



Universidade Estadual de Campinas
Faculdade de Tecnologia

Danilo Covaes Nogarotto

Avaliação de modelos de regressão de trajetórias para a
previsão de poluentes atmosféricos

Limeira
2019

Danilo Covaes Nogarotto

**Avaliação de modelos de regressão de trajetórias para a previsão
de poluentes atmosféricos**

Tese apresentada à Faculdade de Tecnologia da
Universidade Estadual de Campinas como
parte dos requisitos para a obtenção do título
de Doutor em Tecnologia, na área de Ambiente.

Orientadora: Profa. Dra. Simone Andrea Pozza

Este trabalho corresponde à versão final da
Tese defendida por Danilo Covaes Nogarotto
e orientada pela Profa. Dra. Simone Andrea
Pozza.

Limeira
2019

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Tecnologia
Felipe de Souza Bueno - CRB 8/8577

N686a Nogarotto, Danilo Covaes, 1987-
Avaliação de modelos de regressão de trajetórias para a previsão de poluentes atmosféricos / Danilo Covaes Nogarotto. – Limeira, SP : [s.n.], 2019.

Orientador: Simone Andréa Pozza.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Tecnologia.

1. Análise de regressão. 2. Material particulado. 3. Trajetórias reversas. 4. Análise por agrupamento. 5. Análise de componentes principais. I. Pozza, Simone Andréa, 1976-. II. Universidade Estadual de Campinas. Faculdade de Tecnologia. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Evaluation of trajectory regression models forecasting atmospheric pollutants

Palavras-chave em inglês:

Regression analysis

Particulate matter

Backward trajectories

Cluster analysis

Principal components analysis

Área de concentração: Ambiente

Titulação: Doutor em Tecnologia

Banca examinadora:

Simone Andréa Pozza [Orientador]

Leonardo Hoinaski

Antônio Carlos Simões Pião

Johan Hendrik Poker Jr.

Bernardino Ribeiro de Figueiredo

Data de defesa: 21-02-2019

Programa de Pós-Graduação: Tecnologia

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0002-5046-807X>

- Currículo Lattes do autor: <http://lattes.cnpq.br/8803768507431721>

Danilo Covaes Nogarotto

**Avaliação de modelos de regressão de trajetórias para a previsão
de poluentes atmosféricos**

Banca Examinadora:

- Profª. Dra. Simone Andrea Pozza
FT/UNICAMP
- Prof. Dr. Leonardo Hoinaski
Universidade Federal de Santa Catarina
- Prof. Dr. Antônio Carlos Simões Pião
Universidade Estadual Paulista
- Prof. Dr. Johan Hendrik Poker Junior
Universidade Estadual de Campinas
- Prof. Dr. Bernardino Ribeiro de Figueiredo
Universidade Estadual de Campinas

A ata da defesa com as respectivas assinaturas dos membros da banca encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Limeira, 21 de fevereiro de 2019

*Há muitos planos no coração do homem,
mas é a vontade do Senhor que se realiza.*
(Provérbios 19, 21)

Agradecimentos

Gostaria de agradecer imensamente a Deus!

Aos meus pais, minha família e minha namorada, pelo carinho, apoio e incentivo em toda a minha trajetória!

À Simone, minha orientadora, pela confiança, presença e conselhos!

Aos amigos que direta e indiretamente me auxiliaram nesta caminhada!

Ao programa da pós-graduação da FT, pelo suporte nesses anos!

À toda banca examinadora, pelas críticas e sugestões neste trabalho!

Por fim, agradeço a Capes, pelo apoio financeiro.

Resumo

O Material Particulado inalável (MP_{10}) e o Material Particulado inalável fino ($MP_{2,5}$) são dois dos principais poluentes atmosféricos estudados na literatura. A concentração diária de MP_{10} e $MP_{2,5}$ em Piracicaba e Santos, municípios localizados no interior e litoral do Estado de São Paulo, respectivamente, para os anos de 2014 e 2015, foi estudada usando a Análise de Regressão de Trajetória (ART), Análise de Cluster (AC) e Análise de Componentes Principais (ACP). A ART aplicada em estudos anteriores considerou apenas os modelos de regressão com distribuição normal. Esta tese tem por principal objetivo avaliar diferentes modelos na ART. Na ART, as trajetórias backward foram obtidas no modelo HYSPLIT, e usadas como variáveis regressoras nos modelos Normal, Normal inversa, Gama, Log-normal e Auto-Regressivo de Médias Móveis (ARMA). Medidas comparativas (correlação, raiz quadrada do erro quadrático médio, AIC e log-verossimilhança) e uma análise de resíduos foram usadas para comparar cada um dos modelos, mostrando que os ARMA tiveram melhor desempenho. Regiões locais contribuíram com mais de 70% nos níveis do MP, para as duas frações, nas duas cidades. Piracicaba tem influências de emissões industriais, veiculares e queima de cana-de-açúcar. Santos apresenta influências de emissões portuárias e industrial. Por meio da AC, os dias foram agrupados em 4 clusters (em cada cidade e para cada MP), identificando um perfil temporal com os dados obtidos da ART e dados meteorológicos do período. Com uso da ACP, foi possível caracterizar a relação da concentração do MP com variáveis meteorológicas (umidade relativa, velocidade do vento e temperatura), dentro de cada grupo. Dessa análise, percebeu-se maior influência da umidade relativa em Piracicaba e da velocidade do vento para Santos.

Palavras-chave: ART, Material Particulado, Trajetórias de massas de ar, AC, ACP.

Abstract

Coarse (PM_{10}) and fine ($PM_{2.5}$) inhalable Particulate Matter are two of the main atmospheric pollutants studied in the literature. Daily PM_{10} and $PM_{2.5}$ concentrations from Piracicaba and Santos, interior and coastal cities of São Paulo State, respectively, between 2014 and 2015, were studied using Trajectory Regression Analysis (TRA), Cluster Analysis (CA) and Principal Component Analysis (PCA). In previous studies, TRA only considered regression models with normal distribution. This thesis aims, mainly, to evaluate different models in TRA. In TRA, the backward trajectories were obtained by the HYSPLIT model, and used as regressive variables in Normal, Inverse Normal, Gama, Log-normal and Auto-regressive Moving Average (ARMA) models. Comparative measurements (correlation, root mean square error, AIC and log-likelihood) and residual analysis was used to compare all the models. After comparison, the ARMA models performed better. Local region contributed with more than 70% of PM_{10} and $PM_{2.5}$ levels in both cities. Piracicaba has influences of industrial and vehicular emissions and sugarcane burning. Santos presented influences of port and industrial emissions. Through CA, days were grouped in 4 clusters (in each city and for each PM), identifying a temporal profile with obtained data of TRA, and meteorological data of period. Using PCA, it was possible to characterize the relationship between PM concentration and meteorological variables (relative humidity, wind speed and temperature), within each group. This analysis revealed greater influence of relative humidity in Piracicaba and wind speed in Santos.

