos parâmetros. Por exemplo, podemos assumir que os efeitos aleatórios sejam independentes, isto é, que a correlação entre os efeitos aleatórios seja suficientemente pequena e não significativa. Nesse caso, a estrutura da matriz mais indicada seria uma diagonal:

$$\mathbf{D} = \begin{vmatrix} \sigma_{11}^2 & 0 & 0 \\ 0 & \sigma_{22}^2 & 0 \\ 0 & 0 & \sigma_{33}^2 \end{vmatrix}$$

Se os efeitos aleatórios são independentes e têm variâncias iguais, ou seja, a mesma variação, a matriz correspondente seria uma identidade ou um escalar:

$$\mathbf{D} = \begin{vmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{vmatrix}, \text{ onde } a > 0$$

A matriz definida positiva não estruturada é uma estrutura que caracteriza os modelos de efeitos mistos, pressupõe que todos os parâmetros de intercepto e inclinações aleatórios sejam correlacionados dentro de cada subgrupo. A classe *default* da matriz positiva definida

para os efeitos aleatórios pode ser encontrada na biblioteca nlme, um pacote estatístico implementado no software R contendo funções para análise dos modelos de efeitos mistos. Neste pacote existe uma função chamada pdSymm que representa essa matriz e é apresentada como:

$$\mathbf{D} = \begin{vmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22}^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2 \end{vmatrix}$$

Outras classes de estruturas podem também ser consideradas no pacote nlme: como “*compound symmetry*”

$$\mathbf{D} = \begin{vmatrix} \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 & \sigma_{u_0}^2 & \sigma_{u_0}^2 \\ \sigma_{u_0}^2 & \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 & \sigma_{u_0}^2 \\ \sigma_{u_0}^2 & \sigma_{u_0}^2 & \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 \end{vmatrix}$$

E “*Block-diagonal*”

$$\mathbf{D} = \begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{vmatrix}, \text{ onde}$$

$$\mathbf{A} = \begin{vmatrix} \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 & \sigma_{u_0}^2 & \sigma_{u_0}^2 \\ \sigma_{u_0}^2 & \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 & \sigma_{u_0}^2 \\ \sigma_{u_0}^2 & \sigma_{u_0}^2 & \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 \end{vmatrix}$$

$$\mathbf{B} = \begin{vmatrix} \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 & \sigma_{u_0}^2 & \sigma_{u_0}^2 \\ \sigma_{u_0}^2 & \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 & \sigma_{u_0}^2 \\ \sigma_{u_0}^2 & \sigma_{u_0}^2 & \sigma_{u_0}^2 + \sigma_{\varepsilon}^2 \end{vmatrix}$$

Assim, a biblioteca nlme fornece várias classes de matrizes positivas definidas. Através dessas classes (demonstradas no quadro 1) podemos definir estruturas para a matriz \mathbf{D} e obter vários modelos com diferentes estruturas para a matriz de covariância dos efeitos aleatórios (Pinheiro e Bates, 2000)

Quadro 1 - Classes das matrizes de variâncias-covariâncias dos efeitos aleatórios da biblioteca nlme

| pdMat | Classes |
|------------|--------------------------------|
| pdBlocked | Bloco-diagonal |
| pdCompSymm | Estrutura de simetria composta |
| pdDiag | Diagonal |
| pdIdent | Múltiplo de uma identidade |
| pdSymm | Matriz positiva-definida geral |

Pinheiro e Bates (2000). “Padrões das funções variância”. *Mixed-effects models in S and S-PLUS*, 208. New York: Springer.

4.1.7 Adequação do Modelo

A adequação do modelo é verificada, geralmente, através da análise dos gráficos de resíduos e nos valores ajustados do modelo. Os testes de hipótese são alternativas possíveis para a verificação da adequação, mas, uma vez que costumam corroborar com os resultados da análise gráfica, esta última é frequentemente preferida. Num modelo linear de efeitos mistos a não verificação da hipótese de normalidade, homocedasticidade e/ou da independência serial entre as observações do mesmo grupo, são frequentes violações que afetam sua eficiência e precisam ser checadas.

4.1.7.1 Violação da hipótese de normalidade

Os modelos lineares de efeitos mistos dependem de resultados assintóticos assumindo um modelo gaussiano que pode não estar apropriado em algumas situações, produzindo, portanto, desvios-padrão e intervalos de confiança não confiáveis. Quando a adequação do uso de resultados assintóticos é questionada, procedimento como o bootstrap, por exemplo, pode ser usado para estimar essas quantidades. Na configuração dos modelos mistos, o bootstrap paramétrico é uma alternativa que tem importantes referências, como Efron e Tibshirani (1994) e Davison e Hinkley (1997). O procedimento está inserido no pacote lme4, que é um pacote estatístico mais recente em relação ao nlme para análise dos modelos de efeitos mistos.

4.1.7.2 Modelagem da heterocedasticidade

Como foi descrito na seção 4.1.1, o modelo linear de efeitos mistos impõe condições à estrutura de erros aleatórios, que não se verificam em muitos conjuntos de dados. Nos dados longitudinais, em particular, as medições ou observações realizadas sobre o mesmo grupo (ou indivíduos) estão geralmente correlacionadas, comprometendo as inferências no modelo. Contudo, é possível relaxar estas condições adotando o procedimento de modelagem da estrutura de variâncias-covariâncias dos erros aleatórios.

Seguindo Pinheiro e Bates (2000), o modelo linear de efeitos mistos permite a utilização de diferentes funções-variância para modelar a estrutura de variância dos erros. Estas funções têm como objetivo acomodar a heterocedasticidade dos erros aleatórios, onde se verifica, por exemplo, diferentes variâncias em cada grupo ou em que a variância cresce ou decresce com uma determinada variável (como o tempo, por exemplo). Estas funções podem ser aplicadas no modelo de efeitos mistos e estão disponíveis na biblioteca nlme do R.

Na tabela 1 estão apresentados os padrões das funções-variância disponíveis no R.

Tabela 1 - Padrões das funções variância

| | Estruturas | $\text{Var}(\epsilon_{tj})$ |
|---------------|-----------------------------------|--|
| varFixed | Variância fixa | $\sigma^2 x_{tj}$ |
| varIdent | diferentes variâncias por estrato | $\sigma^2 \delta_{stj}^2$ |
| varPower | Potência da variável | $\sigma^2 v_{tj} ^{2\delta}$ |
| varExp | exponencial da variável | $\sigma^2 \exp(2\delta v_{tj})$ |
| varConstPower | constante + potência da variável | $\sigma^2 (\delta_1 + v_{tj} ^{\delta_2})^2$ |
| varComb | combinação de funções-variância | $\sigma^2 g_1^2(s_{tj}, \delta_1) g_2^2(v_{tj}, \delta_2)$ |

Pinheiro e Bates (2000). “Padrões das funções variância”. *Mixed-effects models in S and S-PLUS*, 208. New York: Springer.

A comparação entre as estruturas das variâncias nos modelos não aninhados é feita através do AIC.

4.1.7.3 Modelagem da dependência serial

A tendência de uma série temporal pode ser modelada com o uso de funções polinomiais ou *splines*, que são alternativas para “suavizar” os dados de uma série temporal para explicar flutuações na variável de interesse no modelo, enquanto que os modelos ARIMA (*Autoregressive Integrated Moving Average*) servem para modelar dependência entre as observações ou resíduos no modelo. O Ajuste da autocorrelação em algumas circunstâncias apenas tem o intuito de “acomodar” a dependência serial no contexto maior de realizar inferências sobre a estimação dos parâmetros do modelo, e assim poder realizar as predições com mais eficiência.

4.1.7.3.1 Funções Polinomiais

A utilização de funções polinomiais é uma estratégia para modelar a tendência de uma série temporal. E são definidas por:

$$F_j(t) = \beta_{0j} + \beta_{1j}(t - t_0) + \beta_{2j}(t - t_0)^2 + \dots + \beta_{rj}(t - t_0)^r \quad (18)$$

Onde r é o grau do polinômio (Snijders e Bosker, 1999).

As dificuldades numéricas são frequentemente menores na estimação do modelo quando t_0 tem o valor centrado na média do que quando tem um valor fora do intervalo da série ou em um dos extremos desta faixa. Sendo assim, quando ocorrem problemas de convergência é aconselhável tentar trabalhar com valores de t_0 próximos à média ou à mediana de t (Snijders e Bosker, 1999).

4.1.7.3.2 Splines

As outras classes de funções também adequadas para modelar séries temporais dentro de uma estrutura hierárquica são as *Splines*, que são funções de suavizações polinomiais por

partes. Mais precisamente, a série temporal é dividida em intervalos definidos por “nós”. Em cada parte da série é ajustado um polinômio de grau r que posteriormente serão unidos de forma suaves pelos “nós”, gerando uma função contínua suavizada (Snijders e Bosker, 1999).

Ainda de acordo com Snijders e Bosker (1999), uma *Spline* quadrática com um nó é definida como:

$$\begin{aligned}
 f_1(t) &= t - t_1 \\
 f_2(t) &= \begin{cases} (t - t_1)^2 & (t \leq t_1) \\ 0 & (t \geq t_1) \end{cases} \\
 f_3(t) &= \begin{cases} 0 & (t \leq t_1) \\ (t - t_1)^2 & (t \geq t_1) \end{cases}
 \end{aligned} \tag{19}$$

Onde t_1 é o nó da *Spline*, que é definido pelo ponto médio da série.

A função é modelada como:

$$F_j(t) = \beta_{0j} + \beta_{1j}f_1(t) + \beta_{2j}f_2(t) + \beta_{3j}f_3(t) \tag{20}$$

E a *Spline* cúbica com um nó é determinada por:

$$\begin{aligned}
 f_1(t) &= t - t_1 \\
 f_2(t) &= (t - t_1)^2 \\
 f_3(t) &= \begin{cases} (t - t_1)^3 & (t \leq t_1) \\ 0 & (t \geq t_1) \end{cases} \\
 f_4(t) &= \begin{cases} 0 & (t \leq t_1) \\ (t - t_1)^3 & (t \geq t_1) \end{cases}
 \end{aligned} \tag{21}$$

Para decidir qual o melhor ajuste para tendência é utilizado o AIC

4.1.7.3.3 Modelos autoregressivo, integrados e de médias móveis

Estruturas de correlação são usadas para modelar dependência entre observações. No contexto do modelo linear de efeitos mistos nos é permitido corrigir a dependência temporal intragrupo, isto é, modelar a dependência entre os erros aleatórios ε_{tj} . Dentre as estruturas de correlações seriais mais comuns implementadas na biblioteca nlme, o modelo AR(1) (Autoregressivo de 1º ordem) é o mais simples e um dos mais utilizados na prática (Pinheiro e Bates, 2000). O processo definido por Box *et al.* (2015), representado pela equação (22), fornece um modelo poderoso para as séries temporais não estacionárias com o intuito de corrigir ou de acomodar a dependência temporal. O processo é definido por:

$$w_{tj} = \phi_1 w_{tj-1} + \dots + \phi_p w_{tj-p} + a_{tj} - \theta_1 a_{tj-1} - \dots - \theta_q a_{tj-q} \quad (22)$$

Onde $w_t = \nabla_{Z_t}^d$, é o operador de diferença de ordem d. A equação (22) inclui um misto de modelo estacionário puro autoregressivo com parâmetro ϕ , como também um modelo de médias móveis com parâmetro θ . Como os parâmetros são estimados a partir dos dados, assumir valores para p ou q superiores a 2 ou 3, geralmente origina em problemas de convergência dos algoritmos.

Além desta família de estruturas de correlação, nos modelos de efeitos mistos há outras estruturas de correlação também implementadas na biblioteca nlme, tais como:

4.1.7.3.4 Geral

Essa estrutura representa o outro extremo em complexidade para o composto estrutura de simetria. Cada correlação nos dados é representada por um diferente parâmetro, correspondendo à função de correlação

$$h(k, \rho) = \rho_k, \quad k = 1, 2, \dots, \quad (23)$$

4.1.7.3.5 Simetria Composta

Essa é a estrutura de correlação serial mais simples, que assume igual correlação entre todos os erros dentro do grupo pertencentes ao mesmo grupo. O modelo de correlação correspondente é:

$$\text{Cor}(\varepsilon_{tj}; \varepsilon_{tj'}) = \rho, \forall j \neq j' \text{ com } h(k, \rho) = \rho, \quad k = 1, 2, \dots, \quad (24)$$

Em que o único parâmetro de correlação ρ é geralmente referido como o coeficiente de correlação intraclasse.

4.1.8 Validações - Performance do Modelo

4.1.8.1 Coeficiente de determinação

A avaliação do ajuste total do modelo de regressão é feita através do R^2 , que tem uma propriedade extremamente útil de fornecer informação sobre o ajuste do modelo. É uma estatística sumária que descreve a quantidade da variância explicada definida como:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} = \frac{\text{var}(\hat{y}_i)}{\text{var}(y_i)} \quad (25)$$

Com base também na variância explicada, proposto por (Nakagawa e Schielzeth, 2013), um método simples para obter o R^2 para os modelos lineares de efeitos mistos é ilustrada da seguinte forma:

$$R_m^2 = \frac{\sigma_f^2}{\sigma_f^2 + \sum_{l=1}^u \sigma_l^2 + \sigma_\varepsilon^2} \quad (26)$$

$$\sigma_f^2 = \text{var} \left(\sum_{h=1}^p \beta_h x_{htj} \right)$$

Onde o R_m^2 é o R^2 marginal que indica a variância explicada pelos fatores fixos do modelo; σ_f^2 ³ é o componente da variância dos efeitos fixos; σ_l^2 é o componente da variância do l -ésimo fator aleatório e σ_ε^2 é o componente da variância dos resíduos.

O R^2 condicional, R_c^2 , é a variância explicada pelos fatores fixos e aleatórios do modelo e interpretada, portanto, como a variância explicada por todo o modelo. É expressa da seguinte forma:

$$R_c^2 = \frac{\sigma_f^2 + \sum_{l=1}^u \sigma_l^2}{\sigma_f^2 + \sum_{l=1}^u \sigma_l^2 + \sigma_\varepsilon^2} \quad (27)$$

Tanto o R^2 marginal como o condicional transmitem informações únicas e interessantes, e recomenda-se que ambos sejam apresentados no modelo. O comando para o cálculo de ambos R^2 está disponível na biblioteca *r2glmm* do R e pode ser aplicado no modelo de efeitos mistos utilizando o pacote *nlme*, ou via formula do R_m^2 e R_c^2 .

4.1.8.2 Estatísticas de performance e “Data Splitting”

A avaliação da performance do modelo é feita através da comparação dos dados observados com os valores ajustados. Essa avaliação pode ser feita utilizando-se algumas medidas estatísticas como: RMSE (*Root Mean Square Error*), MAE (*Mean Absolute Error*) e o coeficiente de correlação (r).

Em um modelo de regressão, os valores preditos são avaliados em termos de desempenho e precisão, como é o caso do erro médio quadrático $RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}}$, que é uma medida de avaliação da performance do procedimento de predição em termos de

³ σ_f^2 pode em princípio ser calculado através dos valores ajustados da parte fixa do modelo. Equivalente a multiplicar a matriz de variáveis com o vetor das estimativas dos efeitos fixos

acurácia. Outra medida é o erro médio absoluto, $MAE = \frac{\sum_{i=1}^N |\hat{y}_i - y_i|}{N}$, que é mais sensível e menos influenciado por maiores diferenças entre os valores preditos e observados. O Modelo com valor pequeno do RMSE e do MAE é considerado um modelo com boa performance para previsão. Assim como o coeficiente de correlação de Pearson (r), que é um indicador para avaliação de desempenho do modelo e é dado por:

$$r = \frac{cov(y, \hat{y})}{\sqrt{var(y)var(\hat{y})}}$$

O modelo com o coeficiente de correlação de Pearson alto é considerado um modelo com bom desempenho para previsão.

Dentre outros tipos de procedimentos para validação do modelo, existe o procedimento “*Data Splitting*” ou “*cross validation*”, que consiste em deixar de lado alguns dados originais, os utilizando depois para investigar o desempenho preditivo do modelo.

4.1.9 Valores ajustados e previsões

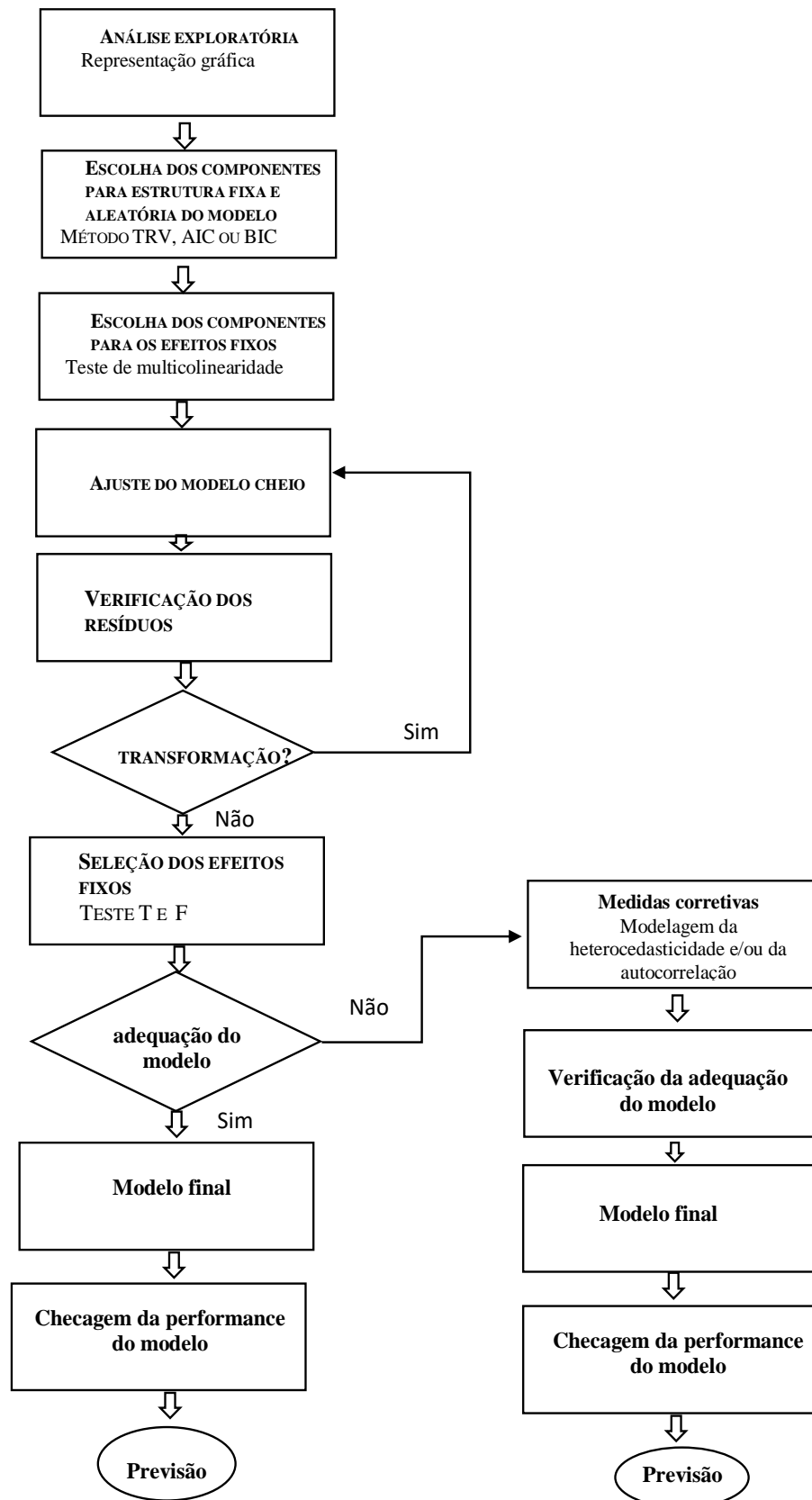
Nos modelos de efeitos mistos os valores ajustados e as previsões para novas observações podem ser obtidos em diferentes níveis: no nível mais agregado (segundo nível) ou no nível menos agregado (primeiro nível). As previsões no primeiro nível estimam o valor marginal esperado da variável resposta (Pinheiro e Bates, 2000). Sendo dado um vetor de variáveis dos efeitos fixos x_h , o valor marginal esperado da variável resposta é representado da seguinte forma:

$$\hat{y}_h = x_h^T \hat{\beta} \tag{28}$$

Já os valores previstos para novas observações estimam a expectativa condicional da variável resposta, dado os efeitos aleatórios do segundo nível. Por exemplo, $Z_h(j)$ denota um vetor de variáveis pertencente ao segundo nível, correspondente aos efeitos aleatórios associados ao j_{th} grupo, a estimativa de previsão do nível 2 é, portanto:

$$\hat{y}_h = x_h^T \hat{\beta} + Z_h(j)^T \hat{b}_j \tag{29}$$

Figura 1 - Esquema de análise do modelo linear de efeitos mistos para previsão



Fonte: Elaboração própria

4.2 Análise Estatística

4.2.1 Fontes de dados

Os dados foram coletados a partir de fontes publicamente disponíveis em sites com o objetivo de construir variáveis relacionadas às variações dos níveis de concentrações de PM₁₀ nas regiões metropolitanas. As variáveis coletadas foram:

Dados sobre consumo de energia: Consumo anual de gasolina, diesel e etanol (em litros). Estes dados estão disponíveis no site da Agência Nacional do Petróleo (ANP). A inclusão de cada tipo de combustível refere-se às variações nos fatores de emissão como atividade econômica, intensidade do uso de combustível e emissões relacionadas ao tráfego em diferentes regiões metropolitanas.

Dados atmosféricos: Foram coletadas informações referentes a 6 variáveis meteorológicas representando as condições climáticas locais: Média anual de nebulosidade (decimais); média anual da pressão barométrica (hPa); média anual da temperatura (Celsius); média anual de umidade relativa (%); média anual da velocidade do vento (m/s) e índice pluviométrico total anual (mm). Estes dados estão disponíveis no site do Instituto Nacional de Meteorologia (INMET).

Dados geográficos: Foram coletadas informações de duas variáveis topográficas: área (km²) e altitude média (em metros), disponíveis no site do Instituto Brasileiro de Geografia e Estatística (IBGE).

Dados demográficos: Estão representados por população e densidade demográfica. Estas variáveis são indicadores de intensidade do problema de poluição em cada região metropolitana. Os dados estão disponíveis no site do IBGE.

Dados socioeconômicos: Produto Interno Bruto (PIB) total, PIB do setor agropecuário, do setor indústria e do setor serviços (em R \$ 1.000), disponíveis no site do IBGE.

Dados sobre transporte: Número de decolagens (aviões): representa as emissões de poluentes que são geradas, de forma predominante, pelas fases em que os motores das aeronaves se encontram em modo de operação de taxiamento na pista durante o ciclo de pouso e decolagem; e taxa de crescimento anual de veículos novos, representada pela fórmula:

$$\frac{F_t - F_{t-1}}{F_{t-1}} \times 100 \quad (30)$$

Onde F_t é o número de frotas de veículos do ano atual e F_{t-1} o número de frotas de veículos do ano anterior. Respectivamente, os dados sobre o número de decolagens e número de veículos estão disponíveis nos sites da Agência Nacional de Aviação Civil – ANAC e do Departamento Nacional de trânsito (DENATRAN).

Dados sobre queimadas: O número de focos de queimadas (total anual de focos) foi extraído no site do Instituto Nacional de Pesquisas Espaciais (INPE). A extração das informações do total de focos em cada ano foi baseada no satélite de referência.

Dados de monitoramento de PM₁₀: Foram calculadas estimativas de médias anuais de PM₁₀ para cada região metropolitana. Os dados horários, diários e mensais de PM₁₀ de monitores automáticos e semiautomáticos, localizados em áreas urbanas, foram obtidos pelos órgãos ambientais competentes (APÊNDICE A). Para a construção das estimativas de médias anuais, foi seguido um protocolo padrão dos órgãos ambientais sobre os dados de monitoramento que estão apresentados no APÊNDICE B.

É comum, mesmo as grandes redes de monitoramentos da qualidade do ar (com controle de qualidade adequado), apresentarem dados faltantes. Um procedimento de análise muito comum diante de “*missing data*” consiste na exclusão das unidades de observação com valores não aferidos de uma ou mais variáveis, também conhecida como análise de dados completos. Mas esse procedimento pode produzir estimativas enviesadas das médias. Uma maneira mais eficiente encontrada para casos de dados faltantes nas estações de monitoramento foi utilizar o procedimento de imputação de dados faltantes em séries temporais multivariado com base no algoritmo EM (*Expectation – Maximization*)(Dempster *et al.*, 1977). Este método propõe estimativas iterativas para a estimação, por máxima verossimilhança, do vetor de médias e da matriz de covariância de uma distribuição normal multivariada com dados faltantes. Assim, o Algoritmo EM é aplicado sob os pressupostos da distribuição normal multivariada. Além das

estruturas de dependência entre as variáveis, o método leva em consideração estruturas de dependência temporal de cada variável. Maiores detalhes são apresentados em Dempster *et al.* (1977); Junger e Ponce De Leon (2015). O método está implementado no R (*R Development Core Team, 2013*) na biblioteca *mtsvdi*. Após o procedimento de imputação, as médias anuais de PM₁₀ foram calculadas em cada região metropolitana.

Para a avaliação da carga de mortalidade nas regiões metropolitanas, os seguintes dados foram coletados:

Dados de mortalidade: A exposição à poluição atmosférica está associada aos efeitos agudos e crônicos na saúde. Os efeitos de saúde associados ao MP incluem câncer de pulmão, doenças respiratórias e cardiovasculares. Apenas dois desses desfechos de saúde foram incluídos neste estudo: doenças respiratórias e doenças cardiovasculares. Foram obtidos dados de mortalidade específica: Cardiovascular (I00–I99) em adultos com idade ≥ 30 anos e respiratória (J00–J99) em crianças menores de um ano e de um a cinco anos, todos compreendidos nos Capítulos IX e X da Décima Revisão da Classificação Internacional de Doenças e Problemas Relacionados à Saúde (CID-10). Todos os dados foram obtidos pelo Sistema de Informações de Mortalidade (SIM) (<http://datasus.saude.gov.br/informacoes-de-saude/tabnet/estatisticas-vitais>).

4.2.2 Estimação das médias anuais de PM₁₀ nas regiões metropolitanas do Brasil

Inicialmente, foi introduzida neste trabalho uma análise exploratória das médias anuais de PM₁₀ e dos fatores contextuais das regiões metropolitanas do Brasil durante o período de 2001 a 2014.

Após a análise exploratória, no intuito de se fazer inferências sobre o modelo de efeitos mistos, a aplicação do procedimento de imputação multivariado foi realizada nas estações de monitoramento com dados faltantes como forma de redução do viés das médias anuais de PM₁₀ nas regiões metropolitanas, assim como nas variáveis explicativas para evitar reduções na amostra. As imputações foram realizadas na proporção limite de 40% de dados faltantes, já que as estações e as variáveis explicativas encontravam-se em condições extremas de dados faltantes. Se fossemos mais rígidos considerando proporções menores como ponto de corte, teríamos muitas perdas de variáveis explicativas e estações de monitoramento o que decorreria em perdas de regiões metropolitanas para o estudo. As estações de monitoramento que

ultrapassaram esse valor limite foram desconsideradas por produzirem correlações espúrias, devido a poucas observações e estimativas irreais das médias anuais de PM₁₀. Em relação às estações de monitoramento que fizeram parte do estudo, as correlações de Spearman foram calculadas e resumidas para cada região metropolitana. O algoritmo EM foi empregado e a trajetória temporal das médias de PM₁₀ foi modelada utilizando splines.

Após o procedimento de imputação, o modelo linear de efeitos mistos foi ajustado para prever médias anuais de PM₁₀ nas regiões metropolitanas brasileiras, condicionadas às características econômicas, demográficas, meteorológicas e de fontes emissoras de poluição. Com base no modelo de Cohen *et al.* (2004), os principais determinantes das variações observadas nas concentrações de PM₁₀ em cada região metropolitana foram estruturados no modelo de efeitos mistos dado por:

$$C_{tj} = \beta_0 + \sum_{f=1}^F \beta_{Ef} E_{tj} + \sum_{g=1}^G \beta_{Mg} M_{tj} + \beta_{DD} DD_{tj} + \sum_{v=1}^V \beta_{Rv} R_{tj} + \sum_{s=1}^S \beta_{Ts} T_{tj} + \beta_{Tempo} Tempo_{tj} + \varepsilon_{tj} + u_{0j} \quad (31)$$

Onde

C_{tj} = log das médias anuais de PM₁₀ na região metropolitana j, no tempo t.

E_{tj} = log das variáveis de consumo de energia para região metropolitana j, no tempo t (f=1, f=2, ..., F)

M_{tj} = log das variáveis climáticas e geográficas para região metropolitana j, no tempo t (m=1, m=2, ..., M)

DD_{tj} = log da densidade demográfica da região metropolitana j, no tempo t

R_{tj} = log das variáveis socioeconômicas da região metropolitana j, no tempo t (v=1, v=2, ..., V)

T_{tj} = log das variáveis de transporte da região metropolitana j, no tempo t (s=1, s=2, ..., S)

$Tempo_{tj}$ = termo de tendência do tempo (2001=1, 2002=2, ..., 2014=13).

Onde $u_{0j} \sim N(0, \sigma_{u_0}^2)$ e $\varepsilon_{tj} \sim N(0, \Omega_\varepsilon)$ variâncias residuais heterocedásticas.

As inferências feitas acerca do modelo foram somente sobre as regiões metropolitanas que possuíam dados. O procedimento utilizado para obter as estimativas dos coeficientes fixos e aleatórios baseou-se no método de estimação REML, considerando a estrutura de variância-

covariância uma matriz definida positiva não estruturada para os efeitos aleatórios. A princípio, foram feitas modelagens para uma série temporal de 14 anos, através de modelos polinomiais e do uso de *splines* com um “*knots*” (em $t_0=7,5$) (Snijders e Bosker, 1999). A avaliação do melhor ajuste dos termos de tendência do modelo foi através da estatística AIC e dos efeitos aleatórios através do TRV. Correlações de Pearson e teste de multicolinearidade, através do VIF (Variance Inflation Factors), foram realizados para verificar a presença de colinearidade entre as variáveis explicativas. Após esta verificação, o modelo foi reespecificado definindo um novo conjunto de variáveis. O modelo completo estimou as médias anuais de PM_{10} explicados por seus determinantes incluindo o termo de tendência temporal (resultante da estatística AIC). Neste modelo, a transformação logarítmica (natural) foi aplicada na variável resposta e nas variáveis independentes, como forma de garantir as condições exigidas para a aplicação de um modelo linear de efeitos mistos como aproximação da distribuição normal e estabilidade da variância (Box e Cox, 1964).

Dada a estrutura dos componentes efeitos fixos e aleatórios com todas as variáveis explicativas definidas, procedeu-se à modelagem para os efeitos fixos através do método de seleção “*Stepwise Regression*”. Os coeficientes dos diferentes modelos foram estimados e os testes T realizados levando ao modelo final, sendo, portanto, considerado o modelo mais parcimonioso, no qual foram feitas as verificações e correções dos pressupostos básicos de um modelo linear com base em métodos gráficos de diagnóstico. Para checar a performance do modelo foram calculados o R_m^2 e R_c^2 proposto por Nakagawa e Schielzeth (2013), RMSE, MAE e o coeficiente de correlação de Pearson (r). As previsões para fora da amostra foram realizadas através do procedimento “*Data Splitting*” com a finalidade de examinar a adequação em extrapolar o modelo. O procedimento consistiu em retirar uma amostra aleatória simples de tamanho n , com base no desfecho, para compor o modelo de treinamento, a fim de predizer estimativas de PM_{10} e comparar com os dados correspondentes de monitoramento de PM_{10} disponíveis.

4.2.3 Estimação da carga de mortalidade e dos óbitos atribuíveis ao material particulado

Em um estudo seccional, funções de C-R foram desenvolvidas para dois desfechos de saúde, a fim de produzir estimativas de risco para os cálculos da carga de mortalidade para o ano de 2014: Mortalidade por doenças cardiovasculares em adultos de 30 anos ou mais e mortalidade em crianças <1 e 1-5 anos para ambos os sexos.

4.2.3.1 Estudo da Associação Americana de Câncer

Pope *et al.* (1995,2002) relataram um estudo de coorte prospectivo iniciado pela ACS em 1982. Mais de um milhão (1,2 milhão) de adultos com idade igual ou superior a 30 anos participaram do estudo. Eram membros de famílias com pelo menos um indivíduo com 45 anos de idade ou mais. Os participantes completaram um questionário confidencial que incluiu informações demográficas, tabagismo, uso de álcool, dieta e educação. As associações entre o MP e mortalidade foram relatadas em três momentos de exposição: 1979-1983: um seguimento de 7 anos (1982-1989) (Pope *et al.*, 1995); 1999-2000: um seguimento de 16 anos (1982-1998) e média dos dois períodos: acompanhamento de 16 anos (1982-1998) (Pope *et al.*, 2002). O risco foi estimado pelo modelo de risco proporcional de Cox de efeitos aleatórios, ajustado pela idade, sexo, raça, tabagismo, educação, estado civil, massa corporal, consumo de álcool, exposição ocupacional e dieta. Praticamente as funções C-R linear e log-linear se adequaram igualmente às concentrações de PM_{2,5} no intervalo entre 10-30 µg/m³ sendo que a função C-R log-linear gerou RR ligeiramente maiores dentro desta faixa e RR menores abaixo e acima deste intervalo (Ostro e Who, 2004)

4.2.3.2 Caso base

Para as análises, as faixas etárias e grupos de morbidades foram escolhidas para a adaptação dos estudos de Pope *et al.* (1995,2002) e Gouveia *et al.* (2018)

No nosso estudo, a estimativa do risco na mortalidade por doenças cardiovasculares em adultos ≥ 30 anos foi baseada na estimativa do RR do estudo do Pope *et al.* (2002), considerando

a média dos dois períodos de exposição do estudo, que correspondem aumentos de 9% na mortalidade por doenças cardiovasculares a cada 10 $\mu\text{g}/\text{m}^3$ na variação das concentrações de $\text{PM}_{2,5}$.

Embora o PM_{10} seja uma medida mais amplamente divulgada, e um indicador de relevância para a maioria dos dados epidemiológicos, as estimativas de PM_{10} adquiridas pelo modelo de efeitos mistos foram convertidas em estimativas de partículas finas de $\text{PM}_{2,5}$. Alternativamente, o valor de 0,5 foi aplicado para a conversão na razão $\text{PM}_{2,5}/\text{PM}_{10}$ definido para as áreas urbanas de países em desenvolvimento, como é o caso do Brasil (Ostro e Who, 2004; WHO, 2006). Esse procedimento foi utilizado para adequação da estimativa de risco para $\text{PM}_{2,5}$ do estudo do Pope *et al.* (2002). O RR foi calculado com uma média anual contrafactual (ou mínimo teórico) igual a 7,5 $\mu\text{g}/\text{m}^3$, a menor concentração observada em Pope *et al.* (1995,2002). A média total dos níveis de concentração de PM_{10} e $\text{PM}_{2,5}$ foi ponderada de acordo com o tamanho da população das regiões metropolitanas em 2014. E a função C-R considerada linear, que leva a seguinte fórmula:

$$RR_{C-C_0} = \frac{\exp(\alpha + \beta C)}{\exp(\alpha + \beta C_0)} = \exp[\beta(C - C_0)] \quad (32)$$

Onde RR_{C-C_0} é o RR da exposição média dos níveis de concentração C ponderada pela população, comparado à exposição de referência C_0 , e β é o coeficiente de inclinação (i.e., o efeito estimado do $\text{PM}_{2,5}$ sobre o desfecho de saúde de interesse).

Em alguns países a utilização de uma função linear gerou resultados improváveis ao extrapolar a função C-R linear para além de 30 $\mu\text{g}/\text{m}^3$ de $\text{PM}_{2,5}$, como foi demonstrado no estudo de coorte da ACS, o resultado poderia ser inadequado, pois sobrestimaria a mortalidade (Pope *et al.*, 2002; Cohen *et al.*, 2004; Ostro e Who, 2004). Neste caso, uma função de exposição não linear seria recomendada, pois sua inclinação somente se achataria nas mais altas concentrações de $\text{PM}_{2,5}$ (Ostro e Who, 2004). A função seria representada pela seguinte fórmula:

$$RR_{C-C_0} = \frac{\exp[\alpha + \beta(\ln(C + 1))]}{\exp[\alpha + \beta(\ln(C_0 + 1))]} = \left[\frac{C + 1}{C_0 + 1} \right]^\beta \quad (33)$$

Com relação à mortalidade por doenças respiratórias em crianças menores de um ano e de um a cinco anos foram calculados os RR para PM_{10} , também através da função C-R linear. As estimativas de risco foram baseados no estudo de Gouveia *et al.* (2018), uma extensão do

projeto ESCALA (“*Estudio de Salud y Contaminación del Aire en Latino América*”), que teve como objetivo avaliar o efeito da poluição do ar sobre a mortalidade respiratória infantil em quatro grandes centros urbanos: Cidade do México, Santiago, no Chile; São Paulo e Rio de Janeiro, no Brasil. Na nossa análise foram utilizadas estimativas das variações percentuais do risco para PM₁₀, estimadas através da metanálise de efeitos fixos do estudo de para crianças <1 e 1-5 anos. O método da metanálise combina resultados específicos das cidades em uma única estimativa de resumo. Em cada cidade e faixas etárias as estimativas das variações percentuais do risco foram estimadas através de modelos aditivos generalizados de Poisson, controlados por tendência, sazonalidade, temperatura e umidade relativa, com o intuito de ajustar a mortalidade diária por doenças respiratórias e PM₁₀ em crianças <1 e 1-5.

Na tabela 2 está o resumo do RR de mortalidade, das inclinações das funções C-R e os tipos de estudo que foram utilizados para o cálculo da carga de mortalidade.

Tabela 2 - Estimativas do RR de mortalidade, das inclinações da função C-R e os tipos de estudo recomendados para o cálculo da carga de mortalidade

| Desfechos de saúde | Fontes de dados | MP | Funções do RR | RR por 10 µg/m ³ (95% IC) | Inclinações das funções C-R por 10 µg/m ³ (95% IC) |
|---|--|-------------------|----------------------------------|--------------------------------------|---|
| Mortalidade por doenças cardiovasculares em adultos >=30 anos | Estudo da ACS (Pope et.al. 2002) | PM _{2,5} | RR=exp [β (X - X ₀)] | 1,09 (1,03 - 1,16) | Linear - média: 0,008933 ^a (0,00322 - 0,01464) |
| Mortalidade por doenças respiratórias em crianças < 1 ano | Estudo de Gouveia <i>et al.</i> (2018). Metanálise de efeitos fixos de 4 estudos de séries temporais sobre mortalidade diária | PM ₁₀ | RR=exp [β (X - X ₀)] | 1,0047 (1,0009 - 1,0085) | Linear 0,000469 ^b (0,00009 - 0,000846) |
| Mortalidade por doenças respiratórias em crianças de 1 a 5 anos | Estudo de Gouveia <i>et al.</i> (2018). Metanálise de efeitos fixos de 4 estudos de séries temporais sobre mortalidade diária | PM ₁₀ | RR=exp [β (X - X ₀)] | 1,0058 (1,0008 - 1,0108) | Linear 0,000578 ^b (0,00008 - 0,0010742) |

Tabela adaptada de Cohen *et al.* (2004)

^a O resultado proveniente do modelo de regressão especificado como média das concentrações de PM_{2,5} medidas no período 1979 - 1983 e 1999 - 2000 (Pope *et al.*, 2002; Cohen *et al.*, 2004).

^b Na metanálise, os autores consideraram mais apropriada para a associação entre mortalidade e PM₁₀ as estimativas das variações percentuais do risco do modelo de efeitos fixos.

Dada as estimativas dos RR, as frações atribuíveis para mortalidades cardiovasculares e respiratórias foram calculadas utilizando a seguinte fórmula (Ostro e Who, 2004):

$$FA = \frac{RR_{C-C_0} - 1}{RR_{C-C_0}} \quad (34)$$

Em sequência, a carga atribuível foi calculada multiplicando o desfecho por causa específica pela fração atribuível, como segue abaixo:

$$CA = FA \times \text{desfecho de saúde} \quad (35)$$

Onde CA é a carga atribuível, FA é a fração atribuível e o desfecho de saúde de interesse, ou seja, óbitos e/ou YLL.

Para cada desfecho e grupos de idade selecionados, os YLL foram calculados utilizando tábuas de mortalidade extraídas do site do IBGE e óbitos do Datasus. Uma vez que não foi possível estimar o impacto do MP sobre a incidência da doença, os DALY foram quantificados apenas com YLL.

4.2.4 Análise de sensibilidade

Uma vez que os efeitos da exposição de longo prazo de MP requerem vários pressupostos, exploramos a sensibilidade dos resultados para a mortalidade cardiovascular desses pressupostos. As análises de sensibilidade foram utilizadas para quantificar as incertezas das estimativas do caso base, em que a carga de mortalidade e óbitos atribuíveis foram estimados através da aplicação da função de C-R linear, considerando um mínimo contrafactual de $7,5 \mu\text{g}/\text{m}^3$ para $\text{PM}_{2,5}$, conforme discutido anteriormente. Então foram explorados cenários alternativos como variação da função C-R (função de exposição linear versus não linear); da concentração contrafactual (7,5, 3,0 ou $10 \mu\text{g}/\text{m}^3$) e da proporção de $\text{PM}_{2,5}$ para PM_{10} (0,50 ou 0,65).

4.2.5 Aspectos computacionais

Toda a programação foi realizada em R (R Core Team, 2013), versão 3.5.0, para fins de cálculos das análises descritivas, da elaboração de gráficos e mapas, da estimação do modelo de efeitos mistos (com a utilização da função lme na biblioteca nlme), e da avaliação de desempenho e precisão do modelo. Todo o cálculo sobre avaliação da carga de mortalidade foi realizado no Excel (2016).

5 RESULTADOS

5.1 O uso de modelos de efeitos mistos para prever a poluição do ar nas regiões metropolitanas do Brasil (manuscrito 1)

The use of mixed effects models to predict air pollution in metropolitan regions of Brazil

Simone de Souza Cardoso

Antônio Ponce de Leon

Resumo

Metodologias para estimar exposições ambientais tornaram-se cruciais nos estudos epidemiológicos sobre poluição do ar, principalmente para fins de cálculos de medidas globais de seus impactos em áreas urbanas. Este trabalho abordou os modelos de efeitos mistos como uma metodologia para previsão das concentrações de material particulado inalável com diâmetro $\leq 10\mu\text{m}$ (PM_{10}) nas regiões metropolitanas do Brasil, com base em informações já existentes de PM_{10} , durante o período de 2001 a 2014 (14 anos). Os dados foram coletados em centrais de monitoramento de poluição do ar e as médias anuais foram calculadas para cada região metropolitana. Como resultado, houve diferenças significativas entre as médias anuais de PM_{10} , bem como, variações ao longo do tempo. O modelo linear de efeitos mistos para dados longitudinais foi ajustado e incorporou preditores importantes, tais como, variáveis de consumo de energia, demográficas, socioeconômicas, ambientais e outras locais, além das variáveis de base temporal. As previsões das médias anuais de PM_{10} para regiões metropolitanas sem informações sobre o poluente foram realizadas sob o modelo mais parcimonioso, em que apenas as variáveis umidade relativa do ar, altitude média, PIB per capita do setor serviços e focos de queimadas explicaram os níveis de concentração de PM_{10} . O modelo apresentou o R^2 de 98,33% da variação, sendo que 27,54% dessa variação foi explicada pelos fatores fixos do modelo (R^2 marginal); o Root Mean Square Error (RMSE) foi de $0,07 \mu\text{g}/\text{m}^3$; a correlação de Pearson (r) foi 0,98; e o Mean Absolute Error (MAE) igual a 0,052. Na validação cruzada, o modelo apresentou, ainda, um R^2 de 0,983 e o R^2 marginal igual a 0,28, o RMSE foi de $0,07 \mu\text{g}/\text{m}^3$ e o MAE de 0,051. As previsões, quando comparadas com os dados observados dos monitores de

PM₁₀, apresentaram uma correlação total significativa de 70% ($\rho(\text{Spearman}) = 0,70$), o RMSE de 0,22 $\mu\text{g}/\text{m}^3$ e o MAE de 2,89.

Conclusão: o modelo demonstrou bom ajuste, robustez, eficiência e boa performance preditiva para realizar previsões.

Palavras-chave: Poluição do ar. Material particulado. Regiões metropolitanas. Modelos de efeitos mistos.

Abstract

Methodologies for estimating environmental exposures have become crucial in epidemiological studies on air pollution, mainly for calculating global measures of their impacts in urban areas. This work has addressed mixed effect models as a methodology for predicting concentrations of inhalable particulate matter with diameter $\leq 10\mu\text{m}$ (PM₁₀) in the metropolitan regions of Brazil, based on information already available from PM₁₀, during the period from 2001 to 2014 (14 years). The data were collected in air pollution monitoring centers and the annual averages were calculated for each metropolitan region. As a result, there were significant differences between the annual PM₁₀ averages, as well as, variations over time. The linear mixed effects model for longitudinal data was adjusted and incorporated important predictors, such as energy consumption, demographic, socioeconomic, environmental and other variables, as well as time base variables. The forecasts of PM₁₀ annual averages for metropolitan regions with no information on the pollutant were carried out under the most parsimonious model, in which only the variables relative air humidity, mean altitude, GDP per capita of the services sector and fires outbreaks explained the levels of concentration of PM₁₀. The model presented the R² of 98.33% of the variation, and 27.54% of this variation was explained by the fixed factors of the model (marginal R²); the Root Mean Square Error (RMSE) was 0.07 $\mu\text{g}/\text{m}^3$; Pearson's correlation (r) was 0.98; and the Mean Absolute Error (MAE) equal to 0.052. In the cross validation, the model also presented an R² of 0.983 and the marginal R² of 0.28, the RMSE was 0.07 $\mu\text{g}/\text{m}^3$ and the MAE of 0.051. The predictions, when compared to PM₁₀ monitors, showed a significant correlation of 70% ($\rho(\text{Spearman}) = 0.70$), the RMSE of 0.22 $\mu\text{g}/\text{m}^3$ and the MAE of 2.89.

Conclusion: the model demonstrated good fit, robustness, efficiency and good predictive performance to make predictions.

Keywords: Air pollution. Particulate matter. Metropolitan regions. Mixed effects models.

Introdução

Grande parte das redes de monitoramento da poluição do ar estão localizadas nos grandes centros urbanos ou em áreas metropolitanas com o intuito de medir emissões de poluição do ar de fontes específicas, tais como indústria local ou tráfego pesado de veículos. Os sistemas nacionais de monitoramento atmosférico existem em todas as partes do mundo, mas o acesso aos dados coletados por eles e a padronização internacional desses métodos de monitoramento são limitados. Alguns países de baixa ou média renda (a exemplo do Brasil) possuem redes de monitoramento, mas não fornecem dados necessários, ou fornecem de maneira precária dados sobre os poluentes em cidades ou regiões. Também frequentemente há regiões em que não existem nem informações sobre poluentes, e estas devem ser estimadas utilizando métodos de previsão (Krzyzanowsk e Schwela, 1999).

Sendo assim, técnicas foram desenvolvidas para lidar com problemas de dados faltantes em áreas geográficas fundamentadas nos métodos de previsão baseados em análise espacial. A maioria dessas técnicas foram aplicadas em cidades utilizando um número limitado de monitores. Além disso, utilizam também para prever níveis de poluição atmosférica em locais não mensurados, com base nos níveis de concentração observados em locais já monitorados, como por exemplo, o desenvolvimento de satélites de sensoriamento remoto e Modelos de Transporte Químicos Global (TM5), que simulam componentes de ozônio e aerossóis, fornecendo uma completa cobertura global em alto nível de resolução espacial, que permite estimações de exposições da poluição do ar (Brauer *et al.*, 2012). Além de satélites, outras metodologias espaciais têm sido exploradas para avaliações e predições dos impactos da exposição interurbanas, como, por exemplo, “*Land Use Regression Models*” (LUR), modelos de dispersão, modelos integrados e modelos híbridos (Jerrett *et al.*, 2005; Hoek *et al.*, 2008).

Ainda que a análise espacial seja justificável para os estudos de previsão, nos estudos de séries temporais, os modelos de efeitos mistos para dados longitudinais tem-se mostrado

como alternativa para estimar exposições de poluentes atmosféricos. Esses modelos capturam estruturas temporais complexas (como as tendências temporais de curto ou longo prazo) juntamente com variações regionais ou locais. Mais especificamente, o modelo consegue estimar médias anuais dos níveis de concentração de Material Particulado (MP) em áreas urbanas sem dados de monitoramento disponíveis, o que possibilita, por exemplo, cálculos de medidas globais de impacto em áreas urbanas. Sendo assim, o modelo não somente é utilizado para abordar questões de associação, quantificando fontes de variabilidade, como também para previsão da poluição do ar. Suas estimativas estão associadas às incertezas levando em consideração o cenário multivariável realista de áreas urbanas onde nem todos os locais de monitoramento medem o mesmo conjunto de poluentes. O método produz estimativas não viesadas e uma distribuição preditiva que engloba vários locais em diferentes pontos do tempo utilizando dados disponíveis, permitindo assim uma previsão simultânea temporal e local.

Desenvolvido pelo banco mundial, o “*Global Model of Ambient Particulates*” (GMAPS) é um modelo linear de efeitos mistos (modelo econométrico) para dados longitudinais que tem como propósito estimar concentrações de MP em cidades com populações maiores que 100.000 hab. e incorpora os principais determinantes das variações de concentrações de MP no ambiente externo (“*outdoors*”), tais como, variáveis de consumo de energia, socioeconômicas, ambientais, demográficas e outras variáveis locais. O modelo gera estimativas com base em medições disponíveis de MP em áreas já monitoradas. Não sendo um modelo causal, apenas prediz níveis de concentrações de poluição em áreas não monitoradas com intuito de apoiar, em termos globais, por exemplo, cálculos para avaliação do impacto dos níveis de concentração de poluição na saúde (Cohen *et al.*, 2004; Pandey *et al.*, 2006). Esse modelo foi utilizado para estimar as concentrações atuais de partículas inaláveis de diâmetro $\leq 10\mu\text{m}$ (PM₁₀) para uso do cálculo da “*Global Burden of Disease*” (GBD) do ano 2000. Cohen *et al.* (2004) obtiveram estimativas atuais de exposição comparáveis de poluição do ar para 14 grupos de países sub-regionais da Organização Mundial da Saúde (OMS) no período de 1985 a 1999. O modelo explicou 88% da variabilidade. A correlação total entre os dados observados e previstos do modelo foi cerca de $r=0,9$ para as observações de PM₁₀ e maior que 0,8 para Partículas Totais em Suspensão (PTS). Outros estudos como de Norman *et al.* (2007) e Brauer *et al.* (2012) também abordaram o modelo estatístico de previsão GMAPS.

Portanto, com base no estudo de Cohen *et al.* (2004), este trabalho teve como objetivo propor um modelo de efeitos mistos como uma metodologia para previsão das médias anuais do poluente PM₁₀ para todas as regiões metropolitanas do Brasil que não possuem dados de monitoramento desse poluente.

Metodologia

Modelo linear de efeitos mistos para dados longitudinais

Nos estudos ecológicos, o modelo linear de efeitos mistos para dados longitudinais é utilizado para descrever, num determinado período, a relação entre uma variável resposta e uma ou mais variáveis explicativas agrupadas em uma única estrutura hierárquica, na qual os níveis mais agregados são grupos, blocos, áreas geográficas ou estratos. Considera-se ainda que, em cada grupo tem-se uma série temporal de observações, que é o menor nível ou o nível menos agregado. Neste modelo, procura-se explorar as diferenças entre os blocos ou estratos captados pelo intercepto β_{0j} , condicionado às variáveis explicativas ou fatores contextuais pertencentes aos estratos, grupos ou áreas geográficas. Incluem, além dos efeitos fixos, termos adicionais de efeitos aleatórios que permitem a adição de mais um erro, que leva em conta a aleatoriedade entre os grupos. No caso de dois níveis, o modelo pode ser escrito como:

$$y_{tj} = \beta_0 + \sum_{h=1}^p \beta_h x_{htj} + u_j + \varepsilon_{tj}, \quad j = 1, 2, 3, \dots, M \text{ e } t = 1, 2, 3, \dots, n_j$$

Em que

$$u_j \sim \mathcal{N}(0, \sigma^2_j)$$

$$\varepsilon_{tj} \sim \mathcal{N}(0, \sigma^2_\varepsilon)$$

Onde u_j e ε_{tj} são independentes; y_{tj} é o valor no n -ésimo tempo pertencente ao j -ésimo grupo da variável resposta; x_{htj} é o t -ésimo valor no tempo pertencente ao j -ésimo grupo para o h -ésimo preditor; β_h é o coeficiente dos efeitos fixos para o h -ésimo preditor; u_j é o termo dos efeitos aleatórios pertencentes aos grupos com distribuição normal, média zero e variância σ^2_j e ε_{tj} é o termo residual com distribuição normal, média zero e variância σ^2_ε . A estimação dos parâmetros, neste modelo, baseia-se no método da máxima verossimilhança - “*Maximum Likelihood*” (ML) - ou no método da máxima verossimilhança restrita - “*Restricted Maximum Likelihood*” (REML). Os estimadores são consistentes, assintoticamente normais, eficientes e não viesados. A modelagem dos termos da estrutura aleatória é feita pelo teste da razão de

verossimilhança (TRV), adequado para comparar modelos aninhados com mesma estrutura fixa dos parâmetros estimados por ML e REML. Sendo este último preferível, pois as estimativas por ML tendem a subestimar as componentes da variância-covariância, tal como σ^2 , por não levar em conta a variação dos parâmetros fixos. Ou seja, a perda de graus de liberdade resultante da estimação dos efeitos fixos não é levada em consideração e, portanto, gera estimadores viesados (Pinheiro e Bates, 2000). Por este motivo, é recomendado a utilização dos testes t e testes F condicionais para a modelagem dos efeitos fixos e avaliação da significância estatística dos seus parâmetros (Pinheiro e Bates, 2000). A verificação dos pressupostos para o modelo linear de efeitos mistos é habitualmente feita através de testes de hipóteses e preferivelmente através de análises gráficas com base nos resíduos e nos valores ajustados. As violações mais frequentes que afetam sua eficiência são: a não verificação da hipótese de normalidade, homocedasticidade e/ou existência de correlação serial. O pressuposto da normalidade depende de resultados assintóticos, assumindo um modelo gaussiano, porém, quando violado pode produzir desvios-padrão e intervalos de confiança não confiáveis. Quanto à homocedasticidade e/ou existência de correlação serial, devido à própria estrutura hierárquica, em particular, as medições ou observações em um mesmo grupo estão geralmente correlacionadas e possuem variâncias e covariâncias residuais heterocedásticas quando uma variável explicativa é adicionada no modelo, comprometendo, portanto, suas inferências. Sendo assim, é possível relaxar estas condições adotando o procedimento de modelagem da estrutura de variâncias e covariâncias dos erros aleatórios. Para acomodar a heterocedasticidade, o modelo permite a utilização de diferentes funções-variância para modelar a estrutura de variância dos erros intragrupo. Além disso, permite também flexibilidade na especificação da estrutura dos efeitos aleatórios. Isto é, permite a adição de termos temporais como splines, funções polinomiais (Snijders e Bosker, 1999) ou utilização de modelos ARIMA (*Autoregressive Integrated Moving Average*) (Box *et al.*, 2015) para a modelagem da série temporal, no intuito de ajustar a autocorrelação (eliminando-a) ou apenas de “acomodar” a dependência serial.