Keywords: TRA, Particulate Matter, Air mass trajectories, CA, PCA.

Lista de Figuras

3.1	Mapa de Piracicaba. Fonte: (DATAGEO, 2019)	38
3.2	Mapa de Santos. Fonte: (DATAGEO, 2019)	39
3.3	Exemplo de uma trajetória reversa para a cidade de Piracicaba	41
3.4	Regiões de residência de massa de ar em Piracicaba	42
4.1	Regiões de residência de massa de ar em Piracicaba	49
4.2	Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo AR(2) - Piracicaba	50
4.3	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP_{10}) - Modelo AR(2) - Piracicaba	51
4.4	Contribuição média do MP_{10} por área - Modelo AR(2) - Piracicaba	52
4.5	Mapa de contribuição das regiões no modelo Normal inversa e AR(2), res- pectivamente - Piracicaba - MP_{10}	52
4.6	Mapa de densidade das trajetórias <i>backward</i> das regiões 1, 2 e 4 - Piracicaba	54
4.7	Ajuste da concentração diária do MP_{10} em 2014 em comparação com valores observados - Piracicaba	55
4.8	Ajuste da concentração diária do MP_{10} em 2015 em comparação com valores observados - Piracicaba	55
4.9	Previsão do MP_{10} e valores reais para o mês de Janeiro de 2016 - Piracicaba	56
4.10	Box-plot do MP_{10} ($\mu g/m^3$) por grupo - MP_{10} - Piracicaba	58
4.11	Box-plot da Temperatura ($^{\circ}C$) por grupo - MP_{10} - Piracicaba	58
4.12	Box-plot da Umidade Relativa (%) por grupo - MP_{10} - Piracicaba	59
4.13	Box-plot da Velocidade do Vento (m/s) por grupo - MP_{10} - Piracicaba . . .	59
4.14	Importância das CP 1 e CP 2 no grupo 1 - MP_{10} - Piracicaba	60
4.15	Importância das CP 1 e CP 2 no grupo 2 - MP_{10} - Piracicaba	60
4.16	Importância das CP 1 e CP 2 no grupo 3 - MP_{10} - Piracicaba	61
4.17	Importância das CP 1 e CP 2 no grupo 4 - MP_{10} - Piracicaba	61
4.18	Regiões de residência de massa de ar em Santos	62
4.19	Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo ARMA(2,2) - Santos	64
4.20	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP_{10}) - Modelo ARMA(2,2) - Santos	65
4.21	Contribuição média do MP_{10} por área - Modelo ARMA(2,2) - Santos . . .	66
4.22	Mapa de contribuição das regiões no modelo Log-Normal e ARMA(2,2), respectivamente - Santos - MP_{10}	66
4.23	Mapa de densidade das trajetórias <i>backward</i> das regiões 1 e 2 - Santos . . .	68
4.24	Ajuste da concentração diária do MP_{10} em 2014 em comparação com valores observados - Santos	68

4.25	Ajuste da concentração diária do MP ₁₀ em 2015 em comparação com valores observados - Santos	69
4.26	Previsão do MP ₁₀ e valores reais para o mês de Janeiro de 2016 - Santos .	69
4.27	Box-plot do MP ₁₀ ($\mu\text{g}/\text{m}^3$) por grupo - MP ₁₀ - Santos	71
4.28	Box-plot da Temperatura ($^{\circ}\text{C}$) por grupo - MP ₁₀ - Santos	71
4.29	Box-plot da Umidade Relativa (%) por grupo - MP ₁₀ - Santos	72
4.30	Box-plot da Velocidade do Vento (m/s) por grupo - MP ₁₀ - Santos	72
4.31	Importância das CP 1 e CP 2 no grupo 1 - MP ₁₀ - Santos	73
4.32	Importância das CP 1 e CP 2 no grupo 2 - MP ₁₀ - Santos	74
4.33	Importância das CP 1 e CP 2 no grupo 3 - MP ₁₀ - Santos	74
4.34	Importância das CP 1 e CP 2 no grupo 4 - MP ₁₀ - Santos	75
4.35	Resíduos padronizados pela ordem das observações (MP _{2,5}) - Modelo ARMA(2,1) - Piracicaba	77
4.36	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP _{2,5}) - Modelo ARMA(2,1) - Piracicaba	78
4.37	Contribuição média do MP _{2,5} por área - Modelo ARMA(2,1) - Piracicaba .	78
4.38	Mapa de contribuição das regiões no modelo Log-Normal e ARMA(2,1), respectivamente - Piracicaba - MP _{2,5}	79
4.39	Ajuste da concentração diária do MP _{2,5} em 2014 em comparação com valores observados - Piracicaba	80
4.40	Ajuste da concentração diária do MP _{2,5} em 2015 em comparação com valores observados - Piracicaba	81
4.41	Previsão do MP _{2,5} e valores reais para o mês de Janeiro de 2016 - Piracicaba	81
4.42	Box-plot do MP _{2,5} ($\mu\text{g}/\text{m}^3$) por grupo - MP _{2,5} - Piracicaba	83
4.43	Box-plot da Temperatura ($^{\circ}\text{C}$) por grupo - MP _{2,5} - Piracicaba	83
4.44	Box-plot da Umidade Relativa (%) por grupo - MP _{2,5} - Piracicaba	84
4.45	Box-plot da Velocidade do Vento (m/s) por grupo - MP _{2,5} - Piracicaba . .	84
4.46	Importância das CP 1 e CP 2 no grupo 1 - MP _{2,5} - Piracicaba	85
4.47	Importância das CP 1 e CP 2 no grupo 2 - MP _{2,5} - Piracicaba	86
4.48	Importância das CP 1 e CP 2 no grupo 3 - MP _{2,5} - Piracicaba	86
4.49	Importância das CP 1 e CP 2 no grupo 4 - MP _{2,5} - Piracicaba	87
4.50	Resíduos padronizados pela ordem das observações (MP _{2,5}) - Modelo ARMA(2,1) - Santos	89
4.51	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP _{2,5}) - Modelo ARMA(2,1) - Santos	89
4.52	Contribuição média do MP _{2,5} por área - Modelo ARMA(2,1) - Santos . . .	90
4.53	Mapa de contribuição das regiões no modelo Gama e ARMA(2,1), respectivamente - Santos - MP _{2,5}	90
4.54	Ajuste da concentração diária do MP _{2,5} em 2014 em comparação com valores observados - Santos	92
4.55	Ajuste da concentração diária do MP _{2,5} em 2015 em comparação com valores observados - Santos	92
4.56	Previsão do MP _{2,5} e valores reais para o mês de Janeiro de 2016 - Santos .	93
4.57	Box-plot do MP _{2,5} ($\mu\text{g}/\text{m}^3$) por grupo - MP _{2,5} - Santos	94
4.58	Box-plot da Temperatura ($^{\circ}\text{C}$) por grupo - MP _{2,5} - Santos	95
4.59	Box-plot da Umidade Relativa (%) por grupo - MP _{2,5} - Santos	95
4.60	Box-plot da Velocidade do Vento (m/s) por grupo - MP _{2,5} - Santos	96
4.61	Importância das CP 1 e CP 2 no grupo 1 - MP _{2,5} - Santos	97

4.62	Importância das CP 1 e CP 2 no grupo 2 - MP _{2,5} - Santos	97
4.63	Importância das CP 1 e CP 2 no grupo 3 - MP _{2,5} - Santos	98
4.64	Importância das CP 1 e CP 2 no grupo 4 - MP _{2,5} - Santos	98
A.1	Diagnóstico dos Resíduos padronizados (MP ₁₀) - Modelo AR(2) - Piracicaba	115
A.2	Resíduos padronizados pela ordem das observações (MP ₁₀) - Modelo Normal - Piracicaba	116
A.3	Resíduos padronizados pela ordem das observações (MP ₁₀) - Modelo Normal Inversa - Piracicaba	116
A.4	Resíduos padronizados pela ordem das observações (MP ₁₀) - Modelo Gama - Piracicaba	117
A.5	Resíduos padronizados pela ordem das observações (MP ₁₀) - Modelo Log-Normal - Piracicaba	117
A.6	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP ₁₀) - Modelo Normal - Piracicaba	118
A.7	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP ₁₀) - Modelo Normal Inversa - Piracicaba	118
A.8	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP ₁₀) - Modelo Gama - Piracicaba	119
A.9	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP ₁₀) - Modelo Log-Normal - Piracicaba	119
A.10	Diagnóstico dos Resíduos padronizados (MP ₁₀) - Modelo ARMA(2,2) - Santos	120
A.11	Resíduos padronizados pela ordem das observações (MP ₁₀) - Modelo Normal - Santos	121
A.12	Resíduos padronizados pela ordem das observações (MP ₁₀) - Modelo Normal Inversa - Santos	121
A.13	Resíduos padronizados pela ordem das observações (MP ₁₀) - Modelo Gama - Santos	122
A.14	Resíduos padronizados pela ordem das observações (MP ₁₀) - Modelo Log-Normal - Santos	122
A.15	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP ₁₀) - Modelo Normal - Santos	123
A.16	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP ₁₀) - Modelo Normal Inversa - Santos	123
A.17	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP ₁₀) - Modelo Gama - Santos	124
A.18	Gráfico normal de probabilidades - <i>qq-plot</i> - (MP ₁₀) - Modelo Log-Normal - Santos	124
A.19	Diagnóstico dos Resíduos padronizados (MP _{2,5}) - Modelo ARMA(2,1) - Piracicaba	125
A.20	Resíduos padronizados pela ordem das observações (MP _{2,5}) - Modelo Normal - Piracicaba	126
A.21	Resíduos padronizados pela ordem das observações (MP _{2,5}) - Modelo Normal Inversa - Piracicaba	126
A.22	Resíduos padronizados pela ordem das observações (MP _{2,5}) - Modelo Gama - Piracicaba	127
A.23	Resíduos padronizados pela ordem das observações (MP _{2,5}) - Modelo Log-Normal - Piracicaba	127

A.24	Gráfico normal de probabilidades - <i>qq-plot</i> - ($MP_{2,5}$) - Modelo Normal - Piracicaba	128
A.25	Gráfico normal de probabilidades - <i>qq-plot</i> - ($MP_{2,5}$) - Modelo Normal Inversa - Piracicaba	128
A.26	Gráfico normal de probabilidades - <i>qq-plot</i> - ($MP_{2,5}$) - Modelo Gama - Piracicaba	129
A.27	Gráfico normal de probabilidades - <i>qq-plot</i> - ($MP_{2,5}$) - Modelo Log-Normal - Piracicaba	129
A.28	Diagnóstico dos Resíduos padronizados ($MP_{2,5}$) - Modelo ARMA(2,1) - Santos	130
A.29	Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Normal - Santos	131
A.30	Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Normal Inversa - Santos	131
A.31	Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Gama - Santos	132
A.32	Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Log-Normal - Santos	132
A.33	Gráfico normal de probabilidades - <i>qq-plot</i> - ($MP_{2,5}$) - Modelo Normal - Santos	133
A.34	Gráfico normal de probabilidades - <i>qq-plot</i> - ($MP_{2,5}$) - Modelo Normal Inversa - Santos	133
A.35	Gráfico normal de probabilidades - <i>qq-plot</i> - ($MP_{2,5}$) - Modelo Gama - Santos	134
A.36	Gráfico normal de probabilidades - <i>qq-plot</i> - ($MP_{2,5}$) - Modelo Log-Normal - Santos	134
B.1	Contribuição média do MP_{10} por área - Modelo Normal - Piracicaba	135
B.2	Contribuição média do MP_{10} por área - Modelo Normal Inversa - Piracicaba	136
B.3	Contribuição média do MP_{10} por área - Modelo Gama - Piracicaba	136
B.4	Contribuição média do MP_{10} por área - Modelo Log-Normal - Piracicaba .	137
B.5	Contribuição média do MP_{10} por área - Modelo Normal - Santos	138
B.6	Contribuição média do MP_{10} por área - Modelo Normal Inversa - Santos .	138
B.7	Contribuição média do MP_{10} por área - Modelo Gama - Santos	139
B.8	Contribuição média do MP_{10} por área - Modelo Log-Normal - Santos . . .	139
B.9	Contribuição média do $MP_{2,5}$ por área - Modelo Normal - Piracicaba	140
B.10	Contribuição média do $MP_{2,5}$ por área - Modelo Normal Inversa - Piracicaba	140
B.11	Contribuição média do $MP_{2,5}$ por área - Modelo Gama - Piracicaba	141
B.12	Contribuição média do $MP_{2,5}$ por área - Modelo Log-Normal - Piracicaba .	141
B.13	Contribuição média do $MP_{2,5}$ por área - Modelo Normal - Santos	142
B.14	Contribuição média do $MP_{2,5}$ por área - Modelo Normal Inversa - Santos .	142
B.15	Contribuição média do $MP_{2,5}$ por área - Modelo Gama - Santos	143
B.16	Contribuição média do $MP_{2,5}$ por área - Modelo Log-Normal - Santos . . .	143
C.1	Dendograma da Análise de Cluster do MP_{10} - Piracicaba	144
C.2	Dendograma da Análise de Cluster do MP_{25} - Piracicaba	145
C.3	Dendograma da Análise de Cluster do MP_{10} - Santos	145
C.4	Dendograma da Análise de Cluster do $MP_{2,5}$ - Santos	145