As medidas de desempenho são cruciais para um bom modelo de previsão. A avaliação do R^2 num modelo de regressão tem uma propriedade de fornecer informação sobre o ajuste. Segundo a proposta de Nakagawa e Schielzeth (2013), nos modelos de efeitos mistos existem dois tipos de R^2 . O R_m^2 , conhecido como R^2 marginal, que indica a variância explicada pelos fatores fixos ilustrado como: $R_m^2 = \frac{\sigma_f^2}{\sigma_f^2 + \sum_{l=1}^u \sigma_l^2 + \sigma_\varepsilon^2}$, Onde, σ_f^2 é o componente da variância dos efeitos fixos, σ_l^2 é o componente da variância do l-ésimo fator aleatório e σ_ε^2 o componente da variância dos resíduos. E o R^2 condicional, que é a variância explicada pelos fatores fixos e

aleatórios e interpretada, portanto, como a variância explicada por todo o modelo, e é expressa como: $R_c^2 = \frac{\sigma_f^2 + \sum_{l=1}^u \sigma_l^2}{\sigma_f^2 + \sum_{l=1}^u \sigma_l^2 + \sigma_\varepsilon^2}$. Tanto o R^2 marginal como o condicional transmitem informações únicas e interessantes, e recomenda-se que ambos sejam apresentados no modelo.

A avaliação da performance do modelo pode ser feita através da comparação dos dados observados com os valores ajustados através da utilização de algumas medidas estatísticas, como: o RMSE (*Root Mean Square Error*), o MAE (*Mean Absolute Error*) e o coeficiente de correlação. Os valores preditos podem ser avaliados em termos de desempenho e precisão, através do erro médio quadrático, $RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}}$. Em termos de acurácia, $MAE = \frac{\sum_{i=1}^N |\hat{y}_i - y_i|}{N}$, que é uma medida mais sensível, menos influenciado por maiores diferenças entre os valores preditos e observados, e o coeficiente de correlação de Pearson (r), um bom indicador para avaliar o desempenho do modelo e expresso como $r = \frac{cov(y, \hat{y})}{\sqrt{var(y)var(\hat{y})}}$. Dentre outros tipos de procedimentos para validação do modelo, o procedimento “*Data Splitting*” ou “*cross validation*”, consiste em deixar de lado alguns dados originais, os utilizando depois para investigar o desempenho preditivo do modelo com base nos valores preditos para novas observações, uma das principais quantidades de interesse para se examinar a adequação de extrapolação do modelo. De acordo com Pinheiro e Bates (2000), os valores ajustados são os valores preditos da variável resposta sob o modelo ajustado que podem ser obtidos no nível menos agregado (primeiro nível), as predições do primeiro nível estimam apenas o valor marginal esperado da variável resposta. Assim, dado um vetor de variáveis dos efeitos fixos x_h , o valor marginal esperado da variável resposta é representado da seguinte forma: $\hat{y}_h = x_h^T \hat{\beta}$. Já os valores previstos estimam a expectativa condicional da variável resposta dado aos efeitos aleatórios do segundo nível. Por exemplo, $Z_h(j)$ denota um vetor de variáveis contextuais correspondente aos efeitos aleatórios associados ao j_{th} , grupo. Sendo assim, a estimativa de previsão para novas observações é: $\hat{y}_h = x_h^T \hat{\beta} + Z_h(j)^T \hat{b}_j$.

Análise Estatística

Dados

A partir das fontes de informações publicamente disponíveis, foram obtidos dados sobre PM₁₀, consumo de energia, transportes, queimadas, dados meteorológicos, geográficos, demográficos e socioeconômicos para 36 regiões metropolitanas brasileiras, no período de 2001 a 2014.

Os dados de poluição (horários, diários e mensais de PM₁₀) resultantes de monitores automáticos e semiautomáticos, localizados nas áreas urbanas, foram obtidos pelos órgãos ambientais competentes para a construção das médias anuais de PM₁₀. No caso de dados faltantes nas estações de monitoramento foi empregado o procedimento de imputação em séries temporais multivariado com base no algoritmo EM (*Expectation – Maximization*) (Dempster *et al.*, 1977). Este método utiliza estimativas iterativas por máxima verossimilhança, para o vetor de médias e da matriz de covariância de uma distribuição normal multivariada com dados faltantes. Assim, o algoritmo EM foi aplicado sob os pressupostos da distribuição normal multivariada. Além das estruturas de dependência entre as variáveis, o método também leva em consideração as estruturas de dependência temporal de cada variável. Maiores detalhes são apresentados em Dempster *et al.* (1977); Junger e Ponce De Leon (2015). O método está implementado no programa R (*R Development Core Team, 2013*) na biblioteca *mtsd*. Após o procedimento de imputação, as médias anuais de PM₁₀ foram calculadas em cada região metropolitana.

Os dados sobre consumo anual de gasolina, diesel e etanol (em litros) estão disponíveis no site da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP. A inclusão de cada tipo de combustível refere-se às variações nos fatores de emissão como atividade econômica, intensidade do uso de combustível e emissões relacionadas ao tráfego em diferentes regiões metropolitanas.

Os dados sobre as variáveis climáticas explicam a dispersão e a capacidade de transporte de emissões de poluentes. Seis variáveis meteorológicas foram utilizadas para representar as condições atmosféricas locais: Média anual de nebulosidade (decimais); média anual da Pressão barométrica (hPa); média anual da temperatura (Celsius); média anual de umidade relativa (%);

média anual da Velocidade do vento (m/s) e índice pluviométrico total anual (mm). Estes dados estão disponíveis no site do Instituto Nacional de Meteorologia - INMET.

As seguintes variáveis topográficas foram consideradas no estudo: área geográfica (km²) e altitude média (em metros). Os dados demográficos foram representados por população e densidade demográfica. Estas variáveis fornecem medidas da intensidade do problema de poluição em cada região metropolitana. Variáveis socioeconômicas como Produto Interno Bruto (PIB) total, PIB do setor agropecuário, PIB do setor indústria e PIB do setor serviços (em R\$ 1.000) também foram consideradas no estudo de poluição. Todas estas variáveis estão disponíveis no site do Instituto Brasileiro de Geografia e Estatística – IBGE.

Duas variáveis de transporte foram representadas no estudo. A primeira delas foi número de decolagens de aviões, que representa as emissões de poluentes geradas pelas fases em que os motores das aeronaves se encontram em modo de operação de taxiamento na pista durante o ciclo de pouso e decolagem. A segunda variável considerada foi a taxa de crescimento anual de veículos novos, que representa as emissões de poluentes geradas pelos carros e é calculada pela seguinte fórmula: $Tx_{veiculos} = \frac{F_t - F_{t-1}}{F_{t-1}} \times 100$, onde F_t é o número de frotas de veículos do ano atual e F_{t-1} , o número de frotas do ano anterior. Respectivamente, as variáveis de transporte estão disponíveis nos sites da Agência Nacional de Aviação Civil – ANAC e do departamento nacional de trânsito DENATRAN.

Também foi considerado no estudo o número de focos de queimadas (total anual de focos). Os dados foram extraídos com base nos registros do satélite de referência do site do Instituto Nacional de Pesquisas Espaciais – INPE.

Estimação das médias anuais de PM₁₀ nas regiões metropolitanas do Brasil

No primeiro passo foi realizada uma análise exploratória das médias anuais de PM₁₀ e dos fatores contextuais das regiões metropolitanas do Brasil, que consistiu de estatísticas descritivas, elaboração de gráficos e mapas.

Na finalidade de reduzir viés das médias anuais de PM₁₀ e evitar reduções na amostra, a aplicação do procedimento de imputação multivariada foi realizada para as estações de monitoramento e para as variáveis explicativas com limite de 40% de dados faltantes, já que as estações e as variáveis explicativas encontravam-se em condições extremas de dados faltantes.

A falta de rigidez no limite foi para evitar perdas de variáveis explicativas e estações de monitoramento o que decorreria em perdas de regiões metropolitanas para o estudo. As estações de monitoramento que ultrapassaram esse limite foram desconsideradas do estudo por produzirem correlações espúrias, devido a poucas observações e estimativas irreais das medidas de PM₁₀. Para as estações que permaneceram no estudo foram calculadas e resumidas correlações de *Spearman* em cada região metropolitana. O algoritmo EM foi empregado para a realização das imputações e a trajetória temporal das médias de PM₁₀ foi modelada utilizando splines.

Após as imputações, o modelo linear de efeitos mistos foi ajustado para prever médias anuais de PM₁₀ nas regiões metropolitanas brasileiras, condicionadas às características econômicas, demográficas, meteorológicas e de fontes emissoras de poluição. Os principais determinantes das variações observadas nas concentrações de PM₁₀ em cada região metropolitana estão estruturadas no modelo a seguir (Cohen *et al.*, 2004):

$$C_{tj} = \beta_0 + \sum_{f=1}^F \beta_{Ef} E_{tj} + \sum_{g=1}^G \beta_{Mg} M_{tj} + \beta_{DD} DD_{tj} + \sum_{v=1}^V \beta_{Rv} R_{tj} + \sum_{s=1}^S \beta_{Ts} T_{tj} + \beta_{Tempo} Tempo_{tj} + \varepsilon_{tj} + u_{0j}$$

Onde

C_{tj} = log das médias anuais de PM₁₀ na região metropolitana j, no tempo t.

E_{tj} = log das variáveis de consumo de energia para região metropolitana j, no tempo t (f=1, f=2, ..., F)

M_{tj} = log das variáveis climáticas e geográficas para região metropolitana j, no tempo t (m=1, m=2, ..., M)

DD_{tj} = log da densidade populacional da região metropolitana j, no tempo t

R_{tj} = log das variáveis socioeconômicas da região metropolitana j, no tempo t (v=1, v=2, ..., V)

T_{tj} = log das variáveis de transporte da região metropolitana j, no tempo t (s=1, s=2, ..., S)

$Tempo_{tj}$ = termo de tendência do tempo (2001=1, 2002=2, ..., 2014=13).

Onde $u_{0j} \sim N(0, \sigma_{u_0}^2)$ e $\varepsilon_{tj} \sim N(0, \Omega_\varepsilon)$ variâncias residuais heterocedásticas.

Para estimar o modelo, a função lme foi aplicada. Esta função pode ser encontrada na biblioteca nlme, um pacote estatístico implementado no software R. O procedimento utilizado

para obter as estimativas dos coeficientes fixos e aleatórios baseou-se no método de estimação REML, considerando a estrutura de variância-covariância sendo uma matriz definida positiva não estruturada para os efeitos aleatórios. Inicialmente, foram feitas modelagens para uma série temporal de 14 anos, através do uso de *splines* com um “*knots*” e modelos polinomiais conforme Snijders e Bosker (1999). A modelagem dos termos temporais do modelo para a escolha do ajuste da tendência foi através da estatística AIC e dos efeitos aleatórios através do teste TRV. Correlações de Pearson e teste de multicolinearidade, VIF (“variance inflation factors”), foram realizados para verificar a presença de colinearidade entre as variáveis explicativas. Dessa forma, o modelo foi reespecificado, definindo um novo conjunto de variáveis. O modelo completo estimou as médias anuais de PM_{10} , explicadas por seus determinantes e pelo termo de tendência temporal linear (obtido através da estatística AIC). As regressões foram realizadas usando o logaritmo natural na variável resposta e nas variáveis independentes, como forma de garantir as condições exigidas para a aplicação de um modelo linear de efeitos mistos, como aproximação da distribuição normal e estabilidade da variância (Box e Cox, 1964). Dada a estrutura dos efeitos fixos, com todas as variáveis definidas, e dos efeitos aleatórios (o modelo completo), procedeu-se à modelagem dos efeitos fixos através do método de seleção “*Stepwise Regression*”. Os coeficientes dos diferentes modelos foram estimados e testes T foram realizados levando ao modelo final. Sendo, portanto, considerado o modelo mais parcimonioso, no qual foram feitas as verificações e correções dos pressupostos básicos de um modelo linear, com base em métodos gráficos de diagnóstico. Na identificação da autocorrelação, ferramentas gráficas foram utilizadas como a função autocorrelação (FAC) e função autocorrelação parcial (FACP). No caso de detecção da autocorrelação, foram realizadas medidas corretivas, como a inclusão de termos dos modelos ARIMA. Já no caso da existência de heterocedasticidade, foram investigadas diferentes funções-variância para modelar a estrutura de variância dos erros.

A performance do modelo foi avaliada através do R_m^2 e R_c^2 , proposto por Nakagawa e Schielzeth (2013), RMSE, MAE e o coeficiente de correlação de Pearson (r). As previsões para fora da amostra foram realizadas através do procedimento “*Data Splitting*” com a finalidade de examinar a adequação do modelo para extrapolação. A partição dos dados foi com base no desfecho empregando cerca de 90% do conjunto de dados original para treinamento e 10% do restante da amostra para previsões. O procedimento foi realizado aplicando a função `sample.split` da biblioteca “*caTools*” implementado no software R. No modelo resultante dos dados de treinamento, foram aplicadas também medidas de desempenho similares tais como, R^2 , RMSE, MAE e correlações de *Pearson*.

Resultados

Descrição dos dados

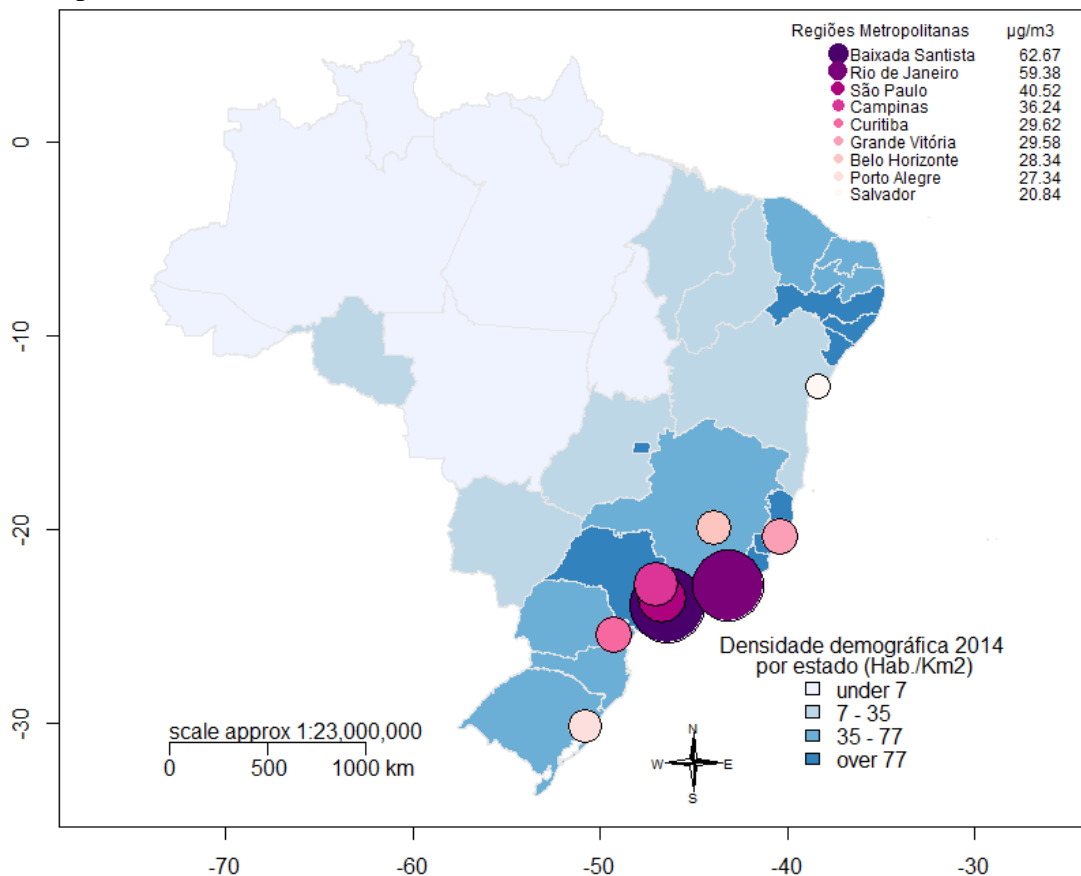
O número total de municípios que possuem redes de monitoramento de PM₁₀ em cada região metropolitana está ilustrado na tabela 3. Durante o período de 2001 a 2014, dentre as 36 regiões metropolitanas do Brasil, apenas nove regiões continham medidas de concentrações de PM₁₀, resultando num total de 149 estações, distribuídas entre 211 municípios. Dentre esses municípios, apenas 48 continham informações sobre os níveis de concentração de PM₁₀. A cobertura das redes nos municípios e populações variaram significativamente nas diferentes regiões metropolitanas. Por exemplo, os dados de PM₁₀ estavam disponíveis em apenas dois municípios da Baixada Santista sendo avaliados por 6 estações. Enquanto que, na região metropolitana do Rio de Janeiro, as redes estavam disponíveis para 11 municípios, totalizados em 52 estações. É importante mencionar que os 48 municípios com redes de monitoramento representaram apenas 8,2% do total de municípios pertencentes às áreas metropolitanas em todo o Brasil. Em 2014, um pouco mais de 42 milhões de habitantes se encontravam em municípios acobertados por essas redes de monitoramento, ou seja, aproximadamente 48% de toda a população metropolitana do Brasil.

Tabela 3 – Municípios e população urbana com redes de monitoramento de PM₁₀, por região metropolitana, período 2001 a 2014

| Regiões Metropolitanas | Municípios | | | | População Urbana em 2014 | | |
|-------------------------|------------|----------------------------|------------|--|--------------------------|----------------------------|---------------|
| | Total | com redes de monitoramento | Estações | Percentual dos municípios com redes de monitoramento | Total | com redes de monitoramento | Percentual |
| Baixada Santista | 9 | 2 | 6 | 22,2% | 1.781.620 | 559.670 | 31,41% |
| Belo Horizonte | 48 | 4 | 11 | 8,3% | 5.767.414 | 3.718.520 | 64,47% |
| Campinas | 19 | 3 | 4 | 15,8% | 3.043.217 | 1.476.808 | 48,53% |
| Curitiba | 26 | 3 | 7 | 11,5% | 3.414.115 | 2.225.644 | 65,19% |
| Grande Vitória | 7 | 4 | 9 | 57,1% | 1.884.096 | 1.673.137 | 88,80% |
| Porto Alegre | 31 | 8 | 11 | 25,8% | 4.161.237 | 2.469.194 | 59,34% |
| Rio de Janeiro | 19 | 11 | 52 | 57,9% | 12.116.616 | 10.405.332 | 85,88% |
| Salvador | 13 | 4 | 13 | 30,8% | 3.919.864 | 3.292.633 | 84,00% |
| São Paulo | 39 | 9 | 36 | 23,1% | 20.935.204 | 16.704.378 | 79,79% |
| Total | 211 | 48 | 149 | 22,7% | 57.023.383 | 42.525.316 | 74,58% |

As medidas sumárias das médias anuais de concentração de PM_{10} , existentes nas nove regiões metropolitanas durante todo o período de estudo, estão apresentadas na tabela 4. Além das variações do número de estações, disponíveis entre as regiões metropolitanas que possuem dados de monitoramento, as médias anuais de PM_{10} também variaram significativamente (teste $F = 53,5$, p -valor=0,0). A Baixada Santista apresentou a mais alta média dos níveis de concentração de PM_{10} , com $62,8 \mu\text{g}/\text{m}^3$ durante o período de 2001 a 2014. Seguida da região metropolitana do Rio de Janeiro, com $59,4 \mu\text{g}/\text{m}^3$ e maior desvio padrão de $14,4 \mu\text{g}/\text{m}^3$ entre os períodos. Além de serem consideradas *outliers* na amostra, pode observar que ambas as regiões violaram os padrões de qualidade do ar estabelecidos pela OMS e pelo Conselho Nacional do Meio Ambiente (CONAMA). Por outro lado, a região metropolitana de Salvador obteve a mais baixa média anual de PM_{10} , com $20,8 \mu\text{g}/\text{m}^3$ e menor variabilidade longo do tempo ($DP=1,8 \mu\text{g}/\text{m}^3$).

Figura 2 - Médias anuais de PM_{10} ($\mu\text{g}/\text{m}^3$) em nove regiões metropolitanas do Brasil durante o período 2001 – 2014



Fonte: Elaboração própria

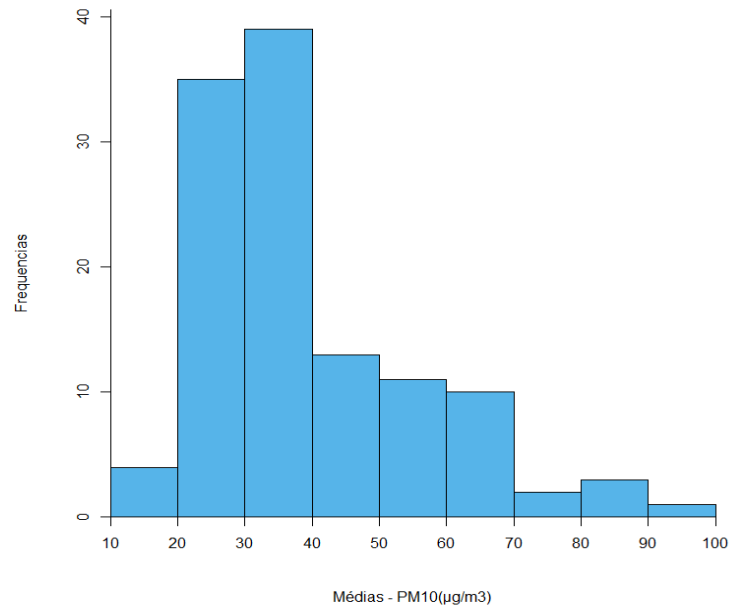
Tabela 4 - Estatística descritiva das médias anuais de PM₁₀ (µg/m³) de acordo com as regiões metropolitanas durante o período de 2001 a 2014

| Regiões Metropolitanas | T (anos) | Média | DP | Min. | 1ºQ | 2ºQ | 3ºQ | Máx. |
|-------------------------------|-----------------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Baixada Santista | 14 | 62,7 | 9,7 | 48,3 | 55,3 | 63,8 | 67,1 | 84,0 |
| Belo Horizonte | 12 | 28,3 | 6,5 | 17,2 | 24,2 | 29,7 | 33,8 | 35,9 |
| Campinas | 14 | 36,2 | 3,8 | 30,8 | 33,1 | 36,8 | 37,8 | 43,5 |
| Curitiba | 9 | 29,6 | 4,0 | 25,9 | 27,1 | 28,3 | 30,8 | 39,0 |
| Grande Vitória | 13 | 29,6 | 2,8 | 25,3 | 28,1 | 29,9 | 31,2 | 34,0 |
| Porto Alegre | 13 | 27,3 | 4,5 | 21,4 | 23,1 | 28,7 | 30,5 | 35,4 |
| Rio de Janeiro | 14 | 59,4 | 14,4 | 43,3 | 49,1 | 55,0 | 62,3 | 90,8 |
| Salvador | 11 | 20,8 | 1,8 | 16,2 | 20,3 | 21,0 | 22,2 | 23,0 |
| São Paulo | 14 | 40,5 | 5,8 | 33,6 | 36,9 | 39,0 | 41,6 | 53,2 |
| Total | 114 | 38,2 | 15,7 | 16,2 | 28,1 | 33,6 | 45,9 | 90,8 |
| | gl | SQ | MQ | F | Prob.>F | | | |
| Between groups | 8 | 22.443,2 | 2.805,4 | 53,5 | 0,0 | | | |
| Within groups | 105 | 5.508,3 | 52,5 | | | | | |
| Total | 113 | 27.951,5 | 247,4 | | | | | |
| Within-groups | DP=7,2 | | | | | | | |
| Between-groups | DP= 15,7 | | | | | | | |

A estatística descritiva das médias de concentração de PM₁₀ foi obtida com base na análise de dados completos. O procedimento pode ter produzido estimativas viesadas das médias, já que grande parte das estações de monitoramento da qualidade do ar apresentaram altos percentuais de dados faltantes.

No gráfico 1, as frequências das regiões metropolitanas estão distribuídas em intervalos de classe das médias anuais de PM₁₀. Durante o período compreendido entre 2001 a 2014, a maior parte das áreas urbanas apresenta níveis de PM₁₀ em torno de 20 a 40 µg/m³ com assimetria à direita indicando que as concentrações urbanas não são normalmente distribuídas.

Gráfico 1 - Histograma das médias anuais de PM₁₀ em áreas metropolitanas durante o período 2001 – 2014



Elaboração própria

A tabela 5 mostra a estatística descritiva das variáveis de estudo pertencentes às 36 regiões metropolitanas durante o período de 2001 a 2014. O conjunto de dados compreendeu 504 valores observados, exceto as variáveis meteorológicas, número de decolagens e número de focos de queimadas, que apresentaram valores faltantes durante o período focalizado nesta pesquisa.

Tabela 5 - Estatística descritiva das características das 36 regiões metropolitanas, 2001 – 2014

| variáveis | N | Média | DP | Min. | Mediana | Max. |
|---|----------|--------------|-------------|-------------|----------------|--------------|
| Consumo anual de gasolina (em litros) | 504 | 362.544,0 | 659.868,5 | 4,3 | 153.436,2 | 5.040.322,7 |
| Consumo anual de óleo diesel (em litros) | 504 | 229.218,6 | 460.881,7 | 2,9 | 11.491,6 | 3.365.170,4 |
| Consumo anual de etanol (em litros) | 504 | 124.771,7 | 363.678,7 | 499,6 | 27.012,1 | 3.434.217,1 |
| Média anual de nebulosidade (decimais) | 419 | 6,0 | 0,9 | 3,6 | 6,1 | 7,8 |
| Média anual da Pressão barométrica (hPa) | 373 | 977,3 | 38,1 | 885,5 | 996,7 | 1.019,3 |
| Média anual da temperatura (Celsius) | 418 | 24,1 | 3,3 | 14,2 | 25,6 | 28,4 |
| Média anual de Umidade relativa do ar (%) | 419 | 76,5 | 6,1 | 54,7 | 77,5 | 98,0 |
| Média anual da velocidade do vento (m/s) | 407 | 1,9 | 0,8 | 0,4 | 1,9 | 4,1 |
| Índice pluviométrico total anual (mm) | 418 | 2.436,1 | 1.293,8 | 104,0 | 2.070,6 | 7.365,9 |
| Altitude média (em metros) | 504 | 272,3 | 291,5 | 17,6 | 121,0 | 982,1 |
| Área (km ²) | 504 | 8.941,4 | 16.355,6 | 865,8 | 5.003,8 | 101.475,0 |
| População (em 10.000 habitantes) | 504 | 229,2 | 358,2 | 32,5 | 109,2 | 2.093,5 |
| Densidade demográfica (Hab./Km ²) | 504 | 458,1 | 549,9 | 16,8 | 233,2 | 2.634,4 |
| PIB (Total em R\$ 100.000) | 504 | 474.116,5 | 1.088.217,1 | 8.343,4 | 154.055,5 | 10.228.665,2 |
| PIB setor serviços (Total em R\$ 100.000) | 504 | 272.273,3 | 662.102,2 | 5.831,8 | 89.261,5 | 6.228.008,3 |
| PIB setor agropecuário (Total em R\$ 100.000) | 504 | 3.170,7 | 2.901,1 | 168,2 | 2.170,2 | 17.809,6 |
| PIB setor industrial (Total em R\$ 100.000) | 504 | 102.221,4 | 201.237,0 | 1.071,3 | 33.906,5 | 1.453.304,6 |
| Número de decolagens (aviões) | 460 | 19.697,8 | 33.990,0 | 1,0 | 7.999,5 | 221.734,0 |
| Total de frotas (veículos) | 504 | 804.508,4 | 1.487.868,0 | 32.232,0 | 364.150,0 | 11.859.681,0 |
| Número de focos de queimadas (total anual) | 493 | 102,0 | 175,9 | 1,0 | 39,0 | 1.396,0 |

Elaboração própria

Dados faltantes

No INMET não houve dados meteorológicos disponíveis para Baixada Santista e Campinas. Um procedimento seguido foi preencher os valores meteorológicos com os dados

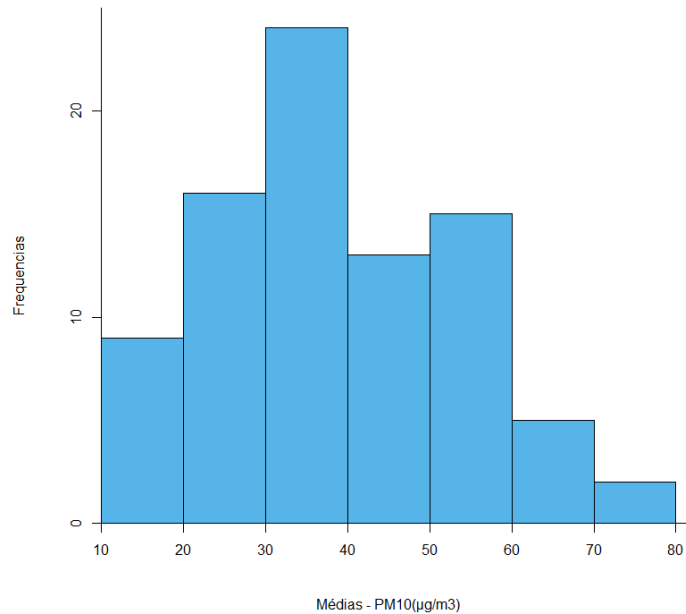
da CETESB localizados nas mesmas regiões. Essa foi uma forma de evitar a eliminação dessas duas importantes regiões metropolitanas. Para realizar a comparabilidade entre as duas fontes, as correlações de Spearman entre as variáveis foram testadas utilizando os dados da região metropolitana de São Paulo. As únicas variáveis meteorológicas que se demonstraram correlacionadas foram umidade e temperatura. Essas variáveis apresentaram uma correlação positiva e significativa (Umidade: $\rho = 0,58$; p-valor = 0.017 e temperatura: $\rho = 0,54$; p-valor = 0.024). As demais variáveis climáticas não apresentaram correlações significativas e, portanto, não foram consideradas na análise multivariada.

Houve, também, dados faltantes para as variáveis número de decolagens e número de focos de queimadas. Neste caso, o procedimento de imputação multivariada foi aplicado. Com base neste procedimento, exceder mais de 40% de dados faltantes é considerado uma condição extrema de informação perdida e, portanto, pode tender a produzir estimativas mais instáveis e má produção da matriz de covariância devido ao alto número de valores faltantes. Sendo assim, a variável número de decolagens, que apresentou em torno de 57% de dados faltantes na Baixada Santista, foi desconsiderada da análise multivariada. Por outro lado, o número de focos de queimadas apresentou apenas 14% de dados faltantes na Baixada Santista e 7% na Grande Vitória, portanto, permanecendo na análise. Os coeficientes de correlação linear de Spearman apresentaram altas e baixas correlações entre os pares de regiões metropolitanas, o que denota uma distribuição heterogênea das ocorrências de queimadas ao longo do tempo. Nessas condições, o método de imputação multivariado é indicado para a previsão das estimativas por apresentar um bom desempenho. As estimativas foram obtidas com ajuste do componente temporal por *splines*, com dois graus de liberdade para cada região metropolitana.

As imputações também foram realizadas nas estações de monitoramento de PM_{10} como uma maneira de reduzir o viés das estimativas, devido à alta proporção de dados faltantes. Todas as estações que apresentaram proporções menores ou iguais a 40% de dados faltantes entram no estudo para o cálculo das médias anuais de PM_{10} . Três regiões metropolitanas foram excluídas da análise multivariada: Região metropolitana de Porto Alegre, Curitiba e Belo Horizonte. Nessas regiões, todas as estações extrapolaram do valor limite e, portanto, não foi possível fazer previsões das medidas de PM_{10} . Em cada uma das regiões que fizeram parte do estudo, os pares de estações apresentaram altas e baixas correlações de *Spearman*, o que, novamente, denota distribuições heterogêneas locais de PM_{10} ao longo do tempo. As previsões das médias anuais faltantes, também, foram obtidas com ajuste de um componente temporal por *splines* com dois graus de liberdade para cada estação.

As médias anuais de PM_{10} , após as imputações, tenderam a ser mais normalmente distribuídas e foram estatisticamente diferentes (*Paired t-test*: $t = -3,019$, $df = 79$, $p\text{-valor} = 0,003$) quando comparadas com as médias anuais calculadas apenas com base nos dados completos do poluente.

Gráfico 2 - Histograma das médias anuais de PM_{10} , após as imputações, em áreas metropolitanas durante o período 2001 a 2014



Elaboração própria

O modelo linear de efeitos mistos

O modelo mais parcimonioso, escolhido pelo método de seleção “*Stepwise Regression*”, foi estimado e está apresentado na tabela 6. É um modelo autoregressivo de primeira ordem – $AR(1)$ - com estrutura de variância identidade (considerando diferentes variâncias por estrato) e distribuição Gaussiana. Na parte fixa do modelo a função linear de base temporal foi introduzida para capturar as tendências das médias anuais de PM_{10} nas regiões metropolitanas e representou 24% da variância explicada. As variáveis locais selecionadas, significativamente associadas aos níveis de concentração de PM_{10} , fizeram diferentes contribuições na variação do poluente. Dentre essas variáveis, altitude média representou 26,95% da variância, umidade relativa do ar 21,51%, PIB per capita do setor serviços 15,24% e focos de queimadas 11,39%. O modelo explicou 98,33% da variação total, sendo que, 27,54% dessa variação foi explicada pelos fatores fixos. A correlação de *Pearson* entre os dados observados e previstos foi cerca de 0,98. As medidas de desempenho do modelo indicaram boa

qualidade dos valores estimados apresentando o RMSE igual a $0,07 \mu\text{g}/\text{m}^3$ e o MAE de 0,052. Na validação cruzada, cerca de 90% do conjunto de dados original foi para treinamento e 10% do restante da amostra para teste, o que correspondeu, do total de 84 observações, 9 observações para previsões. Nesta validação, o modelo de treinamento, com 75 observações, mostrou um R^2 marginal igual a 28%, o R^2 igual a 0,983, o RMSE apresentou um valor de $0,07 \mu\text{g}/\text{m}^3$ e o MAE de 0,051. A tabela 7 apresenta as 9 observações de teste com as médias anuais do monitoramento de PM_{10} provenientes das estações da qualidade do ar para as regiões metropolitanas com as respectivas previsões do modelo de treinamento. A correlação total entre ambas as medidas foi significativamente de 70% ($\rho_{(\text{Spearman})} = 0,70$), o RMSE foi de $0,22 \mu\text{g}/\text{m}^3$ e o MAE de 2,89.

Tabela 6 - Estimativas dos parâmetros para os efeitos fixos e aleatórios, erro-padrão, estatística do teste-t e p-valores: modelo AR(1) com estrutura de variância identidade

(continua)

| Variáveis | Estimativas | Erro Padrão | gl | T | Intervalo de. | | P-valor |
|-------------------------------|-------------|-------------|----|--------|---------------|--------|---------|
| | | | | | confiança 95% | | |
| | | | | | Inf. | Sup. | |
| Resposta: Ln_PM ₁₀ | | | | | | | |
| Parte fixa | | | | | | | |
| constante | 7,719 | 0,891 | 74 | 8,660 | 5,943 | 9,495 | 0,000 |
| tempo (linear) | -0,040 | 0,009 | 74 | -4,624 | -0,058 | -0,023 | 0,000 |
| Ln_ umidade | -0,769 | 0,179 | 74 | -4,295 | -1,126 | -0,412 | 0,000 |
| Ln_ pib serviços per capita | 0,272 | 0,075 | 74 | 3,615 | 0,122 | 0,421 | 0,001 |
| Ln_ altitude media | -0,329 | 0,068 | 4 | -4,809 | -0,519 | -0,139 | 0,009 |
| Ln_focos | 0,031 | 0,010 | 74 | 3,126 | 0,011 | 0,050 | 0,003 |
| Parte aleatória | | | | | | | |
| Nível: RM | | | | | | | |
| s.d. (cons.) | 0,566 | | | | 0,261 | 1,228 | |
| s.d. (trend) | 0,012 | | | | 0,006 | 0,026 | |
| corr.(cons./trend) | -0,993 | | | | -1,000 | 1,000 | |
| Nível: ano | | | | | | | |
| resíduos | 0,087 | | | | 0,053 | 0,143 | |
| R ² (m) | 27,54% | | | | | | |
| R ² (c) | 98,33% | | | | | | |
| Units: RM | 6 | | | | | | |
| Units: ano | 84 | | | | | | |
| Estimation: | REML | | | | | | |
| AIC | -128,792 | | | | | | |
| logLik | 80,396 | | | | | | |

Tabela 6 - Estimativas dos parâmetros para os efeitos fixos e aleatórios, erro-padrão, estatística do teste-t e p-valores: modelo AR(1) com estrutura de variância identidade
(conclusão)

| variáveis | Anova | | | | | | |
|---|-------|---------------------------------------|----|-------|-------------|---------|------------|
| | SSQ | MSQ | gl | DenDF | Valor F | Pr.(>F) | Pct. Expl. |
| tempo (linear) | 0,124 | 0,124 | 1 | 74 | 4,698 | 0,033 | 24,92% |
| Ln_ umidade | 0,107 | 0,107 | 1 | 74 | 21,326 | 0,000 | 21,51% |
| Ln_ pib serviços per capita | 0,076 | 0,076 | 1 | 74 | 5,045 | 0,028 | 15,24% |
| Ln_ altitude media | 0,134 | 0,134 | 1 | 4 | 31,292 | 0,005 | 26,95% |
| Ln_focos | 0,057 | 0,057 | 1 | 74 | 9,769 | 0,003 | 11,39% |
| correlação de Pearson (r) = 0,98 | | RMSE = 0,070 $\mu\text{g}/\text{m}^3$ | | | MAE = 0,052 | | |
| VIF: tempo = 2,62; Ln_ umidade = 1,08; Ln_ pib serviços per capita = 2,92; Ln_ altitude media = 1,23, Ln_focos = 1,22 | | | | | | | |
| cross validation: correlação de Pearson (r) = 0,98 | | RMSE = 0,070 $\mu\text{g}/\text{m}^3$ | | | MAE = 0,051 | | |
| R ² (m) = 28% | | R ² (c) = 98,30% | | | | | |

Tabela 7 - Médias anuais de PM₁₀ das estações de monitoramento da qualidade do ar e as previsões do modelo de efeitos mistos nas regiões metropolitanas do Brasil

| Regiões Metropolitanas | Ano | Médias anuais do monitoramento de PM ₁₀ ($\mu\text{g}/\text{m}^3$) | Previsões de PM ₁₀ do modelo de treinamento ($\mu\text{g}/\text{m}^3$) |
|------------------------|------|---|---|
| São Paulo | 2009 | 32,6 | 35,7 |
| Grande Vitória | 2009 | 29,7 | 30,6 |
| Grande Vitória | 2013 | 28,2 | 26,3 |
| Baixada Santista | 2002 | 48,2 | 57,3 |
| Baixada Santista | 2003 | 53,8 | 53,9 |
| Baixada Santista | 2010 | 49,7 | 49,5 |
| Baixada Santista | 2011 | 53,8 | 48,5 |
| Rio de Janeiro | 2009 | 57,5 | 55,8 |
| Rio de Janeiro | 2010 | 60,7 | 57,0 |

Elaboração própria

As previsões das médias anuais de concentração de PM₁₀ existentes nas 36 regiões metropolitanas no período de 2001 a 2014 estão apresentadas na tabela 8.

Tabela 8 - Médias anuais dos níveis de concentração de PM₁₀ das regiões metropolitanas, estimadas pelo modelo de efeitos mistos no período de 2001 - 2014

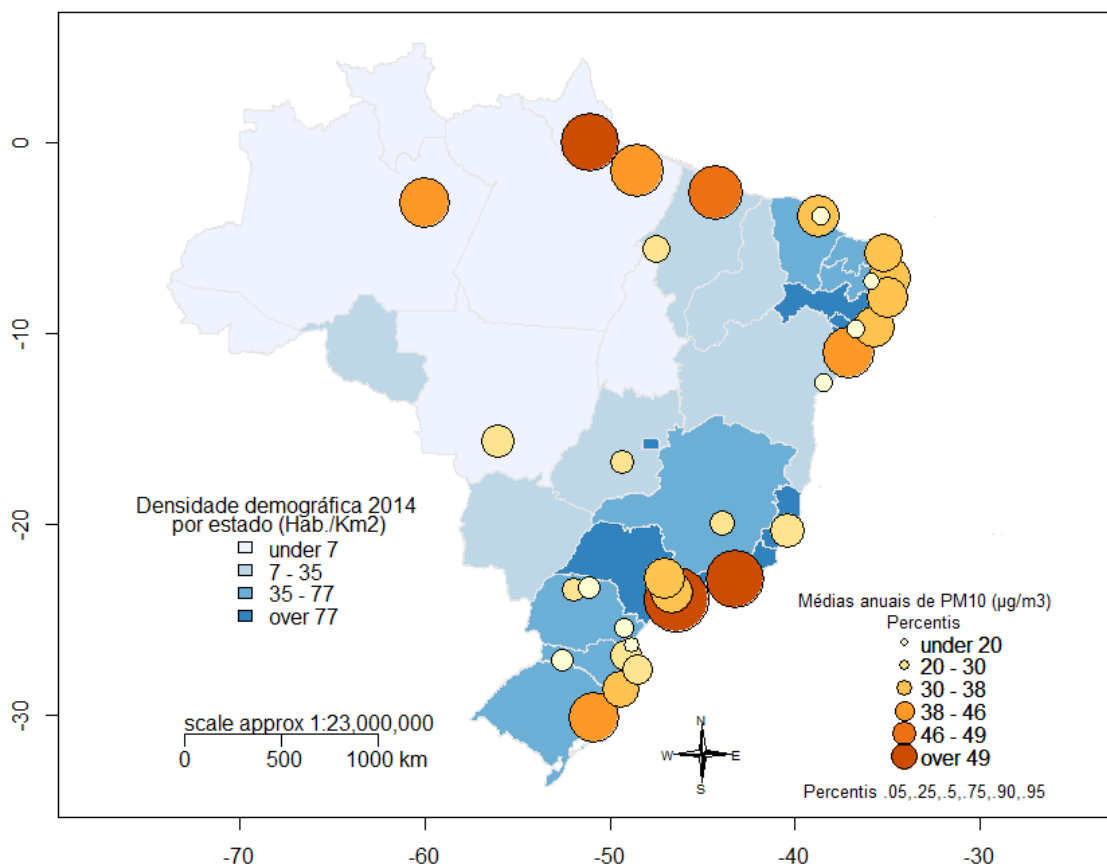
| Regiões metropolitanas | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Agreste | 18,5 | 19,7 | 20,8 | 19,8 | 20,0 | 19,7 | 19,4 | 19,6 | 19,7 | 19,5 | 19,2 | 19,5 | 16,5 | 16,1 |
| Aracaju | - | 56,3 | 55,4 | 52,9 | 52,6 | 50,3 | | 49,4 | 51,1 | 50,3 | 49,4 | 50,0 | 47,1 | 42,7 |
| Baixada Santista | 51,4 | 48,2 | 53,8 | 50,0 | 50,8 | 53,3 | 55,8 | 51,9 | 42,8 | 49,7 | 53,8 | 50,8 | 52,8 | 54,5 |
| Belém | - | 48,8 | 51,2 | 50,5 | 47,2 | 49,0 | 48,5 | 49,2 | 49,9 | 48,3 | 47,4 | 46,5 | 45,1 | 44,5 |
| Belo Horizonte | 21,8 | 21,7 | 22,5 | 20,8 | 19,8 | 20,0 | 22,5 | 20,4 | 19,6 | 21,7 | 21,4 | 20,4 | 18,9 | 20,8 |
| Campina Grande | 17,1 | 17,1 | 17,5 | 17,2 | 16,8 | 17,3 | 17,1 | 16,9 | 17,1 | 16,5 | 15,8 | 16,5 | 14,1 | 14,4 |
| Campinas | 42,0 | 43,5 | 40,0 | 33,9 | 32,0 | 37,8 | 38,7 | 34,0 | 28,5 | 34,0 | 34,0 | 31,5 | 30,5 | 34,0 |
| Carbonífera | | 36,5 | 37,5 | 36,2 | 35,3 | 35,5 | 36,0 | 35,3 | 35,7 | 33,5 | 33,1 | 35,2 | 31,9 | 30,8 |
| Cariri | 23,3 | 21,4 | 21,7 | 19,1 | 19,6 | 18,2 | 17,0 | 17,0 | 16,1 | 17,2 | 15,3 | 17,6 | 15,7 | 15,9 |
| Chapecó | 22,5 | 22,7 | 23,9 | 22,9 | 22,3 | 22,0 | 22,2 | 22,3 | 21,1 | 21,5 | 20,4 | 20,5 | 19,7 | 19,4 |
| Curitiba | 19,6 | 19,8 | 21,1 | 19,6 | 18,9 | 19,6 | 20,6 | 19,6 | 19,1 | 19,1 | 18,9 | 18,4 | 17,6 | 17,3 |
| Florianópolis | 28,8 | 27,5 | 31,0 | 29,2 | 29,0 | 29,3 | 29,8 | 29,1 | 29,1 | 28,3 | 27,7 | 27,2 | 26,0 | 26,4 |
| Fortaleza | 43,2 | 40,1 | 40,4 | 39,3 | 40,2 | 37,9 | 39,9 | 39,0 | 37,7 | 41,0 | 38,3 | 39,0 | 35,5 | 36,0 |
| Foz do Rio Itajaí | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Goiânia | 22,1 | 23,9 | 22,9 | 22,2 | 21,1 | 21,0 | 23,7 | 21,4 | 20,1 | 22,6 | 21,2 | 20,8 | 19,4 | 20,1 |
| Grande São Luís | 56,2 | 57,8 | 56,9 | 56,9 | 56,7 | 53,7 | 51,8 | 54,2 | 54,0 | 52,8 | 50,8 | 53,1 | 47,1 | 46,1 |
| Grande Vitória | 27,6 | 27,4 | 30,0 | 28,7 | 29,6 | 31,5 | 29,6 | 32,8 | 29,7 | 30,4 | 30,6 | 29,9 | 28,2 | 29,1 |
| João Pessoa | 41,6 | 42,4 | 42,9 | 41,1 | 41,7 | 41,8 | 40,9 | 41,6 | 41,7 | 40,4 | 39,2 | 39,3 | 34,3 | 35,9 |
| Lages | 15,6 | 15,2 | 16,4 | 15,4 | 14,8 | 15,4 | 15,7 | 15,0 | 15,2 | 15,1 | 14,7 | 14,8 | 13,4 | 13,4 |
| Londrina | 22,3 | 22,7 | 23,1 | 21,8 | 21,0 | 20,8 | 22,0 | 20,1 | 17,8 | 18,3 | 21,1 | 20,6 | 19,4 | 19,4 |
| Macapá | 70,1 | 68,1 | 65,8 | 65,7 | 61,1 | 61,4 | 63,1 | 62,7 | 63,5 | 59,3 | 60,1 | 60,0 | 49,9 | 49,1 |
| Maceió | 42,7 | 42,7 | 43,4 | 43,4 | 43,1 | 42,0 | 35,1 | 42,3 | 41,1 | 39,4 | 39,2 | 38,8 | 35,3 | 34,4 |
| Manaus | 49,5 | 49,4 | 47,6 | 46,5 | 46,6 | 44,9 | 45,9 | 44,0 | 49,2 | 46,7 | 46,1 | 45,2 | 42,1 | 41,8 |
| Maringá | 24,0 | 24,7 | 25,3 | 24,4 | 24,5 | 24,5 | 24,5 | 22,7 | 21,4 | 22,4 | 22,7 | 22,2 | 21,0 | 20,3 |
| Natal | 38,0 | | 39,2 | 37,3 | 37,2 | 37,3 | 36,4 | 35,5 | 36,4 | 36,8 | 34,5 | 34,5 | 31,5 | 31,8 |
| Norte/Nordeste | | | | | | | | | | | | | | |
| Catarinense | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Porto Alegre | 51,2 | 50,2 | 49,8 | 48,7 | 49,2 | 46,6 | 48,4 | 47,6 | 46,7 | 46,5 | 45,2 | 49,3 | 44,4 | 42,5 |
| Recife | 41,1 | 40,2 | 40,7 | 38,6 | 40,0 | 39,1 | 37,3 | 37,3 | 37,6 | 38,1 | 37,6 | 38,5 | 35,4 | 34,1 |
| Rio de Janeiro | 74,4 | 73,8 | 63,7 | 58,0 | 54,9 | 60,9 | 68,0 | 57,3 | 57,5 | 60,7 | 60,9 | 57,6 | 47,6 | 48,6 |
| Salvador | 19,3 | 19,3 | 19,3 | 20,2 | 20,8 | 16,2 | 18,0 | 19,9 | 18,6 | 21,4 | 22,6 | 23,2 | 18,0 | 15,5 |
| São Paulo | 47,9 | 47,3 | 45,4 | 39,7 | 40,2 | 38,0 | 41,3 | 35,6 | 32,6 | 37,7 | 38,0 | 36,0 | 33,9 | 35,5 |
| Sudoeste Maranhense | 27,5 | 28,2 | 27,5 | 27,3 | 28,6 | 26,2 | 26,2 | 24,9 | 24,9 | 25,1 | 22,2 | 23,2 | 22,8 | 23,1 |
| Tubarão | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Vale do Aço | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Vale do Itajaí | 28,9 | 29,1 | 29,6 | 27,8 | 26,7 | 28,1 | 27,3 | 26,8 | 26,2 | 27,1 | 27,6 | 27,4 | 25,7 | 26,9 |
| Vale do Rio Cuiabá | 33,2 | 32,8 | 31,7 | 32,8 | 33,6 | 29,4 | 33,0 | 30,3 | 29,7 | 32,7 | 30,9 | 29,9 | 28,0 | 27,7 |
| Brasil | 35,2 | 36,1 | 36,2 | 34,6 | 34,2 | 34,0 | 34,1 | 33,6 | 32,9 | 33,6 | 33,1 | 33,0 | 30,3 | 30,3 |

Não estavam disponíveis os dados climáticos para as regiões metropolitanas Foz do Rio Itajaí, Norte/Nordeste Catarinense, Tubarão e Vale, portanto não foram calculados valores previstos da poluição do ar.

Elaboração própria

Referindo-se a um determinado ano da série temporal da tabela 8, a figura 3 apresenta um mapa ilustrativo das previsões das médias anuais de PM_{10} em 2014, resultantes do modelo de efeitos mistos, para as 36 regiões metropolitanas. A amplitude das médias anuais entre as regiões variou de 13 a $54 \mu\text{g}/\text{m}^3$ com média global de $30,3 \mu\text{g}/\text{m}^3$ e desvio padrão de $11,74 \mu\text{g}/\text{m}^3$. Sob a perspectiva do padrão de qualidade do ar no Brasil, com base nos percentis da distribuição das concentrações de PM_{10} , cerca de 95% das regiões metropolitanas apresentaram médias anuais abaixo ou iguais a $49 \mu\text{g}/\text{m}^3$ e, portanto, estiveram dentro dos níveis aceitáveis estabelecidos pelo CONAMA que é de $50 \mu\text{g}/\text{m}^3$, exceto a Baixada Santista, que demonstrou uma média anual de $54 \mu\text{g}/\text{m}^3$. Entretanto, o quadro muda na perspectiva dos padrões publicados pela OMS, que tem como limite a emissão anual de $20 \mu\text{g}/\text{m}^3$. Sob essa ótica, podemos observar que, no Brasil, apenas 25% das regiões metropolitanas tiveram dentro dos níveis aceitáveis de PM_{10} .

Figura 3 - Previsões das médias anuais de PM_{10} do modelo de efeitos mistos nas regiões metropolitanas do Brasil em 2014



Elaboração própria.

Discussão

O objetivo deste artigo foi propor uma estrutura de modelagem hierárquica para fazer previsões robustas para as médias anuais de PM_{10} ao longo de 14 anos. O modelo linear de efeito misto foi desenvolvido para capturar variabilidades inter e intrarregionais das medidas de poluentes particulados específicos em um único modelo. As variáveis locais e de tendência foram implementadas para reduzir as incertezas, minimizar a variação na predição e gerar previsões estáveis. Além da vantagem da flexibilidade com estrutura de dados desbalanceados, o modelo possui uma estrutura flexível, facilitando a incorporação de múltiplos preditores. Por exemplo, o modelo incorporou concentrações médias de longo prazo (intercepto), variável de tendência (funções de base temporal) e preditores de influência local tais como umidade relativa do ar, PIB per capita do setor serviços, altitude média e focos de queimadas. Embora o principal interesse na análise fosse o estudo dos efeitos fixos para previsão, a modelagem dos efeitos aleatórios foi fundamental. A introdução de efeitos aleatórios permitiu acomodar e modelar a variabilidade não explicada pelas variáveis do modelo, aumentando assim a precisão na estimação dos efeitos fixos. As extensões ao modelo linear de efeitos mistos permitiram combinar a introdução de funções de variância para caracterizar a heterocedasticidade e estruturas de correlação para representar a correlação serial entre os erros, e assim modelar de forma flexível a estrutura de variâncias-covariâncias dentro do grupo. O modelo obtido para as previsões revelou bom ajuste, robustez e eficiência. Os coeficientes apresentaram estabilidade, sinais esperados e magnitudes razoáveis. Na validação cruzada o modelo construído através da amostra de treinamento apresentou bom desempenho preditivo.

Vale ressaltar que, neste estudo, apesar do PM_{10} está associado à umidade relativa do ar, outras variáveis climáticas importantes como nebulosidade, pressão atmosférica, índice pluviométrico e velocidade do vento não fizeram parte da análise multivariada por falta de informações no período de estudo. Variáveis explicativas como média anual de geração de energia das usinas termelétricas (MWh), total de rodovias e total de portos também não foram consideradas para compor o modelo, por não informações disponíveis. Contudo, uma das vantagens do modelo é a possibilidade de novas variáveis, quando disponíveis, serem utilizadas na expansão do modelo de previsão para contínuas melhorias.