Lista de Tabelas

2.1	Padrões de qualidade do ar para $MP_{2,5}$ e MP_{10}	22
2.2	Características do mais relevantes artigos sobre Análise de Regressão de Trajetória.	28
2.3	Características de alguns artigos relevantes, associados com AC e ACP. . .	33
3.1	Região, Cidades e Percentual de dados faltantes por variável	40
3.2	Coordenadas das estações automáticas em cada cidade	41
4.1	Descrição das regiões de residência de massa de ar em Piracicaba	48
4.2	Critério de Seleção para os modelos ARMA(p,q) - MP_{10} - Piracicaba . . .	48
4.3	Medidas comparativas - MP_{10} - Piracicaba	50
4.4	Percentual de dias dentro de cada estação do ano por grupo (MP_{10} - Piracicaba)	56
4.5	Percentual de contribuição do grupo dentro de cada região de residência da massa de ar (MP_{10} - Piracicaba)	57
4.6	Resumo da ACP - MP_{10} - Piracicaba	57
4.7	Descrição das regiões de residência de massa de ar em Santos	63
4.8	Critério de Seleção para os modelos ARMA(p,q) - MP_{10} - Santos	63
4.9	Medidas comparativas - MP_{10} - Santos	64
4.10	Percentual de dias dentro de cada estação do ano por grupo (MP_{10} - Santos)	70
4.11	Percentual de contribuição do grupo dentro de cada região de residência da massa de ar (MP_{10} - Santos)	70
4.12	Resumo da ACP - MP_{10} - Santos	73
4.13	Critério de Seleção para os modelos ARMA(p,q) - $MP_{2,5}$ - Piracicaba . . .	76
4.14	Medidas comparativas - $MP_{2,5}$ - Piracicaba	76
4.15	Percentual de dias dentro de cada estação do ano por grupo ($MP_{2,5}$ - Piracicaba)	82
4.16	Percentual de contribuição do grupo dentro de cada região de residência da massa de ar ($MP_{2,5}$ - Piracicaba)	82
4.17	Resumo da ACP - $MP_{2,5}$ - Piracicaba	85
4.18	Critério de Seleção para os modelos ARMA(p,q) - $MP_{2,5}$ - Santos	88
4.19	Medidas comparativas - $MP_{2,5}$ - Santos	88
4.20	Percentual de dias dentro de cada estação do ano por grupo ($MP_{2,5}$ - Santos)	93
4.21	Percentual de contribuição do grupo dentro de cada região de residência da massa de ar ($MP_{2,5}$ - Santos)	94
4.22	Resumo da ACP - $MP_{2,5}$ - Santos	96

Sumário

1	Introdução	16
1.1	Objetivo Geral	18
1.2	Objetivos específicos	18
2	Revisão de Literatura	20
2.1	Material Particulado	20
2.2	Legislação	22
2.3	Poluentes atmosféricos e variáveis meteorológicas	23
2.4	Ferramentas para determinação de fontes de poluentes atmosféricos	24
2.4.1	Modelos receptores	24
2.4.2	Modelos baseados em trajetórias	25
2.5	Análise de Regressão de Trajetória - ART	26
2.6	Uso de AC e ACP no estudo de variáveis meteorológicas e poluentes atmosféricos	31
3	Métodos	37
3.1	Área de estudo	37
3.2	Dados	39
3.3	Trajetoórias das massas de ar	40
3.4	Análise de Regressão de Trajetória	41
3.5	Análise de Cluster	46
3.6	Análise de Componentes Principais	46
4	Resultados e Discussão	47
4.1	Material Particulado inalável - MP ₁₀	47
4.1.1	Piracicaba	47
4.1.2	Santos	62
4.2	Material Particulado inalável fino - MP _{2,5}	75
4.2.1	Piracicaba	76
4.2.2	Santos	87
4.3	Comparação dos resultados de Piracicaba e Santos	100
5	Conclusões	101
6	Atividades Desenvolvidas e Perspectivas Futuras	103
6.1	Artigos publicados	104
6.2	Artigos aceitos para publicação	104
6.3	Trabalhos publicados em anais de eventos (completo)	104

Referências Bibliográficas	105
A Análise dos Resíduos	115
A.1 MP ₁₀ - Piracicaba	115
A.2 MP ₁₀ - Santos	120
A.3 MP _{2,5} - Piracicaba	125
A.4 MP _{2,5} - Santos	130
B Contribuição média por região	135
B.1 MP ₁₀ - Piracicaba	135
B.2 MP ₁₀ - Santos	138
B.3 MP _{2,5} - Piracicaba	140
B.4 MP _{2,5} - Santos	142
C Dendograma	144

Capítulo 1

Introdução

Avaliar, monitorar e prever o nível de poluição atmosférica é importante nos dias atuais, principalmente após o crescimento das cidades, aumento dos veículos automotores, desenvolvimento da indústria, entre outros fatores. Muito se tem discutido a respeito da emissão de poluentes na atmosfera e como buscar soluções para combatê-la. Para isso, o monitoramento e análise de substâncias presentes no ar são importantes para encontrar tal solução e também para entender os impactos na saúde e o bem-estar da população, da fauna e flora. Poluentes do ar afetam diretamente as propriedades da atmosfera, de várias maneiras: redução da visibilidade, formação e composição da chuva, redução da radiação solar, alteração na temperatura e formação de ozônio (SEINFELD; PANDIS, 2006).

Dentre os principais poluentes atmosféricos podemos citar: dióxido de enxofre (SO_2), monóxido de carbono (CO), oxidantes fotoquímicos (como o ozônio, O_3), hidrocarbonetos (HC), óxidos de nitrogênio (NO_x) e Material Particulado Atmosférico (MPA) ou apenas Material Particulado (MP). O MP é todo material sólido e líquido que fica em suspensão no ar, devido ao seu pequeno tamanho (HINDS, 1998; SEINFELD; PANDIS, 2006). Englobam o MP: as Partículas Totais em Suspensão (PTS), Partículas Inaláveis (MP_{10}), Partículas Inaláveis Finas ($\text{MP}_{2,5}$ e MP_1) e Fumaça (FMC) (HINDS, 1998; REVUELTA et al., 2014; CETESB, 2016a).

O MP_{10} e o $\text{MP}_{2,5}$ são dois dos principais poluentes monitorados nos dias de hoje. O MP_{10} são partículas cujo diâmetro aerodinâmico é menor ou igual a $10\text{ }\mu\text{m}$, e que geralmente ficam retidas na parte superior do sistema respiratório, mas que podem penetrar mais profundamente. O $\text{MP}_{2,5}$ são partículas cujo diâmetro aerodinâmico é menor ou igual a $2,5\text{ }\mu\text{m}$, e que devido ao seu tamanho diminuto penetram profundamente no sistema respiratório, podendo alcançar os alvéolos pulmonares e corrente sanguínea (SEINFELD; PANDIS, 2006; CETESB, 2016a; US-EPA, 2017).

O MP é simples de monitorar, entretanto, complicado de saber qual a sua origem. As diversas fontes do MP conferem ao mesmo características e composições diferentes, o que significa uma complexa mistura de substâncias nos estados sólidos e líquidos com diferentes propriedades (VALLERO, 2008). Além disso, pode ser formado a partir de outros poluentes (YNOUE et al., 2017) e há a possibilidade do MP ser transportado a longo alcance (KAVOURAS et al., 2013).

A concentração do MP medido no ponto de amostragem é influenciada pela origem do MP e também pela forma como o MP é transportado pelas massas de ar. Há estudos

em que as fontes emissoras do MP são originadas de distâncias maiores que 500 km do local de medição (RODOPOULOU et al., 2014; DIMITRIOU; KASSOMENOS, 2014a). Alguns estudos analisam a contribuição regional, através das trajetórias das massas de ar, na concentração de poluentes atmosféricos. Uma das ferramentas para isso é a Análise de Regressão de Trajetória (ART) (DIMITRIOU; KASSOMENOS, 2014a; KAVOURAS et al., 2013; CHALBOT; MCELROY; KAVOURAS, 2013). Nesta abordagem supõe-se que a quantidade de tempo que a massa de ar passa sobre uma região afeta linearmente a contribuição da região na concentração do poluente atmosférico medido no local de amostragem. O objetivo da ART é associar a região dominante determinada na análise às fontes localizadas dentro do perímetro desta mesma região. Ou seja, é determinar se a fonte dominante do poluente está localizada numa região próxima ao receptor (como indústrias, rodovias, etc.) ou em regiões continentais ou litorâneas, ou ainda transportadas de longas distâncias (XU et al., 2006).

Variáveis meteorológicas interferem diretamente na concentração dos poluentes atmosféricos. Temperatura, umidade, pluviosidade, velocidade do vento exercem considerável influência nos níveis de MP (QUEROL et al., 2004; ROSA et al., 2010; DOMINICK et al., 2012). A concentração do MP pode aumentar de acordo com as características climáticas, causando prejuízos à saúde e bem-estar das pessoas. Por exemplo, em períodos de baixa umidade do ar e altos índices de poluição, ocorre o aumento de complicações respiratórias (BREITNER et al., 2014). Por outro lado, as chuvas são uma importante ferramenta de remoção de poluentes da atmosfera, assim como a velocidade do vento. Altos níveis de pluviosidade e velocidade do vento resultam em baixas concentrações do poluente atmosférico, devido à dispersão dos mesmos (VALLERO, 2008). Fenômenos que envolvem as massas de ar da atmosfera podem alterar o comportamento das variáveis meteorológicas e por consequência a concentração do MP. Por exemplo, em outubro de 2014, a atuação de anticiclone subtropical do atlântico sul no estado de São Paulo, inibiu a entrada de frentes frias, que provocariam instabilidade, e por consequência chuvas (CETESB, 2014; YNOUE et al., 2017). O que ocorreu, ao contrário, foi baixos níveis de precipitação, e maiores níveis de MP (CETESB, 2014).

O MP está diretamente relacionado com a qualidade do ar e estudar como poluentes atmosféricos se comportam, é de vital importância para buscar melhorias na qualidade de vida humana, da fauna e flora. Além disso, as variáveis meteorológicas exercem forte relação com a concentração dos poluentes atmosféricos, e verificar essa relação é muito importante. Por isso, há muita pesquisa (DOMINICK et al., 2012; LUNA et al., 2014; KWON et al., 2015) em que são analisados os níveis de poluentes na atmosfera relacionando-os com variáveis meteorológicas com o uso da estatística servindo como base na tomada de ações. A busca por soluções é imprescindível para amenizar eventuais problemas do presente e do futuro, tanto para o meio ambiente, quanto para a saúde humana.