Foi observado também que devido à heterogeneidade local, o modelo pode tender a subestimar as estimativas, proporcionando previsões pouco realistas para regiões metropolitanas mais dispersas. Como modelo é baseado em muitos fatores (socioeconômicos,

meteorológicos, demográficos, etc..) e são calculados em termos de média para toda a população. É de conhecimento comum haver diferenças significativas desses fatores nessas áreas, afetando, portanto, suas previsões. Isto é, o modelo fornece estimativas mais robustas em regiões urbanas onde as características locais sejam mais semelhantes ou próximas.

Conclusão

A variável PIB per capita do setor serviços umidade, altitude média e focos de queimadas foram associados aos níveis de concentração de PM₁₀. Além da boa performance preditiva, o modelo linear de efeitos mistos apresentou bom ajuste, robustez e eficiência para realizar previsões.

Suporte de informação

Informações adicionais de suporte podem ser encontradas nos apêndices deste artigo.

Apêndices:

APÊNDICE A - Estruturação das redes de monitoramento de PM₁₀ nas regiões metropolitanas do Brasil

APÊNDICE B - Critérios de validação dos dados das redes de monitoramento

APÊNDICE C - Mapas dos fatores contextuais das regiões metropolitanas do Brasil – 2014

APÊNDICE D – Mapas ilustrativos da distribuição das estações de monitoramento de PM₁₀ nas Regiões metropolitanas do Brasil

APÊNDICE E – Gráficos ilustrativos da evolução dos fatores contextuais nas Regiões metropolitanas brasileiras durante o período 2001 a 2014

APÊNDICE F – Correlações entre fontes de dados dos parâmetros meteorológicos

APÊNDICE G – Sumário de dados faltantes, correlações entre as estações das médias anuais de PM₁₀ e estatística descritiva após a imputação de dados

APÊNDICE H – Modelagem e estimativas dos níveis de concentração de PM10 das regiões metropolitanas do Brasil no período de 2001 a 2014

Referências

BOX, G. E.; COX, D. R. An analysis of transformations. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 211-252, 1964. ISSN 0035-9246.

BOX, G. E. et al. **Time series analysis: forecasting and control**. John Wiley & Sons, 2015. ISBN 1118675029.

BRAUER, M. et al. Exposure Assessment for Estimation of the Global Burden of Disease Attributable to Outdoor Air Pollution. **Environmental Science & Technology**, v. 46, n. 2, p. 652-660, 2012/01/17 2012. ISSN 0013-936X. Disponível em: < <http://dx.doi.org/10.1021/es2025752> >.

COHEN, A. J. et al. Urban air pollution. **Comparative quantification of health risks**, v. 2, p. 1353-1433, 2004.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the royal statistical society. Series B (methodological)**, p. 1-38, 1977. ISSN 0035-9246.

HOEK, G. et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. **Atmospheric Environment**, v. 42, n. 33, p. 7561-7578, 10// 2008. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231008005748> >.

JERRETT, M. et al. A review and evaluation of intraurban air pollution exposure models. **J Expo Anal Environ Epidemiol**, v. 15, n. 2, p. 185-204, Mar 2005. ISSN 1053-4245 (Print) 1053-4245.

JUNGER, W. L.; PONCE DE LEON, A. Imputation of missing data in time series for air pollutants. **Atmospheric Environment**, v. 102, p. 96-104, 2// 2015. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231014009145> >.

KRZYZANOWSK, M.; SCHWELA, D. 7 - Patterns of Air Pollution in Developing Countries A2 - Holgate, Stephen T. In: SAMET, J. M.; KOREN, H. S., et al (Ed.). **Air Pollution and Health**. London: Academic Press, 1999. p.105-113. ISBN 978-0-12-352335-8.

NAKAGAWA, S.; SCHIELZETH, H. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. **Methods in Ecology and Evolution**, v. 4, n. 2, p. 133-142, 2013. ISSN 2041-210X.

NORMAN, R. et al. Estimating the burden of disease attributable to urban outdoor air pollution in South Africa in 2000. **S Afr Med J**, v. 97, n. 8 Pt 2, p. 782-90, Aug 2007. ISSN 0256-9574 (Print).

PANDEY, K. D. et al. Ambient particulate matter concentrations in residential and pollution hotspot areas of world cities: new estimates based on the Global Model of Ambient Particulates (GMAPS). **Washington, DC: The World Bank**, 2006.

PINHEIRO, J.; BATES, D. **Mixed-effects models in S and S-PLUS**. Springer, 2000. ISBN 038798957-9.

SNIJDERS, T.; BOSKER, R. **Multilevel analysis: An introduction to basic and applied multilevel analysis**: London: Sage 1999.

5.2 Avaliação do impacto da poluição do ar na mortalidade em regiões metropolitanas do Brasil (manuscrito 2)

Evaluation of the impact of air pollution on mortality in metropolitan regions of Brazil

Simone de Souza Cardoso

Antônio Ponce de Leon

Resumo

Conhecimento: A exposição à poluição do ar aumenta a morbimortalidade por doenças respiratórias e cardiovasculares contribuindo, portanto, para a carga global de doenças. Impactos significativos são esperados nas grandes cidades onde estão concentradas as maiores fontes de emissão da poluição do ar. Sendo assim, para a saúde pública, o conhecimento da carga global de doenças em áreas urbanas é importante para controle de poluentes atmosféricos específicos e no debate sobre as novas diretrizes da qualidade do ar.

Objetivos: Estimar e avaliar o impacto da exposição prolongada da poluição do ar na mortalidade, restrito às regiões metropolitanas do Brasil, utilizando níveis de concentração de material particulado inalável com diâmetro $\leq 10\mu\text{m}$ (PM₁₀), derivados do modelo de efeitos mistos e níveis de concentração de material particulado fino com diâmetro $\leq 2,5\mu\text{m}$ (PM_{2,5}).

Método: O estudo seguiu uma metodologia de avaliação de riscos comparativos desenvolvida pela Organização Mundial da Saúde (OMS). Os riscos relativos para mortalidade foram calculados empregando estimativas de risco derivada do modelo de Cox na associação entre mortalidade e PM_{2,5}, apresentado no estudo de coorte da Associação Americana de Câncer e da Metanálise dos estudos de associação entre mortalidade e PM₁₀, extraídas do estudo meta-analítico do projeto - Estudo de Saúde e Poluição Atmosférica na América Latina.

Área e ano de estudo: Regiões metropolitanas do Brasil, 2014.

Desfechos: Mortalidade e *Years of Life Lost* (YLL) por doenças respiratórias em crianças < 1 ano e 1 a 5 anos, e por doenças cardiovasculares em adultos com idade ≥ 30 anos.

Resultados: A média global do material particulado nas regiões metropolitanas do Brasil, em 2014, foi estimada em 34,2 $\mu\text{g}/\text{m}^3$ para PM₁₀ e 17,1 $\mu\text{g}/\text{m}^3$ para PM_{2,5}. O impacto da poluição do ar nas regiões metropolitanas para doenças cardiovasculares em adultos ≥ 30 anos foi 12.615 óbitos (IC95%: 4.673 - 20.126) e 465.727 YLL (IC95%: 172.500 - 743.014) e para doenças respiratórias foi 7 óbitos (IC95%: 0 - 13) e 1.286 YLL (IC95%: 25-2.313) em crianças menores

de 1 ano e apenas 5 óbitos (IC95%: 1-9) e 798 YLL (IC95%: 111-1.475) para crianças de 1 a 5 anos.

Conclusão: O estudo demonstrou a possibilidade em utilizar estimativas de concentrações de material particulado, provenientes do modelo de efeitos mistos, na avaliação global do risco de mortalidade atribuível à poluição atmosférica. Além disso, evidenciou que o material particulado é um contribuinte importante nas variações das taxas de mortalidade em áreas urbanas do Brasil, independente das incertezas inerentes às análises globais de mortalidade.

Palavras-chave: Material particulado. Modelos de efeitos mistos. Mortalidade. Carga global de doença.

Abstract

Background: Exposure to air pollution increases morbidity and mortality for respiratory and cardiovascular diseases, thus contributing to the global burden of disease. Significant impacts are expected in great cities where the largest sources of air pollution are concentrated. Thus, for public health, knowledge of the global burden diseases in urban areas is important the control of specific air pollutants and the debate on new air quality guidelines.

Objective: To estimate and evaluate the impacts of prolonged air pollution exposure on mortality, restricted to metropolitan areas of Brazil, using concentration levels of inhalable particulate matter with diameter $\leq 10\mu\text{m}$ (PM₁₀), derived from mixed effects models and concentration levels of fine particulate matter with diameter $\leq 2.5\mu\text{m}$ (PM_{2.5}).

Method: The study followed a comparative risk assessment methodology developed by the World Health Organization (WHO). The relative risks for mortality were calculated using risk estimates derived from Cox model in association between mortality and PM_{2.5}, presented in the American Cancer Association cohort study and Metanalysis study of association between mortality and PM₁₀, extracted from meta-analytical study of the project - Multicity Study of Air Pollution and Mortality in Latin America (the ESCALA Study).

Area and year of study: Metropolitan regions of Brazil, 2014.

Outcomes: Mortality and Years of Life Lost (YLL) for respiratory diseases in children under 1 year and from one to five years, and for cardiovascular diseases in adults aged 30 years and over.

Results: The global average of particulate matter in metropolitan regions of Brazil, in 2014, was estimated at 34.2 $\mu\text{g}/\text{m}^3$ for PM_{10} and 17.1 $\mu\text{g}/\text{m}^3$ for $\text{PM}_{2.5}$. The impact of air pollution in metropolitan regions for cardiovascular diseases in adults from 30 years and over was 12,615 deaths (CI95%: 4,673 – 20,126) and 465,727 YLL (CI95%: 172,500 – 743,014) and for respiratory disease 7 deaths (CI95%: 0 - 13) and 1.286 YLL (CI95%: 25-2,313) in children under one-year e only 5 deaths (CI95%: 1-9) and 798 YLL (CI95%: 111-1,475) for children from one to five years.

Conclusion: The study demonstrated the possibility of using estimates of concentration of particulate matter, from mixed effects models in the global assessment risk of mortality attributable to air pollution. Besides that, showed that particulate matter is an important contributor to changes in mortality rates in urban areas of Brazil, regardless of the uncertainties inherent in the overall mortality analysis.

Keywords: Particulate matter. Mixed effects models. Mortality. Global burden of disease.

Introdução

Elevados níveis de concentração de poluentes causam aumentos na mortalidade, nos atendimentos ambulatoriais e em admissões hospitalares, já comprovados por estudos realizados nos EUA e na Europa em meados do século XX, quando houve eventos maiores de poluição do ar seguidos por aumentos significativos na mortalidade (Logan, 1953; Schwartz, 1994; Brunekreef e Holgate, 2002). Contudo, estudos mais recentes têm demonstrado que os efeitos da poluição sobre a saúde estão presentes mesmo em níveis mais baixos, principalmente no que se refere às exposições ao Material Particulado (MP) (Ostro, 1984; Pope *et al.*, 1992; Dockery e Pope, 1994; Schwartz, 1995; Brunekreef, 1997; Mage *et al.*, 1999; Brunekreef e Holgate, 2002), dando indícios de que não existe um limiar de efeito zero para as partículas de poluição e que os riscos para a saúde podem estar presentes em qualquer nível de exposição.

Grande parte das pesquisas em séries temporais epidemiológicas sobre os efeitos das exposições agudas à poluição atmosférica nos grandes centros urbanos têm se concentrado em doenças respiratórias e cardiovasculares, encontrando evidências consistentes à premissa de que o MP é um fator de risco importante para essas doenças (Dockery e Pope, 1994; Schwartz, 1994; Lin *et al.*, 1999; Pope e Dockery, 1999; Dominici *et al.*, 2000; Samet *et al.*, 2000; Katsouyanni *et al.*, 2001; Dominici *et al.*, 2002; Dominici *et al.*, 2004; Wong *et al.*, 2008;

Romieu *et al.*, 2012). No entanto exposições de longo prazo à poluição atmosférica também afeta as taxas de mortalidade no modo de indução de doenças crônicas (Pope e Dockery, 1999; Zeger *et al.*, 1999; Schwartz, 2000; 2001; Zanobetti *et al.*, 2002; Dominici *et al.*, 2003). Essa é uma questão muito importante para saúde pública no que se refere ao controle de poluentes atmosféricos específicos e ao debate sobre as novas diretrizes da qualidade do ar para a manutenção da saúde (Kunzli *et al.*, 2000).

Em 2016, a Organização Mundial da Saúde (OMS) lançou um relatório sobre a avaliação da poluição do ar e da carga global de doenças em cidades do mundo, para o ano de 2012. Mundialmente, foi estimado cerca de 3 milhões de mortes atribuídas ao MP. Dentro deste total, o MP foi responsável por 2 milhões de mortes por doenças cardiovasculares e 169 mil mortes por doenças respiratórias em crianças de 0-4 anos. Para o Brasil, foi estimado cerca de 26 mil mortes atribuídas ao MP, sendo aproximadamente 23 mil mortes por doença isquêmica do coração e acidente vascular cerebral em adultos de 25 anos ou mais; e 198 mortes por doenças respiratórias inferiores agudas em crianças de 0-4 anos (WHO,2016).

Cohen *et al.* (2004) fizeram uma avaliação do risco comparativo incluindo uma descrição detalhada dos métodos e resultados encontrados no estudo “*Urban air Pollution*”, publicado pela OMS. Neste estudo, a poluição do ar foi estimada apenas em capitais nacionais e áreas urbanas com populações > 100.000 hab. de todo o mundo em 2000. Restrito às áreas urbanas foi necessário o uso de um modelo de efeitos mistos, o “*Global Model of Ambient Particulates*” (GMAPS), para estimar as concentrações ambientais de partículas. A utilização desse modelo proporcionou uma abordagem apropriada para desenvolver estimativas globais de concentrações de MP inalável, com diâmetro $\leq 10\mu\text{m}$ (PM₁₀), e fino, com diâmetro $\leq 2,5\mu\text{m}$ (PM_{2,5}), para a análise da carga de mortes atribuível. Cerca de 3% da mortalidade atribuível ao MP foi por doenças cardiovasculares em adultos, 5% por câncer de traqueia, brônquios e pulmão e 1% por infecções respiratórias agudas em crianças de 0-4 anos. Isso equivaleu a cerca de 0,80 milhões de óbitos prematuros (1,4% do total global) e 6,4 milhões YLL (*Years of Life Lost*) (0,7% do total global).

Baseado na avaliação do risco comparativo de Cohen *et al.* (2004), este trabalho teve como objetivo estimar e avaliar a carga de mortalidade atribuída ao MP para doenças cardiovasculares, em adultos com idade ≥ 30 anos, e respiratória, em crianças menores de 1 ano e de 1-5 anos, nas regiões metropolitanas do Brasil para o ano de 2014. A avaliação foi com base nas estimativas atuais das concentrações médias de PM₁₀ do modelo de efeitos mistos (modelo econométrico), restritas às áreas urbanas. Apenas as estimativas de mortalidade

atribuíveis foram incluídas nesta análise, uma vez que os efeitos da mortalidade da poluição atmosférica são os mais importantes e mais susceptíveis à avaliação global.

Método

Dados

Estimativas das médias anuais do material particulado (PM₁₀)

As informações sobre as estimativas das médias anuais de PM₁₀ basearam-se nos resultados do artigo 1 em que as estimativas de 2014, derivadas do modelo linear de efeitos mistos, foram utilizadas para os cálculos da carga de mortalidade e dos óbitos atribuíveis restritos às regiões metropolitanas do Brasil neste trabalho.

No artigo 1, os dados horários, diários e mensais de PM₁₀ dos locais fixos de monitoramento da qualidade do ar nas regiões metropolitanas foram obtidos pelos órgãos ambientais competentes. Antes da construção das médias anuais de PM₁₀, foi seguido um protocolo padrão desses órgãos ambientais sobre a validação dos dados de monitoramento. O método de imputação dos dados foi realizado no caso de dados faltantes nas estações de monitoramento, aplicando o algoritmo EM (*Expectation – Maximization*) para dados de séries temporais. Após esses procedimentos, as médias anuais foram calculadas para cada região metropolitana. Em seguida, através do modelo linear de efeitos mistos, foram estimadas médias anuais de PM₁₀ para as regiões metropolitanas que não possuíam informações sobre o poluente. No modelo as médias anuais de PM₁₀ foram explicadas por variáveis como consumo de energia, condições climáticas, demográficos, socioeconômicos, transportes e outros. A modelagem foi realizada para determinar quais as variáveis estiveram, de fato, associadas às concentrações de PM₁₀. Como resultado foram umidade relativa do ar, PIB do setor serviços, altitude média e focos de queimadas. O modelo mostrou a correlação entre os dados medidos de PM₁₀ e os ajustados, aproximadamente de 0,98 (correlação de Pearson), o RMSE = 0,07 µg/m³, o MAE = 0,052 e o coeficiente de determinação (R²) igual a 0,98. Na validação para fora da amostra o desempenho preditivo foi de 70% ($\rho_{\text{Spearman}} = 0,70$), o RMSE de 0,22 µg/m³ e o MAE de 2,89.

Dados de mortalidade

A exposição à poluição atmosférica está associada aos efeitos agudos e crônicos na saúde. Os efeitos de saúde associados ao MP incluem câncer de pulmão, doenças respiratórias e cardiovasculares. Apenas dois desses desfechos de saúde foram incluídos neste estudo: doenças respiratórias e doenças cardiovasculares. Foram obtidos dados de mortalidade específica: Cardiovascular (I00–I99) em adultos com idade ≥ 30 anos e respiratória (J00–J99) em crianças < 1 e 1–5 anos, todos compreendidos nos Capítulos IX e X da Décima Revisão da Classificação Internacional de Doenças e Problemas Relacionados à Saúde (CID-10). Todos obtidos por meio do Sistema de Informações de Mortalidade – SIM (<http://datasus.saude.gov.br/informacoes-de-saude/tabnet/estatisticas-vitais>).

Estimação da carga de mortalidade e dos óbitos atribuíveis ao material particulado

Nossa metodologia baseou-se na avaliação do GBD de 2000 da WHO para a poluição do ar em áreas urbanas em Cohen *et al.* (2004).

Em um estudo seccional, foram desenvolvidas funções de concentração-resposta (C-R) para dois desfechos de saúde, afim de produzir estimativas de risco para os cálculos da carga de mortalidade e dos óbitos por doenças cardiovasculares em adultos de 30 anos ou mais e respiratória em crianças de <1 e 1-5 anos devido à poluição atmosférica para o ano de 2014, ambos os sexos.

A estimativa do RR para doenças cardiovasculares em adultos ≥ 30 anos foi baseada nos resultados do “*American Cancer Society*” (ACS).

Em 1982, a ACS iniciou um estudo de coorte prospectivo e 1,2 milhão de adultos com idade igual ou superior a 30 anos participaram do estudo. Eram membros de famílias com pelo menos um indivíduo com 45 anos de idade ou mais. Os participantes completaram um questionário confidencial que incluiu informações demográficas, tabagismo, uso de álcool, dieta e educação. As associações entre $PM_{2,5}$ e mortalidade foram relatadas em três períodos de exposição ao $PM_{2,5}$: O primeiro deles foi o intervalo 1979-1983 que referiu-se ao seguimento de 7 anos (1982-1989) (Pope *et al.*, 1995), o segundo foi o intervalo 1999-2000 que referiu-se ao seguimento de 16 anos (1982-1998) e a média dos dois períodos de exposição que também

referiu-se ao acompanhamento de 16 anos (1982-1998)(Pope *et al.*, 2002). O risco relativo da mortalidade associado ao PM_{2,5} foi estimado pelo modelo de risco proporcional de Cox de efeitos aleatórios, controlado pela idade, sexo, raça, tabagismo, educação, estado civil, massa corporal, consumo de álcool, exposição ocupacional e dieta.

Como caso base para o nosso estudo, consideramos a média dos dois períodos de exposição e a estimativa de risco associado a esse período do estudo da ACS.

As estimativas de PM₁₀, derivadas do modelo de efeitos mistos, foram convertidas para PM_{2,5} no valor de 0,5 da razão PM_{2,5}/PM₁₀.

Como a nossa média global do PM_{2,5} estimada para as áreas urbanas não ultrapassou de 30 µg/m³ e para obter um modelo mais realista, além de ter o mínimo de erro de medição, consideramos a função C-R linear que é representada pela seguinte fórmula:

$$RR_{C-C_0} = \frac{\exp(\alpha + \beta C)}{\exp(\alpha + \beta C_0)} = \exp[\beta(C - C_0)]$$

Onde RR_{C-C_0} é o RR da exposição média dos níveis de concentração C (ponderada pela população) comparada à exposição de referência C_0 , - média anual contrafactual (ou mínimo teórico) considerada igual a 7,5 µg/m³, a menor concentração observada no estudo da ACS - e β é o efeito estimado do PM_{2,5} sobre o desfecho de saúde de interesse.

Ao extrapolar a função C-R linear para além de 30 µg/m³ de PM_{2,5}, sugeriu-se que o uso de uma função linear poderia ser inadequado, pois sobrestimaria a mortalidade (Pope *et al.*, 1995; Pope *et al.*, 2002; Cohen *et al.*, 2004; Ostro e Who, 2004). Assim, o uso de uma função de exposição, log-linear, seria recomendada, pois sua inclinação se achataria nas mais altas concentrações de PM_{2,5} (Cohen *et al.*, 2004; Ostro e Who, 2004).

Para análises referentes às doenças respiratórias em crianças <1 e 1-5 anos, utilizamos a estimativa global de PM₁₀. Foi empregado a função C-R linear para os cálculos dos RR, considerando a média anual contrafactual (ou mínimo teórico) de 15 µg/m³ e as estimativas de risco da metanálise dos efeitos fixos do estudo de Gouveia *et al.* (2018), que é uma extensão do projeto ESCALA (“*Estudio de Salud y Contaminación del Aire en Latino América*”) e teve como objetivo avaliar o efeito da poluição do ar sobre a mortalidade respiratória infantil em quatro grandes centros urbanos: Cidade do México, Santiago, São Paulo e Rio de Janeiro. A metanálise, utilizada como método, combina resultados específicos das cidades em uma única estimativa de resumo. Em cada cidade e faixas etárias as estimativas das variações percentuais

do risco foram estimadas empregando modelos aditivos generalizados Poisson, controlados por tendência, sazonalidade, temperatura e umidade relativa, para ajustar a mortalidade diária por doenças respiratórias e PM₁₀.

Na tabela 9 (abaixo) está o resumo do risco relativo de mortalidade, das inclinações da função C-R e os tipos de estudo que foram utilizados para o cálculo da carga de mortalidade e dos óbitos atribuíveis.

Tabela 9 - Estimativas do RR de mortalidade, das inclinações da função C-R e os tipos de estudo recomendados para o cálculo da carga de mortalidade e dos óbitos atribuíveis

| Desfechos de saúde | Fontes de dados | MP | Funções do RR | RR por 10 µg/m ³ (95% IC) | Inclinações das funções C-R por 10 µg/m ³ (95% IC) |
|---|--|-------------------|----------------------------------|--------------------------------------|---|
| Mortalidade por doenças cardiovasculares em adultos >=30 anos | Estudo da ACS (Pope et.al. 2002) | PM _{2,5} | RR=exp [β (X - X ₀)] | 1,09 (1,03 - 1,16) | Linear - média: 0,008933 ^a (0,00322 - 0,01464) |
| Mortalidade por doenças respiratórias em crianças < 1 ano | Estudo de Gouveia <i>et al.</i> (2018). Metanálise de efeitos fixos de 4 estudos de séries temporais sobre mortalidade diária | PM ₁₀ | RR=exp [β (X - X ₀)] | 1,0047 (1,0009 - 1,0085) | Linear 0,000469 ^b (0,00009 - 0,000846) |
| Mortalidade por doenças respiratórias em crianças de 1 a 5 anos | Estudo de Gouveia <i>et al.</i> (2018). Metanálise de efeitos fixos de 4 estudos de séries temporais sobre mortalidade diária | PM ₁₀ | RR=exp [β (X - X ₀)] | 1,0058 (1,0008 - 1,0108) | Linear 0,000578 ^b (0,00008 - 0,0010742) |

Tabela adaptada de Cohen *et al.* (2004)

^a O resultado proveniente do modelo de regressão especificado como média das concentrações de PM_{2,5} medidas no período 1979 - 1983 e 1999 - 2000 (Pope *et al.*, 2002; Cohen *et al.*, 2004).

^b Na metanálise, os autores consideraram mais apropriada para a associação entre mortalidade e PM₁₀ as estimativas das variações percentuais do risco do modelo de efeitos fixos.

As frações atribuíveis (FA) para mortalidade cardiovasculares e respiratórias foram calculadas utilizando a seguinte fórmula (Ostro e Who, 2004):

$$FA = \frac{RR_{C-C_0} - 1}{RR_{C-C_0}}$$

A carga atribuível foi calculada multiplicando o desfecho por causa específica pela fração atribuível, como segue abaixo:

$$CA = FA \times \text{desfecho de saúde}$$

Onde CA é a carga atribuível, FA é a fração atribuível e o desfecho de saúde de interesse, ou seja, óbitos e/ou YLL. Para cada desfecho e grupos de idade selecionados, os YLL foram calculados utilizando tábuas de mortalidade extraídas do site do Instituto Brasileiro de Geografia e Estatística (IBGE) e óbitos do SIM.

Não foi estimado o impacto do MP sobre a incidência da doença, e, portanto, os anos de vida ajustados por incapacidade (DALY) foram quantificados apenas com YLL.

Análise de sensibilidade

Uma vez que os efeitos da exposição de longo prazo de MP requerem vários pressupostos, exploramos a sensibilidade dos resultados para a mortalidade cardiovascular desses pressupostos. As análises de sensibilidade foram utilizadas para quantificar as incertezas das estimativas do caso base, em que a carga de mortalidade e os óbitos atribuíveis foram estimados através da aplicação da função C-R linear, considerando um mínimo contrafactual de 7,5 $\mu\text{g}/\text{m}^3$ para $\text{PM}_{2,5}$, conforme discutido anteriormente. Então foram explorados cenários alternativos como variação da função C-R (função de exposição linear versus não linear); da concentração contrafactual (7,5, 3,0 ou 10 $\mu\text{g}/\text{m}^3$); da proporção de $\text{PM}_{2,5}$ para PM_{10} (0,50 ou 0,65) e do período de exposição ao MP.

Resultados

A tabela 10 mostra as estimativas das concentrações de material particulado: média anual de PM₁₀, média anual de PM_{2,5}. - Resultante da conversão através razão PM_{2,5}/PM₁₀ - e o total da mortalidade por doenças cardiovasculares e respiratórias em regiões metropolitanas do Brasil em 2014.

Tabela 10 – Estimativa da média anual de PM₁₀ proveniente do modelo de efeitos mistos, média anual de PM_{2,5} e total de óbitos por doenças cardiovasculares e respiratórias em áreas metropolitanas do Brasil, 2014

| Áreas Metrop. | População Metrop. (100.000) | PM ₁₀ * Média anual (µg/m ³) | Razão PM _{2,5} /PM ₁₀ | PM _{2,5} * Média anual (µg/m ³) | Mínimo contrafactual | | Óbitos | | | |
|---------------|-----------------------------|---|---|--|----------------------|------------------|------------------|----------|---------------|--|
| | | | | | PM _{2,5} | PM ₁₀ | Cardiovasculares | | Respiratórias | |
| | | | | | | | adultos ≥30 anos | <1 ano | Crianças | |
| | | | | | | | | 1-5 anos | | |
| Brasil | 864,25 | 34,2 | 0,5 | 17,1 | 7,5 | 15 | 153.436 | 826 | 409 | |

*Média anual ponderada pela população
Elaboração própria

Avaliação da carga de mortalidade atribuível ao material particulado

Entre as regiões metropolitanas, os RR de mortes por doenças cardiovasculares em adultos ≥ 30 anos associados ao PM_{2,5} foram calculados de acordo com o cenário base. As regiões metropolitanas do Rio de Janeiro - 1,162 (IC95%: 1,056-1,279), de Macapá - 1,165 (IC95%: 1,057-1,284) e da Baixada Santista - 1,193 (IC95%: 1,066-1,335) apresentaram as maiores estimativas de RR, com frações atribuíveis de 14% (IC95%: 5%-22%), 14% (IC95%: 5%-22%) e 16% (IC95%: 6%-25%) respectivamente. As regiões metropolitanas que apresentaram os menores RR foram Salvador-1,002 (IC95%: 1,001-1,003), Cariri-1,005 (IC95%: 1,002-1,008) e Agreste-1,005 (IC95%: 1,002-1,008). Com frações atribuíveis de 0,2% (IC95%: 0,1%-0,3%), 0,4% (IC95%: 0,1%-0,7%) e 0,5% (IC95%: 0,2%-0,8%), respectivamente. Na tabela 11, o risco global das áreas metropolitanas no Brasil foi de 1,09 (IC95%: 1,03-1,15), e a fração da mortalidade por doenças cardiovasculares atribuível ao PM_{2,5} foi de 8,0% (IC95%: 3%-13%).

O mesmo comportamento aconteceu com os RR de mortes por doenças respiratórias em crianças menores de 1 ano e de 1 a 5 anos associados ao PM₁₀. Para as crianças menores de 1 ano, as regiões metropolitanas do Rio de Janeiro-1,016 (IC95%: 1-1,029), de Macapá-1,016 (IC95%: 1-1,029) e da Baixada Santista-1,019 (IC95%: 1-1,034) apresentaram as maiores estimativas de RR, com suas respectivas frações atribuíveis de 1,6% (IC95%: 0%-2,8%), 1,6% (IC95%: 0%-2,8%) e 1,8% (IC95%: 0%-3,3%). As regiões metropolitanas com os menores RR e frações atribuíveis, como Salvador, Cariri e Agreste, apresentaram valores muito próximo de um e frações atribuíveis praticamente nulas, concluindo, portanto, impacto ínfimo do PM₁₀ sobre a mortalidade por doenças respiratórias. Para as crianças de 1 a 5 anos, novamente, o Rio de Janeiro-1,02 (IC95%: 1,003-1,037), Macapá-1,02 (IC95%: 1,003-1,037) e Baixada Santista-1,023 (IC95%: 1,003-1,043) apresentaram as maiores estimativas do RR, com frações atribuíveis de 1,9% (IC95%: 0,3%-3,5%), 2,0% (IC95%: 0,3%-3,6%) e 2,3% (IC95%: 0,3%-4,2%), respectivamente. O impacto do PM₁₀ sobre a mortalidade em regiões metropolitanas como Salvador, Cariri e Agreste também foi muito pequeno pois apresentaram RR valores muito próximos de um e frações atribuíveis aproximadamente nulas. Na tabela 11, o risco global das áreas metropolitanas para crianças menores de 1 ano foi de 1,009 (IC95%: 1-1,016) com fração atribuível foi de 0,9% (IC95%: 0%-1,6%) e crianças de 1 a 5 anos o risco foi de 1,011 (IC95%: 1,002-1,021). Com fração atribuível de 1,1% (IC95%: 0,1%-1,9%).

Os resultados indicaram, portanto, que em 2014, o impacto da poluição atmosférica na mortalidade por doenças cardiovasculares e respiratórias foi maior em regiões metropolitanas como Rio de Janeiro e Baixada Santista e Macapá.

Tabela 11 - Risco relativo e fração atribuível da mortalidade global associados ao PM_{2,5} para doenças cardiovasculares e PM₁₀ para doenças respiratórias (IC 95%)

| Áreas Metrop. | RR | | | FA | | |
|---------------|---------------------|--------------------|------------------------|---------------------|--------------------|------------------------|
| | Cardiovasculares | Respiratórias | | Cardiovasculares | Respiratórias | |
| | Adultos ≥30 anos | Crianças < 1 ano | crianças 1-5 anos | Adultos ≥30 anos | Crianças < 1 ano | Crianças 1-5 anos |
| Brasil | 1,09 (1,03-1,15) | 1,009 (1-1,016) | 1,011 (1,002-1,021) | 0,08 (0,03-0,13) | 0,009 (0-0,016) | 0,011 (0,002-0,020) |

Elaboração própria

Na tabela 12, o número de óbitos esperados para adultos ≥30 anos por doenças cardiovasculares atribuídos ao PM_{2,5} em 2014 foi de 12.615 mortes. Para crianças menores de

1 ano, o número de óbitos esperados por doenças respiratórias atribuídos ao PM_{10} foi de 7 mortes. Para crianças de 1 a 5 anos, esse número foi de 5 mortes. Além dos óbitos esperados, as estimativas da carga de mortalidade para as áreas urbanas do país também foram fornecidas por doenças e idade: Para adultos ≥ 30 anos foi de 465.727 YLL. Para crianças menores de 1 ano e 1 a 5 anos foi de 1.286 e 798 YLL, respectivamente.

É importante mencionar que no ponto de vista do risco, as médias anuais de $PM_{2,5}$ e PM_{10} nas regiões metropolitanas de Lages e de Campina Grande ficaram abaixo das respectivas exposições de referência 7,5 e 15 $\mu\text{g}/\text{m}^3$ (mínimo contrafactual), sendo assim, não foram consideradas na análise do risco e da fração atribuível, já que, valores negativos das frações e dos óbitos atribuíveis são esperados.

Tabela 12 - Resumo dos óbitos e da carga de mortalidade atribuíveis à poluição do ar urbana - Brasil, 2014

| Áreas Metrop. | Óbitos atribuíveis | | | YLL atribuíveis | | |
|------------------|---------------------------|---------------------|----------------------|------------------------------|---------------------|----------------------|
| | Cardiovasculares | Respiratórias | | Cardiovasculares | Respiratórias | |
| | Adultos ≥ 30 anos | crianças < 1 ano | Crianças 1-5 anos | Adultos ≥ 30 anos | crianças < 1 ano | Crianças 1-5 anos |
| Brasil | 12.615 (4.673-20.126) | 7 (0 - 13) | 5 (1 - 9) | 465.727 (172.500-743.014) | 1.286 (25-2.313) | 798 (111-1.475) |

Elaboração própria

Análise de sensibilidade

O impacto de três escolhas analíticas foi feito: a forma da função C-R, exposição mínima contrafactual e a variação da razão $PM_{2,5}/PM_{10}$. A tabela 13 resume os óbitos por doenças cardiovasculares em adultos ≥ 30 anos atribuíveis ao $PM_{2,5}$ em 6 cenários considerados.

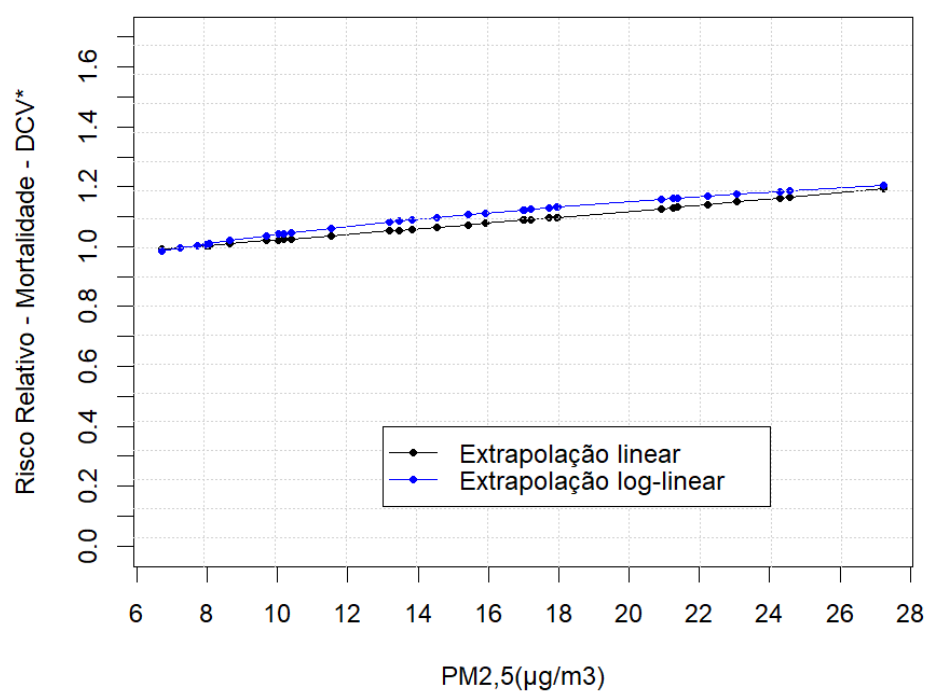
Inicialmente, como podemos observar no gráfico 3, nossas estimativas de $PM_{2,5}$ geraram RR ligeiramente menores na função C-R linear do que na função C-R log-linear.

A função C-R log-linear em comparação à função C-R linear do caso base produziu uma estimativa maior dos óbitos atribuíveis, aumentando em 35% e um intervalo de confiança mais amplo. Por outro lado, reduzindo apenas o valor de referência do caso base de 7,5 $\mu\text{g}/\text{m}^3$ para 3,0 $\mu\text{g}/\text{m}^3$ houve um aumento de 44% na estimativa de mortes atribuíveis. E aumentando o valor para 10 $\mu\text{g}/\text{m}^3$, produziu uma redução de 25%. No caso base consideramos média de dois

períodos de exposição do estudo da ACS (1979-1983 e 1999-2000). Mas, ao considerarmos apenas o período de exposição -1979 a 1983- os óbitos atribuíveis tiveram uma redução de 35%. Ao mudarmos a variação da proporção de 0,5 para 0,65, considerando o mesmo valor de referência do caso base, a estimativa das mortes atribuíveis aumentou para 50%.

Em conclusão, a análise de sensibilidade indicaram que a estimativa dos óbitos do caso base foram mais sensíveis à escolha da proporção de $PM_{2,5}$ para PM_{10} . No geral, os intervalos de confiança foram amplos em todos os 6 cenários. E as estimativas de mortalidade variaram aproximadamente de 8 a 19 mil óbitos esperados, não havendo, portanto, robustez nos resultados.

Gráfico 3 - Comparação das curvas das funções C-R linear e log-linear para o RR às exposições do $PM_{2,5}$ sobre a mortalidade por doenças cardiovasculares



*DCV - Doenças cardiovasculares
Elaboração própria

Tabela 13 - Análise de sensibilidade da mortalidade cardiovascular para adultos ≥ 30 anos relacionada à exposição a longo prazo de $PM_{2,5}$, regiões metropolitanas - Brasil, 2014

| Sensibilidade Cenários | Razão $PM_{2,5}/PM_{10}$ | Forma da curva C-R | Exposições de referência ($\mu g/m^3$) | Doenças Cardiovasculares | |
|-------------------------------------|--------------------------|--------------------|--|--|---------------------|
| | | | | Óbitos atribuíveis (IC 95%) [%variação percentual] | Amplitude do IC 95% |
| Caso base | 0,5 | Linear | 7,5 ^a | 12.611 (4.673-20.126) [100] | 15.453 |
| Escolha da função C-R | 0,5 | Log-linear | 7,5 ^a | 16.984 (6.383-26.820) [35] | 20.437 |
| Mudança da razão $PM_{2,5}/PM_{10}$ | 0,65 | Linear | 7,5 ^a | 18.919 (7.110-29.774) [50] | 22.664 |
| Exposição mínima contrafactual | 0,5 | Linear | 3,0 ^b | 18.158 (6.813-28.626) [44] | 21.813 |
| | 0,5 | Linear | 10 ^c | 9.432 (3.470-15.157) [-25] | 11.687 |
| Período de exposição (1979-1983) | 0,5 | Linear | 7,5 ^a | 8.244 (2.180-14.024) [-35] | 11.844 |

Coefficiente linear (IC 95%): 0,00893 (0,00322-0,01464) (Pope *et al.*, 2002)

Coefficiente log-linear (IC 95%): 0,15515 (0,0562-0,2541) (Pope *et al.*, 2002; Cohen *et al.*, 2004).

^a7,5 $\mu g/m^3$, a menor concentração observada no estudo de (Pope *et al.*, 1995; Pope *et al.*, 2002).

^b3,0 $\mu g/m^3$ é próximo ao nível mínimo de $PM_{2,5}$ observado nos Estados Unidos, e é a concentração anual de $PM_{2,5}$ proposta pelo NAAQS (Environmental Protection Agency 2002).

^c Valor de referência recomendado pela OMS para concentrações ambientais de $PM_{2,5}$

Discussão

Como parte do cálculo da estimativa global da carga de doença atribuível à poluição do ar, obtemos estimativas das exposições de PM_{10} , através dos modelos de efeitos mistos, para todas as regiões metropolitanas do Brasil, permitindo, assim, a obtenção da carga de mortalidade em áreas urbanas. As medidas de PM_{10} do modelo foram convertidas para medidas de $PM_{2,5}$.

Este estudo mostrou que, nas áreas metropolitanas do Brasil, em 2014, as médias globais de PM_{10} foi de 34,2 $\mu g/m^3$ e $PM_{2,5}$ de 17,1 $\mu g/m^3$, bem acima das exposições consideradas sem risco aumentado da mortalidade que foi respectivamente de 15 $\mu g/m^3$ e 7,5 $\mu g/m^3$ retirados do estudo de Cohen *et al.* (2004). Nossas estimativas são conservadoras, porque assumimos que as exposições ocorreram apenas em áreas metropolitanas e não em áreas urbanas mais dispersas (provavelmente algumas contendo polos industriais e outras fontes de emissão de poluentes relevantes) ou em áreas rurais. Mesmo com essa restrição, as estimativas indicaram que a fração atribuível à poluição do ar para doenças cardiovasculares em adultos ≥ 30 anos foi de 8% no total de óbitos nas regiões metropolitanas do Brasil. (Óbitos atribuíveis: 12.615 (IC95%: 4.673

- 20.126) e YLL: 465.727 (IC95%: 172.500 - 743.014)). Quanto ao impacto da poluição do ar sobre a mortalidade por doenças respiratórias nas áreas metropolitanas, apesar das evidências limitadas, utilizamos os resultados de um pequeno número de estudos de séries temporais em países em desenvolvimento como base para estimar o excesso das mortes e YLL em crianças com idade menores de 1 ano e de 1 a 5 anos. As frações atribuíveis em ambos os grupos etários apresentaram aproximadamente 1% das mortes por doenças respiratórias atribuíveis à poluição do ar. Para crianças menores de 1 ano, apenas 7 óbitos e 1.286 YLL foram encontrados como resultados. Para crianças de 1 a 5 anos foram encontrados 5 óbitos e 798 YLL. Em 2014, os óbitos registrados no SIM por doenças respiratórias foram registrados nas regiões metropolitanas com um total de 829 óbitos de crianças menores de 1 ano e 418 óbitos de crianças de 1 a 5 anos.

Embora algumas fontes de incertezas tenham sido quantificadas pela análise de sensibilidade, outras não menos importantes devem ser consideradas, como por exemplo, as incertezas relacionadas às exposições que podem afetar as estimativas da carga de mortalidade. Neste artigo, o MP foi o único indicador de exposição à poluição atmosférica urbana. Como a poluição atmosférica é uma mistura complexa, pode haver, portanto, efeitos de co-poluentes. Isto é, conforme citado por Cohen *et al.* (2004, p. 1415), “Outros poluentes frequentemente medidos, como o ozônio, o monóxido de carbono, os óxidos de enxofre, o nitrogênio, e o chumbo também estão associados à mortalidade e à morbidade, mas não tão consistentemente quando comparado ao MP, mas os efeitos de vários desses poluentes podem ser capturados, ou pelo menos parte deles, através do uso do MP (Sarnat *et al.*, 2001)”. Assim, se fosse possível identificar, medir com precisão e incluir a exposição de todos os poluentes conhecidos no país, a carga de mortalidade atribuível à poluição seria maior.

A utilização de estimativas de exposições, devido à falta de dados, introduz algumas incertezas e tendências nas previsões. O procedimento de imputação de dados faltantes em séries temporais multivariado com base no algoritmo EM apresentou precisão e acurácia em simulações feitas para obtenção das estimativas (Junger e Ponce De Leon, 2015). Este procedimento foi utilizado para estimar dados faltantes dos monitores de poluentes atmosféricos. A utilização de estimativas de exposições pode não representar um problema sério, pois estamos interessados nos efeitos médios a longo prazo da poluição do ar na saúde.

A má classificação da exposição pode levar, também, a uma subestimação da carga de mortalidade atribuível. Assim como os estudos epidemiológicos utilizados para quantificar as estimativas de impacto na saúde, utilizamos a média anual de MP medida a partir de algumas fontes estacionárias de monitoramento da poluição do ar para estimar os níveis médios de

exposição. Diferenças entre exposição individual e concentrações medidas em pontos fixos dependem de quão bem o poluente se mistura no meio ambiente e a eficiência com a qual o poluente penetra no interior dos recintos (Cohen *et al.*, 2004, p.1416). As estimativas do modelo de efeitos mistos são baseadas em dados de monitores populacionais. Quando bem projetados, esses monitores proporcionam níveis representativos de exposição do poluente a toda a cidade, caso o poluente se misture uniformemente no meio ambiente (Cohen *et al.*, 2004; Sarnat *et al.*, 2010). Mas subestimam quando a exposição se distribui de forma heterogênea no ambiente como, por exemplo, para pessoas que vivem ou trabalham perto dos “hotspots” de poluição, como estradas movimentadas ou fontes locais de poluição, o que contribui para a heterogeneidade espacial da exposição nas cidades (Hoek *et al.*, 2002; Cohen *et al.*, 2004, p.1416; Wilson *et al.*, 2005). Esta subestimação provavelmente seja mais pronunciada para cidades em países em desenvolvimento (Hoek *et al.*, 2002), como é o caso do Brasil, onde quase 6% da população reside em aglomerados subnormais (Censo de 2010, IBGE). Sendo assim, é importante que o Brasil desenvolva uma rede nacional de monitoramento da qualidade do ar que cubra todos os assentamentos urbanos significativos, padronizados em matéria de instrumentação, garantia de qualidade de dados e formatos de relatórios. A localização das estações de monitoramento dentro de cada área urbana deve estar em conformidade com a prática internacional aceita: em áreas de alta densidade populacional e próximas às fontes de emissão de poluentes importantes. Isso assegurará uma melhor estimativa da exposição da população e do impacto de fontes de emissão específicas (Norman *et al.*, 2007, p.788).

Outra fonte de má classificação refere-se ao tempo das estimativas de exposição. Como a carga de global de doenças está relacionada com exposições passadas, obtemos apenas os níveis atuais de concentração de MP (no caso, 2014) para informações mais atualizadas sobre a carga global de doenças, apesar do modelo linear de efeitos mistos permitir uma previsão simultânea seccional e temporal do MP. O estudo ACS fornece poucas informações sobre como a função C-R varia no tempo (Krewski *et al.*, 2000; Cohen *et al.*, 2004). E, portanto, não fica claro como essa fonte de má classificação afetaria as estimativas se assumíssemos períodos anteriores para estimar a carga global de doenças.

Neste artigo, as estimativas consideraram apenas o impacto da poluição do ar sobre a mortalidade para que sejam comparáveis aos estudos do Pope *et al.* (2002) e do Gouveia *et al.* (2018), desta forma os YLD não foram quantificados. Uma vez que não estimamos o impacto do MP sobre a incidência da doença, os anos de vida ajustados por incapacidade (DALY) foram quantificados apenas com os YLL. Provavelmente, isto poderá subestimar o verdadeiro valor da carga atribuível, já que existem evidências epidemiológicas sugerindo que a poluição do ar

é relevante na incidência de doenças cardiovasculares e, assim, contribuir para os anos vividos com incapacidades (YLD) (Cohen *et al.*, 2004).

Podemos utilizar diferentes tipos dados e métodos para estimar a carga global de doença devido à poluição do ar. Desta forma, é difícil realizar uma comparação entre as estimativas anuais de mortalidade global atribuível ao PM_{2,5}. Por exemplo, não foi possível comparar as nossas estimativas com as estimativas da carga de mortalidade de 2012 para o Brasil, publicado recentemente no relatório de 2016 pela OMS (WHO,2016). Além da diferença de período dos estudos, nas nossas análises as exposições foram estimadas através dos modelos de efeitos mistos, para os quais não se incluem áreas rurais. Já no relatório de 2016 as estimativas de PM_{2,5} foram capturadas por satélites de profundidade óptica de aerossóis e Modelos de Transporte Químicos (TM5) e incluíram áreas rurais na avaliação. Ambos os estudos não empregaram a mesma função C–R para derivar as estimativas de RR; não empregaram as mesmas estimativas de risco e nem a mesma concentração contrafactual. Além disso, as análises dos grupos de morbidades entre os estudos foram diferentes. No nosso estudo, os grupos de morbidades analisados foram doenças cardiovasculares em adultos com idade ≥ 30 anos e respiratórias em crianças < 1 e 1–5 anos. No relatório de 2016, os grupos analisados foram infecções respiratórias agudas inferiores em crianças menores que 5 anos, câncer de pulmão, obstrução pulmonar crônica, acidente vascular cerebral e doenças isquêmicas do coração em adultos maiores de 25 anos. Os resultados no relatório para o Brasil mostraram que o número de mortes por infecções respiratórias agudas inferiores associados ao PM_{2,5}, em crianças menores que 5 anos, foi de 198 mortes e 17.953 YLL. Em adultos maiores de 25 anos, a exposição ao PM_{2,5} causou 2.901 mortes e 74.360 YLL por câncer de pulmão, 539 mortes e 10.115 YLL por obstrução pulmonar crônica, 12.987 mortes e 354.716 YLL por doenças isquêmicas do coração e por fim 9.616 mortes e 239.367 YLL por acidente vascular cerebral. Resultando em um total de 26.241 (IC95%: 11.292 - 39.983) mortes prematuras e 696.511 (IC95%: 301.813 - 1.053.085) YLL.

A modelagem da qualidade do ar através de modelos estatísticos, como foi apresentado anteriormente no artigo 1, com os modelos de efeitos mistos relacionando dados de monitoramento com características econômicas, demográficas, meteorológicas e fontes de poluição, proporcionaram estimativas de previsão potencialmente precisas da exposição, viabilizando, portanto, medidas globais do risco confiáveis.

A quantificação dos riscos e dos impactos causados pela poluição do ar pode ser usada para conscientizar a gestão em saúde pública sobre os riscos associados à saúde e é um guia

eficaz na tomada de decisões. Essa quantificação fornecerá uma indicação do nível de esforços necessários para o desenvolvimento de políticas e estratégias no controle da poluição do ar.

Conclusão

Este estudo demonstrou a viabilidade em utilizar estimativas de concentrações de poluição do ar em termos de MP derivadas do modelo de efeitos mistos na avaliação global do risco de mortalidade atribuível à exposição ao MP nas regiões metropolitanas. Independente das incertezas intrínsecas às análises globais da mortalidade atribuíveis a poluição do ar, com provável subestimação, o MP é um contribuinte importante para as variações nas taxas de mortalidade por doenças cardiovasculares e respiratórias em áreas urbanas.

Suporte de informação

Informações adicionais de suporte podem ser encontradas nos apêndices deste artigo.

APÊNDICE I – Tabelas ilustrativas de avaliação da carga de mortalidade nas Regiões metropolitanas brasileiras, 2014

Referências

BRUNEKREEF, B. Air pollution and life expectancy: is there a relation? **Occupational and environmental medicine**, v. 54, n. 11, p. 781, 1997.

BRUNEKREEF, B.; HOLGATE, S. T. Air pollution and health. **The Lancet**, v. 360, n. 9341, p. 1233-1242, 10/19/ 2002. ISSN 0140-6736. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0140673602112748> >.

COHEN, A. J. et al. Urban air pollution. **Comparative quantification of health risks**, v. 2, p. 1353-1433, 2004.

DOCKERY, D. W.; POPE, C. A., 3RD. Acute respiratory effects of particulate air pollution. **Annu Rev Public Health**, v. 15, p. 107-32, 1994. ISSN 0163-7525 (Print) 0163-7525.

DOMINICI, F. et al. Air pollution and mortality: estimating regional and national dose-response relationships. **Journal of the American Statistical Association**, v. 97, n. 457, p. 100-111, 2002. ISSN 0162-1459.

DOMINICI, F. et al. Airborne particulate matter and mortality: timescale effects in four US cities. **American Journal of Epidemiology**, v. 157, n. 12, p. 1055-1065, 2003. ISSN 0002-9262.

DOMINICI, F.; SAMET, J. M.; ZEGER, S. L. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. **Journal of the Royal Statistical Society: Series A (Statistics in Society)**, v. 163, n. 3, p. 263-302, 2000. ISSN 1467-985X.

DOMINICI, F. et al. Hierarchical bivariate time series models: a combined analysis of the effects of particulate matter on morbidity and mortality. **Biostatistics**, v. 5, n. 3, p. 341-60, Jul 2004. ISSN 1465-4644 (Print) 1465-4644.

GOUVEIA, N. et al. Effects of air pollution on infant and children respiratory mortality in four large Latin-American cities. **Environmental Pollution**, v. 232, p. 385-391, 2018. ISSN 0269-7491.

HOEK, G. et al. Spatial variability of fine particle concentrations in three European areas. **Atmospheric Environment**, v. 36, n. 25, p. 4077-4088, 9// 2002. ISSN 1352-2310. Disponível em: < [//www.sciencedirect.com/science/article/pii/S1352231002002972](http://www.sciencedirect.com/science/article/pii/S1352231002002972) >.

JUNGER, W. L.; PONCE DE LEON, A. Imputation of missing data in time series for air pollutants. **Atmospheric Environment**, v. 102, p. 96-104, 2// 2015. ISSN 1352-2310. Disponível em: < [//www.sciencedirect.com/science/article/pii/S1352231014009145](http://www.sciencedirect.com/science/article/pii/S1352231014009145) >.

KATSOUYANNI, K. et al. Confounding and Effect Modification in the Short-Term Effects of Ambient Particles on Total Mortality: Results from 29 European Cities within the APHEA2 Project. **Epidemiology**, v. 12, n. 5, p. 521-531, 2001. ISSN 1044-3983. Disponível em: < http://journals.lww.com/epidem/Fulltext/2001/09000/Confounding_and_Effect_Modification_in_the.11.aspx >.

KREWSKI, D. et al. Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of particulate air pollution and mortality. **Health Effects Institut, Cambridge, MA**, 2000.

KUNZLI, N. et al. Public-health impact of outdoor and traffic-related air pollution: a European assessment. **Lancet**, v. 356, n. 9232, p. 795-801, Sep 2 2000. ISSN 0140-6736 (Print) 0140-6736.

LIN, C. A. et al. Air pollution and respiratory illness of children in Sao Paulo, Brazil. **Paediatr Perinat Epidemiol**, v. 13, n. 4, p. 475-88, Oct 1999. ISSN 0269-5022 (Print)

0269-5022.

LOGAN, W. P. D. MORTALITY IN THE LONDON FOG INCIDENT, 1952. **The Lancet**, v. 261, n. 6755, p. 336-338, 1953/02/14 1953. ISSN 0140-6736. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0140673653910125> >.

MAGE, D. et al. Assessment of human exposure to ambient particulate matter. **J Air Waste Manag Assoc**, v. 49, n. 11, p. 1280-91, Nov 1999. ISSN 1096-2247 (Print) 1096-2247.

NORMAN, R. et al. Estimating the burden of disease attributable to urban outdoor air pollution in South Africa in 2000. **S Afr Med J**, v. 97, n. 8 Pt 2, p. 782-90, Aug 2007. ISSN 0256-9574 (Print).

ORGANIZATION, W. H. Ambient air pollution: a global assessment of exposure and burden of disease. In: (Ed.). **Ambient air pollution: a global assessment of exposure and burden of disease**, 2016.

OSTRO, B. A search for a threshold in the relationship of air pollution to mortality: a reanalysis of data on London winters. **Environmental Health Perspectives**, v. 58, p. 397, 1984.

OSTRO, B.; WHO, W. H. O. O. A. E. H. T. Outdoor air pollution: assessing the environmental burden of disease at national and local levels/ Bart Ostro. p. 54, 2004. ISSN 1728-1652. Disponível em: < <http://www.who.int/iris/handle/10665/42909> >.

POPE, C. A., 3RD et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. **Jama**, v. 287, n. 9, p. 1132-41, Mar 06 2002. ISSN 0098-7484 (Print) 0098-7484.

POPE, C. A., 3RD; DOCKERY, D. W. 31 - Epidemiology of Particle Effects. In: HOLGATE, S. T.;SAMET, J. M., et al (Ed.). **Air Pollution and Health**. London: Academic Press, 1999. p.673-705. ISBN 978-0-12-352335-8.

POPE, C. A., 3RD; SCHWARTZ, J.; RANSOM, M. R. Daily mortality and PM₁₀ pollution in Utah Valley. **Arch Environ Health**, v. 47, n. 3, p. 211-7, May-Jun 1992. ISSN 0003-9896 (Print) 0003-9896.

POPE, C. A., 3RD et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. **Am J Respir Crit Care Med**, v. 151, n. 3 Pt 1, p. 669-74, Mar 1995. ISSN 1073-449X (Print) 1073-449x.

ROMIEU, I. et al. Multicity study of air pollution and mortality in Latin America (the ESCALA study). **Res Rep Health Eff Inst**, n. 171, p. 5-86, Oct 2012. ISSN 1041-5505 (Print) 1041-5505.

SAMET, J. M. et al. The National Morbidity, Mortality, and Air Pollution Study. Part I: Methods and methodologic issues. **Res Rep Health Eff Inst**, n. 94 Pt 1, p. 5-14; discussion 75-84, Jun 2000. ISSN 1041-5505 (Print) 1041-5505.

SARNAT, J. A. et al. Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? **Environmental Health Perspectives**, v. 109, n. 10, p. 1053-1061, 2001. ISSN 0091-6765. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1242083/> >.

SARNAT, S. E. et al. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. **Journal of Exposure Science and Environmental Epidemiology**, v. 20, n. 2, p. 135-146, 2010. ISSN 1559-0631.

SCHWARTZ, J. Air pollution and daily mortality: a review and meta analysis. **Environ Res**, v. 64, n. 1, p. 36-52, Jan 1994. ISSN 0013-9351 (Print) 0013-9351.

_____. Short term fluctuations in air pollution and hospital admissions of the elderly for respiratory disease. **Thorax**, v. 50, n. 5, p. 531-538, 1995. ISSN 0040-6376 1468-3296. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1021224/> >.

_____. Harvesting and long term exposure effects in the relation between air pollution and mortality. **Am J Epidemiol**, v. 151, n. 5, p. 440-8, Mar 1 2000. ISSN 0002-9262 (Print) 0002-9262.

_____. Is there harvesting in the association of airborne particles with daily deaths and hospital admissions? **Epidemiology**, v. 12, n. 1, p. 55-61, Jan 2001. ISSN 1044-3983 (Print) 1044-3983.

WILSON, J. G. et al. A review of intraurban variations in particulate air pollution: Implications for epidemiological research. **Atmospheric Environment**, v. 39, n. 34, p. 6444-6462, 11// 2005. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231005006370> >.

WONG, C.-M. et al. Public Health and Air Pollution in Asia (PAPA): a multicity study of short-term effects of air pollution on mortality. **Environmental Health Perspectives**, v. 116, n. 9, p. 1195, 2008. ISSN 0091-6765.

ZANOBETTI, A. et al. The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement. **Epidemiology**, v. 13, n. 1, p. 87-93, Jan 2002. ISSN 1044-3983 (Print) 1044-3983.

ZEGER, S. L.; DOMINICI, F.; SAMET, J. Harvesting-resistant estimates of air pollution effects on mortality. **Epidemiology**, v. 10, n. 2, p. 171-5, Mar 1999. ISSN 1044-3983 (Print) 1044-3983.

CONCLUSÃO

O artigo 1 demonstrou a possibilidade de se fazer previsões de médias anuais de PM_{10} aplicando modelos de efeitos mistos. Apesar das limitações apresentadas, o modelo foi capaz de produzir estimativas robustas e eficientes. Estimou, de forma consistente, os principais determinantes das variações de concentrações de PM_{10} no ambiente externo. Conforme os resultados obtidos, deve-se considerar, a um nível de significância de 5%, que os principais determinantes das variações de PM_{10} foram o PIB per capita do setor serviços e variáveis ambientais, como umidade, altitude média e focos de queimadas. O artigo 2, no qual o estudo foi de caráter epidemiológico, os níveis de concentração da poluição do ar, calculados no artigo 1, foram utilizados na avaliação do impacto que levou a concluir a viabilidade em se utilizar modelos de efeitos mistos na avaliação global do risco de mortalidade atribuível à exposição ao MP, sendo este um importante contribuinte nas taxas de mortalidade. Mesmo que as medidas de impacto estimadas neste artigo não tenham apresentado robustez em seus resultados e possivelmente terem sido subestimadas, a quantificação dessas medidas pode ser usada para conscientizar sobre os riscos associados à saúde e ser um guia eficaz na tomada de decisões, isto é, um indicador dos esforços necessários para o desenvolvimento de políticas e estratégias no controle da poluição do ar em uma determinada cidade ou região.

Vale ressaltar também que ambos os artigos mostraram todo o exercício de vigilância incluindo a coleta de dados (definição de indicadores demográficos, ambientais, socioeconômicos e de saúde), aplicação de metodologias para análise e interpretação de resultados demonstrando, ao mesmo tempo, o potencial que os dados de domínio público e irrestrito podem proporcionar para vigilância em saúde. Característicos dos estudos ecológicos em séries temporais epidemiológicas, os dados secundários são facilmente acessados, produzem resultados rápidos e apresentam baixo custo de execução. Combinados com os avanços metodológicos e computacionais, permitem produzir informações a fim de recomendar e adotar medidas de prevenção e controle dos agravos à saúde devido à exposição de poluentes atmosféricos.

REFERÊNCIAS

ABBEY, D. E. et al. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. **Am J Respir Crit Care Med**, v. 159, n. 2, p. 373-82, Feb 1999. ISSN 1073-449X (Print) 1073-449x.

ANENBERG, S. C. et al. An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modeling. **Environ Health Perspect**, v. 118, n. 9, p. 1189-95, Sep 2010. ISSN 0091-6765.

BOX, G. E.; COX, D. R. An analysis of transformations. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 211-252, 1964. ISSN 0035-9246.

BOX, G. E. et al. **Time series analysis: forecasting and control**. John Wiley & Sons, 2015. ISBN 1118675029.

BRAUER, M. et al. Exposure Assessment for Estimation of the Global Burden of Disease Attributable to Outdoor Air Pollution. **Environmental Science & Technology**, v. 46, n. 2, p. 652-660, 2012/01/17 2012. ISSN 0013-936X. Disponível em: < <http://dx.doi.org/10.1021/es2025752> >.

BRAUER, M. et al. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. **Environ Sci Technol**, v. 50, n. 1, p. 79-88, Jan 5 2016. ISSN 0013-936x.

BRAUER, M. et al. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. **Epidemiology**, v. 14, n. 2, p. 228-39, Mar 2003. ISSN 1044-3983 (Print) 1044-3983.

BRIGGS, D. J. et al. Mapping urban air pollution using GIS: a regression-based approach. **International Journal of Geographical Information Science**, v. 11, n. 7, p. 699-718, 1997/10/01 1997. ISSN 1365-8816. Disponível em: < <http://dx.doi.org/10.1080/136588197242158> >.

BRUNEKREEF, B. Air pollution and life expectancy: is there a relation? **Occupational and environmental medicine**, v. 54, n. 11, p. 781, 1997.

BRUNEKREEF, B.; HOLGATE, S. T. Air pollution and health. **The Lancet**, v. 360, n. 9341, p. 1233-1242, 10/19/ 2002. ISSN 0140-6736. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0140673602112748> >.

CHANG, H. H.; HU, X.; LIU, Y. Calibrating MODIS aerosol optical depth for predicting daily PM2.5 concentrations via statistical downscaling. **J Expos Sci Environ Epidemiol**, v. 24, n. 4, p. 398-404, 07//print 2014. ISSN 1559-0631. Disponível em: < <http://dx.doi.org/10.1038/jes.2013.90> >.

CHUDNOVSKY, A. A. et al. Fine particulate matter predictions using high resolution Aerosol Optical Depth (AOD) retrievals. **Atmospheric Environment**, v. 89, p. 189-198, 6// 2014. ISSN 1352-2310. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S1352231014001150> >.

COHEN, A. J. et al. Urban air pollution. **Comparative quantification of health risks**, v. 2, p. 1353-1433, 2004.

COHEN, A. J. et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. **The Lancet**, v. 389, n. 10082, p. 1907-1918, 2017. ISSN 0140-6736.

DAVISON, A. C.; HINKLEY, D. V. **Bootstrap methods and their application**. Cambridge university press, 1997. ISBN 0521574714.

DE LEON, A. P.; JACOBSON, L. V.; HISHINUMA, H. P-008: Brazilian Model of Ambient Particulates (BMAPS). **Epidemiology**, v. 23, p. 1, 2012. ISSN 1044-3983.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the royal statistical society. Series B (methodological)**, p. 1-38, 1977. ISSN 0035-9246.

DOCKERY, D. W.; POPE, C. A., 3RD. Acute respiratory effects of particulate air pollution. **Annu Rev Public Health**, v. 15, p. 107-32, 1994. ISSN 0163-7525 (Print) 0163-7525.

DOMINICI, F. et al. Air pollution and mortality: estimating regional and national dose-response relationships. **Journal of the American Statistical Association**, v. 97, n. 457, p. 100-111, 2002. ISSN 0162-1459.

DOMINICI, F. et al. Airborne particulate matter and mortality: timescale effects in four US cities. **American Journal of Epidemiology**, v. 157, n. 12, p. 1055-1065, 2003. ISSN 0002-9262.

DOMINICI, F.; SAMET, J. M.; ZEGER, S. L. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. **Journal of the Royal Statistical Society: Series A (Statistics in Society)**, v. 163, n. 3, p. 263-302, 2000. ISSN 1467-985X.

DOMINICI, F. et al. Hierarchical bivariate time series models: a combined analysis of the effects of particulate matter on morbidity and mortality. **Biostatistics**, v. 5, n. 3, p. 341-60, Jul 2004. ISSN 1465-4644 (Print) 1465-4644.

EEFTENS, M. et al. Development of Land Use Regression models for PM_{2.5}, PM_{2.5} absorbance, PM₁₀ and PM_{coarse} in 20 European study areas; results of the ESCAPE project. **Environ Sci Technol**, v. 46, n. 20, p. 11195-205, Oct 16 2012. ISSN 0013-936x.

EFRON, B.; TIBSHIRANI, R. J. **An introduction to the bootstrap**. CRC press, 1994. ISBN 0412042312.

ELLIOT, P. et al. **Spatial epidemiology: methods and applications**. Oxford University Press, 2000. ISBN 0192629417.

EMILI, E. et al. PM10 remote sensing from geostationary SEVIRI and polar-orbiting MODIS sensors over the complex terrain of the European Alpine region. **Remote Sensing of Environment**, v. 114, n. 11, p. 2485-2499, 11/15/ 2010. ISSN 0034-4257. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S003442571000180X> >.

EVANS, J. et al. Estimates of global mortality attributable to particulate air pollution using satellite imagery. **Environmental Research**, v. 120, p. 33-42, 2013/01/01/ 2013. ISSN 0013-9351. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S001393511200240X> >.

FIRKET, J. Fog along the Meuse valley. **Transactions of the Faraday Society**, v. 32, p. 1192-1196, 1936.

GOLDSTEIN, H. **Multilevel statistical models**. John Wiley & Sons, 2011. ISBN 111995682X.

GOUVEIA, N. et al. Effects of air pollution on infant and children respiratory mortality in four large Latin-American cities. **Environmental Pollution**, v. 232, p. 385-391, 2018. ISSN 0269-7491.

GRANT, L. D.; SHOAF, C. R.; DAVIS, J. M. 42 - United States and International Approaches to Establishing Air Standards and Guidelines A2 - Holgate, Stephen T. In: SAMET, J. M.; KOREN, H. S., *et al* (Ed.). **Air Pollution and Health**. London: Academic Press, 1999. p.947-982. ISBN 978-0-12-352335-8.

HOEK, G. et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. **Atmospheric Environment**, v. 42, n. 33, p. 7561-7578, 10// 2008. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231008005748> >.

HOEK, G. et al. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. **Journal of Exposure Science and Environmental Epidemiology**, v. 11, n. 6, p. 459, 2001. ISSN 1559-0631.

HOEK, G. et al. Spatial variability of fine particle concentrations in three European areas. **Atmospheric Environment**, v. 36, n. 25, p. 4077-4088, 9// 2002. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231002002972> >.

HOFF, R. M.; CHRISTOPHER, S. A. Remote sensing of particulate pollution from space: have we reached the promised land? **Journal of the Air & Waste Management Association**, v. 59, n. 6, p. 645-675, 2009. ISSN 1096-2247.

HOLMAN, C. 8 - Sources of Air Pollution A2 - Holgate, Stephen T. In: SAMET, J. M.; KOREN, H. S., *et al* (Ed.). **Air Pollution and Health**. London: Academic Press, 1999. p.115-148. ISBN 978-0-12-352335-8.

HRUBA, F. et al. Childhood respiratory symptoms, hospital admissions, and long-term exposure to airborne particulate matter. **J Expo Anal Environ Epidemiol**, v. 11, n. 1, p. 33-40, Jan-Feb 2001. ISSN 1053-4245 (Print) 1053-4245.

IEMA. **1º Diagnóstico da rede de monitoramento da qualidade do ar no Brasil**: Instituto de Energia e Meio Ambiente 2014.

JACOBSON, L. V.; DE LEON, A. P.; HISHINUMA, H. P-007: Health Impact Assessment of Particulate Matter Air Pollution in Brazilian Metropolitan Regions. **Epidemiology**, v. 23, p. 1, 2012. ISSN 1044-3983.

JANSSEN, N. A. H. et al. Assessment of exposure to traffic related air pollution of children attending schools near motorways. **Atmospheric Environment**, v. 35, n. 22, p. 3875-3884, 8// 2001. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231001001443> >.

JERRETT, M. et al. A review and evaluation of intraurban air pollution exposure models. **J Expo Anal Environ Epidemiol**, v. 15, n. 2, p. 185-204, Mar 2005. ISSN 1053-4245 (Print) 1053-4245.

JUNGER, W. L.; PONCE DE LEON, A. Imputation of missing data in time series for air pollutants. **Atmospheric Environment**, v. 102, p. 96-104, 2// 2015. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231014009145> >.

KATSOUYANNI, K. et al. Confounding and Effect Modification in the Short-Term Effects of Ambient Particles on Total Mortality: Results from 29 European Cities within the APHEA2 Project. **Epidemiology**, v. 12, n. 5, p. 521-531, 2001. ISSN 1044-3983. Disponível em: < http://journals.lww.com/epidem/Fulltext/2001/09000/Confounding_and_Effect_Modification_in_the.11.aspx >.

KLOOG, I. et al. Estimating daily PM_{2.5} and PM₁₀ across the complex geo-climate region of Israel using MAIAC satellite-based AOD data. **Atmospheric Environment**, v. 122, p. 409-416, 12// 2015. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231015304234> >.

KREWSKI, D. et al. Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of particulate air pollution and mortality. **Health Effects Institut, Cambridge, MA**, 2000.

KRZYZANOWSK, M.; SCHWELA, D. 7 - Patterns of Air Pollution in Developing Countries A2 - Holgate, Stephen T. In: SAMET, J. M.; KOREN, H. S., et al (Ed.). **Air Pollution and Health**. London: Academic Press, 1999. p.105-113. ISBN 978-0-12-352335-8.

KUNZLI, N. et al. Public-health impact of outdoor and traffic-related air pollution: a European assessment. **Lancet**, v. 356, n. 9232, p. 795-801, Sep 2 2000. ISSN 0140-6736 (Print) 0140-6736.

KUNZLI, N. et al. Assessment of deaths attributable to air pollution: should we use risk estimates based on time series or on cohort studies? **Am J Epidemiol**, v. 153, n. 11, p. 1050-5, Jun 01 2001. ISSN 0002-9262 (Print) 0002-9262.

LEBRET, E. et al. Small area variations in ambient NO₂ concentrations in four European areas. **Atmospheric Environment**, v. 34, n. 2, p. 177-185, 1// 2000. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231099002927> >.

LI, Y. et al. Assessing Long-Term Trend of Particulate Matter Pollution in the Pearl River Delta Region Using Satellite Remote Sensing. **Environ Sci Technol**, v. 49, n. 19, p. 11670-8, Oct 6 2015. ISSN 0013-936x.

LIM, S. S. et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. **The Lancet**, v. 380, n. 9859, p. 2224-2260, 2012/12/15/ 2012. ISSN 0140-6736. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0140673612617668> >.

LIN, C. A. et al. Air pollution and respiratory illness of children in Sao Paulo, Brazil. **Paediatr Perinat Epidemiol**, v. 13, n. 4, p. 475-88, Oct 1999. ISSN 0269-5022 (Print) 0269-5022.

LOGAN, W. P. D. MORTALITY IN THE LONDON FOG INCIDENT, 1952. **The Lancet**, v. 261, n. 6755, p. 336-338, 1953/02/14 1953. ISSN 0140-6736. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0140673653910125> >.

MAGE, D. et al. Assessment of human exposure to ambient particulate matter. **J Air Waste Manag Assoc**, v. 49, n. 11, p. 1280-91, Nov 1999. ISSN 1096-2247 (Print) 1096-2247.

MCMICHAEL, A. J. et al. Inappropriate use of daily mortality analyses to estimate longer-term mortality effects of air pollution. **Int J Epidemiol**, v. 27, n. 3, p. 450-3, Jun 1998. ISSN 0300-5771 (Print) 0300-5771.

MULHOLLAND, J. A. et al. Temporal and spatial distributions of ozone in Atlanta: regulatory and epidemiologic implications. **J Air Waste Manag Assoc**, v. 48, n. 5, p. 418-26, May 1998. ISSN 1096-2247 (Print) 1096-2247.

NAKAGAWA, S.; SCHIELZETH, H. A general and simple method for obtaining R² from generalized linear mixed-effects models. **Methods in Ecology and Evolution**, v. 4, n. 2, p. 133-142, 2013. ISSN 2041-210X.

NAVIDI, W.; LURMANN, F. Measurement error in air pollution exposure assessment. **Journal of exposure analysis and environmental epidemiology**, v. 5, n. 2, p. 111-124, 1994. ISSN 1053-4245.

NORMAN, R. et al. Estimating the burden of disease attributable to urban outdoor air pollution in South Africa in 2000. **S Afr Med J**, v. 97, n. 8 Pt 2, p. 782-90, Aug 2007. ISSN 0256-9574 (Print).

ORGANIZATION, W. H. **Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide**. World Health Organization, 2006. ISBN 9289021926.

_____. Ambient air pollution: a global assessment of exposure and burden of disease. In: (Ed.). **Ambient air pollution: a global assessment of exposure and burden of disease**, 2016.

OSTRO, B. A search for a threshold in the relationship of air pollution to mortality: a reanalysis of data on London winters. **Environmental Health Perspectives**, v. 58, p. 397, 1984.

_____. Methodology for estimating air pollution health effects. In: (Ed.). **Methodology for estimating air pollution health effects**: OMS, 1996.

OSTRO, B.; WHO, W. H. O. O. A. E. H. T. Outdoor air pollution: assessing the environmental burden of disease at national and local levels/ Bart Ostro. p. 54, 2004. ISSN 1728-1652. Disponível em: < <http://www.who.int/iris/handle/10665/42909> >.

PANDEY, K. D. et al. Ambient particulate matter concentrations in residential and pollution hotspot areas of world cities: new estimates based on the Global Model of Ambient Particulates (GMAPS). **Washington, DC: The World Bank**, 2006.

PINHEIRO, J.; BATES, D. **Mixed-effects models in S and S-PLUS**. Springer, 2000. ISBN 038798957-9.

POPE, C. A., 3RD et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. **Jama**, v. 287, n. 9, p. 1132-41, Mar 06 2002. ISSN 0098-7484 (Print) 0098-7484.

POPE, C. A., 3RD; DOCKERY, D. W. 31 - Epidemiology of Particle Effects. In: HOLGATE, S. T.;SAMET, J. M., *et al* (Ed.). **Air Pollution and Health**. London: Academic Press, 1999. p.673-705. ISBN 978-0-12-352335-8.

POPE, C. A., 3RD; DOCKERY, D. W. Health effects of fine particulate air pollution: lines that connect. **J Air Waste Manag Assoc**, v. 56, n. 6, p. 709-42, Jun 2006. ISSN 1096-2247 (Print) 1096-2247.

POPE, C. A., 3RD; SCHWARTZ, J.; RANSOM, M. R. Daily mortality and PM10 pollution in Utah Valley. **Arch Environ Health**, v. 47, n. 3, p. 211-7, May-Jun 1992. ISSN 0003-9896 (Print) 0003-9896.

POPE, C. A., 3RD et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. **Am J Respir Crit Care Med**, v. 151, n. 3 Pt 1, p. 669-74, Mar 1995. ISSN 1073-449X (Print) 1073-449x.

ROMIEU, I. et al. Multicity study of air pollution and mortality in Latin America (the ESCALA study). **Res Rep Health Eff Inst**, n. 171, p. 5-86, Oct 2012. ISSN 1041-5505 (Print) 1041-5505.

SAMET, J. M. et al. The National Morbidity, Mortality, and Air Pollution Study. Part I: Methods and methodologic issues. **Res Rep Health Eff Inst**, n. 94 Pt 1, p. 5-14; discussion 75-84, Jun 2000. ISSN 1041-5505 (Print) 1041-5505.

SARNAT, J. A. et al. Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? **Environmental Health Perspectives**, v. 109, n. 10, p. 1053-1061, 2001. ISSN 0091-6765. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1242083/> >.

SARNAT, S. E. et al. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. **Journal of Exposure Science and Environmental Epidemiology**, v. 20, n. 2, p. 135-146, 2010. ISSN 1559-0631.

SCHWARTZ, J. Air pollution and daily mortality: a review and meta analysis. **Environ Res**, v. 64, n. 1, p. 36-52, Jan 1994. ISSN 0013-9351 (Print) 0013-9351.

_____. Short term fluctuations in air pollution and hospital admissions of the elderly for respiratory disease. **Thorax**, v. 50, n. 5, p. 531-538, 1995. ISSN 0040-6376 1468-3296. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1021224/> >.

_____. Harvesting and long term exposure effects in the relation between air pollution and mortality. **Am J Epidemiol**, v. 151, n. 5, p. 440-8, Mar 1 2000. ISSN 0002-9262 (Print) 0002-9262.

_____. Is there harvesting in the association of airborne particles with daily deaths and hospital admissions? **Epidemiology**, v. 12, n. 1, p. 55-61, Jan 2001. ISSN 1044-3983 (Print) 1044-3983.

SCHWARTZ, J.; MARCUS, A. Mortality and air pollution in London: a time series analysis. **Am J Epidemiol**, v. 131, n. 1, p. 185-94, Jan 1990. ISSN 0002-9262 (Print) 0002-9262.

SNIJDERS, T.; BOSKER, R. **Multilevel analysis: An introduction to basic and applied multilevel analysis**: London: Sage 1999.

STIEB, D. M. et al. Estimated public health impacts of changes in concentrations of fine particle air pollution in Canada, 2000 to 2011. **Can J Public Health**, v. 106, n. 6, p. e362-8, Sep-Oct 2015. ISSN 0008-4263.

VAN DONKELAAR, A. et al. Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application. **Environmental Health Perspectives**, v. 118, n. 6, p. 847-855, 03/16 10/26/received 02/25/accepted 2010. ISSN 0091-6765

1552-9924. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2898863/> >.

VERBEKE, G. Linear mixed models for longitudinal data. In: (Ed.). **Linear mixed models in practice**: Springer, 1997. p.63-153.

WILSON, J. G. et al. A review of intraurban variations in particulate air pollution: Implications for epidemiological research. **Atmospheric Environment**, v. 39, n. 34, p. 6444-6462, 11// 2005. ISSN 1352-2310. Disponível em: < [//www.sciencedirect.com/science/article/pii/S1352231005006370](http://www.sciencedirect.com/science/article/pii/S1352231005006370) >.

WOLFINGER, R. Covariance structure selection in general mixed models. **Communications in statistics-Simulation and computation**, v. 22, n. 4, p. 1079-1106, 1993. ISSN 0361-0918.

WONG, C.-M. et al. Public Health and Air Pollution in Asia (PAPA): a multicity study of short-term effects of air pollution on mortality. **Environmental Health Perspectives**, v. 116, n. 9, p. 1195, 2008. ISSN 0091-6765.

ZANOBETTI, A. et al. The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement. **Epidemiology**, v. 13, n. 1, p. 87-93, Jan 2002. ISSN 1044-3983 (Print) 1044-3983.

ZEGER, S. L.; DOMINICI, F.; SAMET, J. Harvesting-resistant estimates of air pollution effects on mortality. **Epidemiology**, v. 10, n. 2, p. 171-5, Mar 1999. ISSN 1044-3983 (Print) 1044-3983.

ZEGER, S. L. et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. **Environ Health Perspect**, v. 108, n. 5, p. 419-26, May 2000. ISSN 0091-6765 (Print) 0091-6765.

APÊNDICE A - Estruturação das redes de monitoramento de PM₁₀ nas regiões metropolitanas do Brasil

Apenas 9 das 36 regiões metropolitanas possuem redes de monitoramento da qualidade do ar com estações automáticas, semiautomáticas e manuais que monitoram poluentes como o PM₁₀ estabelecidos pela resolução CONAMA 03/1990, são eles: as regiões metropolitanas de São Paulo (Baixada Santista, Campinas e São Paulo) (rede de estações operada pela Companhia de Tecnologia de Saneamento Ambiental – CETESB, <http://www.cetesb.sp.gov.br>), Belo Horizonte (Fundação Estadual de Meio Ambiente – FEAM, <http://www.feam.br>), Rio de Janeiro (Instituto Estadual do Ambiente – INEA, <http://www.inea.rj.gov.br/fma/qualidade-ar.asp>), Grande Vitória (Instituto Estadual de Meio Ambiente e Recursos Hídricos – IEMA, <http://www.meioambiente.es.gov.br/default.asp>), Curitiba (Instituto Ambiental do Paraná – IAP, <http://www.iap.pr.gov.br>), Porto Alegre (Fundação Estadual de Proteção Ambiental – FEPAM, <http://www.fepam.rs.gov.br>), Salvador (Instituto de Meio Ambiente e Recursos Hídricos – INEMA, <http://www.inema.ba.gov.br/>). Na maioria dos trabalhos sobre MP são realizados apenas nessas regiões metropolitanas devido à existência de estações que monitoram o poluente.

Quadro 2 - Órgãos Estaduais de Meio Ambiente – OEMA

| |
|---|
| Companhia de Tecnologia de Saneamento Ambiental- São Paulo - CETESB http://www.cetesb.sp.gov.br |
| Fundação Estadual de Meio Ambiente – FEAM http://www.feam.br - |
| Secretaria Municipal de Meio Ambiente – SMAC http://www.rio.rj.gov.br/web/smac |
| Instituto Estadual de Meio Ambiente e Recursos Hídricos - IEMA http://www.meioambiente.es.gov.br/default.asp |
| Instituto Ambiental do Paraná – IAP http://www.iap.pr.gov.br ; |
| Instituto Estadual do Ambiente – INEA http://www.inea.rj.gov.br/fma/qualidade-ar.asp ; |
| Fundação Estadual de Proteção Ambiental – FEPAM http://www.fepam.rs.gov.br |
| Instituto de Meio Ambiente e Recursos Hídricos – INEMA http://www.inema.ba.gov.br/ |

APÊNDICE B - Critérios de validação dos dados das redes de monitoramento

A adoção de critérios de representatividade de dados é de extrema importância em sistemas de monitoramento. O não cumprimento destes critérios implica em interpretações errôneas dos resultados obtidos devido às ocorrências de falhas de medições. Entre os órgãos ambientais, tais critérios podem diferir.

Quadro 3 - Critérios de representatividade temporal dos dados utilizados pela CETESB, SMAC, FEAM e IEMA

| |
|--|
| Redes automáticas |
| Média horária: 3/4 das medidas válidas na hora |
| Média diária: 2/3 das médias horárias válidas no dia |
| Média mensal: 2/3 das médias diárias válidas no mês |
| Média anual: 1/2 das médias diárias válidas para os quadrimestres janeiro-abril, maio-agosto e setembro-dezembro |
| Redes Manuais |
| Média diária: pelo menos 22 horas de amostragem |
| Média mensal: 2/3 das médias diárias válidas no mês |
| Média anual: 1/2 das médias diárias válidas para os quadrimestres janeiro-abril, maio-agosto e setembro-dezembro |

Quadro 4 - Critérios de representatividade temporal dos dados utilizados pela IAP

| |
|--|
| Horária: Pelo menos uma média de 30 minutos válida |
| Oito horas: Pelo menos seis médias horárias válidas |
| Diária (24 horas): Pelo menos 16 médias horárias válidas |
| Mensal: Pelo menos 2/3 das médias diárias válidas |
| Quadrimestral: Pelo menos a metade das médias diárias válidas |
| Anual: Todas as três médias quadrimestrais válidas (janeiro-abril, maio-agosto, setembro-dezembro) |

Quadro 5 - Critérios de representatividade de dados utilizados pela FEPAM

| |
|--|
| Redes automáticas |
| Horária: 3/4 das medidas válidas na hora |
| Diária: 2/3 das médias horárias válidas no dia |
| Anual: 1/2 das médias válidas no ano |

Quadro 6 - Critérios de representatividade de dados utilizados pela INEA

| |
|--|
| Rede Automática |
| Média horária 75% das medidas válidas em 1 hora |
| Média diária 75% das médias horárias válidas em 24 horas |
| Média anual 50% das médias diárias válidas para os períodos: janeiro-abril; maio-agosto; setembro-dezembro |
| Rede Semiautomática |
| Diária: 24 horas \pm 15 minutos |
| Mensal: 2/3 das médias diárias válidas no mês |
| Anual: 1/2 das médias diárias válidas no ano |

APÊNDICE C - Mapas dos fatores contextuais das regiões metropolitanas do Brasil – 2014

Figura 4 - Consumo anual per capita de etanol (litros/hab.) nas regiões metropolitanas, 2014

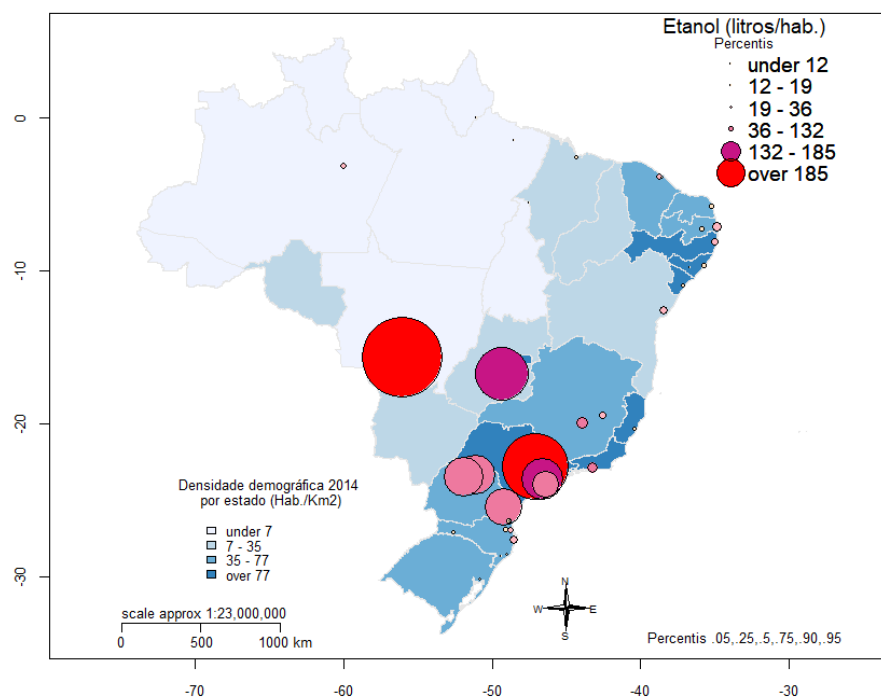
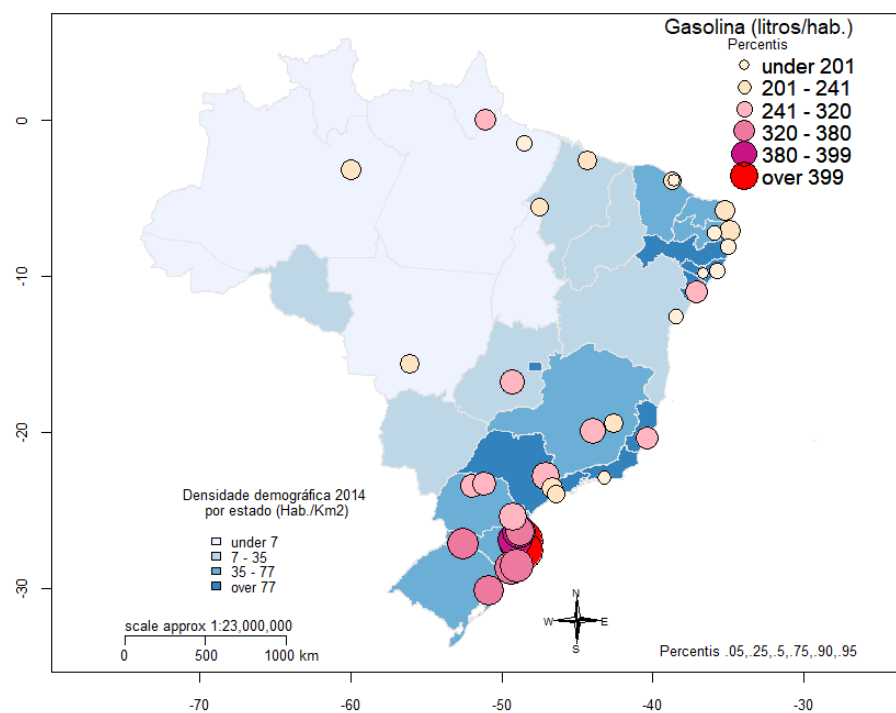


Figura 5 - Consumo anual per capita de gasolina (litros/hab.) nas regiões metropolitanas, 2014



Elaboração própria

Figura 6 - Consumo anual per capita de óleo diesel (litros/hab.) nas regiões metropolitanas, 2014

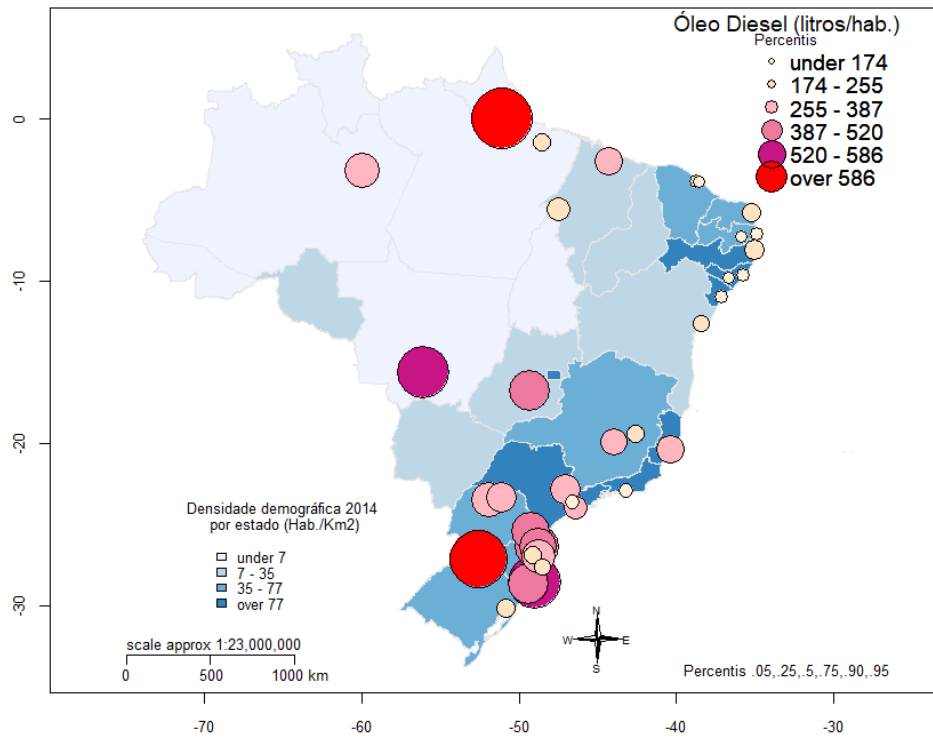
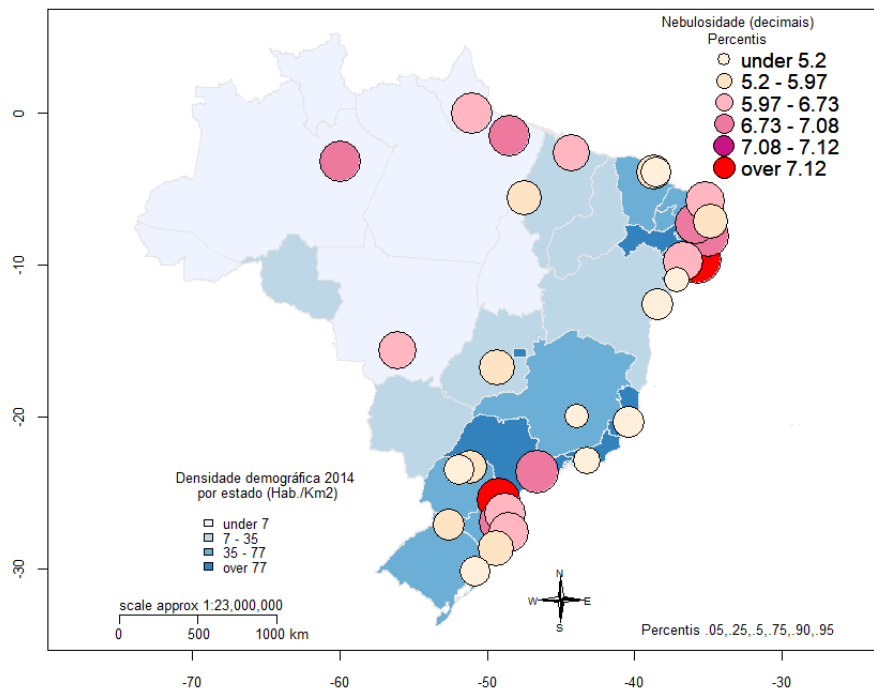


Figura 7 - Média anual de nebulosidade (decimais) nas regiões metropolitanas, 2014



Elaboração própria

Figura 8 - Média anual da pressão atmosférica (hPa) nas regiões metropolitanas, 2014

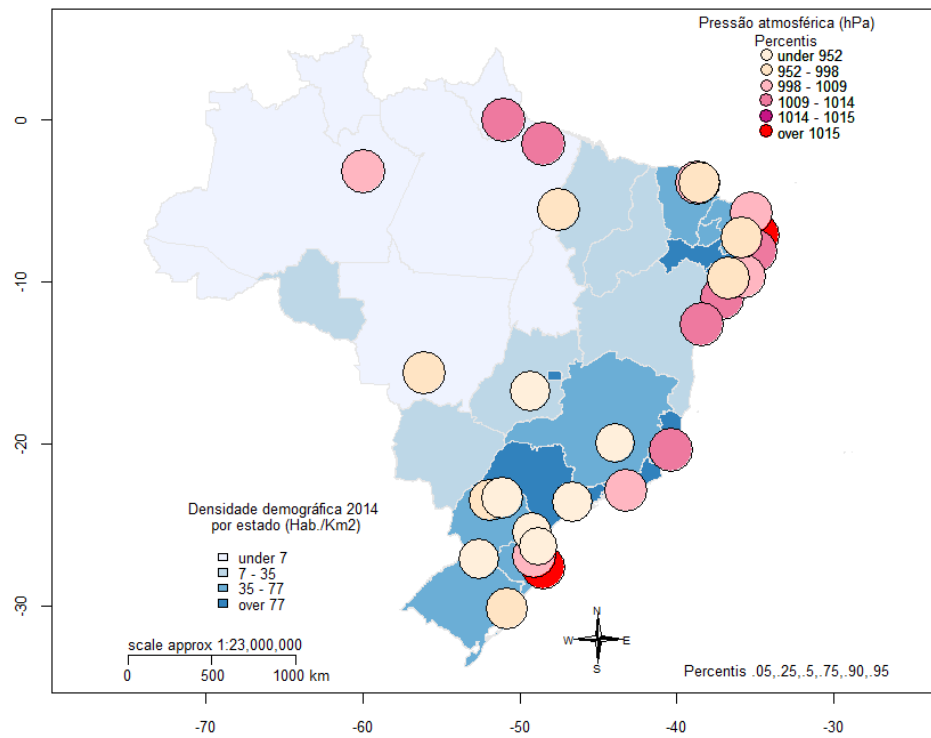
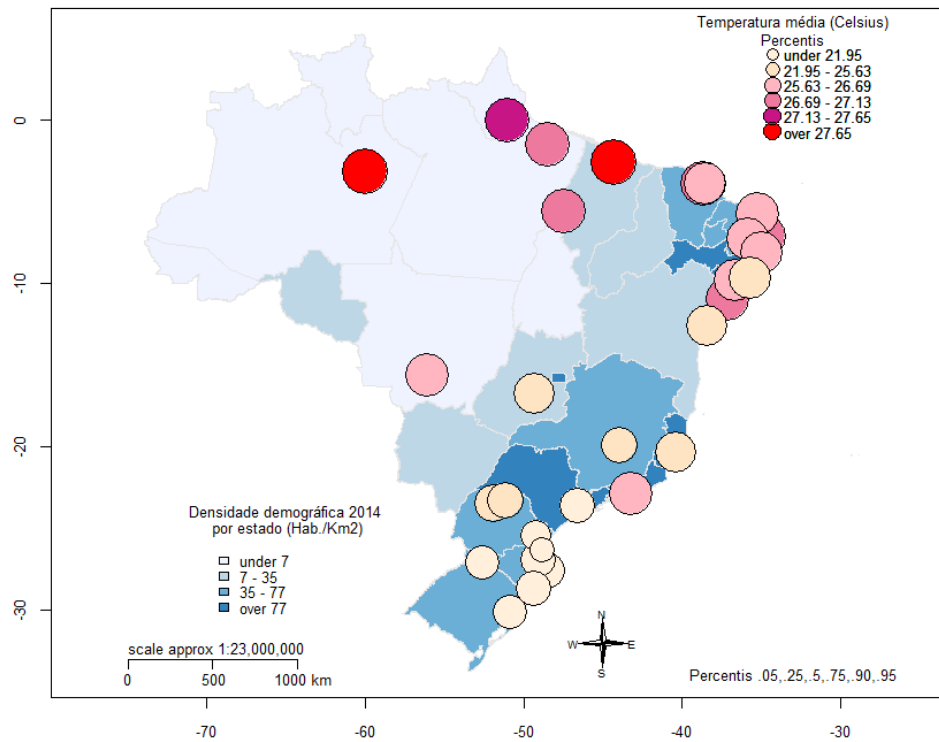


Figura 9 - Média anual da temperatura (Celsius) nas regiões metropolitanas, 2014



Elaboração própria

Figura 10 - Média anual de umidade relativa (%) nas regiões metropolitanas, 2014

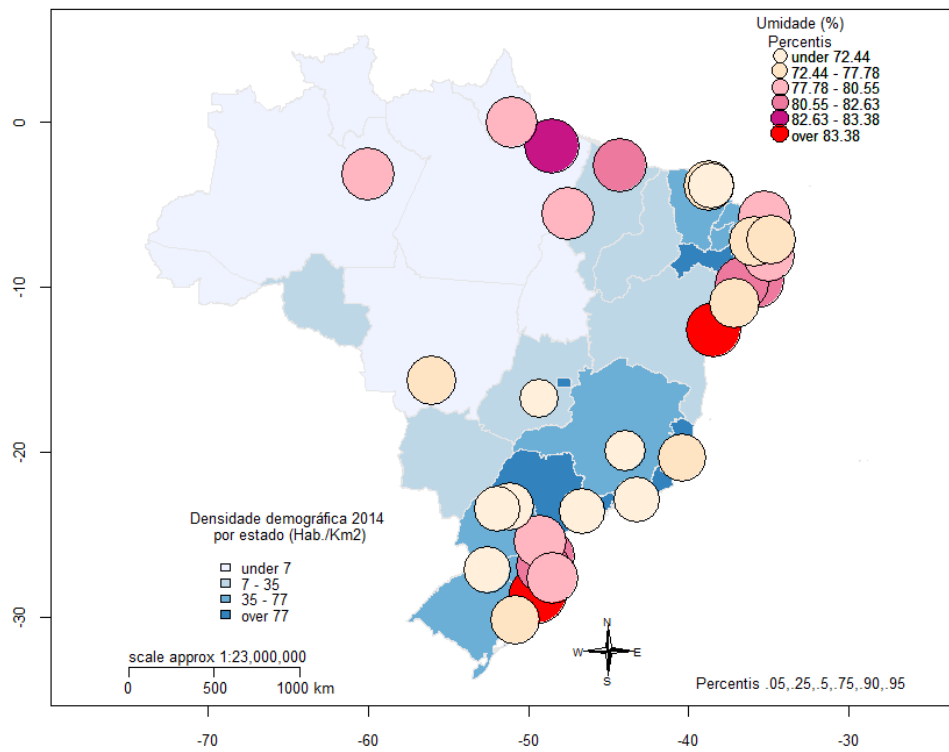
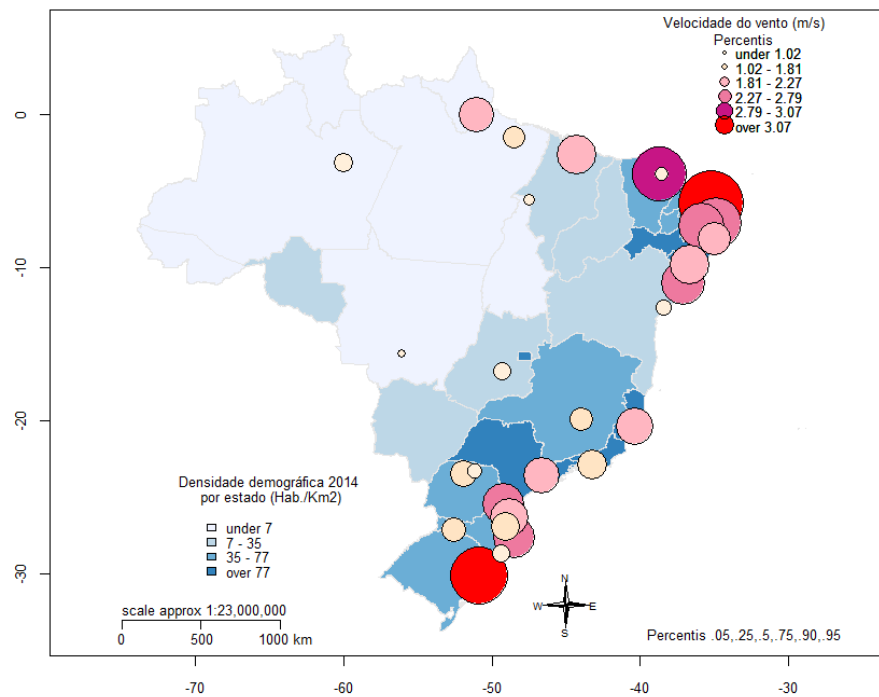


Figura 11 - Média anual da velocidade do vento (m/s) nas regiões metropolitanas, 2014



Elaboração própria

Figura 12 - Total anual do índice pluviométrico (mm) nas regiões metropolitanas, 2014

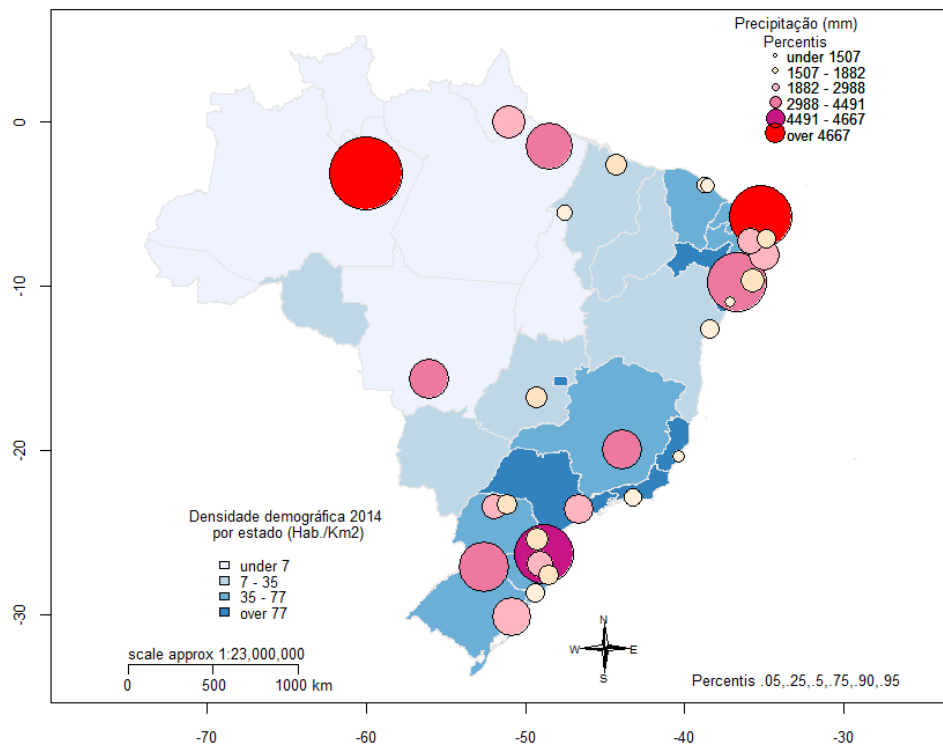
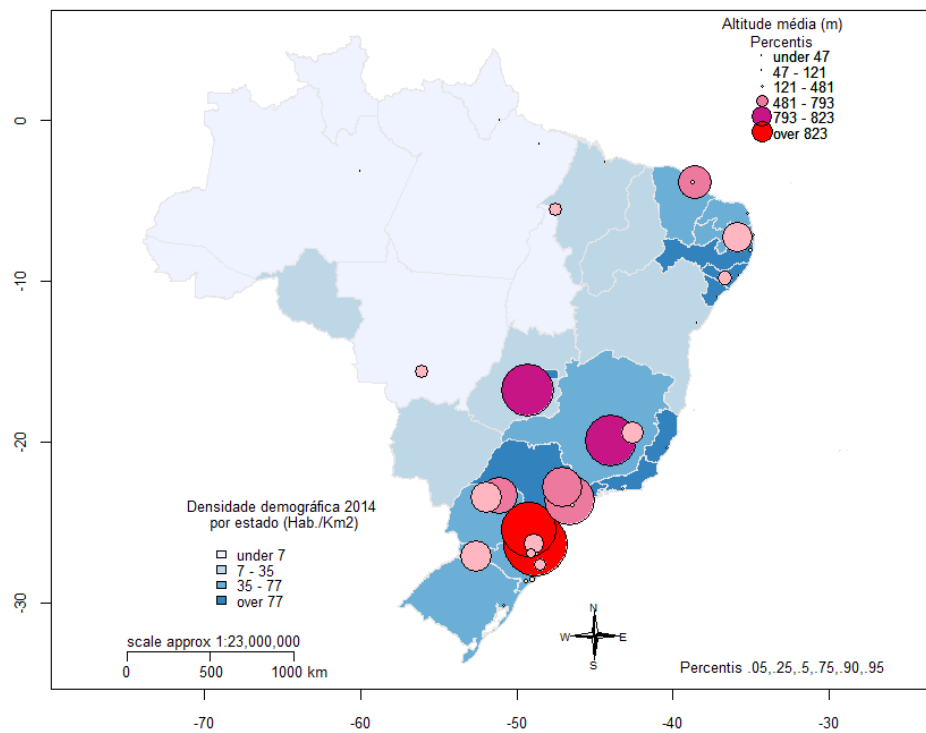