Este trabalho pretende aplicar a ART para esse fim, usando o modelo de regressão múltipla empregado em trabalhos anteriores (DIMITRIOU; KASSOMENOS, 2014a; RODOPOULOU et al., 2014). Entretanto, também se abordará a ART de duas outras maneiras. A primeira, será usando um modelo linear generalizado e um modelo generalizado de locação escala e forma. E a segunda será usando modelos de séries temporais. Uma comparação entre as abordagens será feita, descrevendo a diferença entre elas. Pela

revisão de literatura conhecida até o momento, esta é a primeira tentativa de utilizar esses modelos para esse fim, além de ser a primeira tentativa de usar a ART em cidades do Estado de São Paulo. Os trabalhos anteriores não consideraram a estrutura temporal dos dados, visto que os modelos foram ajustados com os dados das concentrações diárias dos poluentes (XU et al., 2006; DIMITRIOU; KASSOMENOS, 2014a; RODOPOULOU et al., 2014; KAVOURAS et al., 2013; CHALBOT; MCELROY; KAVOURAS, 2013). Kavouras et al. (2013) apresenta algumas limitações dos modelos ART, incluindo erros nas localizações das trajetórias, conduzindo a grandes erros no cálculo da contribuição de uma região para dias com altas (ou baixas) concentrações, e também na instabilidade da contribuição individual diária. O modelo de séries temporais sugerido nesta tese pretende melhorar estas limitações.

Além da ART, este estudo aplicará outras duas técnicas estatísticas multivariadas para entender a relação entre o MP (MP_{10} e $MP_{2,5}$) e as variáveis meteorológicas (umidade relativa, temperatura e velocidade do vento). São elas, a Análise de Cluster (AC) e a Análise de Componentes Principais (ACP). A AC será utilizada para identificar padrões temporais do comportamento das variáveis meteorológicas e o MP, agrupando dias do ano com comportamentos similares. A ACP será utilizada para quantificar a principal relação entre as variáveis (meteorológicas e MP), em cada agrupamento obtido. Essas análises multivariadas ajudarão num melhor entendimento da relação existente entre as variáveis.

Diante disso, são apresentados os objetivos desta tese, que direcionarão à análise realizada nesse estudo. Eles estão divididos em um objetivo geral e 4 objetivos específicos.

1.1 Objetivo Geral

Avaliar os modelos de regressão de trajetórias para previsão do MP_{10} e $MP_{2,5}$.

1.2 Objetivos específicos

- Identificar a localização das fontes emissoras e estimar as contribuições de fontes regionais na concentração dos poluentes atmosféricos (MP_{10} e $MP_{2,5}$) usando Análise de Regressão de Trajetória;
- Comparar três abordagens para a ART: os modelos usuais de regressão, modelos lineares generalizados e modelos de séries temporais;
- Identificar padrões temporais de qualidade do ar, com uso da Análise de Cluster;
- Dimensionar as relações entre as variáveis meteorológicas e poluentes, utilizando Análise de Componentes Principais.

Além disso, esse trabalho apresenta uma revisão de literatura (Capítulo 2), focada em estudos anteriores sobre poluentes atmosféricos, em especial o MP, e variáveis meteorológicas, com uso da ART, AC e ACP. No Capítulo 3 é descrito como serão obtidos os dados, além de uma breve exposição das três análises multivariadas: ART, AC e ACP. Enquanto que no Capítulo 4 são apresentados os resultados e discussão obtidos da análise de dados

de duas cidades do Estado de São Paulo, Piracicaba e Santos. Finalmente, no Capítulo 5 tem-se a conclusão deste trabalho, além de algumas sugestões para trabalhos futuros, e no Capítulo 6 tem-se as atividades científicas desenvolvidas durante o doutorado.

Capítulo 2

Revisão de Literatura

Neste capítulo temos uma descrição do estado da arte referente ao Material Particulado, outros poluentes atmosféricos, variáveis meteorológicas e análises multivariadas usadas no estudo dessas variáveis. A Seção 2.1 apresenta características do MP, incluindo alguns estudos sobre os efeitos na saúde humana. Em seguida, na Seção 2.2 são apresentadas alguns limites nacionais e internacionais para o MP. Na Seção 2.3 são apresentados alguns estudos que relacionam a poluição atmosférica com parâmetros meteorológicos.

Na sequência, na Seção 2.4 são apresentados alguns modelos receptores e modelos baseados em trajetórias, usados na determinação de fontes emissoras de poluição. A Seção 2.5 discute sobre trabalhos que abordam a ART (principal objeto de estudo desta tese) como ferramenta para identificar fontes emissoras do MP e de outros poluentes atmosféricos.

Por fim, a Seção 2.6 mostra alguns trabalhos que utilizam análises multivariadas, especificamente a AC e ACP, num contexto de poluição atmosférica e variáveis meteorológicas.

2.1 Material Particulado

O Material Particulado (MP) é um dos poluentes mais estudados na literatura. Diferente de outros poluentes, que são classificados pela propriedades químicas, o MP é, principalmente, caracterizado fisicamente. A classificação é desta forma, devido a complexidade química das partículas, sendo uma mistura de diferentes fontes, tamanhos, composições e propriedades (VALLERO, 2008; SEINFELD; PANDIS, 2006).

O MP_{10} e o $MP_{2,5}$ são dois dos principais tipos do MP. A formação da partícula determinam o seu tamanho. Partículas mais grossas são geralmente formadas por processos mecânicos, já partículas mais finas podem ser geradas pela combustão. Se as partículas são suficientemente pequenas, elas podem ficar suspensas no ar por um longo período de tempo (HINDS, 1998; VALLERO, 2008). Esse tempo de permanência pode variar, a depender do tamanho do MP e da precipitação. O processo de deposição seca ocorre quando as partículas se depositam na superfície, e a deposição úmida está associada a processos com água (YNOUE et al., 2017).

O MP também provoca impactos no clima, sendo que os efeitos são divididos em duas classes: efeitos diretos e indiretos. Os efeitos diretos são resultantes do espalhamento e

absorção da radiação solar, enquanto que os efeitos indiretos ocorrem quando o MP age como núcleo de condensação de nuvens (YNOUE et al., 2017).

O MP é proveniente de diversas fontes emissoras, sendo as principais a re-suspensão do solo, queima de biomassa, processos industriais, emissões de veículos automotores, entre outros (VALLERO, 2008).