Figura 13 - Altitude média (em metros) das regiões metropolitanas, 2014



Elaboração própria

Figura 14 - PIB per capita (em R\$ 1.000) nas regiões metropolitanas, 2014

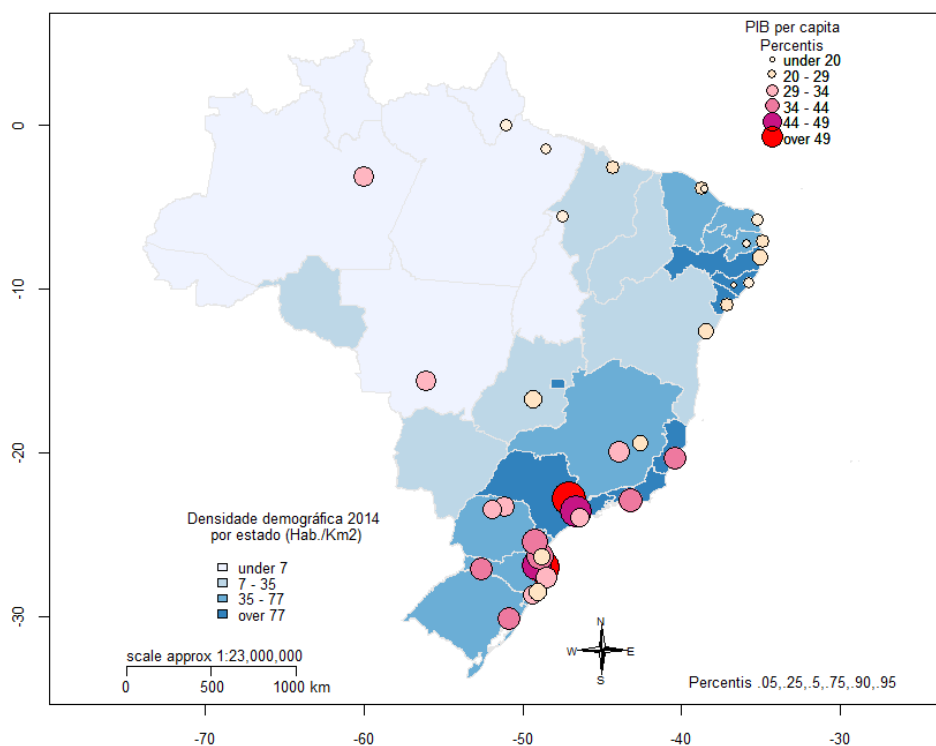
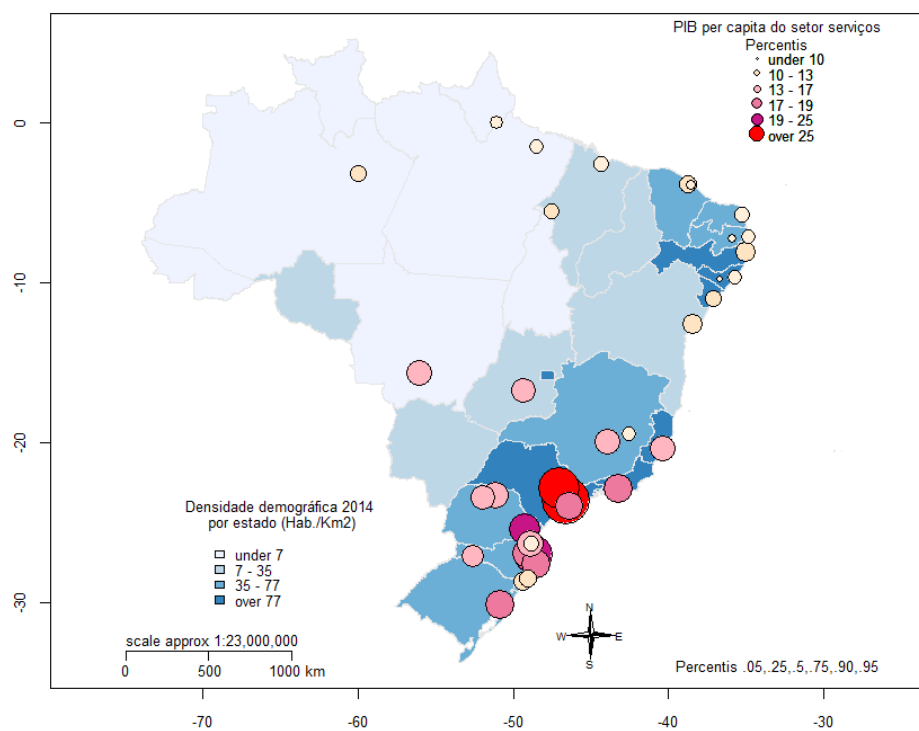


Figura 15 - PIB per capita do setor serviços (em R\$ 1.000) nas regiões metropolitanas, 2014



Elaboração própria

Figura 16 - PIB per capita do setor agropecuário (em R\$ 1.000) nas regiões metropolitanas, 2014

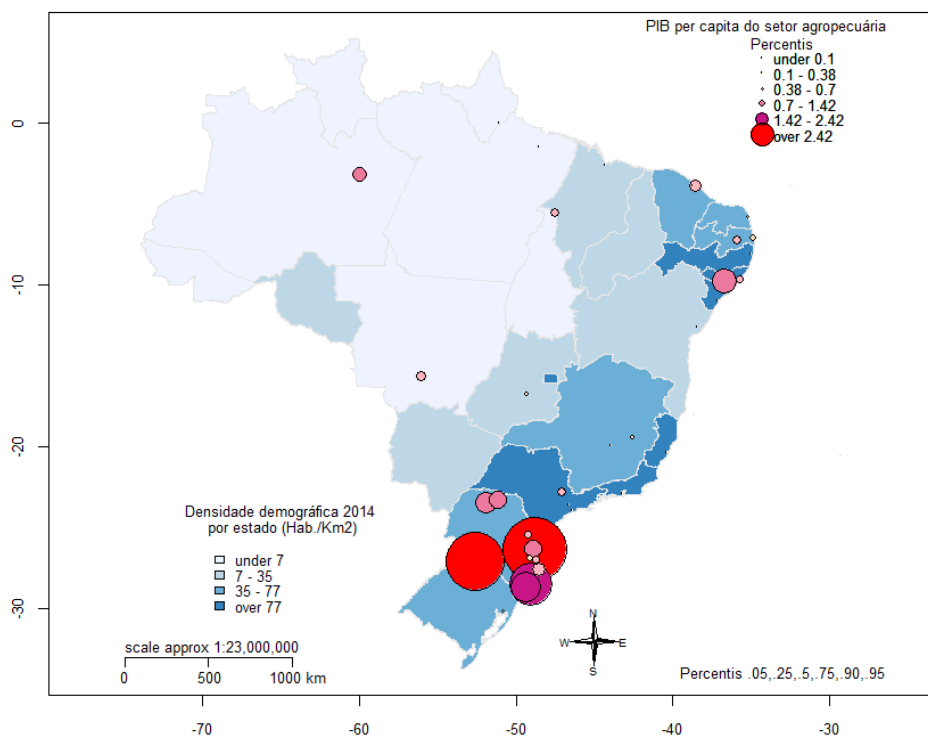
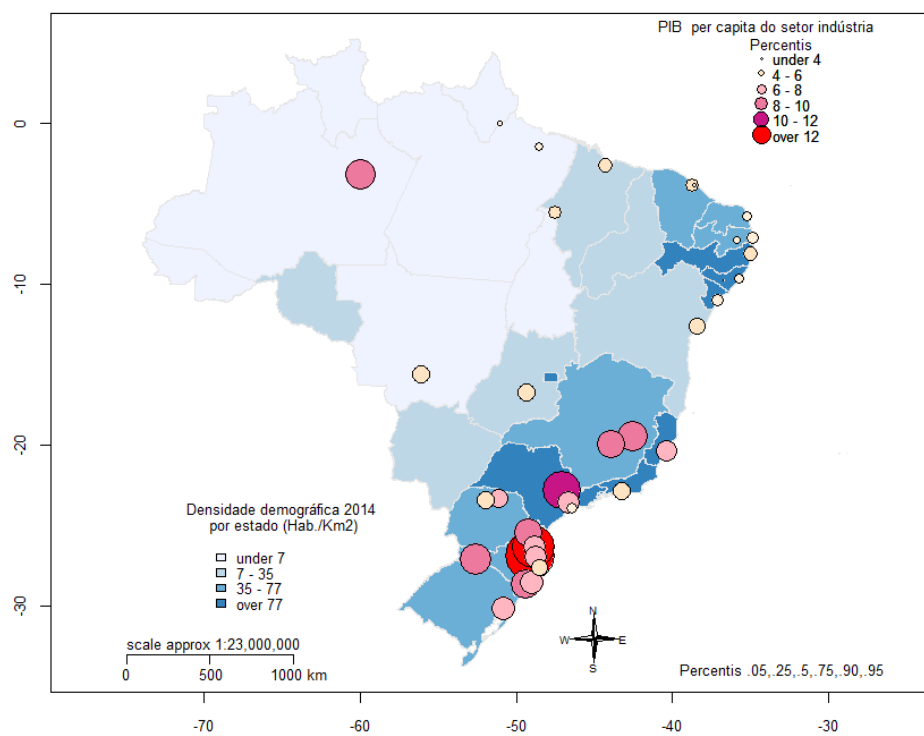


Figura 17 - PIB per capita do setor indústria (em R\$ 1.000) nas regiões metropolitanas, 2014



Elaboração própria

Figura 18 - Total de decolagens (aviões) nas regiões metropolitanas, 2014

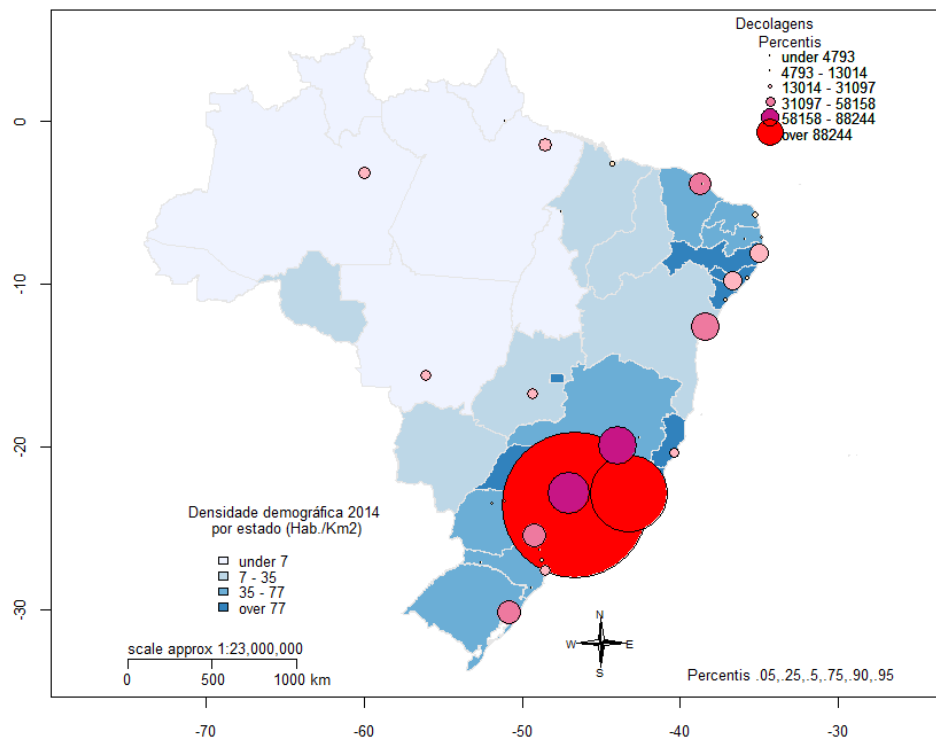
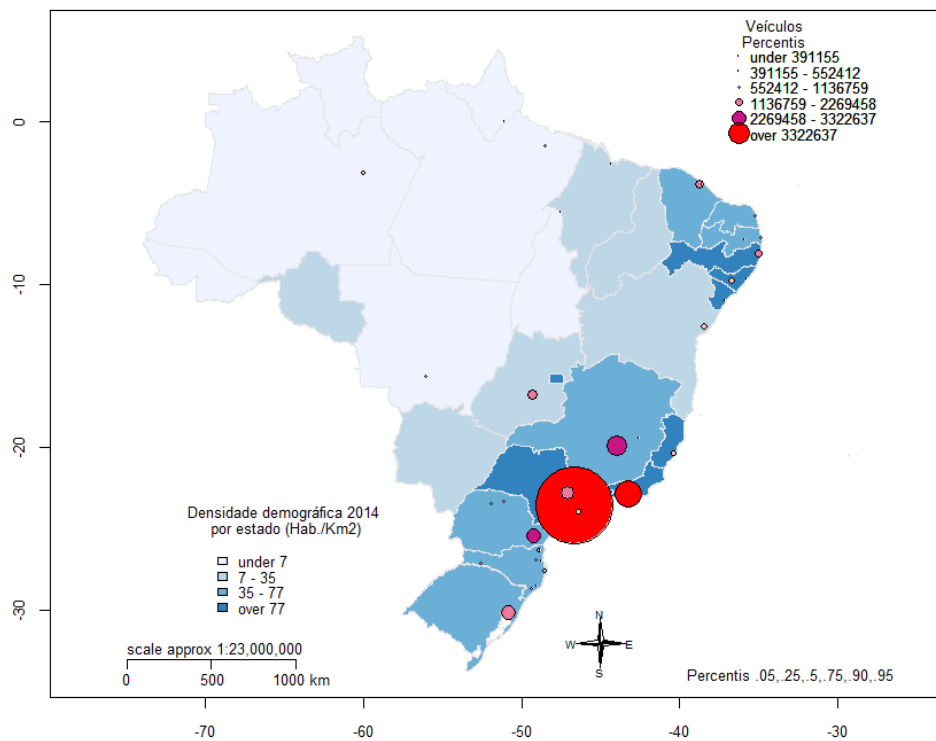


Figura 19 - Total de frotas de veículos nas regiões metropolitanas, 2014



Elaboração própria

Figura 20 - Taxa de crescimento anual de veículos novos (%) nas regiões metropolitanas, 2014

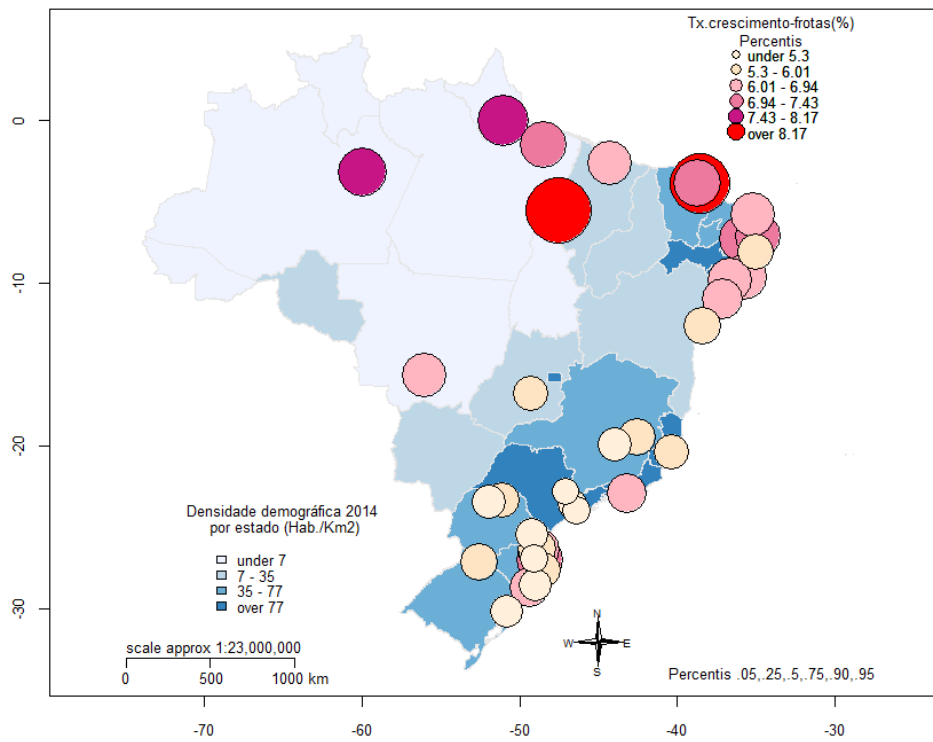
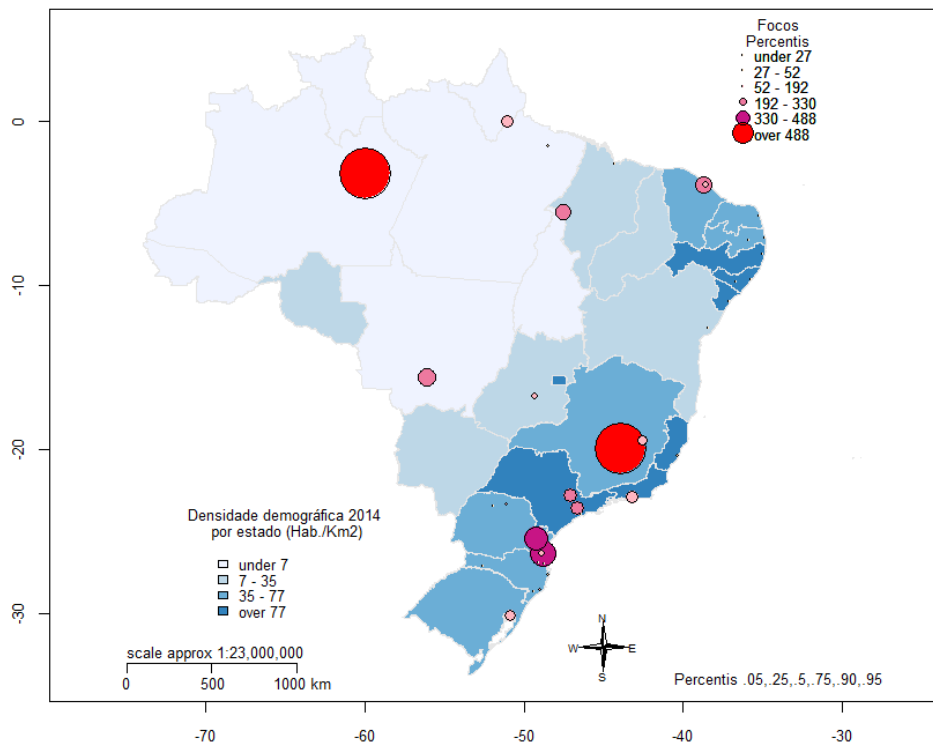


Figura 21 - Total de focos de queimadas nas regiões metropolitanas, 2014



Elaboração própria

APÊNDICE D – Mapas ilustrativos da distribuição das estações de monitoramento de PM_{10} nas Regiões metropolitanas do Brasil

Figura 22 - Estações de monitoramento de PM_{10} com informações disponíveis durante o período de 2001 a 2014 na região metropolitana de São Paulo – SP

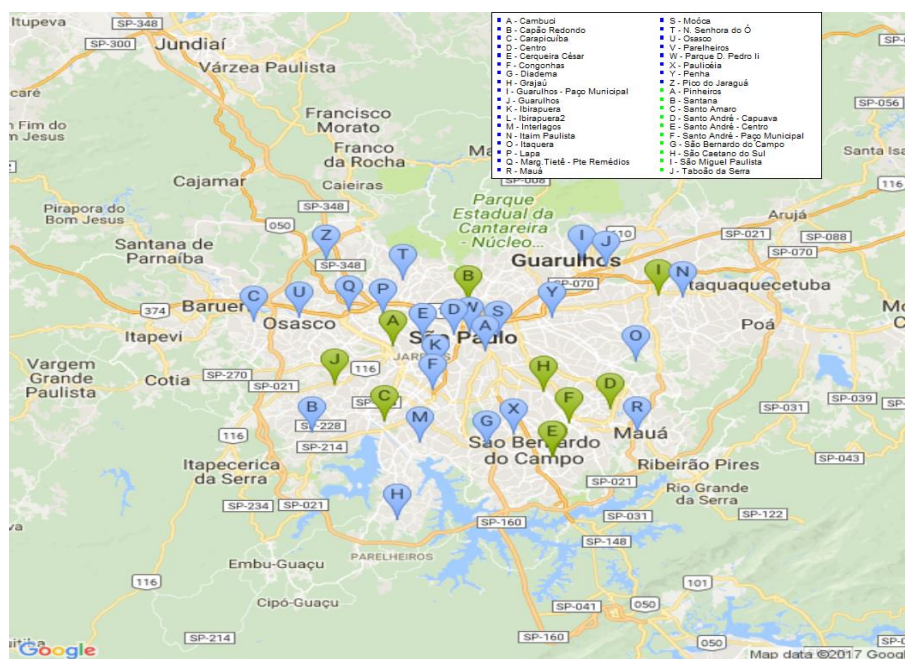
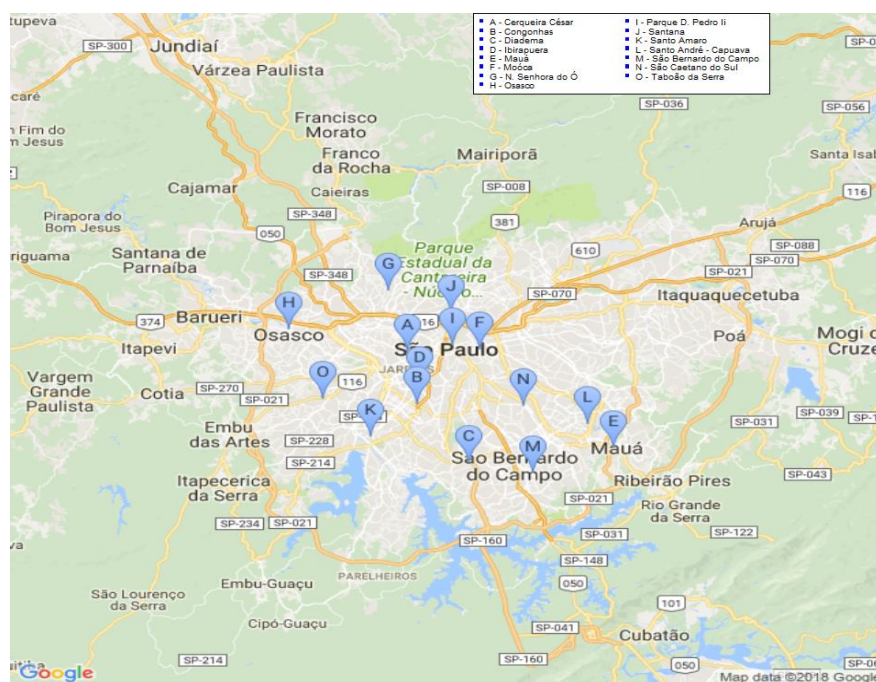


Figura 23 - Estações de monitoramento da região metropolitana de São Paulo - SP com menos de 40% de dados faltantes de PM_{10} durante o período de 2001 a 2014



Elaboração própria

Figura 24 - Estações de monitoramento de PM₁₀ com informações disponíveis durante o período de 2001 a 2014 na região metropolitana do Rio de Janeiro – RJ

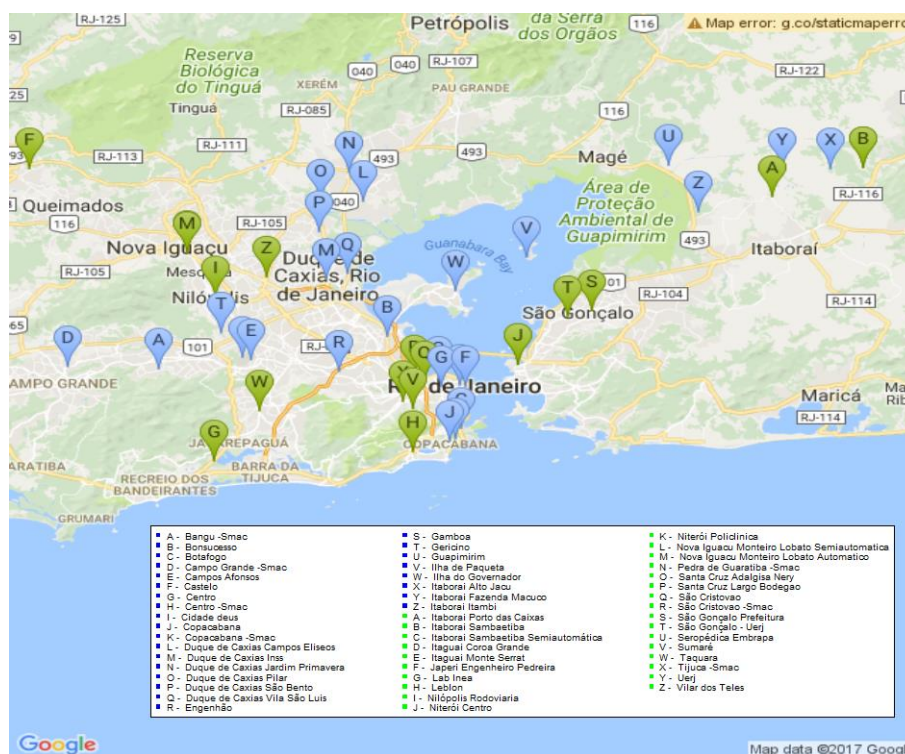
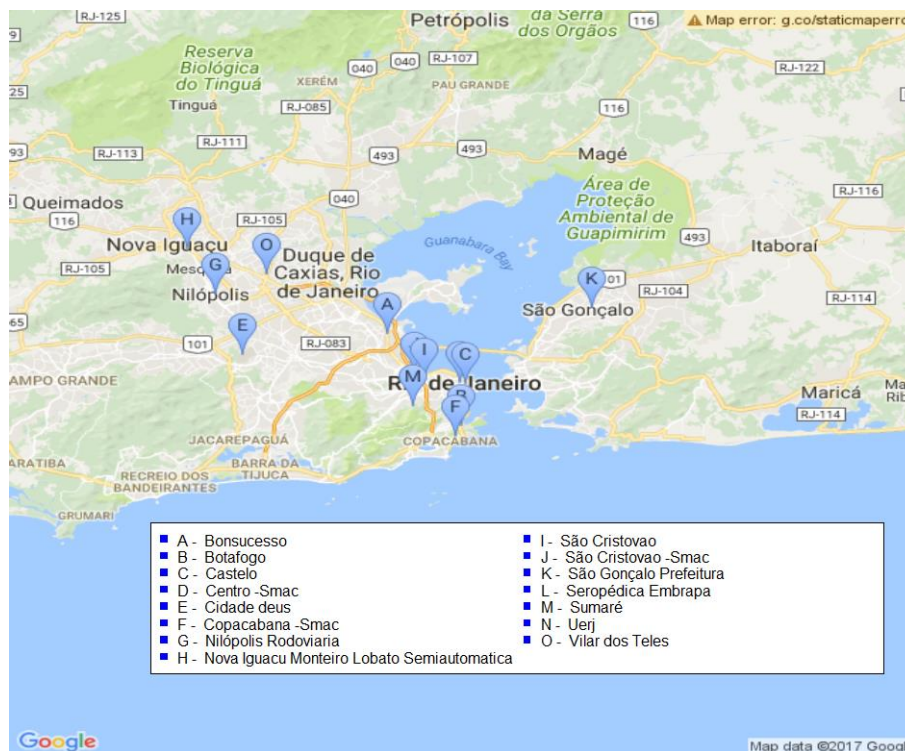


Figura 25 - Estações de monitoramento da região metropolitana do Rio de Janeiro - RJ com menos de 40% de dados faltantes de PM₁₀ durante o período de 2001 a 2014



Elaboração própria

Figura 26 - Estações de monitoramento de PM_{10} com informações disponíveis durante o período de 2001 a 2014 na Baixada Santista – SP

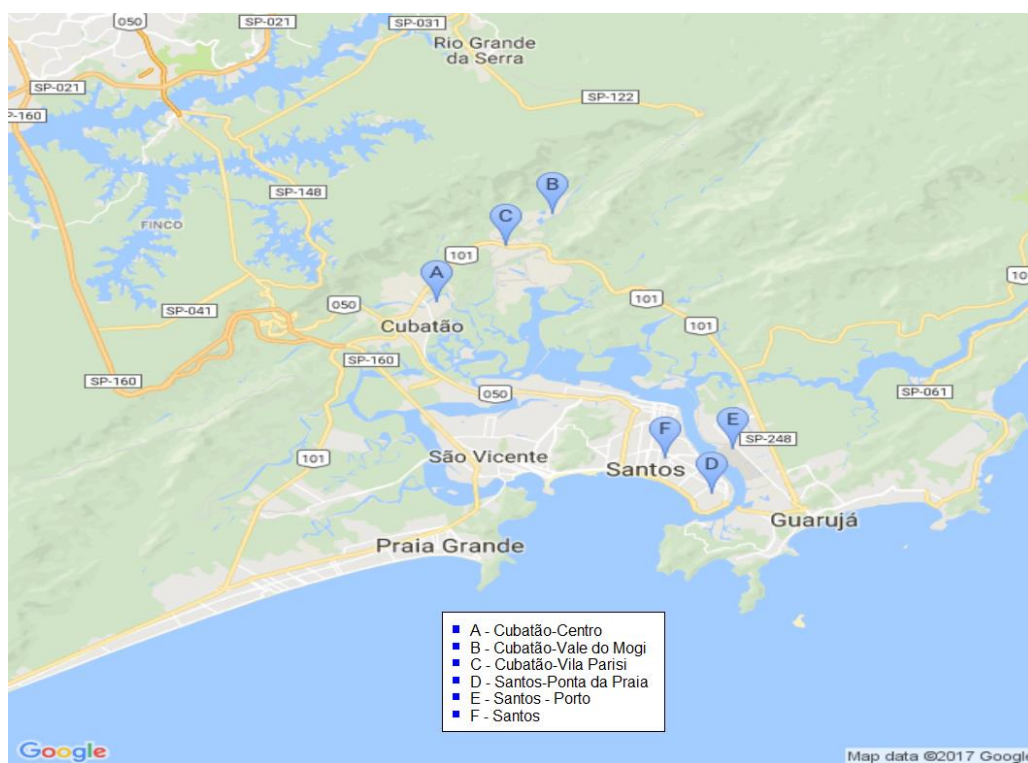
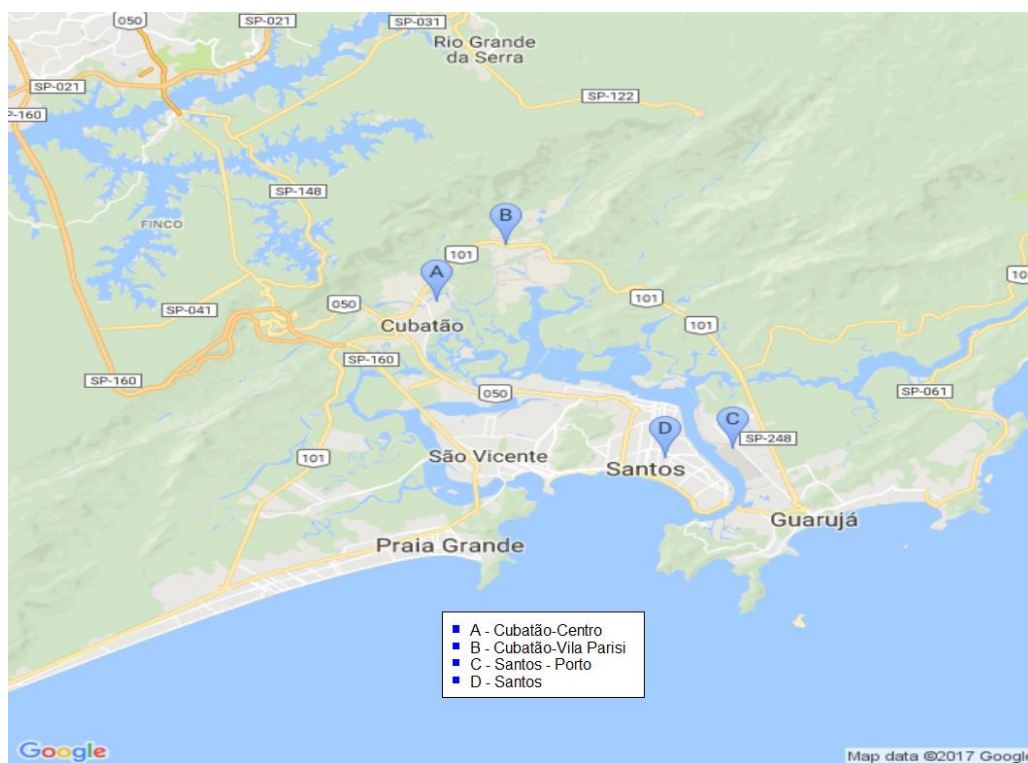
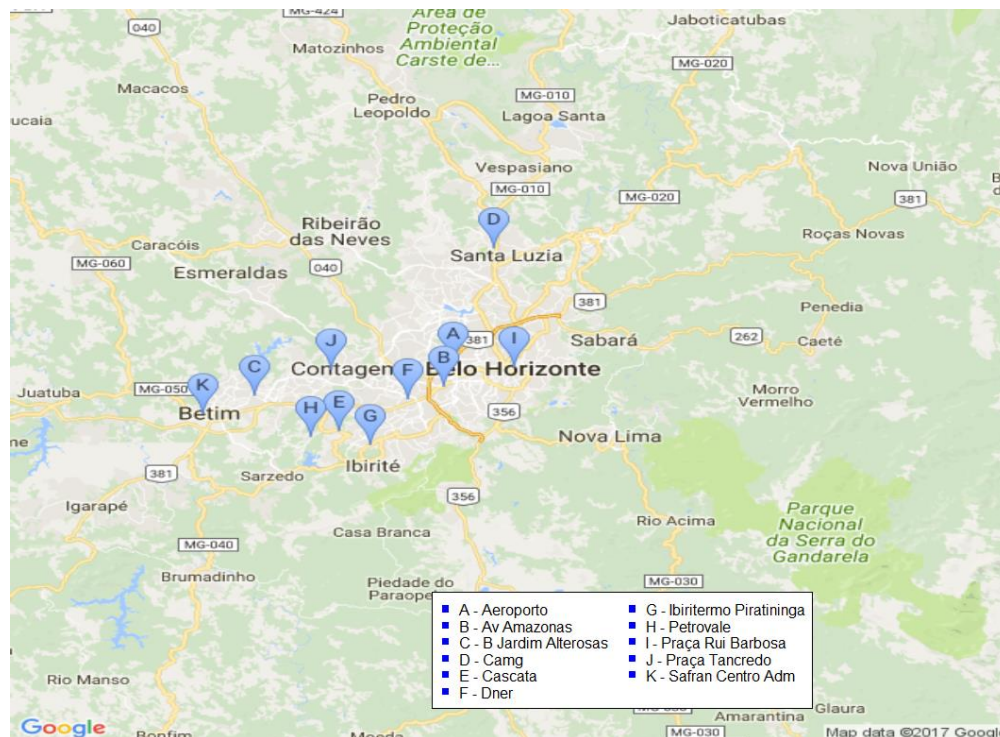


Figura 27 - Estações de monitoramento da Baixada Santista - SP com menos de 40% de dados faltantes de PM_{10} durante o período de 2001 a 2014



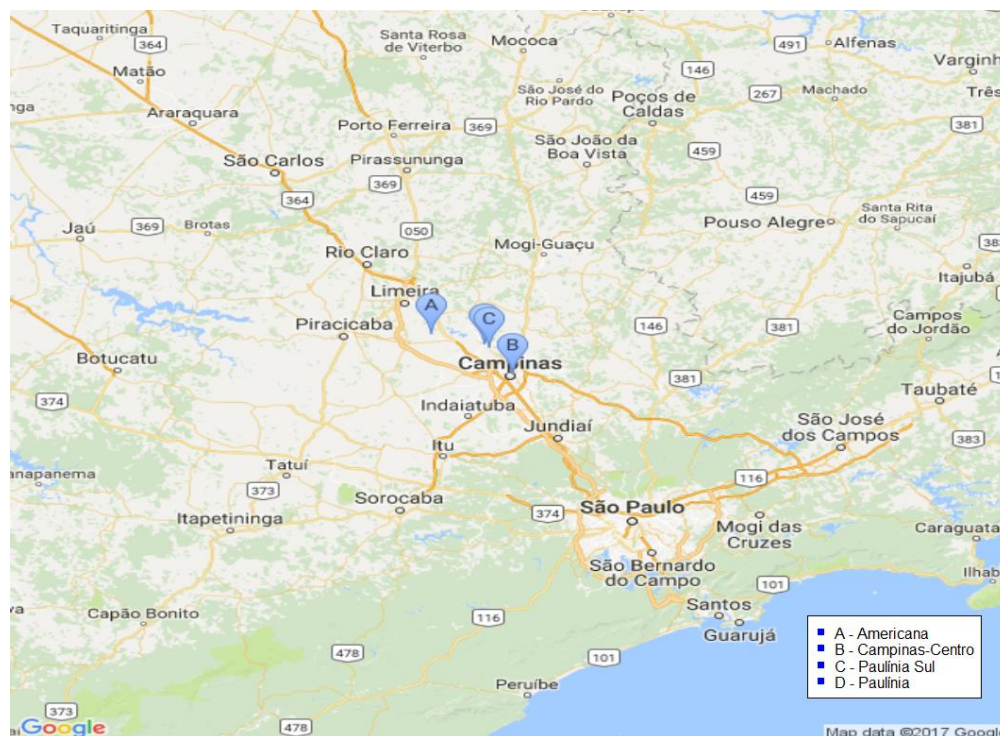
Elaboração própria

Figura 28 - Estações de monitoramento de PM_{10} com informações disponíveis durante o período de 2001 a 2014 em Belo Horizonte – MG



*Não houve estações de monitoramento com menos de 40% de dados faltantes de PM_{10} durante o período de 2001 a 2014.

Figura 29 - Estações de monitoramento de PM_{10} com informações disponíveis durante o período de 2001 a 2014 em Campinas – SP



Elaboração própria

Figura 30 - Estações de monitoramento de Campinas - SP com menos de 40% de dados faltantes de PM_{10} durante o período de 2001 a 2014

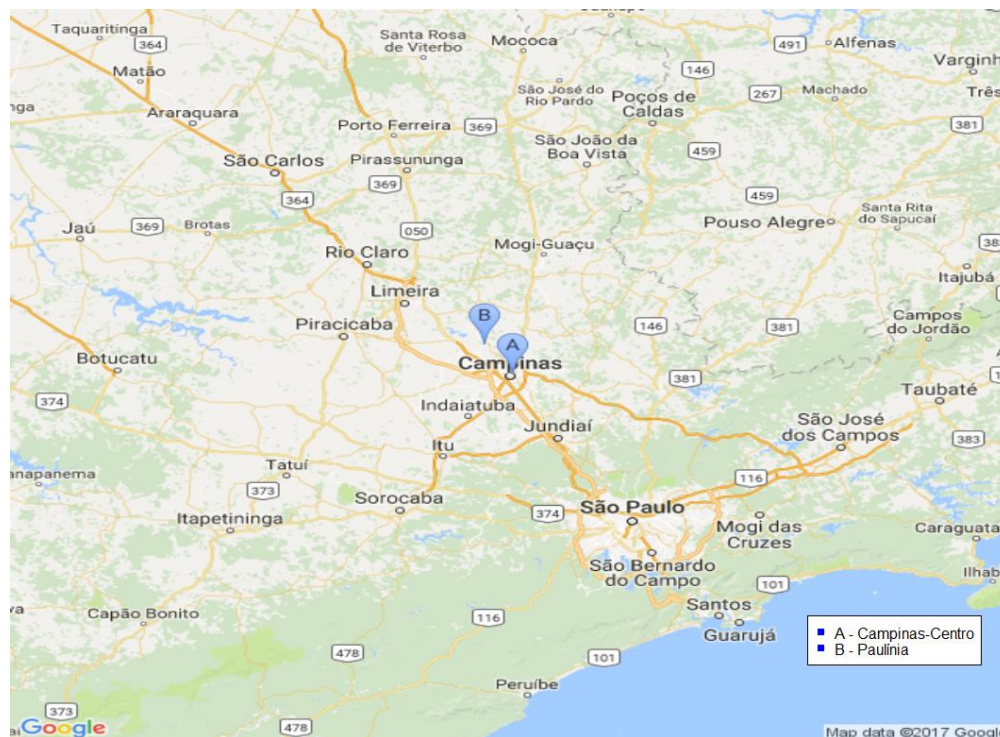
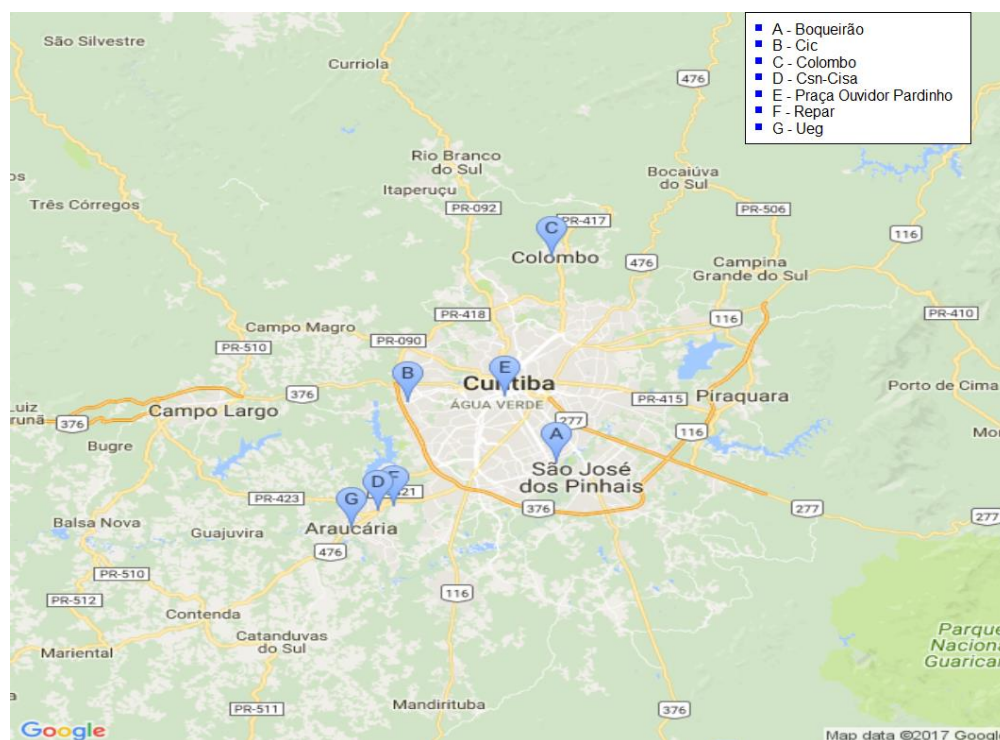


Figura 31 - Estações de monitoramento de PM_{10} com informações disponíveis durante o período de 2001 a 2014 em Curitiba – PR



*Apenas uma estação de monitoramento com menos de 40% de dados faltantes de PM_{10} durante o período de 2001 a 2014.
Elaboração própria

Figura 32 - Estações de monitoramento de Curitiba - PR com menos de 50% de dados faltantes de PM_{10} durante o período de 2001 a 2014

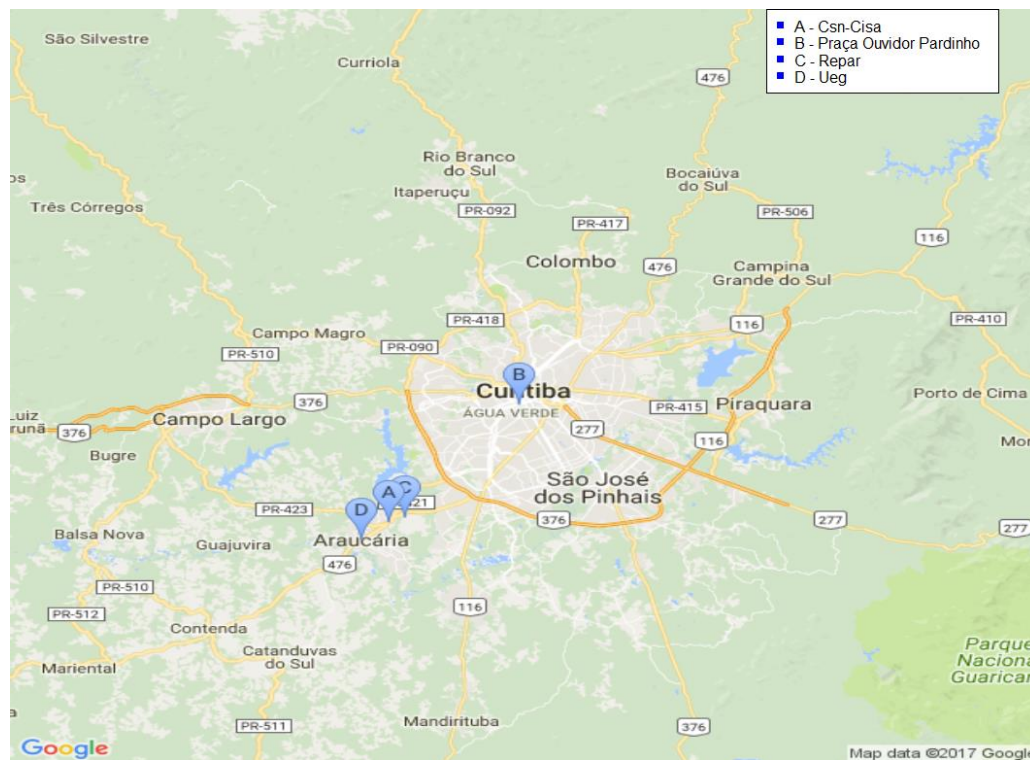
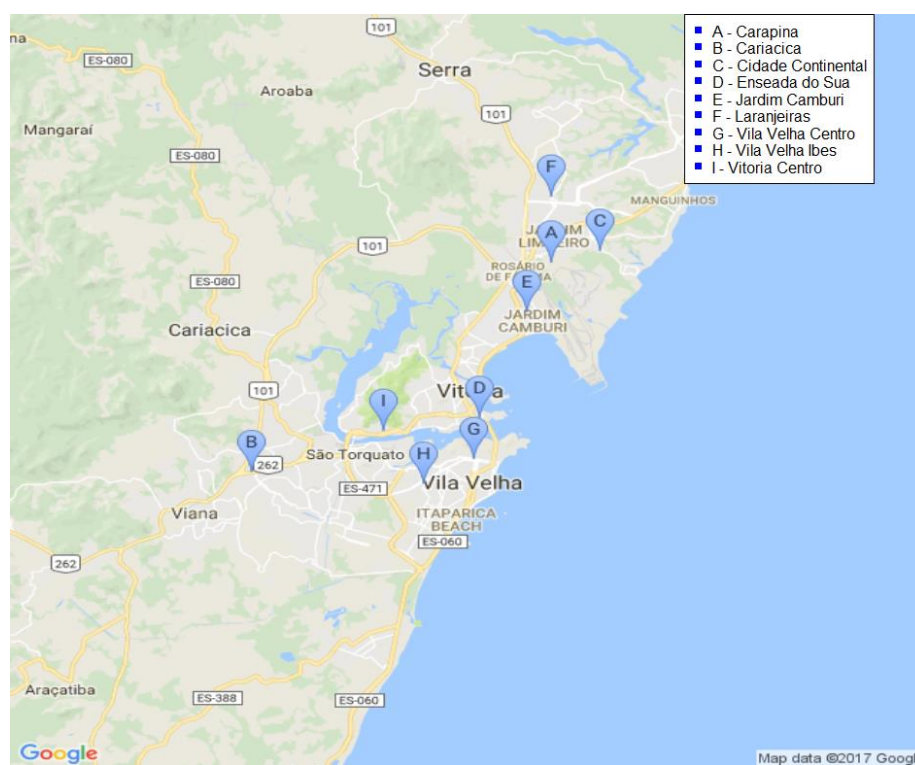


Figura 33 - Estações de monitoramento de PM_{10} com informações disponíveis durante o período de 2001 a 2014, Grande Vitória – ES



Elaboração própria

Figura 34 - Estações de monitoramento da Grande Vitória - ES com menos de 40% de dados faltantes de PM_{10} durante o período de 2001 a 2014

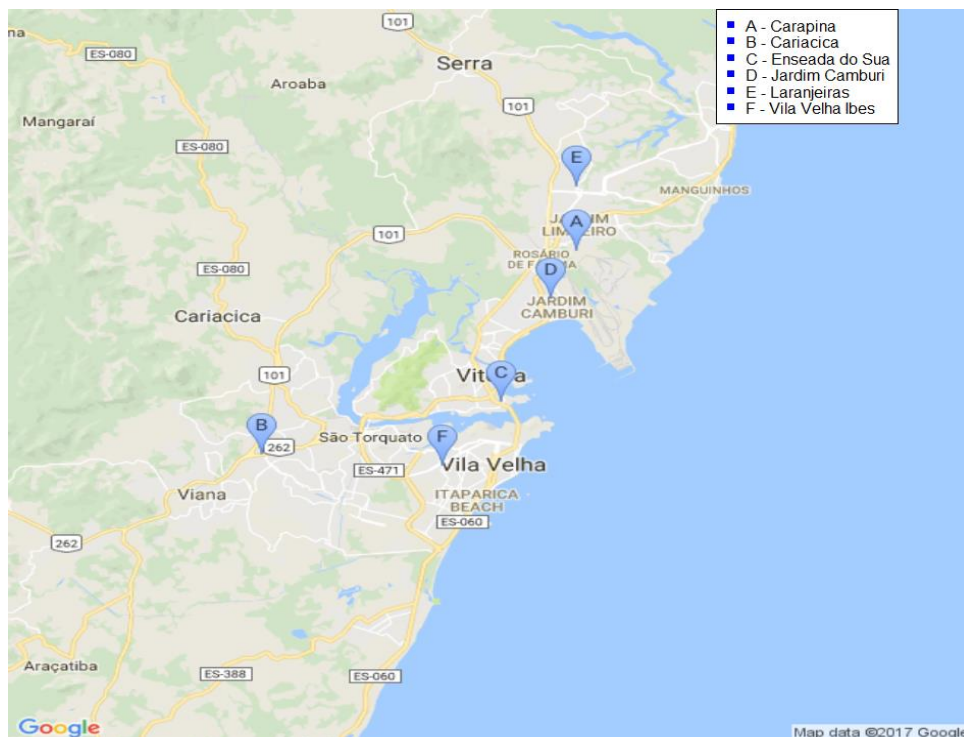
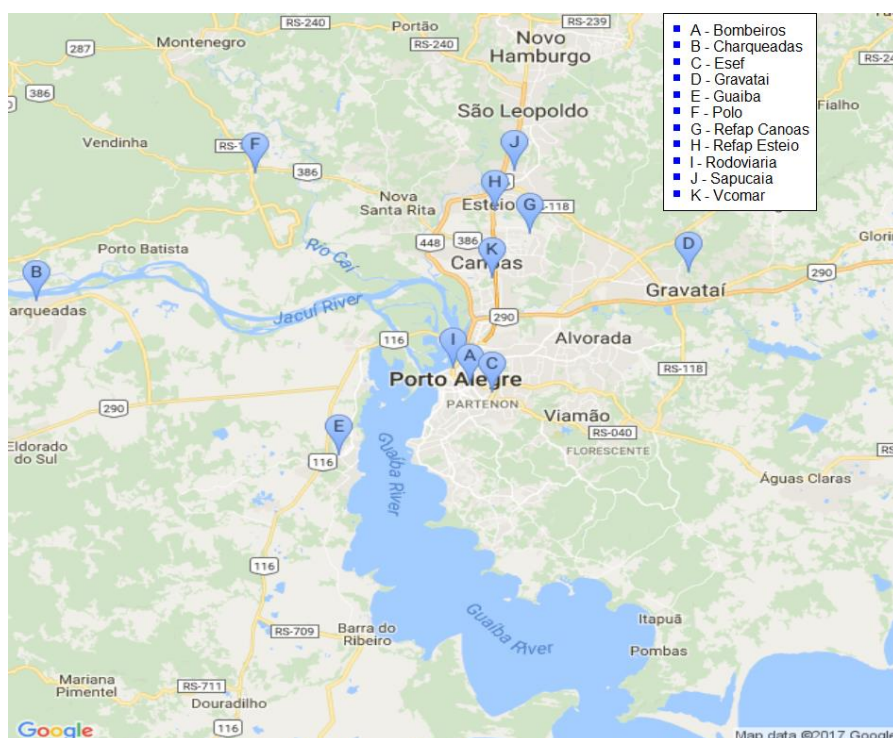


Figura 35 - Estações de monitoramento de PM_{10} com informações disponíveis durante o período de 2001 a 2014 na Região metropolitana de Porto Alegre – RS



*Não houve estações de monitoramento com menos de 40% de dados faltantes de PM_{10} durante o período de 2001 a 2014.

Elaboração própria

Figura 36 - Estações de monitoramento de PM₁₀ com informações disponíveis durante o período de 2001 a 2014 na Região metropolitana de Salvador – BA

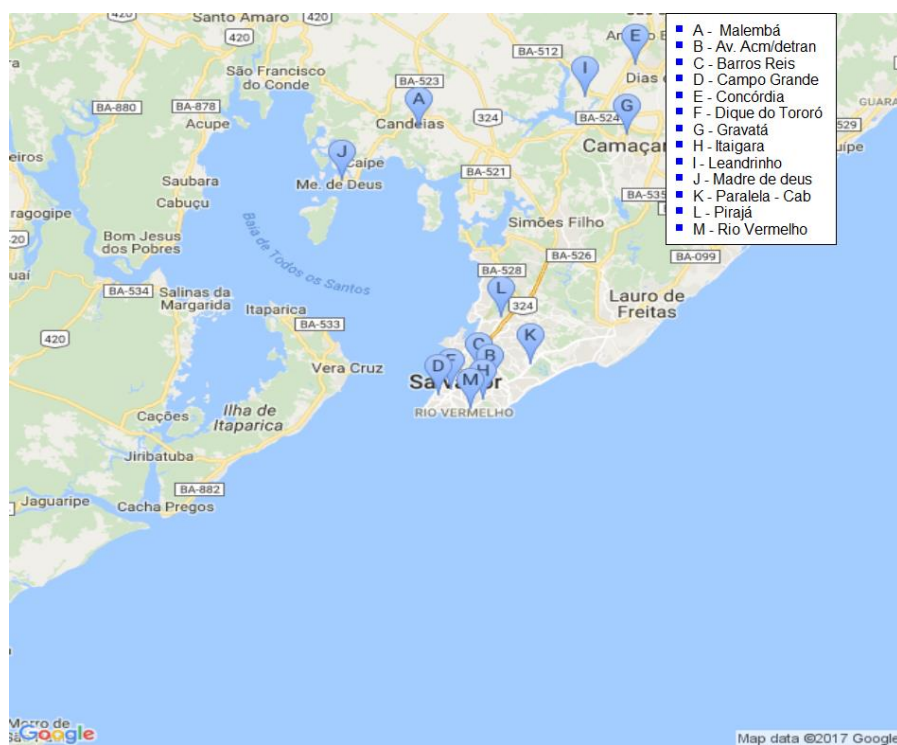
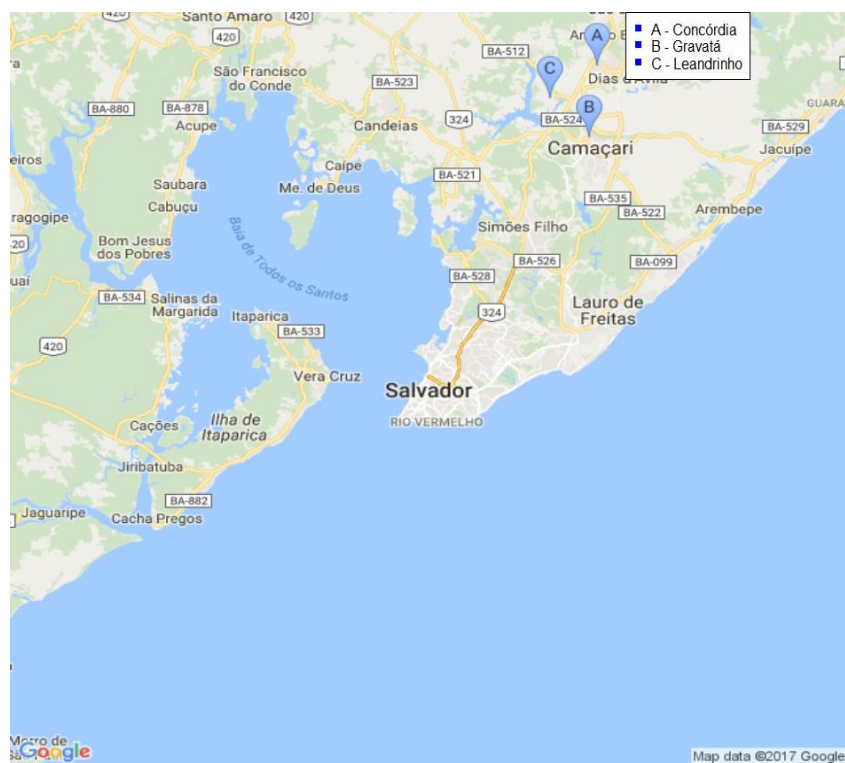


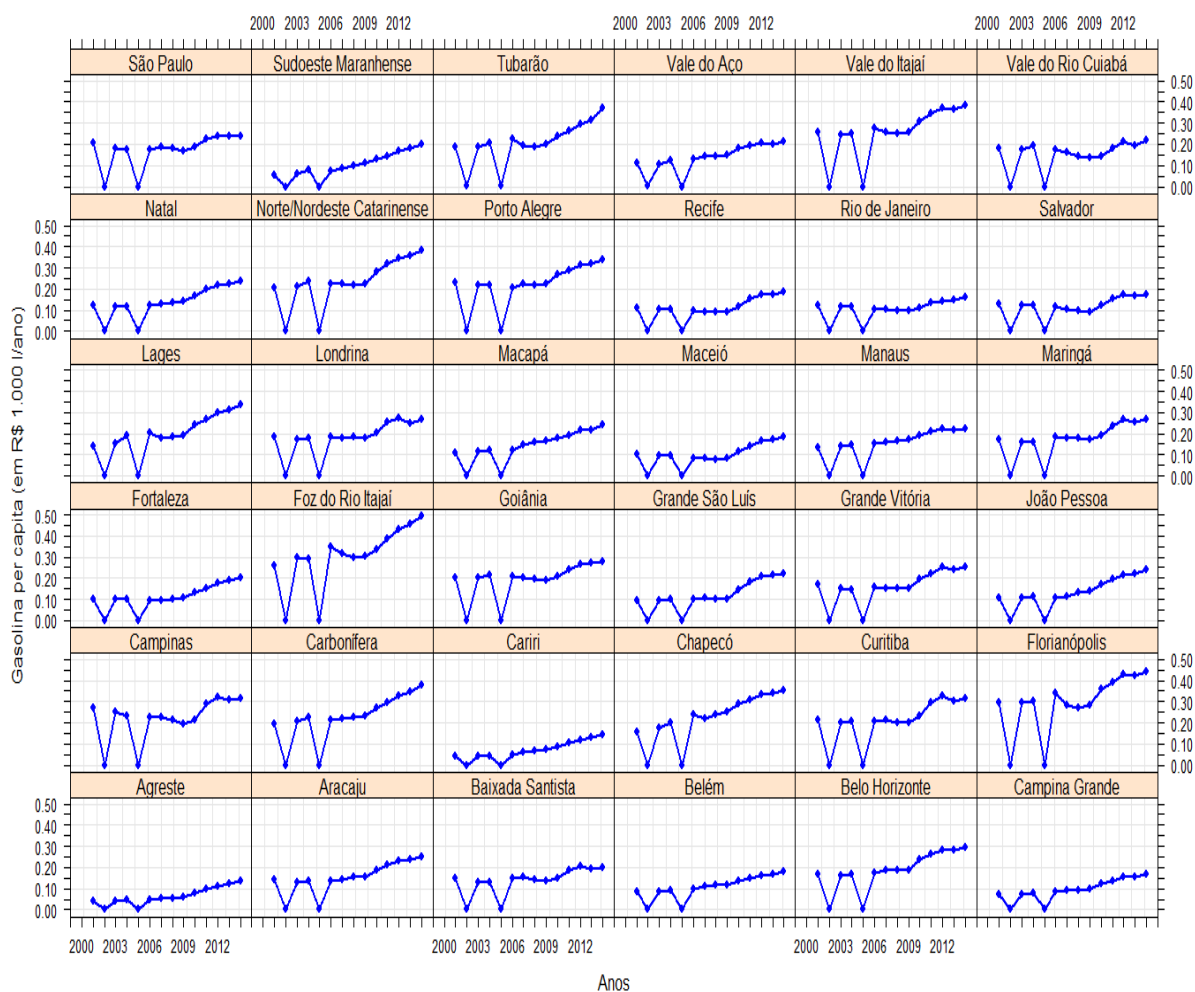
Figura 37 - Estações de monitoramento da Região metropolitana de Salvador – BA com menos de 40% de dados faltantes de PM₁₀ durante o período de 2001 a 2014



Elaboração própria

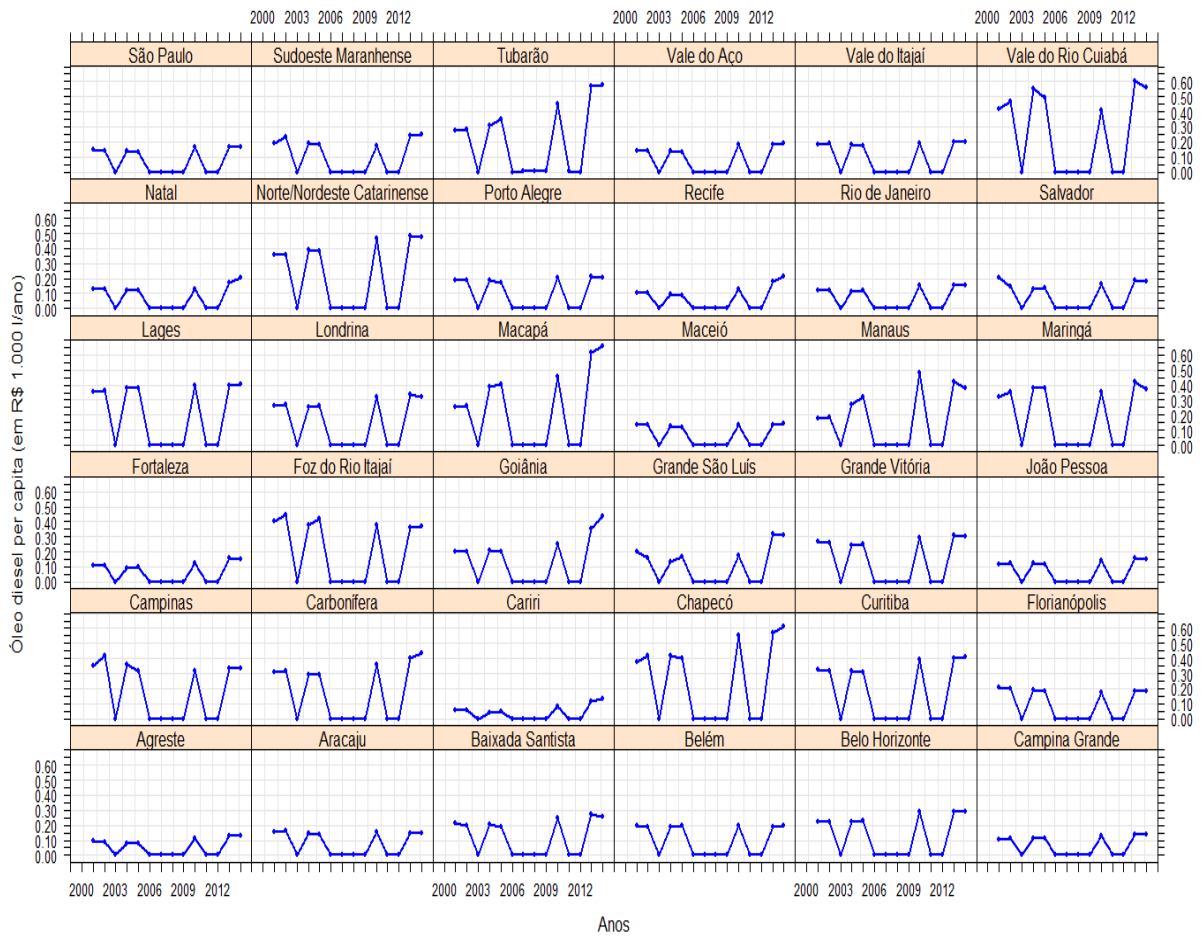
APÊNDICE E – Gráficos ilustrativos da evolução dos fatores contextuais nas Regiões metropolitanas brasileiras durante o período 2001 a 2014

Gráfico 4 - Evolução do consumo anual de gasolina per capita (em R\$ 1.000 l/ano) - Regiões metropolitanas, 2001 a 2014



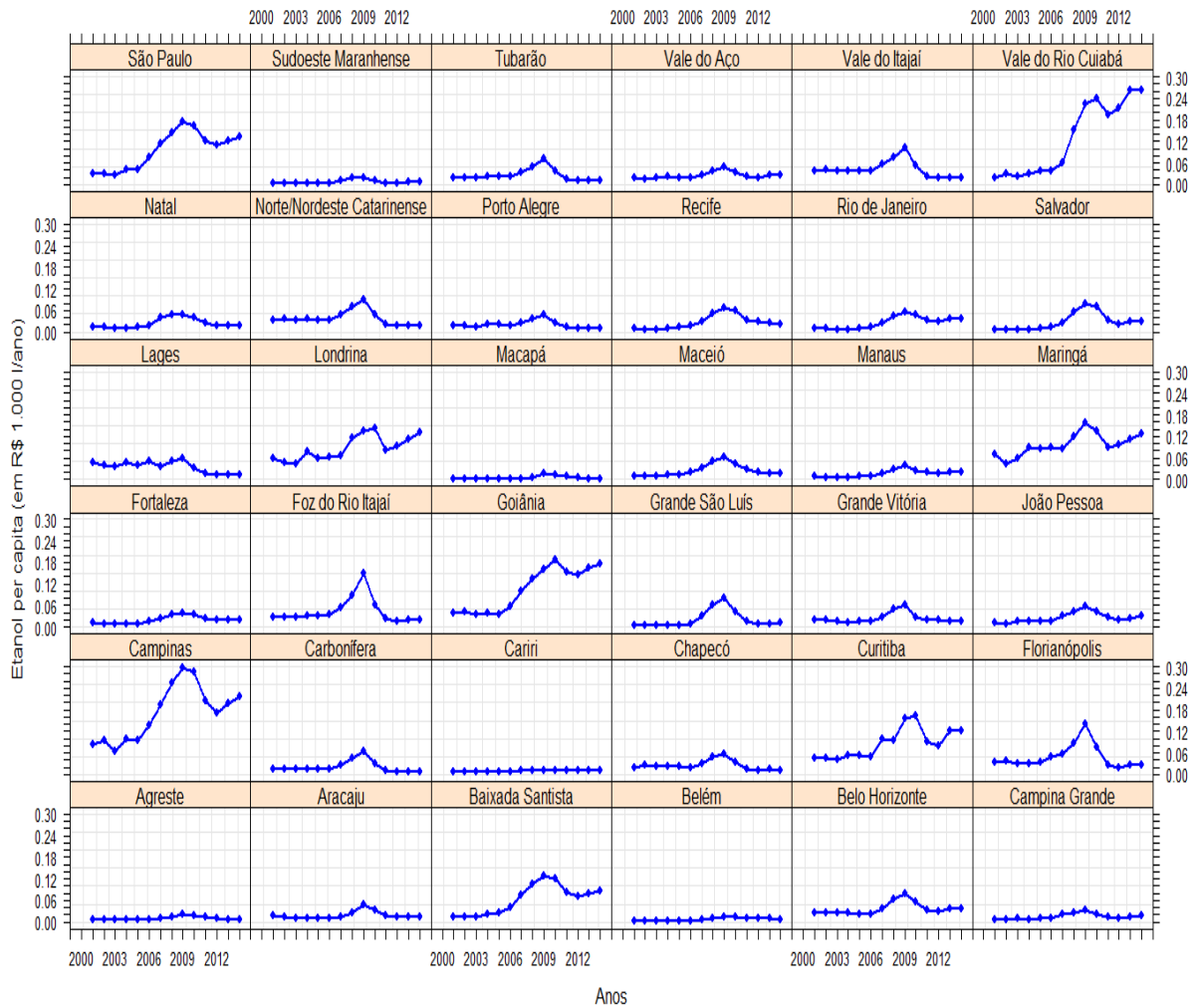
Elaboração própria

Gráfico 5 - Evolução do consumo anual de óleo diesel per capita (em R\$ 1.000 l/ano) - Regiões metropolitanas, 2001 a 2014



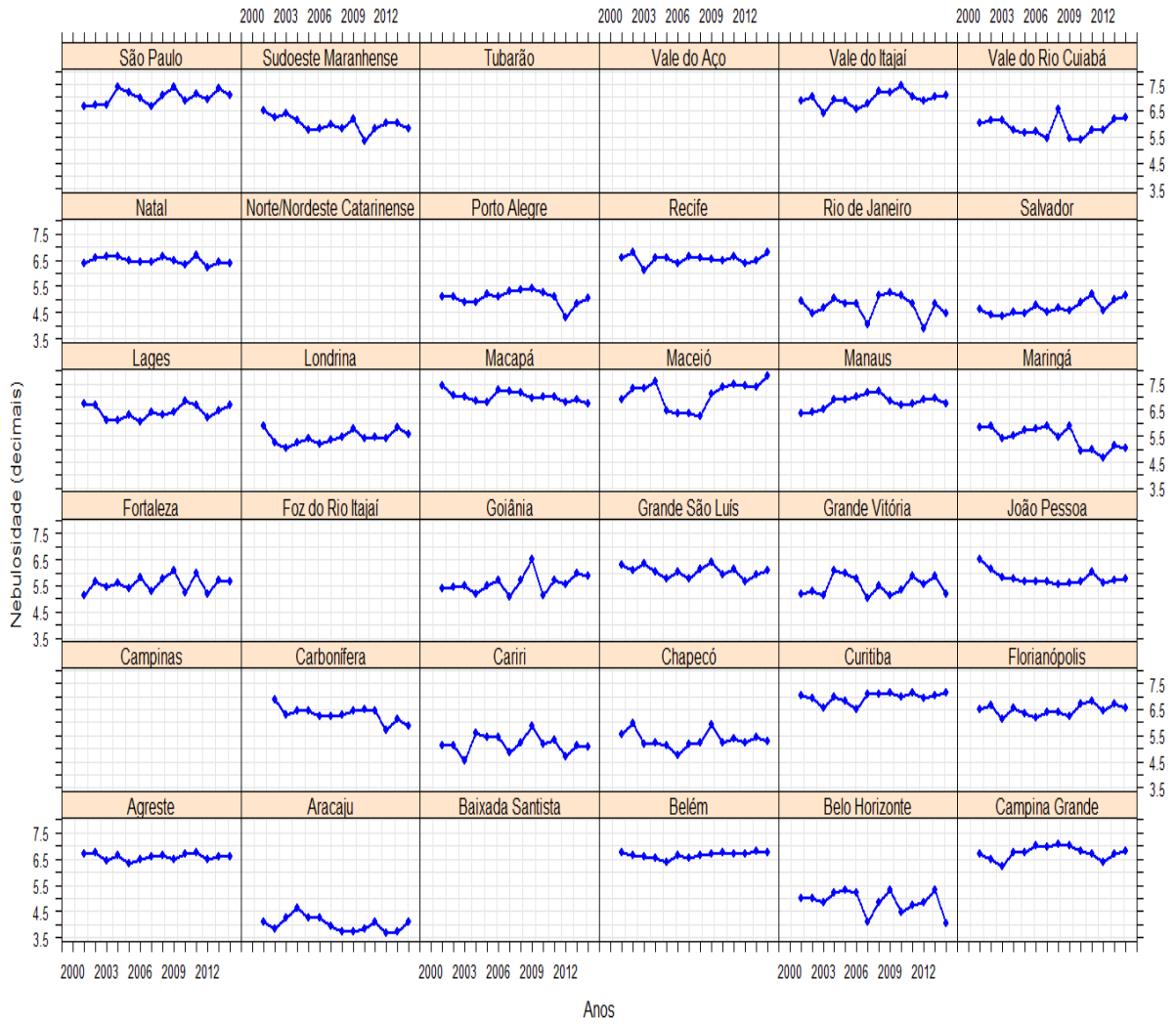
Elaboração própria

Gráfico 6 - Evolução do consumo anual de etanol per capita (em R\$ 1.000 l/ano) - Regiões metropolitanas - 2001 a 2014



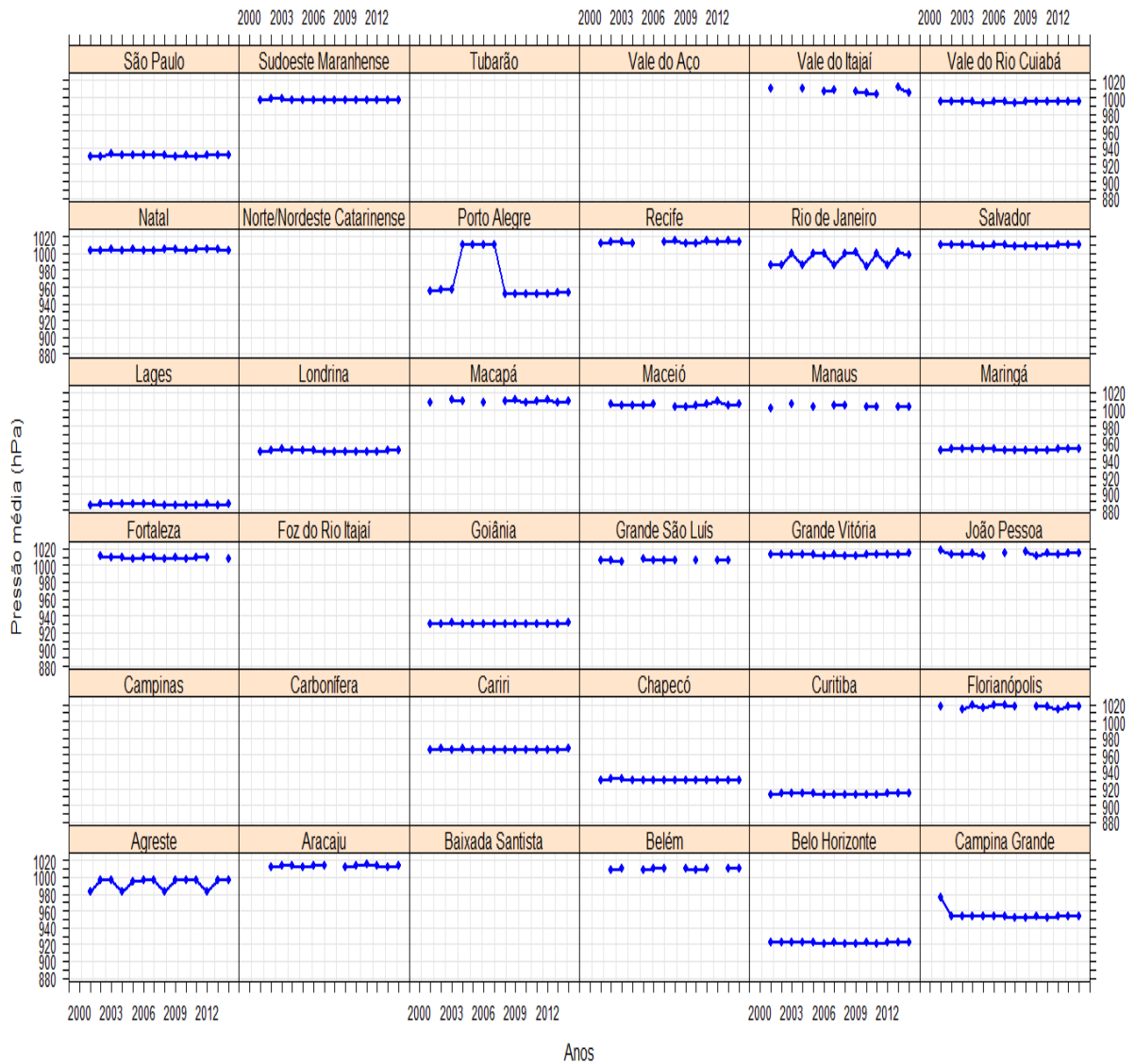
Elaboração própria

Gráfico 7 - Evolução da média anual de nebulosidade (decimais) – Regiões metropolitanas, 2001 a 2014



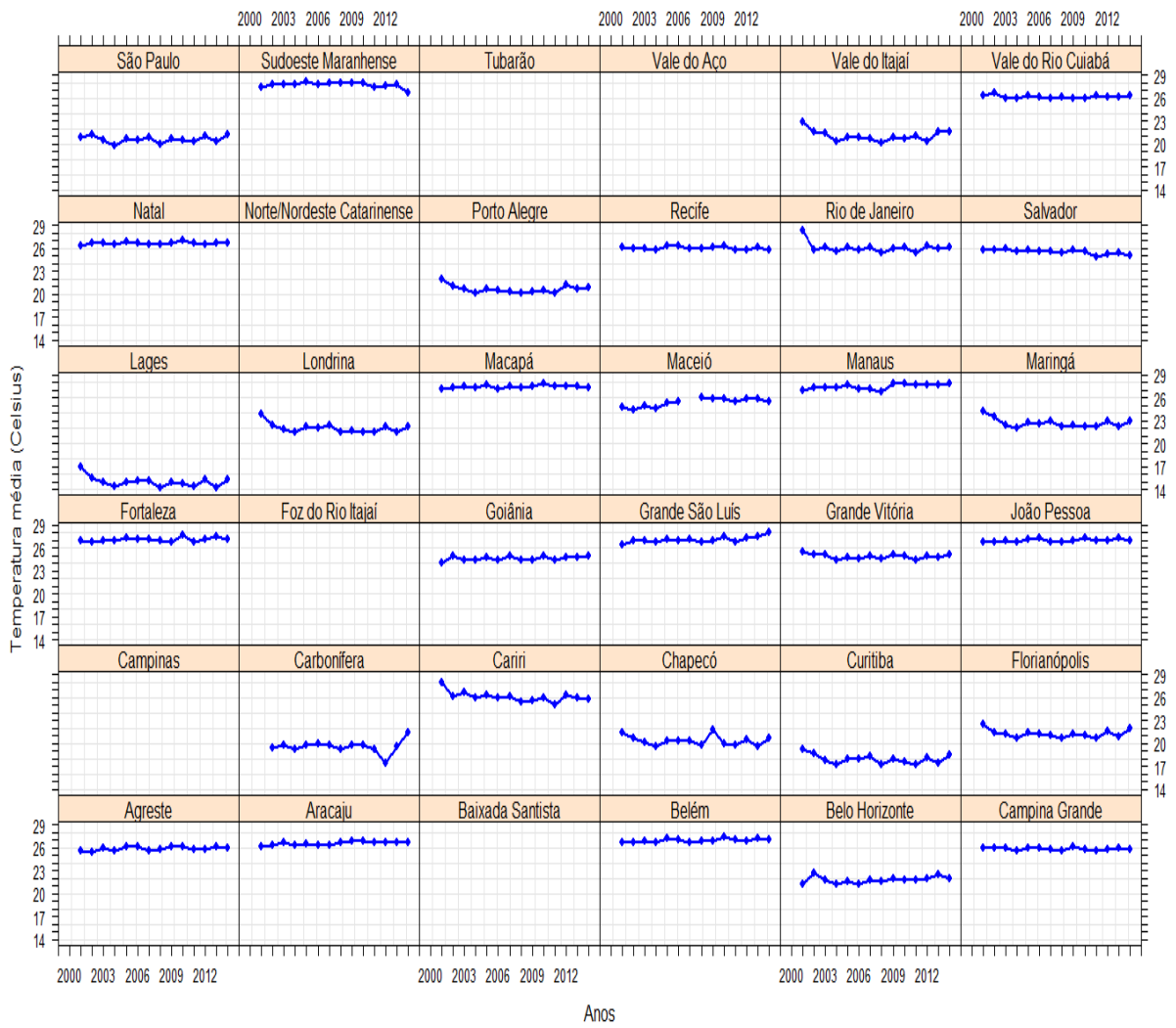
Elaboração própria

Gráfico 8 - Evolução da média anual da pressão barométrica (hPa) - Regiões metropolitanas, 2001 a 2014



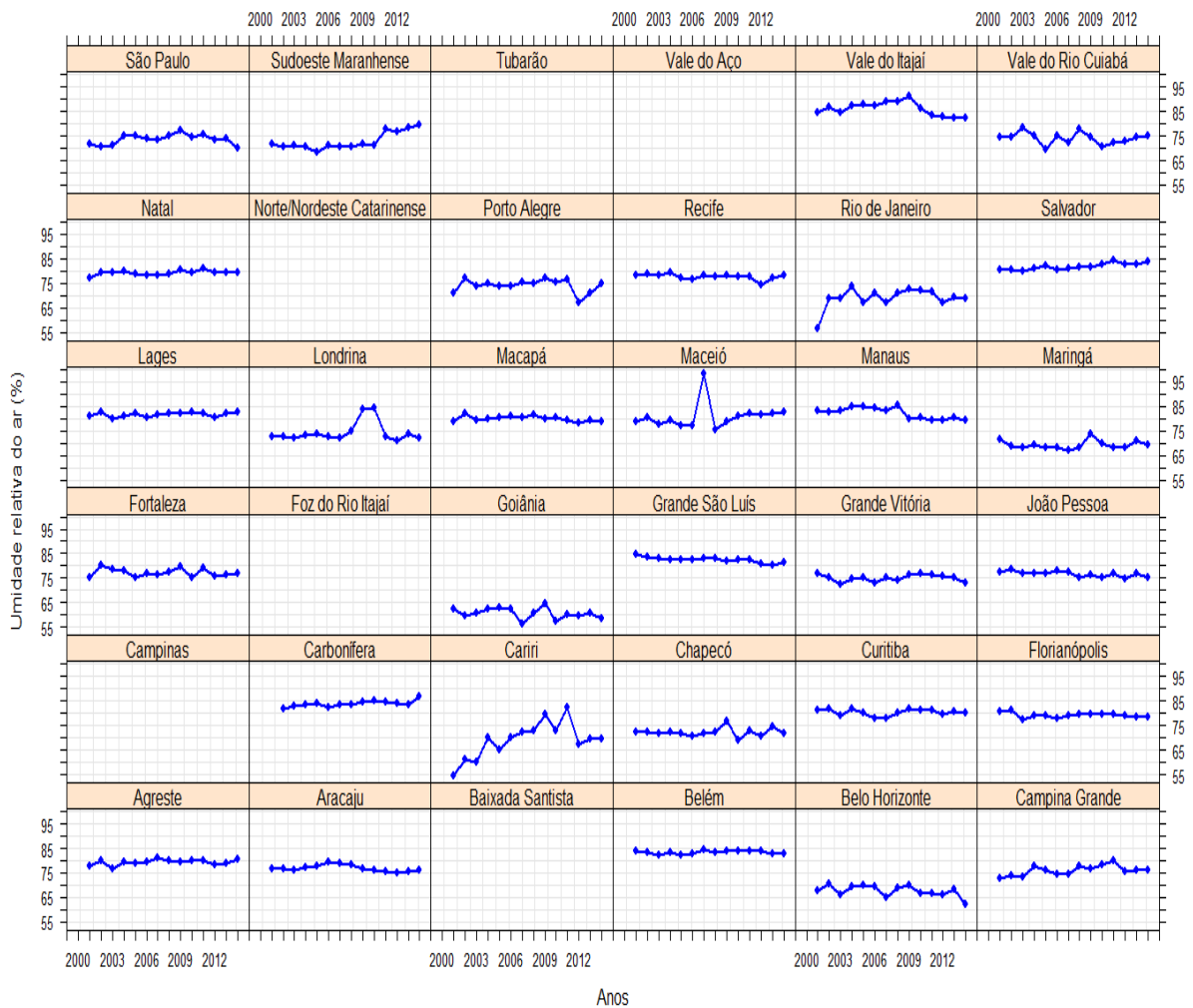
Elaboração própria

Gráfico 9 - Evolução da média anual da temperatura média (Celsius) – Regiões metropolitanas, 2001 a 2014



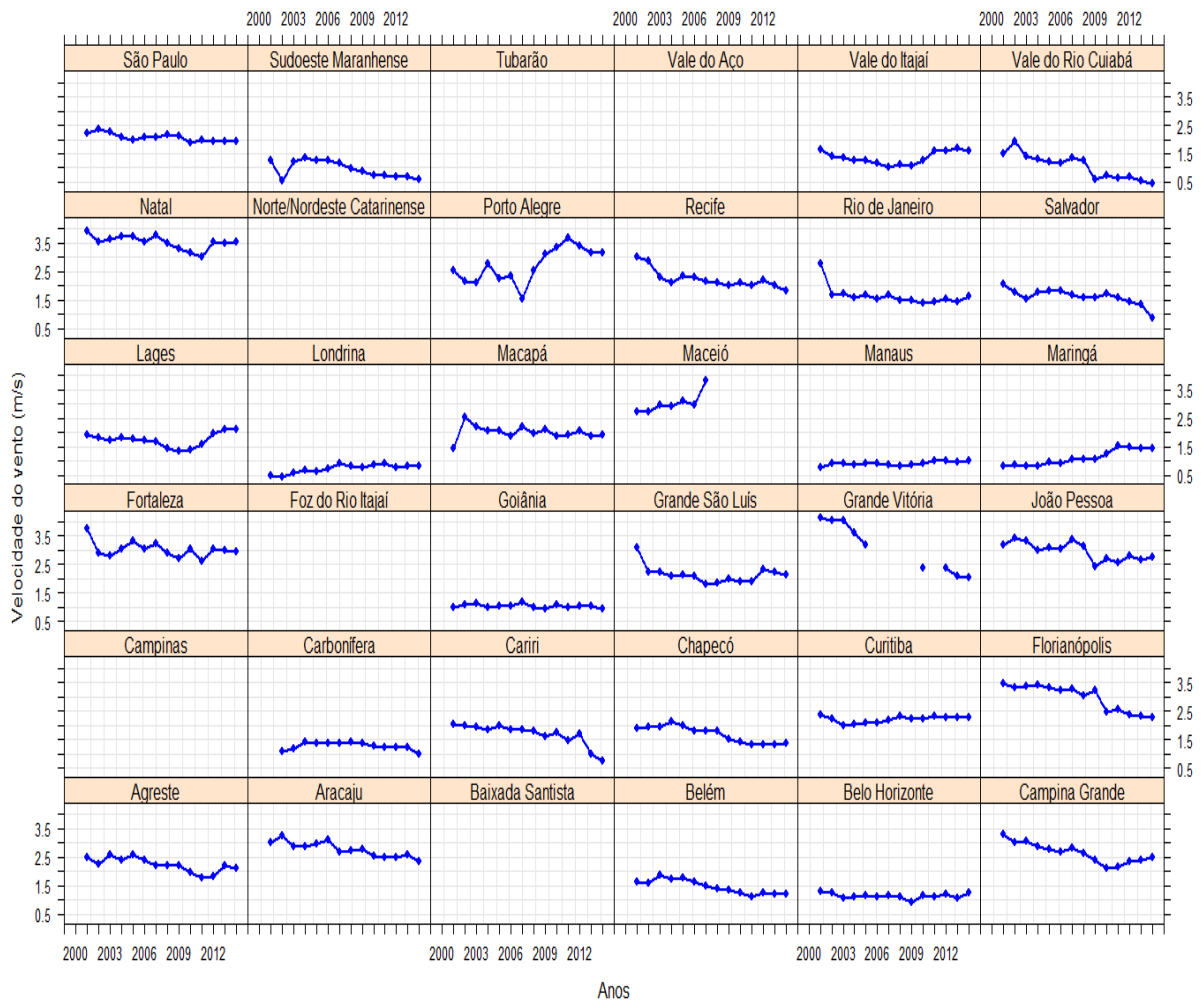
Elaboração própria

Gráfico 10 - Evolução da média anual da umidade relativa do ar (%) - Regiões metropolitanas, 2001 a 2014



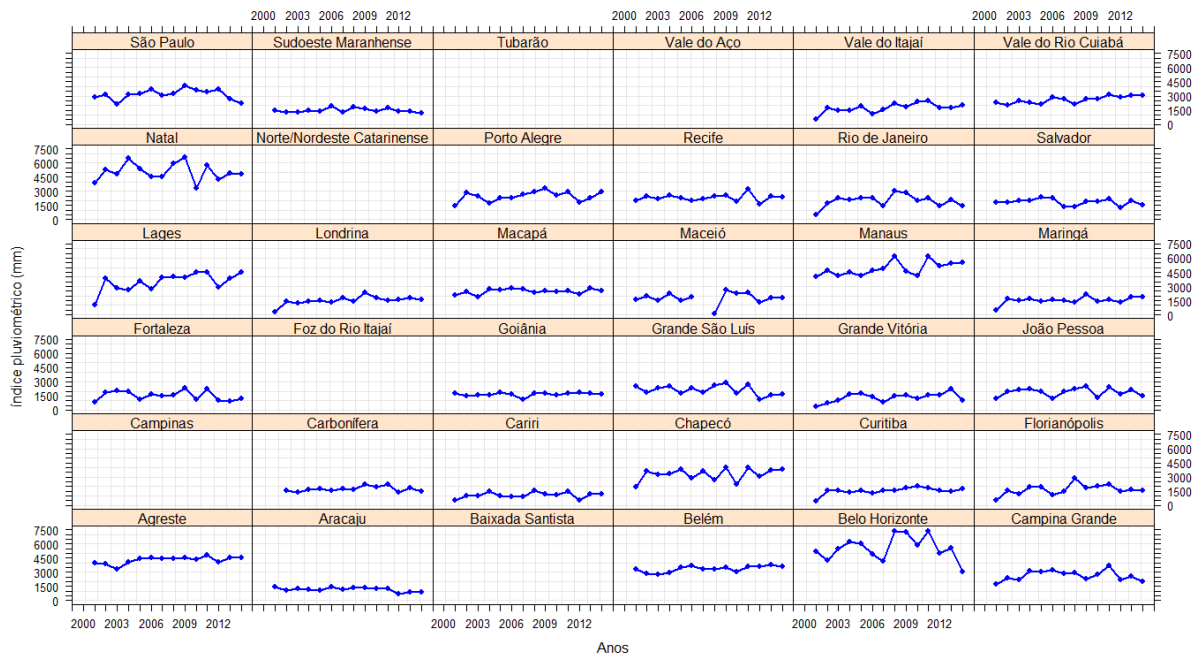
Elaboração própria

Gráfico 11 - Evolução da média anual da velocidade do vento (m/s) - Regiões metropolitanas, 2001 a 2014



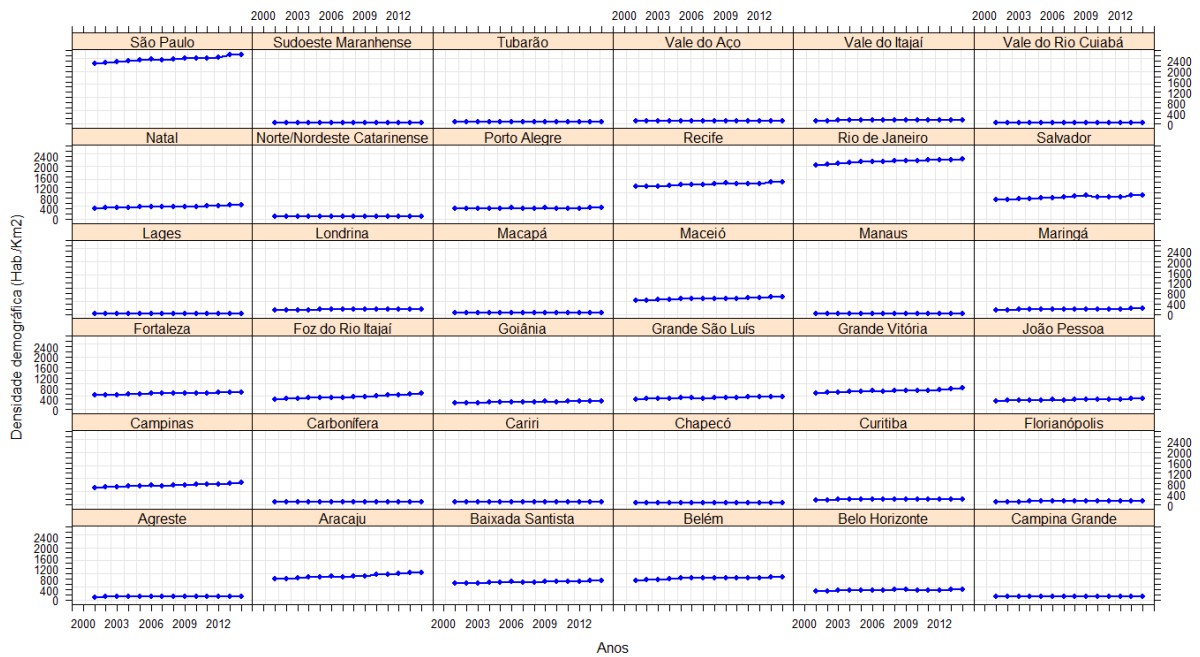
Elaboração própria

Gráfico 12 - Evolução da média anual do índice pluviométrico (mm) - Regiões metropolitanas, 2001 a 2014



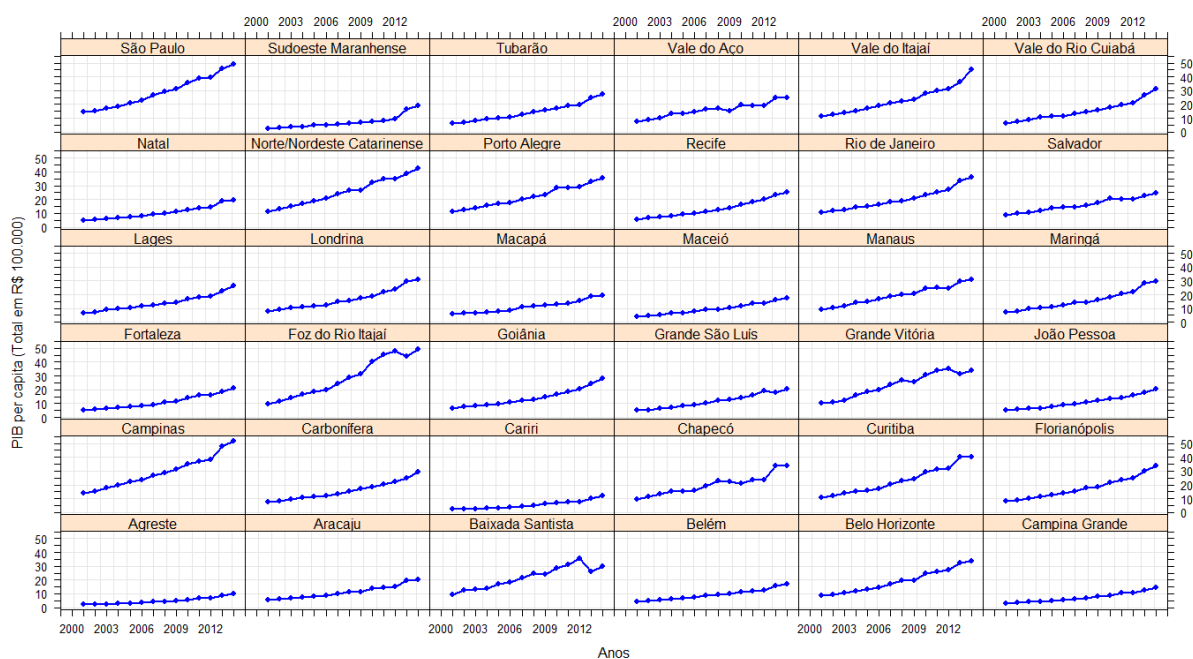
Elaboração própria

Gráfico 13 - Evolução anual da densidade demográfica (Hab./Km2) - Regiões metropolitanas, 2001 a 2014



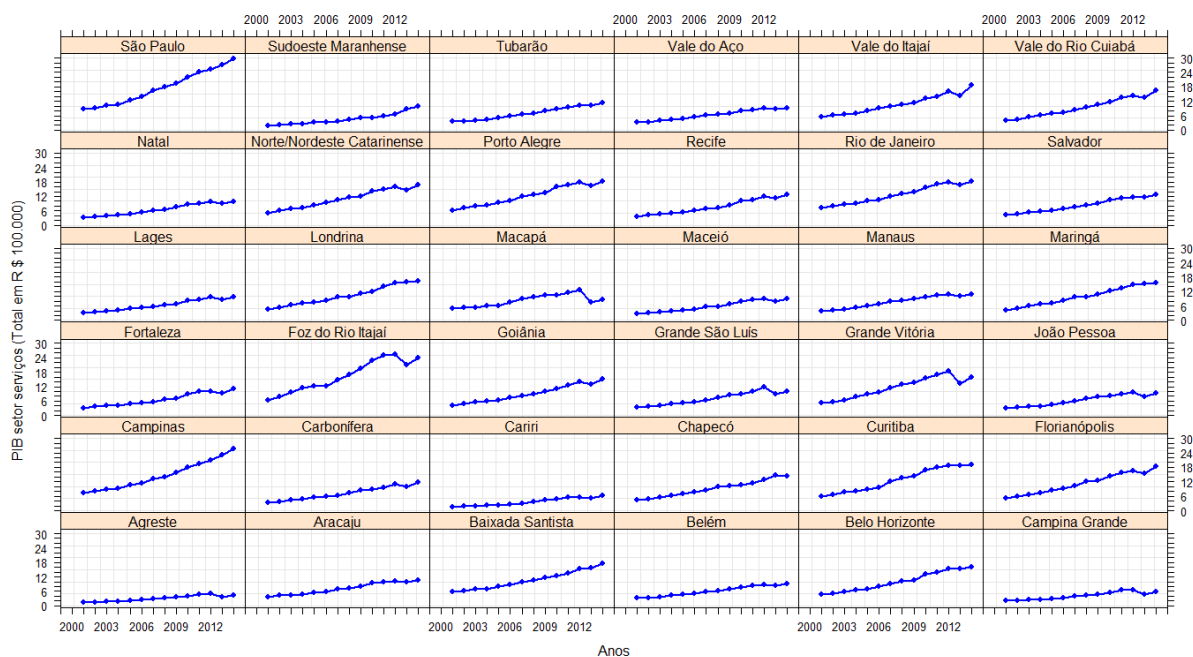
Elaboração própria

Gráfico 14 - Evolução anual do PIB per capita (Total em R \$ 100.000) - Regiões metropolitanas, 2001 a 2014



Elaboração própria

Gráfico 15 - Evolução anual do PIB setor serviços (Total em R\$ 100.000) - Regiões metropolitanas, 2001 a 2014



Elaboração própria

Gráfico 16 - Evolução anual do PIB setor agropecuária (total em R\$ 100.000) - Regiões metropolitanas, 2001 a 2014

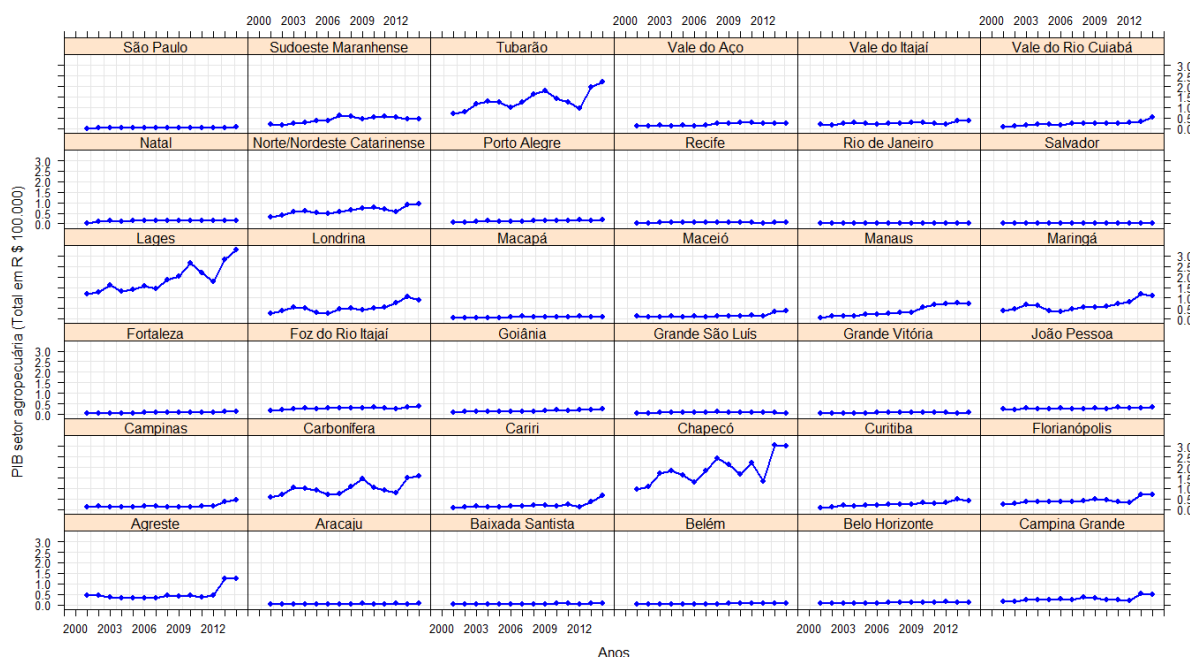


Gráfico 17 - Evolução anual do PIB setor indústria (total em R\$ 100.000) - Regiões metropolitanas, 2001 a 2014

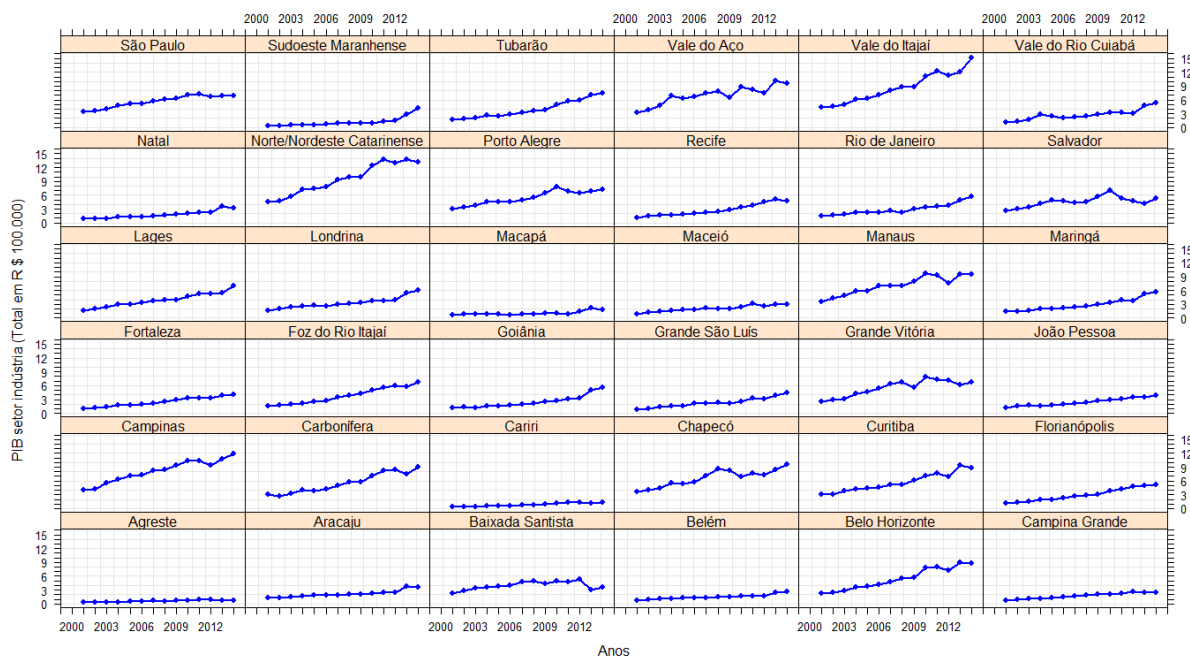
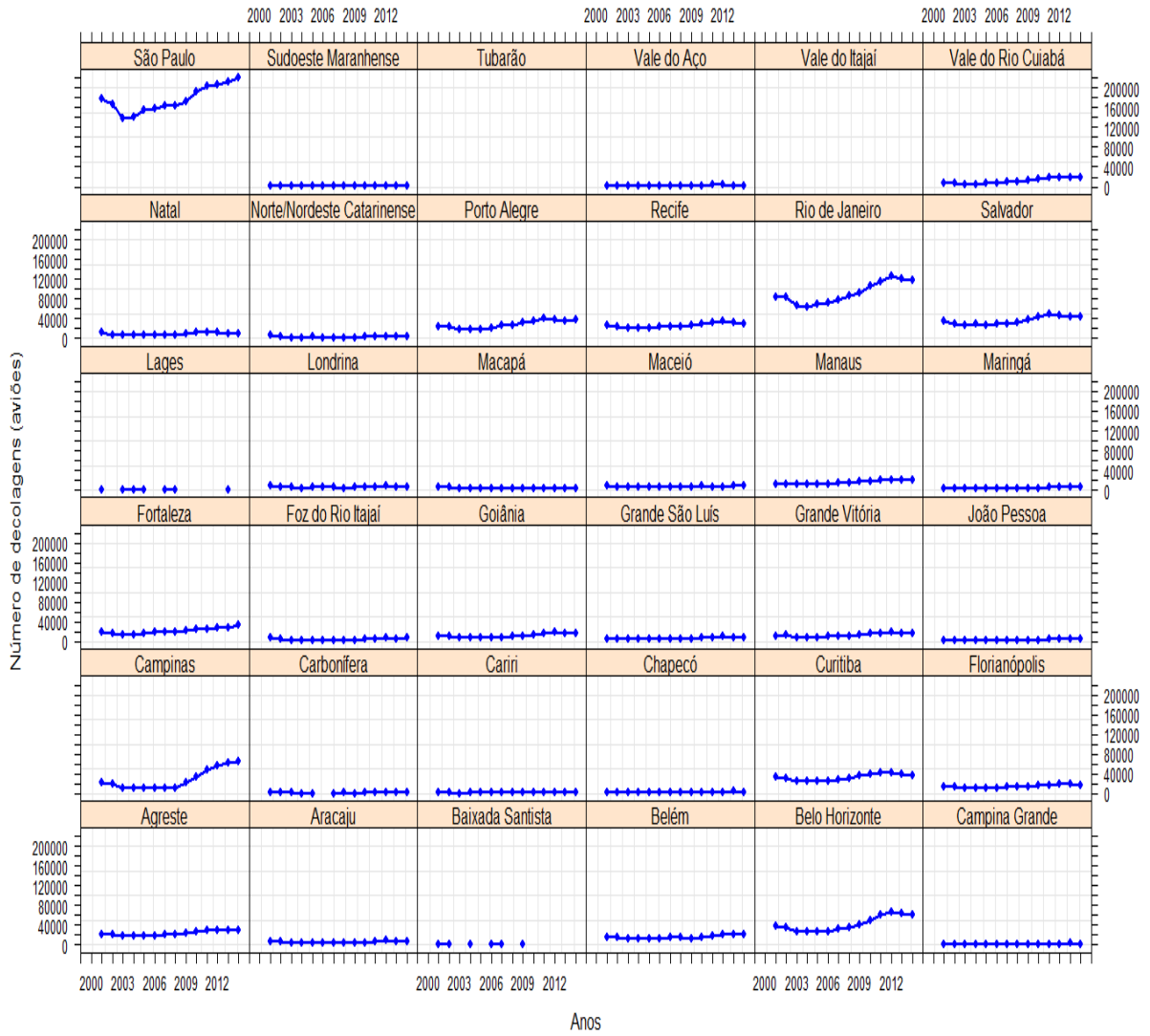
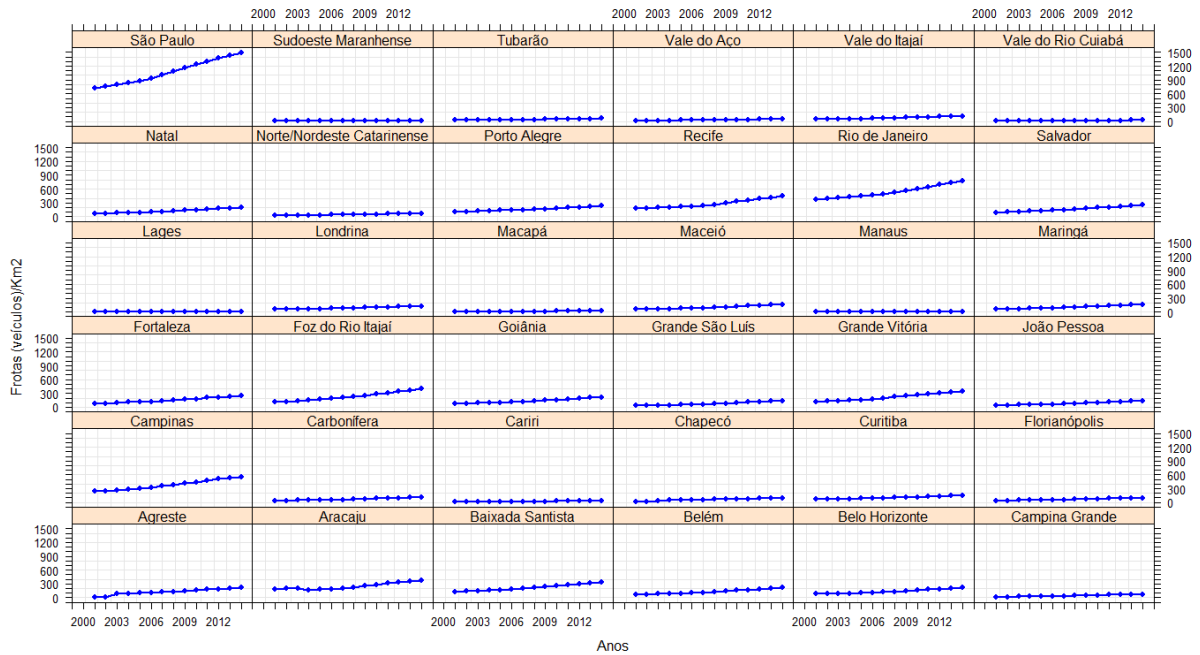


Gráfico 18 - Evolução anual do número de decolagens (aviões) - Regiões metropolitanas, 2001 a 2014