Em alguns estudos feitos no Brasil, o interesse foi identificar possíveis fontes de emissão do MP e outros poluentes atmosféricos. Níveis de chumbo e MP foram analisados na cidade de São Paulo durante 4 dias no inverno de 2005, e Gioia et al. (2010) concluíram que as emissões veiculares são uma das principais fontes desses poluentes. Estudos realizados em seis capitais de estados brasileiros indicaram que os principais fatores da concentração do MP_{2,5} foram a re-suspensão da poeira do solo, emissão e queima de biomassa, e combustão de óleo em indústrias (ANDRADE et al., 2012b; MIRANDA et al., 2012).

Andrade et al. (2012a) encontraram que emissões locais de compostos orgânicos voláteis (COV) e óxidos de nitrogênio (NO_x) têm papel significativo sobre a formação do ozônio da região metropolitana de São Paulo em estações de monitoramento urbanas e industriais. Isto é, devido a alta emissão de NO_x emitidos pela queima de diesel da frota veicular, há menor concentração do O₃ e COV na região urbana. Entretanto, em indústrias localizadas distantes de rodovias (com baixo tráfego de veículos, e consequentemente, pouca influência do NO_x), a emissão de COVs, e a concentração do O₃ são mais altas.

Dimitriou (2016) analisou a concentração do MP₁₀ na cidade de Birmingham, Inglaterra, durante janeiro de 2011 e julho 2012. Considerou-se um modelo de regressão múltipla em que a variável dependente foi a concentração do MP₁₀, com o intuito de estimar a contribuição devido a fontes locais, a dispersão do vento e ao transporte de longo alcance. Encontrou-se que a intensa emissão de veículos foi a principal contribuição nos níveis do MP₁₀.

Diversas pesquisas já foram feitas para previsão de níveis de poluição e sua relação com a saúde humana (CARMO et al., 2010; MATYASOVSKY et al., 2011; RODOPOULOU et al., 2015). Spiezia et al. (2014), na cidade de Pádua, Itália, mostraram que indivíduos expostos a elevados níveis de diferentes poluentes (MP₁₀, CO, O₃, SO₂, entre outros) aumentaram o risco de desenvolvimento de embolia pulmonar aguda, sendo o risco 4 vezes maior devido a exposição ao MP₁₀. Na cidade Doña Ana County, nos Estados Unidos, investigou-se os efeitos do MP₁₀, MP_{2,5} e O₃ em doenças respiratórias e cardiovasculares influenciadas por partículas minerais e queimadas. As estimativas de aumento em internações devido ao MP₁₀ e MP_{2,5} foram de 3,1% e 2,8%, respectivamente. (RODOPOULOU et al., 2014).

Carmo et al. (2010) investigaram os efeitos da exposição de curto prazo ao MP_{2,5} decorrente de queimadas na floresta amazônica. Um incremento de 2,9% e 2,6% nas internações depois dos 6º e 7º dias subsequentes da exposição, foi associado a um aumento de 10 µg/m³ na concentração do MP_{2,5}. Já em Vitória, no Espírito Santo, um acréscimo de 10 µg/m³ na concentração do MP₁₀ acarretou num aumento de 9,67% no risco relativo de hospitalização por doenças respiratórias (FREITAS et al., 2016).

Como visto, a poluição devido ao MP está muito relacionada a problemas de saúde humana, principalmente problemas respiratórios e cardiovasculares. Por essa razão, inú-

meros são os esforços para, cada vez mais, entender seus efeitos na saúde e controlar as fontes emissoras. Órgãos ambientais, tanto nacionais, quanto internacionais, buscam estabelecer limites para controle da poluição do ar.

2.2 Legislação

Há uma grande preocupação, em termos legais, de se estabelecer limites para a concentração do MP. Existem os limites estabelecidos pelo padrão federal do Conselho Nacional do Meio Ambiente (CONAMA), Resolução 491/2018 (BRASIL, 2018). Em São Paulo, há o Decreto Estadual N° 59.113 de 23/04/2013 (São Paulo, 2013), que determina os limites vigentes no estado. Tanto o CONAMA, quanto o decreto estabeleceram 3 metas intermediárias (MI) ou padrões intermediários (PI), até que se estabeleça os padrões finais (PF) baseado nas diretrizes da Organização Mundial da Saúde (OMS) (WHO – World Health Organization, 2006). A ideia é que a poluição atmosférica seja reduzida gradativamente e progressivamente, passando da MI-1 (ou PI-1) até atingir os padrões finais. Atualmente, está em vigor a MI-1 no Estado de São Paulo, desde o estabelecimento do Decreto Estadual.

Internacionalmente, dentre os principais, há os limites estabelecidos pela OMS (WHO – World Health Organization, 2006), pela Agência Ambiental Europeia (EEA, 2016) e os limites da Agência Ambiental dos Estados Unidos (US-EPA, 2017). Os limites estaduais, nacionais e internacionais são apresentados de forma comparativa na Tabela 2.1.

Tabela 2.1: Padrões de qualidade do ar para $MP_{2,5}$ e MP_{10}

Padrões de Qualidade do ar		$MP_{2,5}$ ($\mu g/m^3$)		MP_{10} ($\mu g/m^3$)	
		24 horas ¹	Anual ²	24 horas ¹	Anual ²
CONAMA	PI-1	60	20	120	40
	PI-2	50	17	100	35
	PI-3	37	15	75	30
Decreto 59113/13	MI-1	60	20	120	40
	MI-2	50	17	100	35
	MI-3	37	15	75	30
OMS		25	10	50	20
União Europeia		-	25 ³	50	40 ³
EUA		35 ⁴	12 ⁵	150	-

¹: Média de 24 horas consecutivas. Não deve ser excedido mais que uma vez no ano

²: Média aritmética anual.

³: Não deve ser excedido.

⁴: Média aritmética da média anual de 3 anos.

⁵: Média do 98° percentil de 3 anos.

Em relação aos níveis estabelecidos no Estado de São Paulo, nota-se que para o MP_{10} , o máximo estipulado pela MI-1 ($120 \mu g/m^3$) é menor que o vigente em padrões americanos. Esse valor, deverá chegar progressivamente até o limite de $50 \mu g/m^3$, igual ao determinado

pela OMS e também aos padrões atuais da Europa. Já em relação ao $MP_{2,5}$, o valor de média anual estabelecido pela MI-1 ($20 \mu\text{g}/\text{m}^3$) é menor que o do padrão europeu ($25 \mu\text{g}/\text{m}^3$), entretanto quase o dobro do limite americano ($12 \mu\text{g}/\text{m}^3$). Para prevenção de episódios críticos de poluição do ar, se faz necessário o estudo e previsão das condições meteorológicas desfavoráveis a dispersão da poluição atmosférica.