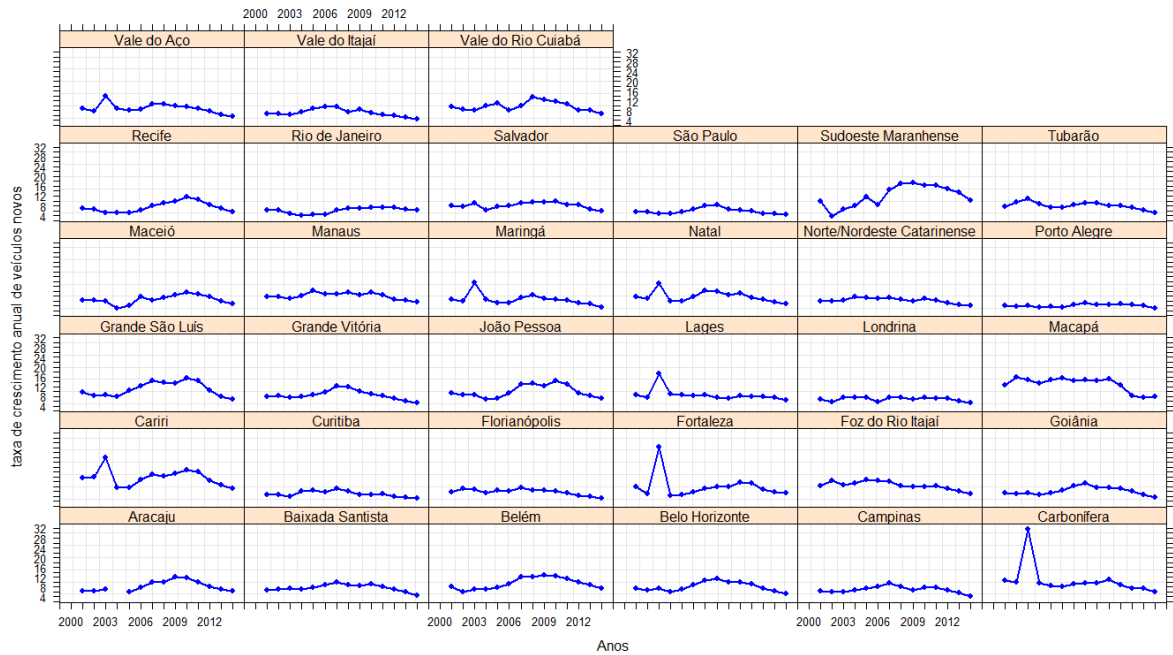
Elaboração própria

Gráfico 19 - Evolução anual de frotas de veículos por km2 - Regiões metropolitanas, 2001 a 2014



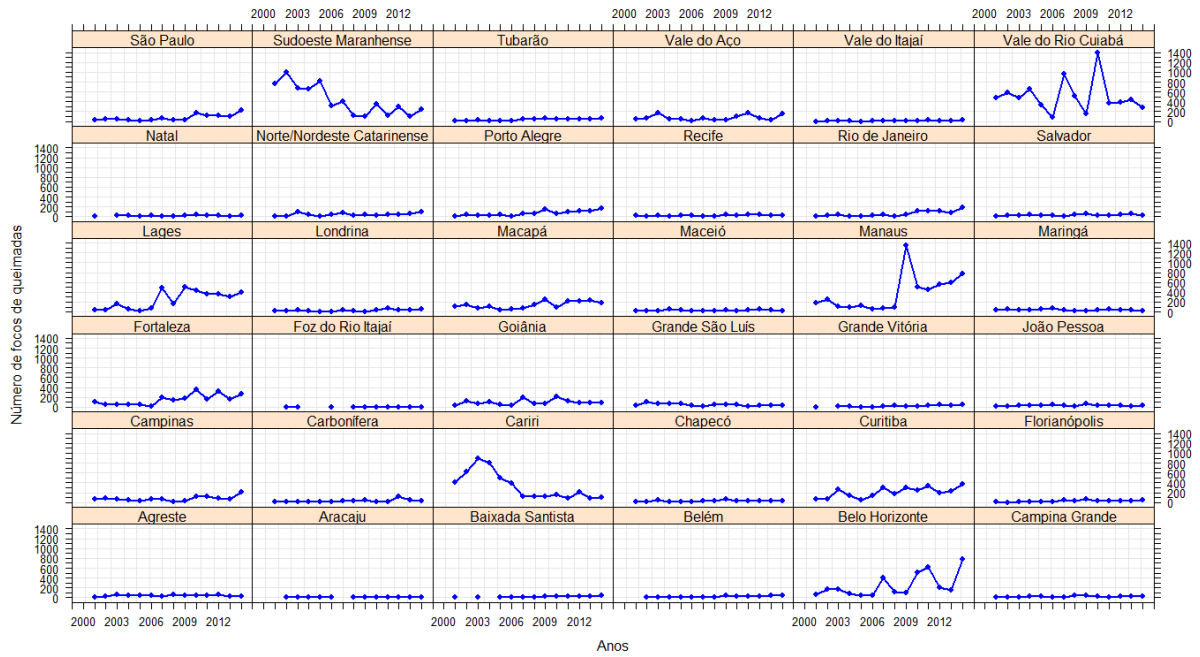
Elaboração própria

Gráfico 20 - Evolução da taxa de crescimento anual de veículos novos - Regiões metropolitanas, 2001 a 2014



Elaboração própria

Gráfico 21 - Evolução do total anual do número de focos de queimadas - Regiões metropolitanas, 2001 a 2014



Elaboração própria

APÊNDICE F – Correlações entre fontes de dados dos parâmetros meteorológicos

Tabela 14 - Correlação de Spearman dos parâmetros meteorológico entre CETESB e INMET na Região metropolitana de São Paulo, 2001 -2014

| Hipóteses | Temperatura | | | Umidade | | | velocidade | | | Pressão barométrica | | |
|--------------------|--------------|--------|---------|--------------|--------|---------|--------------|--------|---------|---------------------|--------|---------|
| | $\sum d_i^2$ | ρ | p-valor | $\sum d_i^2$ | ρ | p-valor | $\sum d_i^2$ | ρ | p-valor | $\sum d_i^2$ | ρ | p-valor |
| $H_1: \rho \neq 0$ | 208 | 0.543 | 0.048 | 192 | 0.578 | 0.033 | 460.010 | -0.011 | 0.970 | 470 | -0.033 | 0.916 |
| $H_1: \rho > 0$ | 208 | 0.543 | 0.024 | 192 | 0.578 | 0.017 | 460.010 | -0.011 | 0.515 | 470 | -0.033 | 0.548 |
| $H_1: \rho < 0$ | 208 | 0.543 | 0.977 | 192 | 0.578 | 0.984 | 460.010 | -0.011 | 0.485 | 470 | -0.033 | 0.458 |

Elaboração própria

APÊNDICE G – Sumário de dados faltantes, correlações entre as estações das médias anuais de PM10 e estatística descritiva após a imputação de dados

Tabela 15 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana da Baixada Santista

| Baixada Santista | | | |
|----------------------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| Cubatão Vila Parisi | 14 | 0 | 0 |
| Santo André-Capuava | 14 | 0 | 0 |
| Cubatão Centro | 13 | 1 | 7,14 |
| Mauá | 12 | 2 | 14,29 |
| São Bernardo do Campo | 11 | 3 | 21,43 |
| Cubatão Vale do Mogi | 6 | 8 | 57,14 |
| Santo André -Centro | 6 | 8 | 57,14 |
| Santo André-Paço Municipal | 4 | 10 | 71,43 |
| Santos | 3 | 11 | 78,57 |
| Santos Ponta da Praia | 3 | 11 | 78,57 |
| Santos Porto | 0 | 14 | 100 |

Critério de exclusão acima de 40% de *missing*

Tabela 16 - Correlações de Spearman entre as estações para as médias anuais de PM₁₀ na Baixada Santista

| | E1 | E2 | E3 | E4 | E5 |
|-----------|-----------|-----------|-----------|-----------|-----------|
| E1 | E1 | 0,597 | 0,549 | 0,642 | 0,425 |
| E2 | 0,597 | E2 | 0,117 | 0,232 | 0,269 |
| E3 | 0,549 | 0,117 | E3 | 0,789 | 0,737 |
| E4 | 0,642 | 0,232 | 0,789 | E4 | 0,365 |
| E5 | 0,425 | 0,269 | 0,737 | 0,365 | E5 |

E1 - Cubatão - Centro

E2 - Cubatão – Vila Parisi

E3 - Mauá

E4 - Santo André Capuava

E5 - São Bernardo do Campo

Tabela 17 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana de Belo Horizonte

| Belo Horizonte | | | |
|------------------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| aeroporto | 8 | 6 | 42,86 |
| Av. Amazonas | 8 | 6 | 42,86 |
| Praça Rui Barbosa | 8 | 6 | 42,86 |
| B. Jardim Alterosas | 5 | 9 | 64,29 |
| Cascata | 5 | 9 | 64,29 |
| Petrovale | 5 | 9 | 64,29 |
| Ibiritermo Piratininga | 4 | 10 | 71,43 |
| DNER | 3 | 11 | 78,57 |
| Safran Centro Adm. | 2 | 12 | 85,71 |
| CAMG | 0 | 14 | 100 |
| Praça Tancredo | 0 | 14 | 100 |

Critério de exclusão acima de 40% de *missing*

Tabela 18 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana de Campinas

| Campinas | | | |
|-----------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| Campinas Centro | 14 | 0 | 0 |
| Paulínia | 11 | 3 | 21,43 |
| Paulínia Sul | 6 | 8 | 57,14 |
| Americana | 4 | 10 | 71,43 |

Critério de exclusão acima de 40% de *missing*

Tabela 19 - Correlações de Spearman entre as estações para as médias anuais de PM₁₀ na região metropolitana de Campinas

| | Campinas Centro | Paulínia |
|-----------------|-----------------|-------------|
| Campinas Centro | 1,00 | 0,71 |
| Paulínia | 0,71 | 1,00 |

Tabela 20 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana de Curitiba

| Curitiba | | | |
|---------------------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| Praça do Ouvidor Pardinho | 9 | 5 | 35,71 |
| REPAR | 8 | 6 | 42,86 |
| UEG | 8 | 6 | 42,86 |
| CSN CISA | 7 | 7 | 50,00 |
| Boqueirão | 4 | 10 | 71,43 |
| Colombo | 4 | 10 | 71,43 |
| CIC | 1 | 13 | 92,86 |

Critério de exclusão acima de 40% de *missing*

Tabela 21 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana do Espírito Santo

| Espírito Santo | | | |
|--------------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| Carapina | 13 | 1 | 7,14 |
| Enseada do Suá | 13 | 1 | 7,14 |
| laranjeiras | 13 | 1 | 7,14 |
| Jardim Camburi | 11 | 3 | 21,43 |
| Vila velha ibes | 11 | 3 | 21,43 |
| Cariacica | 9 | 5 | 35,71 |
| Vila Velha Centro | 7 | 7 | 50,00 |
| Vitória Centro | 7 | 7 | 50,00 |
| Cidade Continental | 2 | 12 | 85,71 |

Critério de exclusão acima de 40% de *missing*

Tabela 22 - Correlações de Spearman entre as estações para as médias anuais de PM₁₀ na região metropolitana do Espírito Santo

| | E1 | E2 | E3 | E4 | E5 | E6 |
|----|-----------|-----------|-----------|-----------|-----------|-----------|
| E1 | E1 | -0,13 | 0,35 | 0,28 | -0,27 | -0,08 |
| E2 | -0,13 | E2 | 0,72 | -0,25 | 0,67 | 0,75 |
| E3 | 0,35 | 0,72 | E3 | -0,45 | 0,36 | 0,45 |
| E4 | 0,28 | -0,25 | -0,45 | E4 | -0,11 | 0,12 |
| E5 | -0,27 | 0,67 | 0,36 | -0,11 | E5 | 0,63 |
| E6 | -0,08 | 0,75 | 0,45 | 0,12 | 0,63 | E6 |

E1 = Carapina

E2 = Cariacica

E3 = Enseada do Suá

E4 = Jardim Camburi

E5 = Laranjeiras

E6 = Vila Velha ibes

Tabela 23 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana de Porto Alegre

| Porto Alegre | | | |
|--------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| Refap Canoas | 9 | 5 | 35,71 |
| Refap Esteio | 7 | 7 | 50,00 |
| Bombeiros | 5 | 9 | 64,29 |
| Sapucaia | 5 | 9 | 64,29 |
| Polo | 4 | 10 | 71,43 |
| Rodoviária | 4 | 10 | 71,43 |
| Vcomar | 4 | 10 | 71,43 |
| Charqueadas | 3 | 11 | 78,57 |
| Esef | 3 | 11 | 78,57 |
| Gravataí | 1 | 13 | 92,86 |
| Guaíba | 0 | 14 | 100,00 |

Critério de exclusão acima de 40% de *missing*

Tabela 24 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana do Rio de Janeiro
(Continua)

| Rio de Janeiro | | | |
|----------------------------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| Botafogo | 13 | 1 | 7,14 |
| Castelo | 13 | 1 | 7,14 |
| Seropédica Embrapa | 13 | 1 | 7,14 |
| Bonsucesso | 12 | 2 | 14,29 |
| São Cristóvão | 12 | 2 | 14,29 |
| São Gonçalo prefeitura | 12 | 2 | 14,29 |
| Vilar dos Teles | 12 | 2 | 14,29 |
| Cidade Deus | 11 | 3 | 21,43 |
| Copacabana SMAC | 10 | 4 | 28,57 |
| Nilópolis rodoviária | 10 | 4 | 28,57 |
| Nova Iguaçu monteiro lobato | 10 | 4 | 28,57 |
| São Cristóvão SMAC | 10 | 4 | 28,57 |
| Sumaré | 10 | 4 | 28,57 |
| Uerj | 10 | 4 | 28,57 |
| Centro SMAC | 9 | 5 | 35,71 |
| Duque de Caxias INSS | 8 | 6 | 42,86 |
| Duque de Caxias Jardim primavera | 8 | 6 | 42,86 |
| Duque de Caxias - São bento | 8 | 6 | 42,86 |
| Niterói centro | 8 | 6 | 42,86 |
| Tijuca SMAC | 8 | 6 | 42,86 |
| Duque de Caxias campos Elíseos | 7 | 7 | 50,00 |
| Santa Cruz Adalgisa Nery | 6 | 8 | 57,14 |
| Santa Cruz largo bodegão | 6 | 8 | 57,14 |

Tabela 24 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana do Rio de Janeiro
(Conclusão)

| Rio de Janeiro | | | | |
|--|---------|----------------|--------------------|--|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) | |
| Taquara | 6 | 8 | 57,14 | |
| Guapimirim | 5 | 9 | 64,29 | |
| Itaboraí fazenda macuco | 5 | 9 | 64,29 | |
| Itaboraí Itambi | 5 | 9 | 64,29 | |
| Itaguaí Monte Serrat | 5 | 9 | 64,29 | |
| Centro | 4 | 10 | 71,43 | |
| Copacabana | 4 | 10 | 71,43 | |
| Duque de Caxias vila São Luís | 4 | 10 | 71,43 | |
| Itaboraí alto jacu | 4 | 10 | 71,43 | |
| Campo Grande SMAC | 3 | 11 | 78,57 | |
| Nova Iguaçu Monteiro Lobato automático | 3 | 11 | 78,57 | |
| Pedra de Guaratiba SMAC | 3 | 11 | 78,57 | |
| São Gonçalo Uerj | 3 | 11 | 78,57 | |
| Bangu SMAC | 2 | 12 | 85,71 | |
| Gamboa | 2 | 12 | 85,71 | |
| Ilha de Paquetá | 2 | 12 | 85,71 | |
| Ilha do governador | 2 | 12 | 85,71 | |
| Itaboraí Sambaetiba | 2 | 12 | 85,71 | |
| Japeri engenheiro Pedreira | 2 | 12 | 85,71 | |
| Lab. INEA | 2 | 12 | 85,71 | |
| Leblon | 2 | 12 | 85,71 | |
| Campos Afonsos | 1 | 13 | 92,86 | |
| Gericinó | 1 | 13 | 92,86 | |
| Itaboraí Porto das Caixas | 1 | 13 | 92,86 | |
| Itaguaí Coroa Grande | 1 | 13 | 92,86 | |
| Niterói policlínica | 1 | 13 | 92,86 | |
| Duque de Caxias Pilar | 0 | 14 | 100,00 | |
| Engenhão | 0 | 14 | 100,00 | |
| Itaboraí Sambaetiba | 0 | 14 | 100,00 | |

Critério de exclusão acima de 40% de *missing*

Tabela 25 - Correlações de Spearman entre as estações para as médias anuais de PM₁₀ na região metropolitana do Rio de Janeiro

| | E1 | E2 | E3 | E4 | E5 | E6 | E7 | E8 | E9 | E10 | E11 | E12 | E13 | E14 | E15 |
|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|
| E1 | E1 | 0,731 | 0,555 | 0,664 | 0,483 | 0,573 | 0,580 | 0,836 | 0,491 | 0,321 | 0,636 | 0,539 | 0,467 | 0,733 | 0,467 |
| E2 | 0,731 | E2 | 0,544 | 0,745 | 0,049 | 0,147 | 0,524 | 0,382 | 0,552 | 0,491 | 0,333 | 0,406 | 0,224 | 0,612 | 0,117 |
| E3 | 0,555 | 0,544 | E3 | 0,800 | -0,112 | 0,587 | 0,769 | 0,545 | 0,503 | 0,806 | 0,721 | 0,697 | 0,794 | 0,794 | 0,567 |
| E4 | 0,664 | 0,745 | 0,800 | E4 | -0,358 | 0,576 | 0,952 | 0,733 | 0,619 | 0,714 | 0,905 | 0,738 | 0,238 | 0,905 | 0,571 |
| E5 | 0,483 | 0,049 | -0,112 | -0,358 | E5 | 0,189 | -0,049 | 0,364 | 0,067 | -0,133 | -0,018 | -0,030 | 0,467 | -0,042 | -0,033 |
| E6 | 0,573 | 0,147 | 0,587 | 0,576 | 0,189 | E6 | 0,720 | 0,736 | 0,067 | 0,617 | 0,952 | 0,770 | 0,588 | 0,770 | 0,950 |
| E7 | 0,580 | 0,524 | 0,769 | 0,952 | -0,049 | 0,720 | E7 | 0,800 | 0,539 | 0,967 | 0,927 | 0,661 | 0,515 | 0,952 | 0,733 |
| E8 | 0,836 | 0,382 | 0,545 | 0,733 | 0,364 | 0,736 | 0,800 | E8 | 0,624 | 0,643 | 0,818 | 0,467 | 0,455 | 0,833 | 0,500 |
| E9 | 0,491 | 0,552 | 0,503 | 0,619 | 0,067 | 0,067 | 0,539 | 0,624 | E9 | 0,452 | 0,367 | -0,067 | 0,233 | 0,550 | -0,200 |
| E10 | 0,321 | 0,491 | 0,806 | 0,714 | -0,133 | 0,617 | 0,967 | 0,643 | 0,452 | E10 | 0,857 | 0,476 | 0,714 | 0,976 | 0,607 |
| E11 | 0,636 | 0,333 | 0,721 | 0,905 | -0,018 | 0,952 | 0,927 | 0,818 | 0,367 | 0,857 | E11 | 0,817 | 0,517 | 0,952 | 0,952 |
| E12 | 0,539 | 0,406 | 0,697 | 0,738 | -0,030 | 0,770 | 0,661 | 0,467 | -0,067 | 0,476 | 0,817 | E12 | 0,600 | 0,617 | 0,833 |
| E13 | 0,467 | 0,224 | 0,794 | 0,238 | 0,467 | 0,588 | 0,515 | 0,455 | 0,233 | 0,714 | 0,517 | 0,600 | E13 | 0,417 | 0,476 |
| E14 | 0,733 | 0,612 | 0,794 | 0,905 | -0,042 | 0,770 | 0,952 | 0,833 | 0,550 | 0,976 | 0,952 | 0,617 | 0,417 | E14 | 0,619 |
| E15 | 0,467 | 0,117 | 0,567 | 0,571 | -0,033 | 0,950 | 0,733 | 0,500 | -0,200 | 0,607 | 0,952 | 0,833 | 0,476 | 0,619 | E15 |

E1-Botafogo

E2-Castelo

E3-Seropédica Embrapa

E4-Bonsucesso

E5-São Cristóvão

E6-São Gonçalo prefeitura

E7-Vilar dos Teles

E8-Cidade Deus

E9-Copacabana SMAC

E10-Nilópolis rodoviária

E11-Nova Iguaçu monteiro lobato

E12-São Cristóvão SMAC

E13-Sumaré

E14-Uerj

E15-Centro SMAC

Tabela 26 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana de Salvador

| Salvador | | | |
|-----------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| Concórdia | 11 | 3 | 21,43 |
| Gravatá | 11 | 3 | 21,43 |
| Leandrinho | 11 | 3 | 21,43 |
| Malembá | 7 | 7 | 50,00 |
| Paralela - CAB | 4 | 10 | 71,43 |
| Campo Grande | 3 | 11 | 78,57 |
| Dique do Tororó | 3 | 11 | 78,57 |
| Pirajá | 3 | 11 | 78,57 |
| Rio Vermelho | 3 | 11 | 78,57 |
| Madre de Deus | 2 | 12 | 85,71 |
| Av. ACM DETRAN | 2 | 12 | 85,71 |
| Barros Reis | 2 | 12 | 85,71 |
| Itaigara | 2 | 12 | 85,71 |

Critério de exclusão acima de 40% de *missing*

Tabela 27 - Correlações de Spearman entre as estações para as médias anuais de PM₁₀ na região metropolitana de Salvador

| | E1 | E2 | E3 |
|-----------|-----------|-----------|-----------|
| E1 | E1 | 0,260 | -0,445 |
| E2 | 0,260 | E2 | -0,310 |
| E3 | -0,445 | -0,310 | E3 |

E1-Concórdia
E2-Gravatá
E3-Leandrinho

Tabela 28 - Percentual de *missing* nos dados de médias anuais de PM₁₀ das estações de monitoramento da poluição do ar na região metropolitana de São Paulo

| São Paulo | | | |
|------------------------------|---------|----------------|--------------------|
| Estações | Nº obs. | <i>missing</i> | <i>missing</i> (%) |
| Santo André - Capuava | 14 | 0 | 0,00 |
| Diadema | 13 | 1 | 7,14 |
| Osasco | 13 | 1 | 7,14 |
| Ibirapuera | 13 | 1 | 7,14 |
| Mauá | 12 | 2 | 14,29 |
| São Caetano do Sul | 12 | 2 | 14,29 |
| Cerqueira César | 12 | 2 | 14,29 |
| Congonhas | 12 | 2 | 14,29 |
| Santo Amaro | 12 | 2 | 14,29 |
| São Bernardo do Campo | 11 | 3 | 21,43 |
| N Senhora do Ó | 11 | 3 | 21,43 |
| Parque D Pedro II | 11 | 3 | 21,43 |
| Taboão da Serra | 11 | 3 | 21,43 |
| Santana | 10 | 4 | 28,57 |
| Mooca | 9 | 5 | 35,71 |
| Parelheiros | 7 | 7 | 50,00 |
| Santo André Centro | 6 | 8 | 57,14 |
| Cambuci | 6 | 8 | 57,14 |
| Centro | 6 | 8 | 57,14 |
| Guarulhos | 5 | 9 | 64,29 |
| Pinheiros | 5 | 9 | 64,29 |
| Santo André - Paço Municipal | 4 | 10 | 71,43 |
| Interlagos | 3 | 11 | 78,57 |
| São Miguel Paulista | 3 | 11 | 78,57 |
| Guarulhos - Paço Municipal | 2 | 12 | 85,71 |
| Paulicéia | 2 | 12 | 85,71 |
| Capão Redondo | 2 | 12 | 85,71 |
| Carapicuíba | 2 | 12 | 85,71 |
| Lapa | 2 | 12 | 85,71 |
| Marg. Tietê - Pte Remédios | 2 | 12 | 85,71 |
| Ibirapuera | 1 | 13 | 92,86 |
| Itaim Paulista | 1 | 13 | 92,86 |
| Itaquera | 1 | 13 | 92,86 |
| Penha | 0 | 14 | 100,00 |
| Pico do Jaraguá | 0 | 14 | 100,00 |
| Panorama | 0 | 14 | 100,00 |

Critério de exclusão acima de 40% de *missing*

Tabela 29 - Correlações de Spearman entre as estações para as médias anuais de PM₁₀ na região metropolitana de São Paulo

| | E1 | E2 | E3 | E4 | E5 | E6 | E7 | E8 | E9 | E10 | E11 | E12 | E13 | E14 | E15 |
|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|
| E1 | E1 | 0,436 | 0,506 | 0,641 | 0,789 | 0,820 | 0,558 | 0,392 | 0,482 | 0,365 | 0,433 | 0,585 | 0,442 | 0,276 | 0,026 |
| E2 | 0,436 | E2 | 0,784 | 0,662 | 0,733 | 0,844 | 0,721 | 0,655 | 0,710 | 0,830 | 0,538 | 0,768 | 0,745 | 0,438 | 0,267 |
| E3 | 0,506 | 0,784 | E3 | 0,775 | 0,515 | 0,629 | 0,747 | 0,731 | 0,947 | 0,414 | 0,838 | 0,817 | 0,638 | 0,115 | 0,625 |
| E4 | 0,641 | 0,662 | 0,775 | E4 | 0,639 | 0,762 | 0,681 | 0,736 | 0,888 | 0,347 | 0,724 | 0,849 | 0,542 | -0,043 | 0,289 |
| E5 | 0,789 | 0,733 | 0,515 | 0,639 | E5 | 0,855 | 0,407 | 0,412 | 0,546 | 0,737 | 0,097 | 0,383 | 0,693 | 0,641 | 0,091 |
| E6 | 0,820 | 0,844 | 0,629 | 0,762 | 0,855 | E6 | 0,655 | 0,576 | 0,681 | 0,709 | 0,439 | 0,822 | 0,777 | 0,377 | 0,043 |
| E7 | 0,558 | 0,721 | 0,747 | 0,681 | 0,407 | 0,655 | E7 | 0,777 | 0,899 | 0,543 | 0,424 | 0,872 | 0,388 | -0,402 | 0,633 |
| E8 | 0,392 | 0,655 | 0,731 | 0,736 | 0,412 | 0,576 | 0,777 | E8 | 0,924 | 0,445 | 0,477 | 0,826 | 0,748 | -0,075 | 0,378 |
| E9 | 0,482 | 0,710 | 0,947 | 0,888 | 0,546 | 0,681 | 0,899 | 0,924 | E9 | 0,270 | 0,793 | 0,908 | 0,685 | 0,050 | 0,864 |
| E10 | 0,365 | 0,830 | 0,414 | 0,347 | 0,737 | 0,709 | 0,543 | 0,445 | 0,270 | E10 | -0,279 | 0,422 | 0,964 | 0,553 | -0,088 |
| E11 | 0,433 | 0,538 | 0,838 | 0,724 | 0,097 | 0,439 | 0,424 | 0,477 | 0,793 | -0,279 | E11 | 0,773 | 0,144 | -0,103 | 0,305 |
| E12 | 0,585 | 0,768 | 0,817 | 0,849 | 0,383 | 0,822 | 0,872 | 0,826 | 0,908 | 0,422 | 0,773 | E12 | 0,716 | -0,309 | -0,493 |
| E13 | 0,442 | 0,745 | 0,638 | 0,542 | 0,693 | 0,777 | 0,388 | 0,748 | 0,685 | 0,964 | 0,144 | 0,716 | E13 | 0,466 | -0,248 |
| E14 | 0,276 | 0,438 | 0,115 | -0,043 | 0,641 | 0,377 | -0,402 | -0,075 | 0,050 | 0,553 | -0,103 | -0,309 | 0,466 | E14 | 0,164 |
| E15 | 0,026 | 0,267 | 0,625 | 0,289 | 0,091 | 0,043 | 0,633 | 0,378 | 0,864 | -0,088 | 0,305 | -0,493 | -0,248 | 0,164 | E15 |

E1-Santo André - Capuava

E2-Diadema

E3-Osasco

E4-Ibirapuera

E5-Mauá

E6-São Caetano do Sul

E7-Cerqueira César

E8-Congonhas

E9-Santo Amaro

E10-São Bernardo do Campo

E11-N Senhora do Ó

E12-Parque D Pedro II

E13-Taboão da Serra

E14-Santana

E15-Mooça

Tabela 30 - Estatística descritiva das médias anuais de PM₁₀, após as imputações de dados, de acordo com as regiões metropolitanas durante 2001 a 2014

| Regiões Metropolitanas | N | Média | DP | Min. | 1ºQ | 2ºQ | 3ºQ | Máx. |
|------------------------|----|-------|------|------|------|------|------|------|
| Baixada Santista | 14 | 51,4 | 3,2 | 42,8 | 50,2 | 51,6 | 53,6 | 55,8 |
| Campinas | 14 | 35,3 | 4,4 | 28,5 | 32,5 | 34,0 | 38,5 | 43,5 |
| Grande Vitória | 14 | 29,6 | 1,4 | 27,4 | 28,8 | 29,6 | 30,3 | 32,8 |
| Rio de Janeiro | 14 | 60,3 | 7,9 | 47,6 | 57,4 | 59,4 | 63,0 | 74,4 |
| Salvador | 14 | 19,5 | 2,2 | 15,5 | 18,2 | 19,3 | 20,6 | 23,2 |
| São Paulo | 14 | 39,2 | 4,8 | 32,6 | 35,7 | 38,0 | 41,1 | 47,9 |
| Total | 84 | 39,2 | 14,2 | 15,5 | 29,6 | 36,9 | 50,7 | 74,4 |

Elaboração própria

APÊNDICE H – Modelagem e estimativas dos níveis de concentração de PM₁₀ das regiões metropolitanas do Brasil no período de 2001 a 2014

Análise exploratória

A estatística descritiva das médias anuais de PM₁₀ durante todo o período de estudo estão apresentados na tabela 31. A Baixada Santista apresentou a mais alta média dos níveis de concentração de PM₁₀ com 62,7 µg/m³ durante o período de 2001 a 2014. Por outro lado, a região metropolitana de Salvador obteve a mais baixa média dos níveis de concentração de PM₁₀, com 20,8 µg/m³ e menor variabilidade ao longo do tempo. No presente caso, foi usado um procedimento de "grandes amostras" para testes de significância. Para testar se há diferenças entre as médias das regiões metropolitanas, tradicionalmente foi realizado um teste F que apresentou significância estatística. A estatística de teste foi de 53,5 com p-valor=0.00 ou seja o teste indicou que há diferenças entre as médias regionais no nível de significância de 0,05, com um desvio de 15,7 µg/m³. Essas diferenças podem ser explicadas pelos fatores contextuais de cada região. Além das diferenças das médias anuais de PM₁₀ entre as regiões metropolitanas apresentaram também variações ao longo do tempo.

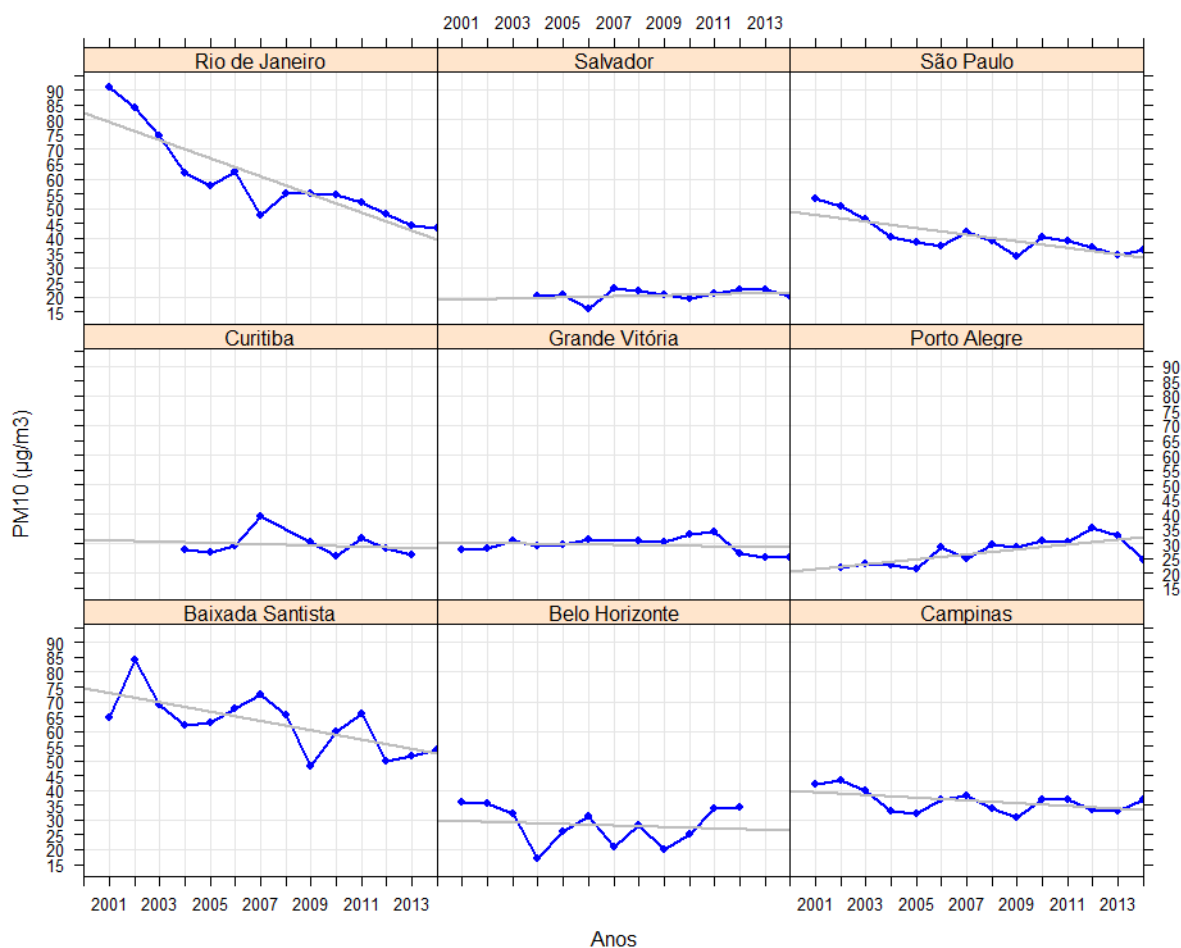
Tabela 31 - Estatísticas descritivas das médias anuais de PM₁₀ de acordo com as regiões metropolitanas durante 2001 - 2014

| Regiões Metropolitanas | N | Média | DP | Min. | 1ºQ | 2ºQ | 3ºQ | Máx. |
|-------------------------------|-----------|--------------|-----------|-------------|-------------------|------------|------------|-------------|
| Baixada Santista | 14 | 62,7 | 9,7 | 48,3 | 55,3 | 63,8 | 67,1 | 84,0 |
| Belo Horizonte | 12 | 28,3 | 6,5 | 17,2 | 24,2 | 29,7 | 33,8 | 35,9 |
| Campinas | 14 | 36,2 | 3,8 | 30,8 | 33,1 | 36,8 | 37,8 | 43,5 |
| Curitiba | 9 | 29,6 | 4,0 | 25,9 | 27,1 | 28,3 | 30,8 | 39,0 |
| Grande Vitória | 13 | 29,6 | 2,8 | 25,3 | 28,1 | 29,9 | 31,2 | 34,0 |
| Porto Alegre | 13 | 27,3 | 4,5 | 21,4 | 23,1 | 28,7 | 30,5 | 35,4 |
| Rio de Janeiro | 14 | 59,4 | 14,4 | 43,3 | 49,1 | 55,0 | 62,3 | 90,8 |
| Salvador | 11 | 20,8 | 1,8 | 16,2 | 20,3 | 21,0 | 22,2 | 23,0 |
| São Paulo | 14 | 40,5 | 5,8 | 33,6 | 36,9 | 39,0 | 41,6 | 53,2 |
| Total | 114 | 38,2 | 15,7 | 16,2 | 28,1 | 33,6 | 45,9 | 90,8 |
| | gl | SQ | MQ | F | Prob.>F | | | |
| Between groups | 8 | 22443,2 | 2805,4 | 53,5 | 0,000 | | | |
| Within groups | 105 | 5508,3 | 52,5 | | | | | |
| Total | 113 | 27951,5 | 247,4 | | | | | |
| Within-groups | DP=7,2 | | | | | | | |
| Between-groups | DP= 15,7 | | | | | | | |

A estatística descritiva das médias de concentração de PM_{10} foi com base na análise de dados completos. O procedimento pode ter produzido estimativas viesadas das médias. Já que grande parte de estações de monitoramento da qualidade do ar apresentou altos percentuais de dados faltantes.

A figura 38 apresenta a evolução das médias anuais de PM_{10} ao longo do tempo por região metropolitana.

Figura 38 - Evolução das médias anuais de PM_{10} nas regiões metropolitanas durante o período 2001 a 2014



Elaboração própria

Escolha das variáveis para a estrutura fixa e aleatória do modelo

A ideia da introdução dos termos temporais como um dos efeitos aleatórios no modelo foi realizada com base na figura 38 e no desvio padrão temporal das médias anuais de PM_{10} nas regiões metropolitanas apresentado na tabela 31 (DP=7,2). Na análise apresentada a variabilidade das médias anuais de PM_{10} ao longo do tempo permitiu considerar o termo do tempo como um possível componente dos efeitos aleatório. Neste sentido, as estimativas dos parâmetros do modelo linear de efeitos mistos e testes também constituem um meio de auxiliar nessa decisão. Por motivos de convergência, o termo do tempo foi centrado em $t_0 = 7,5$, o ponto médio da série de 14 anos. É importante modelar a tendência temporal das médias anuais de PM_{10} apresentada para um melhor ajuste no modelo. Na tabela 32 encontra-se o resumo dos critérios de seleção entre os tipos de função para modelagem da série temporal. Os AIC realizados levaram à seleção de um modelo com polinômio de grau 1 (AIC= 507,5) por possuir o menor AIC. E as inclinações podem ser testadas, como é ilustrado na tabela 34. Conforme o TRV o melhor modelo é o modelo 1.1 que considera além das inclinações do intercepto, as inclinações do tempo como um termo aleatório. Na tabela 35, é ilustrada a conclusão do modelo final, o modelo 1.1, que introduz na parte fixa e aleatória o termo do intercepto e do tempo.

Tabela 32 - Modelos considerados na seleção dos efeitos fixos dos termos temporais para a análise das médias anuais de PM_{10}

| Variáveis/termos | Parâmetros | Modelos | | | | |
|--|------------|------------------------------|-------|-------|----------------------|------------------|
| | | Funções Polinomiais Graus | | | Spline quadrática | Spline cúbica |
| | | 1 | 2 | 3 | 4 | 5 |
| Intercepto | β_0 | √ | √ | √ | √ | √ |
| Função linear (t-7,5) | β_1 | √ | √ | √ | √ | √ |
| Função quadrática (t-7,5) ² | β_2 | | √ | √ | | √ |
| Função cúbica (t-7,5) ³ | β_3 | | | √ | | |
| Spline quadrática a esquerda de 7,5 | β_4 | | | | √ | |
| Spline quadrática a direita de 7,5 | β_5 | | | | √ | |
| Spline cúbica a esquerda de 7,5 | β_6 | | | | | √ |
| Spline cúbica a direita de 7,5 | β_7 | | | | | √ |
| AIC | | 507,5 | 513,8 | 521,2 | 515,8 | 527,1 |

Tabela 33 - Modelos considerados na seleção dos efeitos aleatórios para a análise das médias anuais de PM₁₀

| | Nível | Termos | Parâmetros | Modelos | |
|--------------------|-------|------------|---------------|----------------|----------------|
| | | | | 1 | 1.1 |
| Efeitos aleatórios | 2 | Intercepto | σ_{0j} | $\sqrt{\quad}$ | $\sqrt{\quad}$ |
| | | (t-7,5) | σ_{1j} | | $\sqrt{\quad}$ |
| | 1 | | σ_{e0} | | |

Tabela 34 - Modelos considerados na modelagem dos efeitos aleatórios para a análise das médias anuais de PM₁₀

| Modelo | gl | AIC | BIC | Test | TRV | P-valor |
|------------|----|-------|-------|--------|------|---------|
| 1 | 4 | 507,5 | 517,1 | | | |
| 1.1 | 6 | 488,5 | 503,0 | 1 vs 2 | 23,0 | <.0001 |

Tabela 35 - Estimativas dos parâmetros para os efeitos fixos, erro-padrão, estatística do teste-t e p-valores do modelo 1.1

| Variáveis | Estimativas | Erro Padrão | gl | T | Intervalo de. | | P-valor |
|-------------------------------------|-------------|-------------|----|--------|---------------|--------|---------|
| | | | | | confiança 95% | | |
| | | | | | Inf. | Sup. | |
| Resposta: Ln_PM₁₀ | | | | | | | |
| Parte fixa | | | | | | | |
| constante | 39,222 | 6,027 | 77 | 6,508 | 27,220 | 51,223 | 0,000 |
| tempo (linear)* | -0,492 | 0,266 | 77 | -1,852 | -1,022 | 0,037 | 0,068 |
| Parte aleatória | | | | | | | |
| Nível: RM | | | | | | | |
| s.d. (cons.) | 14,737 | | | | 7,911 | 27,452 | |
| s.d. (tempo) | 0,614 | | | | 0,306 | 1,233 | |
| corr.(cons./tempo) | -0,591 | | | | -0,928 | 0,277 | |
| Nível: ano | | | | | | | |
| resíduos | 3,280 | | | | 2,786 | 3,862 | |
| Units: RM | 6 | | | | | | |
| Units: ano | 84 | | | | | | |
| Estimation: | REML | | | | | | |
| AIC | 488,542 | | | | | | |
| logLik | -238,271 | | | | | | |

*A variável de tendência (t-7,5) passou a ser denominada apenas como tempo.

Escolhas dos componentes dos efeitos fixos

Diagnóstico de multicolinearidade

Matriz de correlação

| | V1 | V2 | V3 | V4 | V5 | V6 | V7 | V8 | V9 | V10 | V11 | V12 | V13 | V14 |
|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|------------|------------|------------|------------|
| V1 | V1 | | | | | | | | | | | | | |
| V2 | -0,16 | V2 | | | | | | | | | | | | |
| V3 | -0,09 | -0,05 | V3 | | | | | | | | | | | |
| V4 | -0,1 | 0,5* | 0,02 | V4 | | | | | | | | | | |
| V5 | 0,03 | -0,38 | 0,04 | -0,52* | V5 | | | | | | | | | |
| V6 | -0,26* | -0,19 | -0,04 | -0,27* | 0,01 | V6 | | | | | | | | |
| V7 | -0,07 | 0,37* | 0,07 | 0,64* | -0,81* | -0,41* | V7 | | | | | | | |
| V8 | 0,43* | -0,06 | -0,22 | -0,02 | -0,31* | -0,42* | 0,4* | V8 | | | | | | |
| V9 | -0,09 | 0,66* | 0,04 | 0,63* | -0,42* | -0,23* | 0,41* | 0,17 | V9 | | | | | |
| V10 | 0,03 | 0,6* | -0,04 | 0,59* | -0,45* | -0,31* | 0,47* | 0,41* | 0,95* | V10 | | | | |
| V11 | -0,21 | 0,49* | 0,32* | 0,55* | -0,07 | -0,3* | 0,31* | -0,38* | 0,5* | 0,35* | V11 | | | |
| V12 | -0,44* | 0,62* | 0,07 | 0,76* | -0,4* | -0,21 | 0,52* | -0,18 | 0,78* | 0,65* | 0,69* | V12 | | |
| V13 | -0,34* | -0,12 | -0,31* | -0,02 | 0,21 | 0,34* | -0,36* | -0,51* | -0,18 | -0,3* | -0,05 | 0,11 | V13 | |
| V14 | 0,01 | 0,43* | 0,11 | 0,44* | -0,22 | -0,45* | 0,43* | 0,37* | 0,66* | 0,72* | 0,34* | 0,48* | -0,34* | V14 |

* As correlações são significativas no nível de 5%

| | |
|-----------------------------|-----------------------------------|
| V1 - PM ₁₀ | V8 - Densidade demográfica |
| V2 - Gasolina per capita | V9 - Pib per capita |
| V3 - Óleo diesel per capita | V10 - Pib serviços per capita |
| V4 - Álcool per capita | V11 - Pib agropecuária per capita |
| V5 - Temperatura | V12 - Pib indústria per capita |
| V6 - Umidade | V13 - Taxa de veículos novos |
| V7 - Altitude media | V14 - Focos |

De acordo com a matriz de correlação de *Pearson*, podemos observar uma forte presença de multicolinearidade entre os pares de algumas variáveis independentes, apresentando correlações significativas com p-valores menores que 0,05.

Primeiramente foi feito um teste de diagnóstico de multicolinearidade das variáveis independentes através do VIF, isso é útil pois fornece informações sobre quais os regressores estão envolvidos na multicolinearidade. O teste está ilustrado na tabela 36.

Tabela 36 - VIF das variáveis independentes

| Variáveis | Painel 1 | Painel 2 |
|---------------------------------|----------|----------|
| Tempo(linear) | 8,35 | 6,42 |
| Ln_ gasolina per capita | 1,37 | 1,32 |
| Ln_ óleo diesel per capita | 1,67 | 1,65 |
| Ln_ álcool per capita | 2,22 | 2,23 |
| Ln_ temperatura | 1,45 | 1,31 |
| Ln_ umidade | 1,29 | 1,22 |
| Ln_ altitude media | 2,00 | 2,08 |
| Ln_ densidade demográfica | 1,81 | 1,66 |
| Ln_ pib per capita | 14,77 | |
| Ln_ pib serviços per capita | 7,22 | 6,63 |
| Ln_ pib agropecuária per capita | 1,83 | 1,51 |
| Ln_ pib indústria per capita | 4,89 | 2,05 |
| Ln_ taxa de veículos novos | 3,14 | 3,11 |
| Ln_ focos | 1,36 | 1,34 |
| Máximo VIF | 14,77 | 6,63 |

A tabela 36 mostra as variáveis independentes em situações de multicolinearidade, no painel 1, o VIF máximo foi de 14,77, referente à variável PIB per capita, que excedeu de 10, valor limite para indicação de multicolinearidade. De acordo com a matriz de correlação, o PIB per capita é altamente correlacionado com o PIB serviços per capita. Assim foi decidido eliminar a variável PIB per capita para melhorar a multicolinearidade. E recalculamos os valores do VIF das variáveis, na situação 2, o VIF máximo foi de 6,63, uma situação aceitável de multicolinearidade. Assim, as variáveis que fizeram parte dos efeitos fixos estão no painel 2. Uma transformação logarítmica foi aplicada na variável resposta e nas variáveis independentes, de forma a garantir as condições exigidas para a aplicação do modelo linear de efeitos mistos. A princípio optou-se por todas as variáveis de estudo como componentes dos efeitos fixos apenas como uma forma ilustrativa do modelo cheio. Na tabela 37, encontram-se as estimativas dos parâmetros, os erros-padrão, os valores da estatística de teste-t e os respectivos valores-p para o modelo cheio.

Tabela 37 – Estimativas dos parâmetros para os efeitos fixos, erro-padrão, estatística do teste-t e p-valores do modelo 1.2

| Variáveis | Estimativas | Erro Padrão | gl | T | Intervalo de. | | P-valor |
|-------------------------------------|-------------|-------------|----|--------|-----------------------|--------|---------|
| | | | | | confiança 95% Inf. | Sup. | |
| Resposta: Ln PM₁₀ | | | | | | | |
| Parte fixa | | | | | | | |
| constante | 11,392 | 1,957 | 66 | 5,822 | 7,485 | 15,298 | 0,000 |
| tempo (linear) | -0,036 | 0,014 | 66 | -2,534 | -0,064 | -0,008 | 0,014 |
| Ln_ gasolina per capita | -0,002 | 0,003 | 66 | -0,614 | -0,007 | 0,004 | 0,541 |
| Ln_ óleo diesel per capita | 0,000 | 0,002 | 66 | 0,152 | -0,003 | 0,004 | 0,880 |
| Ln_ álcool per capita | -0,034 | 0,021 | 66 | -1,596 | -0,077 | 0,009 | 0,115 |
| Ln_ temperatura | -0,502 | 0,295 | 66 | -1,700 | -1,091 | 0,088 | 0,094 |
| Ln_ umidade | -0,664 | 0,201 | 66 | -3,313 | -1,065 | -0,264 | 0,002 |
| Ln_ altitude media | -0,335 | 0,113 | 4 | -2,977 | -0,648 | -0,023 | 0,041 |
| Ln_ densidade demográfica | -0,408 | 0,243 | 66 | -1,681 | -0,892 | 0,077 | 0,098 |
| Ln_ pib serviços per capita | 0,403 | 0,133 | 66 | 3,030 | 0,137 | 0,668 | 0,004 |
| Ln_ pib agropecuária per capita | 0,042 | 0,036 | 66 | 1,187 | -0,029 | 0,114 | 0,239 |
| Ln_ pib indústria per capita | -0,158 | 0,057 | 66 | -2,796 | -0,272 | -0,045 | 0,007 |
| Ln_ taxa de veículos novos | 0,145 | 0,068 | 66 | 2,120 | 0,008 | 0,282 | 0,038 |
| Ln_focos | 0,035 | 0,012 | 66 | 2,847 | 0,010 | 0,059 | 0,006 |
| Parte aleatória | | | | | | | |
| Nível: RM | | | | | | | |
| s.d. (cons.) | 0,721 | | | | 0,315 | 1,650 | |
| s.d. (trend) | 0,013 | | | | 0,006 | 0,026 | |
| corr.(cons./trend) | -0,996 | | | | -1,000 | 0,875 | |
| Nível: ano | | | | | | | |
| resíduos | 0,069 | | | | 0,058 | 0,082 | |
| R²(m) | 28,01% | | | | | | |
| R²(c) | 99,35% | | | | | | |
| Units: RM | 6 | | | | | | |
| Units: ano | 84 | | | | | | |
| Estimation: | REML | | | | | | |
| AIC | -76,961 | | | | | | |
| logLik | 56,481 | | | | | | |

Shapiro-Wilk normality test: W = 0,982, p-value = 0,327

Seleção dos efeitos fixos

Com base no modelo cheio, modelo 1.2, se deu início à modelagem dos efeitos fixos. O método de estimação para os coeficientes dos diferentes modelos foi através do REML. A modelagem foi por seleção “*Stepwise Regression*”⁴. O procedimento de seleção terminou apenas com umidade relativa do ar, PIB setor serviços, altitude média e focos de queimadas como modelo final, conforme é demonstrado na tabela 38.

Tabela 38 - Estimativas dos parâmetros para os efeitos fixos e aleatórios, erro-padrão, estatística do teste-t e p-valores, modelo 1.3

(continua)

| Variáveis | Estimativas | Erro Padrão | gl | T | Intervalo de. | | P-valor |
|-------------------------------------|-------------|-------------|----|--------|-----------------------|--------|---------|
| | | | | | confiança 95% Inf. | Sup. | |
| Resposta: Ln_PM₁₀ | | | | | | | |
| Parte fixa | | | | | | | |
| constante | 6,839 | 1,053 | 74 | 6,498 | 4,742 | 8,936 | 0,000 |
| tempo (linear) | -0,052 | 0,012 | 74 | -4,450 | -0,075 | -0,029 | 0,000 |
| Ln_ umidade | -0,649 | 0,202 | 74 | -3,216 | -1,051 | -0,247 | 0,002 |
| Ln_ pib serviços per capita | 0,401 | 0,116 | 74 | 3,450 | 0,169 | 0,633 | 0,001 |
| Ln_ altitude media | -0,320 | 0,105 | 4 | -3,044 | -0,612 | -0,028 | 0,038 |
| Ln_focos | 0,031 | 0,012 | 74 | 2,511 | 0,006 | 0,056 | 0,014 |
| Parte aleatória | | | | | | | |
| Nível: RM | | | | | | | |
| s.d. (cons.) | 0,543 | | | | 0,212 | 1,393 | |
| s.d. (trend) | 0,012 | | | | 0,006 | 0,024 | |
| corr.(cons./trend) | -0,894 | | | | -0,998 | 0,600 | |
| Nível: ano | | | | | | | |
| resíduos | 0,073 | | | | 0,062 | 0,086 | |
| R²(m) | 25,76% | | | | | | |
| R²(c) | 98,69% | | | | | | |
| Units: RM | 6 | | | | | | |
| Units: ano | 84 | | | | | | |
| Estimation: | REML | | | | | | |
| AIC | -118,69 | | | | | | |
| logLik | 69,34 | | | | | | |

⁴ Por ser muito extenso, todo o passo a passo do método “*Stepwise Regression*” encontra-se com o autor

Tabela 38 - Estimativas dos parâmetros para os efeitos fixos e aleatórios, erro-padrão, estatística do teste-t e p-valores, modelo 1.3

(Conclusão)

| Variáveis | Anova | | | | | | |
|-----------------------------|-------|-------|----|-------|---------|--------|------------|
| | SSQ | MSQ | gl | DenDF | Valor F | Pr(>F) | Pct. Expl. |
| tempo (linear) | 0,105 | 0,105 | 1 | 74 | 4,405 | 0,039 | 34,37% |
| Ln_ umidade | 0,055 | 0,055 | 1 | 74 | 10,020 | 0,002 | 17,95% |
| Ln_ pib serviços per capita | 0,063 | 0,063 | 1 | 74 | 6,615 | 0,012 | 20,66% |
| Ln_ altitude media | 0,049 | 0,049 | 1 | 4 | 11,526 | 0,027 | 16,08% |
| Ln_focos | 0,033 | 0,033 | 1 | 74 | 6,305 | 0,014 | 10,94% |

correlação de Pearson (r) = 0,98

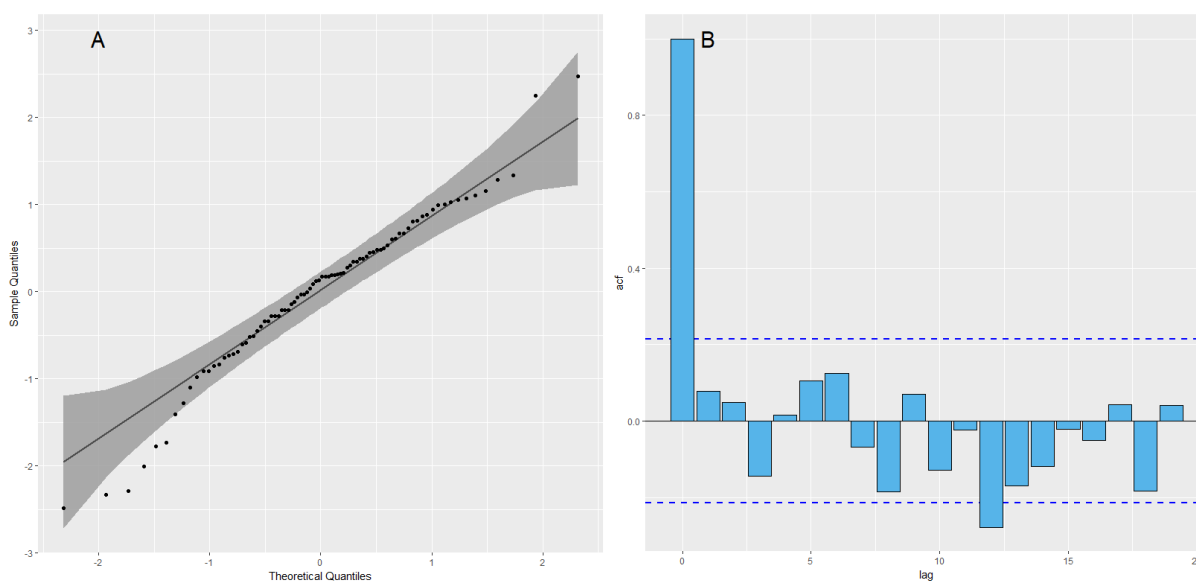
RMSE = 0,067 $\mu\text{g}/\text{m}^3$

MAE = 0,051

VIF: tempo = 5,19; Ln_ umidade = 1,04; Ln_ pib serviços per capita = 5,36; Ln_ altitude media = 1,20, Ln_focos = 1,21

Avaliação da adequação do modelo

Uma vez obtido o modelo final, modelo 1.3, verificou-se a qualidade do ajuste no sentido de validar os pressupostos assumidos para o modelo linear de efeitos mistos através de métodos gráficos de diagnóstico. O pressuposto da normalidade dos erros dentro dos grupos pode ser avaliado através do gráfico de Probabilidade dos resíduos padronizados (figura 39-A). Neste caso não permitiu admitir uma distribuição gaussiana. Na figura 39-B e C, representam a função de autocorrelação e autocorrelação parcial dos resíduos e se observa autocorrelações significativas após a defasagem 12. Na figura 39-D, representa o gráfico dos resíduos padronizados que demonstra presença de heterocedasticidade. E na figura 39-E, podemos observar discrepâncias dos resíduos. Com base nestas avaliações, o modelo ainda pode ser melhorado com medidas corretivas.



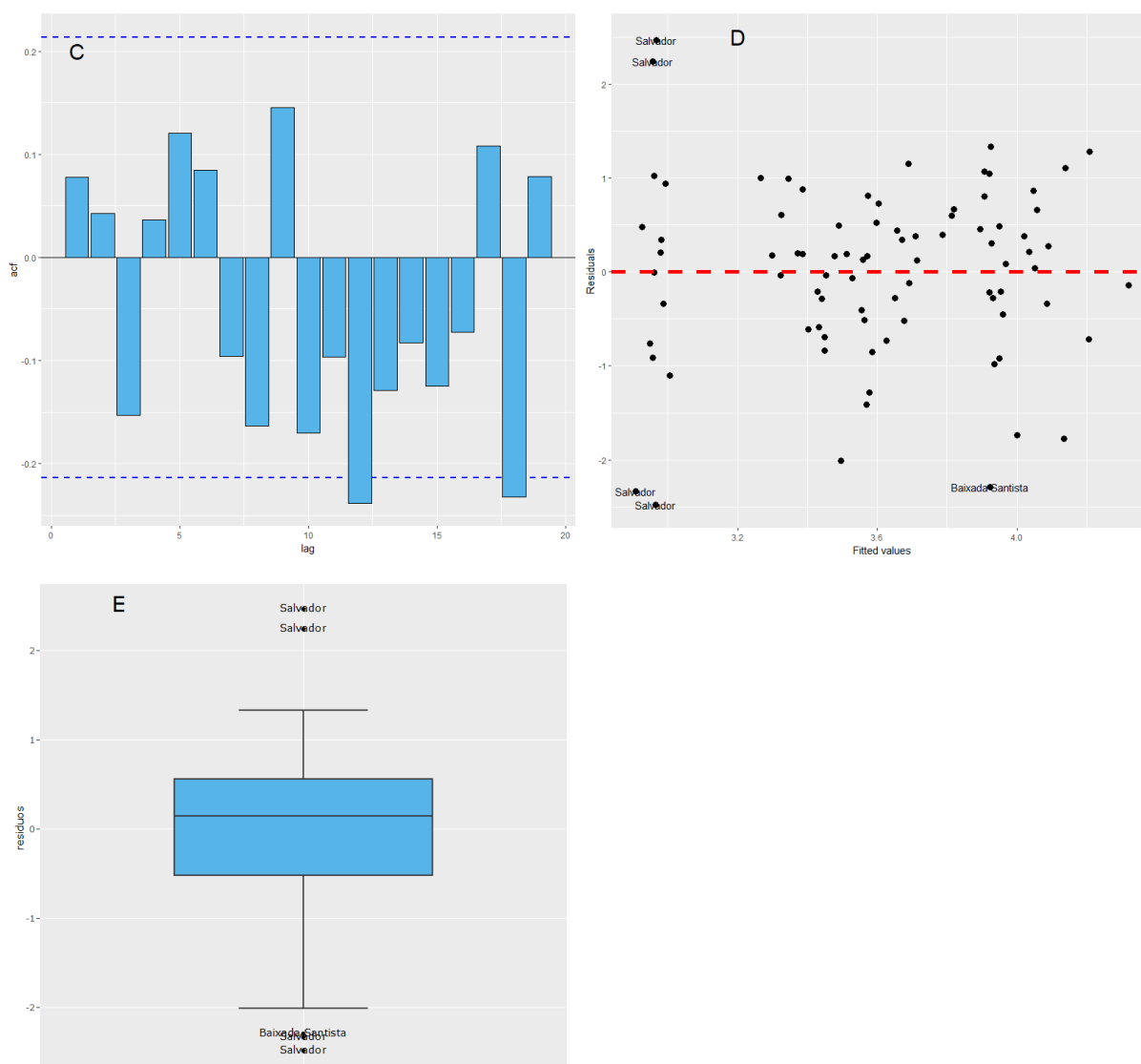


Figura 39 - Probabilidade dos resíduos padronizados (A), função de autocorrelação dos resíduos (B), função de autocorrelação parcial dos resíduos (C) e gráfico dos resíduos padronizados versus valores ajustados às médias anuais de PM_{10} ($*Y=\log(Y)$) (D) e box plot dos resíduos padronizados do modelo 1.3 (E)

Medidas Corretivas

Modelagem da autocorrelação

Na figura 39-B e C, a função de autocorrelação (FAC) é significativamente diferente de 0 na defasagem 12 e a função de autocorrelação parcial (FACP) é significativamente diferente de 0 na defasagem 12 e 18 ou seja picos são significativos através das defasagens. Isto sugere, por exemplo, que um modelo AR(1) pode ser adequado para acomodar correlação intragrupo.

O p-valor significativo para o TRV indica que o AR(1) fornece um ajuste substancialmente melhor dos dados do que o modelo de erros independentes, modelo 1.3, sugerindo que a correlação serial está presente nas regiões metropolitanas. Portanto, o modelo escolhido foi o Modelo 1.4, que assume uma estrutura de correlação serial autoregressiva de ordem 1, tabela 39.

Tabela 39 - Modelos considerados para o teste de estruturas de correlação dos erros aleatórios

| Mod. | corStruct | gl | AIC | BIC | logLik | Teste | TRV | p-value |
|-------------------|--------------------|----|---------|--------|--------|--------|------|---------|
| Modelo 1.3 | | 10 | -118,69 | -95,12 | 69,34 | | | |
| Modelo 1.4 | corARMA (p=1, q=0) | 11 | -123,21 | -97,29 | 72,61 | 1.3 vs | 6,53 | 0,01 |

Modelagem da heterocedasticidade

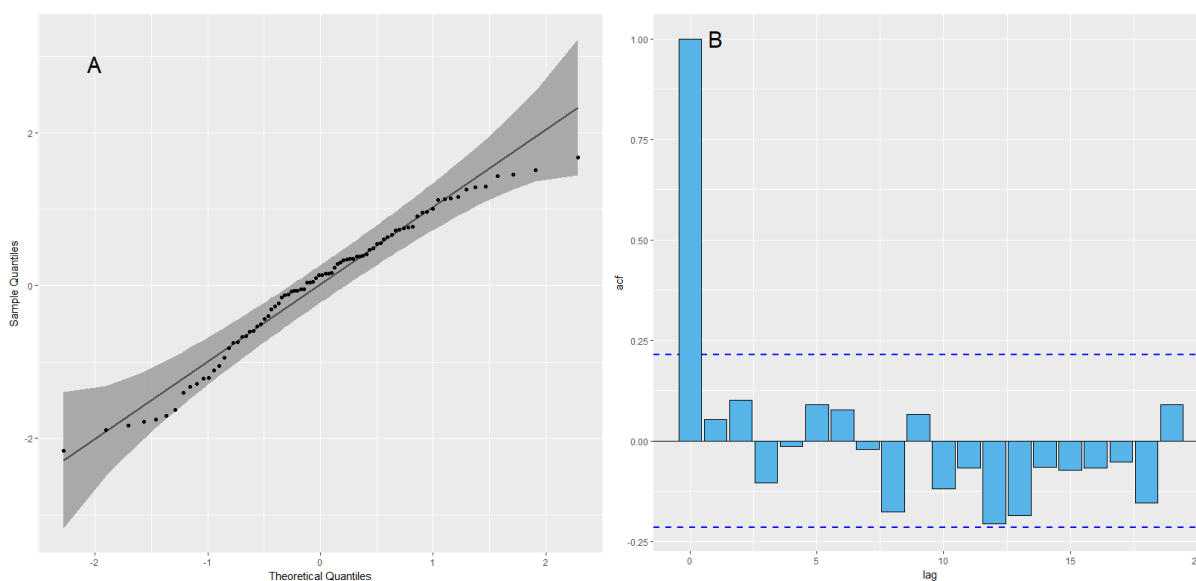
Quanto ao pressuposto da homocedasticidade, referente aos resultados da tabela 40 concluiu que somente o modelo que incorporou a função identidade explicou os dados significativamente melhor em comparação ao modelo que assume a homocedasticidade (Modelo 1.4). A comparação dos modelos com a estrutura de variância modelada foi feita através dos valores de AIC. O Modelo 1.4.4 foi o modelo que apresentou um menor valor de AIC, sendo, portanto, o modelo escolhido.

Tabela 40 - Resumo dos critérios utilizados na modelagem da estrutura de variância-covariância dos erros aleatórios, assumindo diferentes funções variância

| Modelos | VarFunc | gl | AIC | BIC | logLik | Test | L.Ratio | p-value |
|---------------------|----------------------|----|---------|--------|--------|------|---------|---------|
| Modelo 1.4 | | 11 | -123,21 | -97,29 | 72,61 | | | |
| Modelo 1.4.1 | <i>varPower</i> | 17 | -127,24 | -87,17 | 80,62 | 2 vs | 16,03 | 0,01 |
| Modelo 1.4.2 | <i>varConstPower</i> | 23 | -119,69 | -65,48 | 82,84 | 2 vs | 20,47 | 0,06 |
| Modelo 1.4.3 | <i>varExp</i> | 17 | -120,96 | -80,89 | 77,48 | 2 vs | 9,74 | 0,14 |
| Modelo 1.4.4 | <i>varIdent</i> | 16 | -128,79 | -91,08 | 80,40 | 2 vs | 15,58 | 0,01 |

Verificação da adequação do modelo

Após as medidas corretivas para o atendimento dos pressupostos do modelo, houve uma certa melhora nas suposições do modelo. De acordo com a figura 40-A, o pressuposto da normalidade dos erros dentro dos grupos não é rejeitado. Na figura 40-B e C, que representa função de autocorrelação empírica dos resíduos, não se observam autocorrelações significativas, indicando que os resíduos padronizados se comportam como ruído branco. E com relação a heterocedasticidade verificamos melhora significativas nos resíduos após as correções. Com a melhora dos pressupostos após as correções, o modelo se tornou mais eficiente para realizar as previsões.



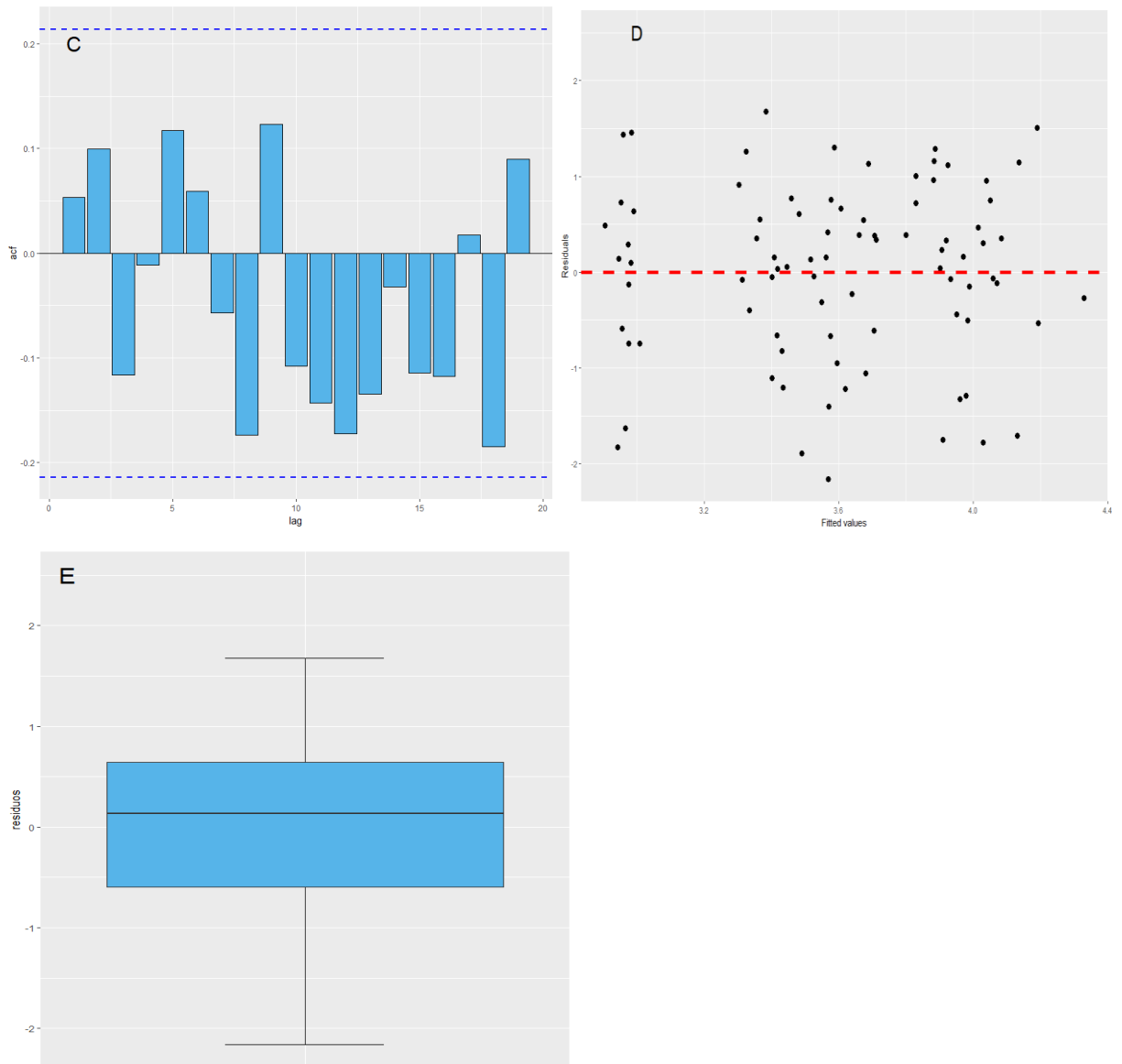


Figura 40 - Probabilidade dos resíduos padronizados (A), função de autocorrelação dos resíduos (B), função de autocorrelação parcial dos resíduos (C), gráfico dos resíduos padronizados versus valores ajustados às médias anuais de PM_{10} ($*Y=\log(Y)$) (D) e box plot dos resíduos padronizados (E) do Modelo 1.4.4

Modelo final

Descrição do modelo final para previsão

As médias anuais de PM_{10} ao longo do tempo foi modelada por um modelo linear de efeitos mistos AR(1) com estrutura da função variância *varIdent* (diferentes variâncias por estrato), dado por:

$$\begin{aligned} \ln[PM10_{tj}] = & \beta_0 \\ & + \beta_1 \text{tempo}_j + \beta_2 \ln(\text{umidade}_j) + \beta_3 \ln(\text{PIB serviços per capita}_j) \\ & + \beta_4 \ln(\text{altitude media}) + \beta_5 \ln(\text{focos}) + u_{0j} + u_{1j} + \varepsilon_{tj} \end{aligned}$$

Onde $u_{0j} \sim N(0, \sigma_{u_0}^2)$; $u_{1j} \sim N(0, \sigma_{u_1}^2)$; $\varepsilon_{tj} \sim N(0, \sigma_\varepsilon^2)$.

Tabela 41 - Estimativas dos parâmetros para os efeitos fixos e aleatórios, erro-padrão, estatística do teste-t e p-valores: modelo AR(1) com estrutura da função variância *varIndent*

| Variáveis | Estimativas | Erro Padrão | gl | T | Intervalo de. | | P-valor |
|-------------------------------------|-------------|-------------|----|--------|---------------|--------|---------|
| | | | | | confiança 95% | | |
| | | | | | Inf. | Sup. | |
| Resposta: Ln_PM₁₀ | | | | | | | |
| Parte fixa | | | | | | | |
| constante | 7,719 | 0,891 | 74 | 8,660 | 5,943 | 9,495 | 0,000 |
| tempo (linear) | -0,040 | 0,009 | 74 | -4,624 | -0,058 | -0,023 | 0,000 |
| Ln_ umidade | -0,769 | 0,179 | 74 | -4,295 | -1,126 | -0,412 | 0,000 |
| Ln_ pib serviços per capita | 0,272 | 0,075 | 74 | 3,615 | 0,122 | 0,421 | 0,001 |
| Ln_ altitude media | -0,329 | 0,068 | 4 | -4,809 | -0,519 | -0,139 | 0,009 |
| Ln_focos | 0,031 | 0,010 | 74 | 3,126 | 0,011 | 0,050 | 0,003 |
| Parte aleatória | | | | | | | |
| Nível: RM | | | | | | | |
| s.d. (cons.) | 0,566 | | | | 0,261 | 1,228 | |
| s.d. (trend) | 0,012 | | | | 0,006 | 0,026 | |
| corr.(cons./trend) | -0,993 | | | | -1,000 | 1,000 | |
| Nível: ano | | | | | | | |
| resíduos | 0,087 | | | | 0,053 | 0,143 | |
| <i>R</i> ² (m) | 27,54% | | | | | | |
| <i>R</i> ² (c) | 98,33% | | | | | | |
| <i>Units: RM</i> | 6 | | | | | | |
| <i>Units: ano</i> | 84 | | | | | | |
| <i>Estimation:</i> | REML | | | | | | |
| <i>AIC</i> | -128,792 | | | | | | |
| <i>logLik</i> | 80,396 | | | | | | |

| Anova | | | | | | | |
|-----------------------------|-------|-------|----|-------|---------|--------|------------|
| variáveis | SSQ | MSQ | gl | DenDF | Valor F | Pr(>F) | Pct. Expl. |
| tempo (linear) | 0,113 | 0,113 | 1 | 74 | 4,405 | 0,039 | 24,92% |
| Ln_ umidade | 0,098 | 0,098 | 1 | 74 | 10,020 | 0,002 | 21,51% |
| Ln_ pib serviços per capita | 0,069 | 0,069 | 1 | 74 | 6,615 | 0,012 | 15,24% |
| Ln_ altitude media | 0,123 | 0,123 | 1 | 4 | 11,526 | 0,027 | 26,95% |
| Ln_focos | 0,052 | 0,052 | 1 | 74 | 6,305 | 0,014 | 11,39% |

correlação de Pearson (r) = 0,98

RMSE = 0,070 µg/m³

MAE = 0,052

VIF: tempo = 2,62; Ln_ umidade = 1,08; Ln_ pib serviços per capita = 2,92; Ln_ altitude media = 1,23, Ln_focos = 1,22

Tabela 42 - Médias anuais dos níveis de concentração de PM₁₀ das regiões metropolitanas, estimadas pelo modelo de efeitos mistos no período de 2001 - 2014

| Regiões metropolitanas | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|-------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Agreste | 18,5 | 19,7 | 20,8 | 19,8 | 20,0 | 19,7 | 19,4 | 19,6 | 19,7 | 19,5 | 19,2 | 19,5 | 16,5 | 16,1 |
| Aracaju | - | 56,3 | 55,4 | 52,9 | 52,6 | 50,3 | | 49,4 | 51,1 | 50,3 | 49,4 | 50,0 | 47,1 | 42,7 |
| Baixada Santista | 51,4 | 48,2 | 53,8 | 50,0 | 50,8 | 53,3 | 55,8 | 51,9 | 42,8 | 49,7 | 53,8 | 50,8 | 52,8 | 54,5 |
| Belém | - | 48,8 | 51,2 | 50,5 | 47,2 | 49,0 | 48,5 | 49,2 | 49,9 | 48,3 | 47,4 | 46,5 | 45,1 | 44,5 |
| Belo Horizonte | 21,8 | 21,7 | 22,5 | 20,8 | 19,8 | 20,0 | 22,5 | 20,4 | 19,6 | 21,7 | 21,4 | 20,4 | 18,9 | 20,8 |
| Campina Grande | 17,1 | 17,1 | 17,5 | 17,2 | 16,8 | 17,3 | 17,1 | 16,9 | 17,1 | 16,5 | 15,8 | 16,5 | 14,1 | 14,4 |
| Campinas | 42,0 | 43,5 | 40,0 | 33,9 | 32,0 | 37,8 | 38,7 | 34,0 | 28,5 | 34,0 | 34,0 | 31,5 | 30,5 | 34,0 |
| Carbonífera | - | 36,5 | 37,5 | 36,2 | 35,3 | 35,5 | 36,0 | 35,3 | 35,7 | 33,5 | 33,1 | 35,2 | 31,9 | 30,8 |
| Cariri | 23,3 | 21,4 | 21,7 | 19,1 | 19,6 | 18,2 | 17,0 | 17,0 | 16,1 | 17,2 | 15,3 | 17,6 | 15,7 | 15,9 |
| Chapecó | 22,5 | 22,7 | 23,9 | 22,9 | 22,3 | 22,0 | 22,2 | 22,3 | 21,1 | 21,5 | 20,4 | 20,5 | 19,7 | 19,4 |
| Curitiba | 19,6 | 19,8 | 21,1 | 19,6 | 18,9 | 19,6 | 20,6 | 19,6 | 19,1 | 19,1 | 18,9 | 18,4 | 17,6 | 17,3 |
| Florianópolis | 28,8 | 27,5 | 31,0 | 29,2 | 29,0 | 29,3 | 29,8 | 29,1 | 29,1 | 28,3 | 27,7 | 27,2 | 26,0 | 26,4 |
| Fortaleza | 43,2 | 40,1 | 40,4 | 39,3 | 40,2 | 37,9 | 39,9 | 39,0 | 37,7 | 41,0 | 38,3 | 39,0 | 35,5 | 36,0 |
| Foz do Rio Itajaí | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Goiânia | 22,1 | 23,9 | 22,9 | 22,2 | 21,1 | 21,0 | 23,7 | 21,4 | 20,1 | 22,6 | 21,2 | 20,8 | 19,4 | 20,1 |
| Grande São Luís | 56,2 | 57,8 | 56,9 | 56,9 | 56,7 | 53,7 | 51,8 | 54,2 | 54,0 | 52,8 | 50,8 | 53,1 | 47,1 | 46,1 |
| Grande Vitória | 27,6 | 27,4 | 30,0 | 28,7 | 29,6 | 31,5 | 29,6 | 32,8 | 29,7 | 30,4 | 30,6 | 29,9 | 28,2 | 29,1 |
| João Pessoa | 41,6 | 42,4 | 42,9 | 41,1 | 41,7 | 41,8 | 40,9 | 41,6 | 41,7 | 40,4 | 39,2 | 39,3 | 34,3 | 35,9 |
| Lages | 15,6 | 15,2 | 16,4 | 15,4 | 14,8 | 15,4 | 15,7 | 15,0 | 15,2 | 15,1 | 14,7 | 14,8 | 13,4 | 13,4 |
| Londrina | 22,3 | 22,7 | 23,1 | 21,8 | 21,0 | 20,8 | 22,0 | 20,1 | 17,8 | 18,3 | 21,1 | 20,6 | 19,4 | 19,4 |
| Macapá | 70,1 | 68,1 | 65,8 | 65,7 | 61,1 | 61,4 | 63,1 | 62,7 | 63,5 | 59,3 | 60,1 | 60,0 | 49,9 | 49,1 |
| Maceió | 42,7 | 42,7 | 43,4 | 43,4 | 43,1 | 42,0 | 35,1 | 42,3 | 41,1 | 39,4 | 39,2 | 38,8 | 35,3 | 34,4 |
| Manaus | 49,5 | 49,4 | 47,6 | 46,5 | 46,6 | 44,9 | 45,9 | 44,0 | 49,2 | 46,7 | 46,1 | 45,2 | 42,1 | 41,8 |
| Maringá | 24,0 | 24,7 | 25,3 | 24,4 | 24,5 | 24,5 | 24,5 | 22,7 | 21,4 | 22,4 | 22,7 | 22,2 | 21,0 | 20,3 |
| Natal | 38,0 | | 39,2 | 37,3 | 37,2 | 37,3 | 36,4 | 35,5 | 36,4 | 36,8 | 34,5 | 34,5 | 31,5 | 31,8 |
| Norte/Nordeste Catarinense | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Porto Alegre | 51,2 | 50,2 | 49,8 | 48,7 | 49,2 | 46,6 | 48,4 | 47,6 | 46,7 | 46,5 | 45,2 | 49,3 | 44,4 | 42,5 |
| Recife | 41,1 | 40,2 | 40,7 | 38,6 | 40,0 | 39,1 | 37,3 | 37,3 | 37,6 | 38,1 | 37,6 | 38,5 | 35,4 | 34,1 |
| Rio de Janeiro | 74,4 | 73,8 | 63,7 | 58,0 | 54,9 | 60,9 | 68,0 | 57,3 | 57,5 | 60,7 | 60,9 | 57,6 | 47,6 | 48,6 |
| Salvador | 19,3 | 19,3 | 19,3 | 20,2 | 20,8 | 16,2 | 18,0 | 19,9 | 18,6 | 21,4 | 22,6 | 23,2 | 18,0 | 15,5 |
| São Paulo | 47,9 | 47,3 | 45,4 | 39,7 | 40,2 | 38,0 | 41,3 | 35,6 | 32,6 | 37,7 | 38,0 | 36,0 | 33,9 | 35,5 |
| Sudoeste Maranhense | 27,5 | 28,2 | 27,5 | 27,3 | 28,6 | 26,2 | 26,2 | 24,9 | 24,9 | 25,1 | 22,2 | 23,2 | 22,8 | 23,1 |
| Tubarão | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Vale do Aço | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Vale do Itajaí | 28,9 | 29,1 | 29,6 | 27,8 | 26,7 | 28,1 | 27,3 | 26,8 | 26,2 | 27,1 | 27,6 | 27,4 | 25,7 | 26,9 |
| Vale do Rio Cuiabá | 33,2 | 32,8 | 31,7 | 32,8 | 33,6 | 29,4 | 33,0 | 30,3 | 29,7 | 32,7 | 30,9 | 29,9 | 28,0 | 27,7 |
| Brasil | 35,2 | 36,1 | 36,2 | 34,6 | 34,2 | 34,0 | 34,1 | 33,6 | 32,9 | 33,6 | 33,1 | 33,0 | 30,3 | 30,3 |

Não estavam disponíveis os dados climáticos para as regiões metropolitanas Foz do Rio Itajaí, Norte/Nordeste Catarinense, Tubarão e Vale, portanto não foram calculados valores previstos da poluição do ar.

APÊNDICE I – Tabelas ilustrativas de avaliação da carga de mortalidade nas Regiões metropolitanas brasileiras, 2014

Tabela 43 - Previsão das médias anuais de PM₁₀ do modelo de efeitos mistos, estimativas das médias anuais de PM_{2,5} e óbitos por doenças cardiovasculares e respiratórias em áreas metropolitanas do Brasil, 2014

| Regiões Metropolitanas | População (100.000) | PM ₁₀ (µg/m ³) | PM _{2,5} (µg/m ³) | Óbitos | | |
|----------------------------|---------------------|---------------------------------------|--|------------------|------------------|-------------------|
| | | | | Cardiovascular | Respiratórias | |
| | | | | Adultos ≥30 anos | Crianças < 1 ano | Crianças 1-5 anos |
| Agreste | 6,4 | 16,1 | 8,1 | 801 | 2 | 3 |
| Aracaju | 9,1 | 42,7 | 21,4 | 1.106 | 12 | 2 |
| Baixada Santista | 17,8 | 54,5 | 27,3 | 4.430 | 18 | 7 |
| Belém | 21,9 | 44,5 | 22,2 | 3.002 | 32 | 23 |
| Belo Horizonte | 57,7 | 20,8 | 10,4 | 7.659 | 20 | 25 |
| Campina Grande | 7,2 | 14,4 | 7,2 | 1.233 | 3 | 3 |
| Campinas | 30,4 | 34,0 | 17,0 | 4.640 | 12 | 17 |
| Carbonífera | 5,7 | 30,8 | 15,4 | 1.102 | 3 | 1 |
| Cariri | 5,9 | 15,9 | 8,0 | 988 | 8 | 1 |
| Chapecó | 4,3 | 19,4 | 9,7 | 583 | 1 | 1 |
| Curitiba | 34,1 | 17,3 | 8,7 | 5.331 | 21 | 6 |
| Florianópolis | 11,1 | 26,4 | 13,2 | 1.672 | 6 | 2 |
| Fortaleza | 38,2 | 36,0 | 18,0 | 4.854 | 31 | 9 |
| Foz do Rio Itajaí | 6,1 | - | - | 800 | 1 | 2 |
| Goiânia | 23,8 | 20,1 | 10,0 | 3.475 | 22 | 6 |
| Grande São Luís | 14,0 | 46,1 | 23,1 | 2.209 | 22 | 8 |
| Grande Vitória | 18,8 | 29,1 | 14,5 | 2.803 | 10 | 5 |
| João Pessoa | 12,8 | 35,9 | 17,9 | 2.281 | 17 | 6 |
| Lages | 3,6 | 13,4 | 6,7 | 686 | 1 | 2 |
| Londrina | 8,2 | 19,4 | 9,7 | 1.942 | - | 2 |
| Macapá | 5,6 | 49,1 | 24,6 | 467 | 17 | 13 |
| Maceió | 12,5 | 34,4 | 17,2 | 2.135 | 18 | 7 |
| Manaus | 23,6 | 41,8 | 20,9 | 1.968 | 44 | 21 |
| Maringá | 6,6 | 20,3 | 10,2 | 1.277 | 1 | 3 |
| Natal | 14,7 | 31,8 | 15,9 | 1.981 | 10 | 5 |
| Norte/Nordeste Catarinense | 11,9 | - | - | 1.860 | 8 | 4 |
| Porto Alegre | 41,6 | 42,5 | 21,3 | 8.149 | 30 | 12 |
| Recife | 38,9 | 34,1 | 17,0 | 6.919 | 37 | 18 |
| Rio de Janeiro | 121,2 | 48,6 | 24,3 | 27.817 | 129 | 50 |
| Salvador | 39,2 | 15,5 | 7,7 | 5.382 | 34 | 25 |
| São Paulo | 209,4 | 35,5 | 17,7 | 39.108 | 226 | 109 |
| Sudoeste Maranhense | 3,5 | 23,1 | 11,5 | 463 | 7 | - |
| Tubarão | 3,7 | - | - | 725 | 5 | 2 |
| Vale do Aço | 6,5 | - | - | 1.161 | 7 | 4 |
| Vale do Itajaí | 7,6 | 26,9 | 13,5 | 1.025 | 2 | 1 |
| Vale do Rio Cuiabá | 8,7 | 27,7 | 13,8 | 1.402 | 12 | 13 |
| Brasil | 892,4 | 34,2* | 17,1* | 153.436 | 829 | 418 |

Mortalidade: <http://datasus.saude.gov.br/informacoes-de-saude/tabnet/estatisticas-vitais>;

População: Disponíveis no site do Instituto Brasileiro de Geografia e Estatística – IBGE.

* Média ponderada pela população.

Tabela 44 - Risco relativo da mortalidade associada com PM_{2,5} para doenças cardiovasculares e PM₁₀ para doenças respiratórias de acordo com o cenário base (IC95%)

| Regiões Metropolitanas | Risco Relativo | | |
|-----------------------------------|----------------------------|------------------------|----------------------------|
| | Cardiovascular (IC95%) | Respiratórias (IC95%) | |
| | adultos >=30 anos | crianças < 1 ano | crianças 1-5 anos |
| Agreste | 1,005 (1,002-1,008) | 1,001 (1-1,001) | 1,001 (1-1,001) |
| Aracaju | 1,132 (1,046-1,225) | 1,013 (1-1,024) | 1,016 (1,002-1,03) |
| Baixada Santista | 1,193 (1,066-1,335) | 1,019 (1-1,034) | 1,023 (1,003-1,043) |
| Belém | 1,141 (1,049-1,241) | 1,014 (1-1,025) | 1,017 (1,002-1,032) |
| Belo Horizonte | 1,026 (1,009-1,043) | 1,003 (1-1,005) | 1,003 (1-1,006) |
| Campina Grande | 0,997 (0,999-0,996) | 1 (1-1) | 1 (1-0,999) |
| Campinas | 1,089 (1,031-1,149) | 1,009 (1-1,016) | 1,011 (1,002-1,021) |
| Carbonífera | 1,073 (1,026-1,123) | 1,007 (1-1,013) | 1,009 (1,001-1,017) |
| Cariri | 1,004 (1,001-1,007) | 1 (1-1,001) | 1,001 (1-1,001) |
| Chapecó | 1,02 (1,007-1,032) | 1,002 (1-1,004) | 1,003 (1-1,005) |
| Curitiba | 1,01 (1,004-1,017) | 1,001 (1-1,002) | 1,001 (1-1,002) |
| Florianópolis | 1,052 (1,019-1,087) | 1,005 (1-1,01) | 1,007 (1,001-1,012) |
| Fortaleza | 1,098 (1,034-1,166) | 1,01 (1-1,018) | 1,012 (1,002-1,023) |
| Foz do Rio Itajaí | | | |
| Goiânia | 1,023 (1,008-1,038) | 1,002 (1-1,004) | 1,003 (1-1,005) |
| Grande São Luís | 1,149 (1,051-1,256) | 1,015 (1-1,027) | 1,018 (1,002-1,034) |
| Grande Vitória | 1,065 (1,023-1,108) | 1,007 (1-1,012) | 1,008 (1,001-1,015) |
| João Pessoa | 1,098 (1,034-1,165) | 1,01 (1-1,018) | 1,012 (1,002-1,023) |
| Lages | 0,993 (0,997-0,989) | 0,999 (1-0,999) | 0,999 (1-0,998) |
| Londrina | 1,02 (1,007-1,032) | 1,002 (1-1,004) | 1,003 (1-1,005) |
| Macapá | 1,165 (1,057-1,284) | 1,016 (1-1,029) | 1,02 (1,003-1,037) |
| Maceió | 1,09 (1,032-1,152) | 1,009 (1-1,017) | 1,011 (1,002-1,021) |
| Manaus | 1,127 (1,044-1,217) | 1,013 (1-1,023) | 1,016 (1,002-1,029) |
| Maringá | 1,024 (1,009-1,04) | 1,002 (1-1,005) | 1,003 (1-1,006) |
| Natal | 1,078 (1,027-1,131) | 1,008 (1-1,014) | 1,01 (1,001-1,018) |
| Norte/Nordeste Catarinense | | | |
| Porto Alegre | 1,131 (1,045-1,223) | 1,013 (1-1,024) | 1,016 (1,002-1,03) |
| Recife | 1,089 (1,031-1,15) | 1,009 (1-1,016) | 1,011 (1,002-1,021) |
| Rio de Janeiro | 1,162 (1,056-1,279) | 1,016 (1-1,029) | 1,02 (1,003-1,037) |
| Salvador | 1,002 (1,001-1,003) | 1 (1-1) | 1 (1-1,001) |
| São Paulo | 1,096 (1,034-1,162) | 1,01 (1-1,017) | 1,012 (1,002-1,022) |
| Sudoeste Maranhense | 1,037 (1,013-1,061) | 1,004 (1-1,007) | 1,005 (1,001-1,009) |
| Tubarão | | | |
| Vale do Aço | | | |
| Vale do Itajaí | 1,055 (1,019-1,091) | 1,006 (1-1,01) | 1,007 (1,001-1,013) |
| Vale do Rio Cuiabá | 1,058 (1,021-1,097) | 1,006 (1-1,011) | 1,007 (1,001-1,014) |
| Brasil | 1,090 (1,031-1,151) | 1,009 (1-1,016) | 1,011 (1,002-1,021) |

Elaboração própria

Tabela 45 - Fração atribuível da mortalidade associada com PM_{2,5} para doenças cardiovasculares e PM₁₀ para doenças respiratórias de acordo com o cenário base (IC 95%)

| Regiões Metropolitanas | Fração atribuível | | |
|----------------------------|----------------------------|------------------------|----------------------------|
| | Cardiovascular (IC95%) | Respiratórias (IC95%) | |
| | adultos ≥30 anos | crianças < 1 ano | crianças 1-5 anos |
| Agreste | 0,005 (0,002-0,008) | 0,001 (0-0,001) | 0,001 (0-0,001) |
| Aracaju | 0,117 (0,044-0,184) | 0,013 (0-0,023) | 0,016 (0,002-0,029) |
| Baixada Santista | 0,162 (0,062-0,251) | 0,018 (0-0,033) | 0,023 (0,003-0,042) |
| Belém | 0,123 (0,046-0,194) | 0,014 (0-0,025) | 0,017 (0,002-0,031) |
| Belo Horizonte | 0,026 (0,009-0,042) | 0,003 (0-0,005) | 0,003 (0-0,006) |
| Campina Grande | -0,003 (-0,001--0,004) | 0 (0-0) | 0 (0--0,001) |
| Campinas | 0,081 (0,03-0,13) | 0,009 (0-0,016) | 0,011 (0,002-0,02) |
| Carbonífera | 0,068 (0,025-0,109) | 0,007 (0-0,013) | 0,009 (0,001-0,017) |
| Cariri | 0,004 (0,001-0,007) | 0 (0-0,001) | 0,001 (0-0,001) |
| Chapecó | 0,019 (0,007-0,031) | 0,002 (0-0,004) | 0,003 (0-0,005) |
| Curitiba | 0,01 (0,004-0,017) | 0,001 (0-0,002) | 0,001 (0-0,002) |
| Florianópolis | 0,05 (0,018-0,08) | 0,005 (0-0,01) | 0,007 (0,001-0,012) |
| Fortaleza | 0,089 (0,033-0,142) | 0,01 (0-0,018) | 0,012 (0,002-0,022) |
| Foz do Rio Itajaí | - | - | - |
| Goiânia | 0,022 (0,008-0,036) | 0,002 (0-0,004) | 0,003 (0-0,005) |
| Grande São Luís | 0,13 (0,049-0,204) | 0,014 (0-0,026) | 0,018 (0,002-0,033) |
| Grande Vitória | 0,061 (0,022-0,098) | 0,007 (0-0,012) | 0,008 (0,001-0,015) |
| João Pessoa | 0,089 (0,033-0,142) | 0,01 (0-0,018) | 0,012 (0,002-0,022) |
| Lages | -0,007 (-0,003--0,012) | -0,001 (0--0,001) | -0,001 (0--0,002) |
| Londrina | 0,019 (0,007-0,031) | 0,002 (0-0,004) | 0,003 (0-0,005) |
| Macapá | 0,141 (0,053-0,221) | 0,016 (0-0,028) | 0,02 (0,003-0,036) |
| Maceió | 0,083 (0,031-0,132) | 0,009 (0-0,016) | 0,011 (0,002-0,021) |
| Manaus | 0,113 (0,042-0,178) | 0,012 (0-0,022) | 0,015 (0,002-0,028) |
| Maringá | 0,023 (0,009-0,038) | 0,002 (0-0,004) | 0,003 (0-0,006) |
| Natal | 0,072 (0,027-0,116) | 0,008 (0-0,014) | 0,01 (0,001-0,018) |
| Norte/Nordeste Catarinense | - | - | - |
| Porto Alegre | 0,116 (0,043-0,182) | 0,013 (0-0,023) | 0,016 (0,002-0,029) |
| Recife | 0,082 (0,03-0,13) | 0,009 (0-0,016) | 0,011 (0,002-0,02) |
| Rio de Janeiro | 0,139 (0,053-0,218) | 0,016 (0-0,028) | 0,019 (0,003-0,035) |
| Salvador | 0,002 (0,001-0,003) | 0 (0-0) | 0 (0-0,001) |
| São Paulo | 0,087 (0,032-0,139) | 0,01 (0-0,017) | 0,012 (0,002-0,022) |
| Sudoeste Maranhense | 0,035 (0,013-0,057) | 0,004 (0-0,007) | 0,005 (0,001-0,009) |
| Tubarão | - | - | - |
| Vale do Aço | - | - | - |
| Vale do Itajaí | 0,052 (0,019-0,084) | 0,006 (0-0,01) | 0,007 (0,001-0,013) |
| Vale do Rio Cuiabá | 0,055 (0,02-0,089) | 0,006 (0-0,011) | 0,007 (0,001-0,014) |
| Brasil | 0,082 (0,030-0,131) | 0,009 (0-0,016) | 0,011 (0,002-0,020) |

Elaboração própria

Tabela 46 - Óbitos atribuíveis associados com PM_{2,5} para doenças cardiovasculares e PM₁₀ para doenças respiratórias de acordo com o cenário base (IC 95%)

| Regiões Metropolitanas | Óbitos atribuíveis | | |
|----------------------------|------------------------------|-----------------------|----------------------|
| | Cardiovascular (IC95%) | Respiratórias (IC95%) | |
| | adultos >=30 anos | Crianças < 1 ano | Crianças 1-5 anos |
| Agreste | 4 (1-6) | 0 (0-0) | 0 (0-0) |
| Aracaju | 129 (48-203) | 0 (0-0) | 0 (0-0) |
| Baixada Santista | 717 (273-1.112) | 0 (0-1) | 0 (0-0) |
| Belém | 371 (139-583) | 0 (0-1) | 0 (0-1) |
| Belo Horizonte | 196 (71-319) | 0 (0-0) | 0 (0-0) |
| Campina Grande | 0 (0-0) | 0 (0-0) | 0 (0-0) |
| Campinas | 378 (140-602) | 0 (0-0) | 0 (0-0) |
| Carbonífera | 75 (28-121) | 0 (0-0) | 0 (0-0) |
| Cariri | 4 (1-7) | 0 (0-0) | 0 (0-0) |
| Chapecó | 11 (4-18) | 0 (0-0) | 0 (0-0) |
| Curitiba | 55 (20-89) | 0 (0-0) | 0 (0-0) |
| Florianópolis | 83 (30-134) | 0 (0-0) | 0 (0-0) |
| Fortaleza | 434 (161-691) | 0 (0-1) | 0 (0-0) |
| Foz do Rio Itajaí | | | |
| Goiânia | 78 (28-127) | 0 (0-0) | 0 (0-0) |
| Grande São Luís | 287 (108-450) | 0 (0-1) | 0 (0-0) |
| Grande Vitória | 171 (63-274) | 0 (0-0) | 0 (0-0) |
| João Pessoa | 203 (75-324) | 0 (0-0) | 0 (0-0) |
| Lages | 0 (0-0) | 0 (0-0) | 0 (0-0) |
| Londrina | 37 (14-61) | | 0 (0-0) |
| Macapá | 66 (25-103) | 0 (0-0) | 0 (0-0) |
| Maceió | 177 (66-282) | 0 (0-0) | 0 (0-0) |
| Manaus | 222 (83-351) | 1 (0-1) | 0 (0-1) |
| Maringá | 30 (11-49) | 0 (0-0) | 0 (0-0) |
| Natal | 143 (53-230) | 0 (0-0) | 0 (0-0) |
| Norte/Nordeste | | | |
| Catarinense | | | |
| Porto Alegre | 942 (353-1486) | 0 (0-1) | 0 (0-0) |
| Recife | 565 (209-902) | 0 (0-1) | 0 (0-0) |
| Rio de Janeiro | 3.878 (1.465-6.067) | 2 (0-4) | 1 (0-2) |
| Salvador | 11 (4-18) | 0 (0-0) | 0 (0-0) |
| São Paulo | 3.417 (1.268-5.441) | 2 (0-4) | 1 (0-2) |
| Sudoeste Maranhense | 16 (6-27) | 0 (0-0) | |
| Tubarão | | | |
| Vale do Aço | | | |
| Vale do Itajaí | 53 (19-86) | 0 (0-0) | 0 (0-0) |
| Vale do Rio Cuiabá | 77 (28-124) | 0 (0-0) | 0 (0-0) |
| Brasil | 12.615 (4.673-20.126) | 7 (0-13) | 5 (1-9) |

Elaboração própria

APÊNDICE J – Características das redes de monitoramento das regiões metropolitanas do Brasil

Tabela 47 - RM São Paulo: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|--------|--------|-------------------|
| Região metropolitana de São Paulo | | | |
| Diadema | CETESB | Urbana | Automático |
| Guarulhos | CETESB | Urbana | Automático |
| Guarulhos - Paço Municipal | CETESB | Urbana | Automático |
| Mauá | CETESB | Urbana | Automático |
| Osasco | CETESB | Urbana | Automático/Manual |
| Santo André - Capuava | CETESB | Urbana | Automático/Manual |
| Santo André - Centro | CETESB | Urbana | Automático |
| Santo André - Paço Municipal | CETESB | Urbana | Automático |
| São Bernardo do Campo-Paulicéia | CETESB | Urbana | Automático/Manual |
| São Bernardo do Campo | CETESB | Urbana | Automático/Manual |
| São Caetano do Sul | CETESB | Urbana | Automático/Manual |
| Cambuci | CETESB | Urbana | Automático |
| Capão Redondo | CETESB | Urbana | Automático |
| Carapicuíba | CETESB | Urbana | Automático |
| Centro | CETESB | Urbana | Automático |
| Cerqueira César | CETESB | Urbana | Automático/Manual |
| Congonhas | CETESB | Urbana | Automático/Manual |
| Grajaú-Parelheiros | CETESB | Urbana | Automático |
| Ibirapuera | CETESB | Urbana | Automático/Manual |
| Interlagos | CETESB | Urbana | Automático |
| Itaim Paulista | CETESB | Urbana | Automático |
| Itaquera | CETESB | Urbana | Automático |
| Lapa | CETESB | Urbana | Automático |
| Marg. Tietê - Pte Remédios | CETESB | Urbana | Automático |
| Mooca | CETESB | Urbana | Automático |
| N. Senhora do Ó | CETESB | Urbana | Automático |
| Parelheiros | CETESB | Urbana | Automático |
| Parque D. Pedro II | CETESB | Urbana | Automático |
| Penha | CETESB | Urbana | Automático |
| Pico do Jaraguá | CETESB | Urbana | Automático |
| Pinheiros | CETESB | Urbana | Automático/Manual |
| Santana | CETESB | Urbana | Automático |
| Santo Amaro | CETESB | Urbana | Automático/Manual |
| São Miguel Paulista | CETESB | Urbana | Automático |
| Taboão da Serra | CETESB | Urbana | Automático |

Elaboração própria

Tabela 48 - RM Baixada Santista: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|--------|--------|-------------------|
| Região metropolitana Baixada Santista | | | |
| Cubatão-Centro | CETESB | Urbana | Automático |
| Cubatão-Vale do Mogi | CETESB | Urbana | Automático |
| Cubatão-Vila Parisi | CETESB | Urbana | Automático/Manual |
| Santos | CETESB | Urbana | Automático |
| Santos - Porto | CETESB | Urbana | Manual |
| Santos-Ponta da Praia | CETESB | Urbana | Automático/Manual |

Elaboração própria

Tabela 49 - RM Campinas: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|--------|--------|-----------------|
| Região metropolitana Campinas | | | |
| Americana | CETESB | Urbana | Automático |
| Campinas-Centro | CETESB | Urbana | Automático |
| Paulínia | CETESB | Urbana | Automático |
| Paulínia Sul | CETESB | Urbana | Automático |

Elaboração própria

Tabela 50 - RM Rio de Janeiro: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|------|--------|-----------------|
| Região metropolitana Rio de Janeiro | | | |
| Bangu | SMAC | Urbana | Automática |
| Bonsucesso | INEA | Urbana | Semiautomática |
| Botafogo | INEA | Urbana | Semiautomática |
| Campo Grande | SMAC | Urbana | Automática |
| Campos Afonsos | INEA | Urbana | Semiautomática |
| Castelo | INEA | Urbana | Semiautomática |
| Centro | INEA | Urbana | Automática |
| Centro | SMAC | Urbana | Automática |
| Cidade de Deus | INEA | Urbana | Semiautomática |
| Copacabana | INEA | Urbana | Semiautomática |
| Copacabana | SMAC | Urbana | Automática |
| Duque de Caxias Campos Elíseos | INEA | Urbana | Automática |
| Duque de Caxias INSS | INEA | Urbana | Semiautomática |
| Duque de Caxias Jardim Primavera | INEA | Urbana | Automática |
| Duque de Caxias Pilar | INEA | Urbana | Automática |
| Duque de Caxias São Bento | INEA | Urbana | Automática |
| Duque de Caxias Vila São Luís | INEA | Urbana | Automática |
| Engenhão | INEA | Urbana | Automática |
| Gamboa | INEA | Urbana | Semiautomática |
| Gericinó | INEA | Urbana | Semiautomática |
| Guapimirim | INEA | Urbana | Semiautomática |

Elaboração própria

Tabela 50 - RM Rio de Janeiro: Características das estações de monitoramento PM₁₀
(Conclusão)

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|------|--------|-----------------|
| Região metropolitana Rio de Janeiro | | | |
| Ilha de Paquetá | INEA | Urbana | Automática |
| Ilha do Governador | INEA | Urbana | Automática |
| Itaboraí Alto Jacu | INEA | Urbana | Semiautomática |
| Itaboraí Fazenda Macuco | INEA | Urbana | Semiautomática |
| Itaboraí Itambi | INEA | Urbana | Semiautomática |
| Itaboraí Porto das Caixas | INEA | Urbana | Automática |
| Itaboraí Sambaetiba | INEA | Urbana | Automática |
| Itaboraí Sambaetiba | INEA | Urbana | Automática |
| Itaguaí Coroa Grande | INEA | Urbana | Automática |
| Itaguaí Monte Serrat | INEA | Urbana | Automática |
| Japeri Engenheiro Pedreira | INEA | Urbana | Automática |
| Lab. INEA | INEA | Urbana | Automática |
| Leblon | INEA | Urbana | Semiautomática |
| Nilópolis Rodoviária | INEA | Urbana | Semiautomática |
| Niterói Centro | INEA | Urbana | Semiautomática |
| Niterói Policlínica | INEA | Urbana | Semiautomática |
| Nova Iguaçu Monteiro Lobato | INEA | Urbana | Automática |
| Nova Iguaçu Monteiro Lobato | INEA | Urbana | Semiautomática |
| Pedra de Guaratiba | SMAC | Urbana | Automática |
| Santa Cruz Adalgisa Nery | INEA | Urbana | Automática |
| Santa Cruz Largo Bodegão | INEA | Urbana | Automática |
| São Cristóvão | INEA | Urbana | Semiautomática |
| São Cristóvão | SMAC | Urbana | Automática |
| São Gonçalo - Uerj | INEA | Urbana | Automática |
| São Gonçalo Prefeitura | INEA | Urbana | Semiautomática |
| Seropédica Embrapa | INEA | Urbana | Semiautomática |
| Sumaré | INEA | Urbana | Semiautomática |
| Taquara | INEA | Urbana | Automática |
| Tijuca | SMAC | Urbana | Automática |
| Uerj | INEA | Urbana | Semiautomática |
| Vilar dos Teles | INEA | Urbana | Semiautomática |

Elaboração própria

Tabela 51 - RM Salvador: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|-------|--------|-----------------|
| Região metropolitana Salvador | | | |
| Malembá | INEMA | Urbana | Automática |
| Av. ACM/Detran | INEMA | Urbana | Automática |
| Barros Reis | INEMA | Urbana | Automática |
| Campo Grande | INEMA | Urbana | Automática |
| Concórdia | INEMA | Urbana | Automática |
| Dique do Tororó | INEMA | Urbana | Automática |
| Gravatá | INEMA | Urbana | Automática |
| Itaigara | INEMA | Urbana | Automática |
| Leandrino | INEMA | Urbana | Automática |
| Madre de deus | INEMA | Urbana | Automática |
| Paralela - Cab. | INEMA | Urbana | Automática |
| Pirajá | INEMA | Urbana | Automática |
| Rio Vermelho | INEMA | Urbana | Automática |

Elaboração própria

Tabela 52 - RM Belo Horizonte: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|------|--------|-----------------|
| Região metropolitana Belo Horizonte | | | |
| Aeroporto | FEAM | Urbana | Automática |
| Av. Amazonas | FEAM | Urbana | Automática |
| B Jardim Alterosas | FEAM | Urbana | Automática |
| CAMG | FEAM | Urbana | Automática |
| Cascata | FEAM | Urbana | Automática |
| DNER | FEAM | Urbana | Automática |
| Ibiritermo Piratininga | FEAM | Urbana | Automática |
| Petrovale | FEAM | Urbana | Automática |
| Praça Rui Barbosa | FEAM | Urbana | Automática |
| Praça Tancredo | FEAM | Urbana | Automática |
| Safran Centro Adm. | FEAM | Urbana | Automática |

Elaboração própria

Tabela 53 - RM Curitiba: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|------|--------|-----------------|
| Região metropolitana Curitiba | | | |
| Boqueirão | IAP | Urbana | Automático |
| CIC | IAP | Urbana | Automático |
| Colombo | IAP | Urbana | Manual |
| CSN-Cisa | IAP | Urbana | Automático |
| Praça Ouvidor Pardini | IAP | Urbana | Automático |
| REPAR | IAP | Urbana | Automático |
| UEG | IAP | Urbana | Automático |

Elaboração própria

Tabela 54 - RM Porto Alegre: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|-------|--------|-----------------|
| Região metropolitana Porto Alegre | | | |
| Bombeiros | FEPAM | Urbana | Automático |
| Charqueadas | FEPAM | Urbana | Automático |
| Esef | FEPAM | Urbana | Automático |
| Gravataí | FEPAM | Urbana | Automático |
| Guaíba | FEPAM | Urbana | Automático |
| Polo | FEPAM | Urbana | Automático |
| Refap Canoas | FEPAM | Urbana | Automático |
| Refap Esteio | FEPAM | Urbana | Automático |
| Rodoviária | FEPAM | Urbana | Automático |
| Sapucaia | FEPAM | Urbana | Automático |
| Vcomar | FEPAM | Urbana | Automático |

Elaboração própria

Tabela 55 - RM Grande Vitória: Características das estações de monitoramento PM₁₀

| Regiões metropolitanas/Estações de monitoramento | OEMA | Rede | Tipo de monitor |
|--|------|--------|-----------------|
| Região metropolitana Grande Vitória | | | |
| Carapina | IEMA | Urbana | Automático |
| Cariacica | IEMA | Urbana | Automático |
| Cidade Continental | IEMA | Urbana | Automático |
| Enseada do Sua | IEMA | Urbana | Automático |
| Jardim Camburi | IEMA | Urbana | Automático |
| Laranjeiras | IEMA | Urbana | Automático |
| Vila Velha Centro | IEMA | Urbana | Automático |
| Vila Velha Ibes | IEMA | Urbana | Automático |
| Vitoria Centro | IEMA | Urbana | Automático |

Elaboração própria