2.3 Poluentes atmosféricos e variáveis meteorológicas

O entendimento das condições meteorológicas da atmosfera auxilia na prevenção de situações críticas de poluição atmosférica. Inúmeros trabalhos procuram analisar a relação entre as variáveis meteorológicas e poluentes (STATHEROPOULOS; VASSILIADIS; PAPPA, 1998; BANERJEE; SINGH; SRIVASTAVA, 2011; DOMINICK et al., 2012).

Gonçalves et al. (2005) estudaram os efeitos da poluição do ar e parâmetros meteorológicos durante o verão da cidade de São Paulo, encontrando que dias quentes e úmidos favoreceram a diminuição da concentração de SO_2 , MP_{10} e O_3 .

Na Índia, foram analisadas estatisticamente, por modelos de regressão, variáveis ambientais do ar (SO_2 , NO_2 e PTS) em função das variáveis meteorológicas: umidade, velocidade do vento, precipitação e temperatura. Destas variáveis meteorológicas, a velocidade do vento foi a de maior influência na equação de regressão (BANERJEE; SINGH; SRIVASTAVA, 2011). Tiwari et al. (2013), também na Índia, especificamente em Nova Delhi, obtiveram que *black carbon* foi correlacionado negativamente com velocidade do vento e temperatura, enquanto que a precipitação não foi correlacionada com o *black carbon*.

A direção do vento foi o fator mais importante na variabilidade dos níveis de MP em Beijing, na China durante o período de Março de 2012 a Abril de 2013. Ventos da direção sul foram associados com os dias mais severos de poluição. Modelos de regressão múltipla explicaram 62% da variação total do MP (ZHOU et al., 2015). Lyra, Souza e Viola (2011) também utilizaram de modelos de regressão lineares para estimar a concentração de MP_{10} na cidade do Rio de Janeiro, entre os anos de 2002 e 2003. As variáveis meteorológicas da Umidade relativa e Precipitação pluvial foram as mais relevantes, explicando mais de 50% da variabilidade do MP. Já Monte, Albuquerque e Reisen (2016) usaram dos modelos Logit para investigar a relação das variáveis meteorológicas na chance de ocorrência de qualidade do ar classificada como "não boa". Precipitação e velocidade do vento altas contribuíram para a redução da chance de ocorrência de dias com baixa qualidade do ar.

Fatores meteorológicos como baixos valores de temperatura e vento, em Birmingham, e estagnação das massas de ar, em Atenas, influenciaram na concentração do MP_{10} nessas cidades, sendo a ACP utilizada para tais conclusões (VARDOULAKIS; KASSOMENOS, 2008). Estudos realizados por Toh, Lim e Glasow (2013) analisaram os efeitos meteorológicos sobre variações diurna, mensais e sazonais das concentrações de ozônio troposférico (O_3) e material particulado (MP_{10}) na cidade de Tanah Rata, Malásia. Para a formação do O_3 , contribuíram positivamente a temperatura e negativamente, a umidade e pressão das massas de ar, além de forte associação com as monções asiáticas.

Além da relação das variáveis meteorológicas com a concentração de poluentes atmosféricos, também há interesse de se identificar fontes emissoras de poluentes do ar. Há

algumas ferramentas descritas na Seção 2.4, que também tem como objetivos identificar as fontes emissoras de poluentes atmosféricos.

2.4 Ferramentas para determinação de fontes de poluentes atmosféricos

Existem diversas abordagens possíveis para se estudar quais as possíveis fontes emissoras dos poluentes atmosféricos. Pode-se destacar dois tipos de modelos, os modelos receptores (subseção 2.4.1), e os modelos baseados em trajetórias de massas de ar (subseção 2.4.2).

2.4.1 Modelos receptores

Os modelos receptores são usados para identificar as possíveis fontes emissoras de poluentes atmosféricos. Nesses modelos a composição química do poluente é analisada, e inferências a respeito das fontes são feitas. Apesar da complexidade da composição química do MP, saber da sua composição nos diz muito a respeito de sua fonte, e os modelos receptores usam desta composição química para esse fim (VALLERO, 2008).

Em Santa Catarina, nas cidades de Tubarão e Capivari de Baixo entre dezembro de 2008 e dezembro de 2009, identificou-se que influências locais na concentração do MP_{10} foram a ressuspensão da poeira do solo/emissão de poeira da estrada (contribuição entre 65 a 75%) e queima de combustíveis (contribuição de 15 a 25%) (HOINASKI et al., 2013). Foram utilizados os modelos receptores para tal conclusão, o UNMIX e o PMF (*Positive Matrix Factorization*). Tais modelos são usados para identificar e quantificar fontes emissoras através das características químicas e físicas das partículas, medidas num local receptor (US-EPA, 2017).

Com dados foram coletados na ilha de Beihuangcheng, no norte da China, entre agosto de 2014 e setembro de 2015, Zong et al. (2018) aplicaram o modelo PMF e observaram que as principais fontes do $MP_{2,5}$ foram, principalmente, combustão de carvão e queima de biomassa (48,21%) e emissão veicular (30,33%).

Canteras, Moreira e Faria (2013) utilizaram da Análise de Componentes Principais e Análise de Cluster para identificar quais foram as principais fontes emissoras do MP_{10} na cidade de Limeira, São Paulo, durante Agosto de 2008 a Janeiro de 2010. Em relação ao material particulado grosso ($MP_{2,5-10}$), a ressuspensão do solo (57%), emissões veiculares (30%) e emissões industriais (13%) foram as principais fontes emissoras, enquanto que para o material particulado fino ($MP_{2,5}$) foram a ressuspensão do solo (79%), emissões veiculares (13%) e emissões industriais (8%).

Jain et al. (2018) analisaram o MP_{10} medidos em Delhi, Índia, com uso do UNMIX, ACP e PMF. Os modelos receptores identificaram cinco fontes, sendo as principais: emissões veiculares, aerossóis marinhos, queima de biomassa e ressuspensão do solo.

Os estudos apresentados nesta subseção 2.4.1 utilizaram da composição química para identificar as possíveis fontes emissoras do MP. Porém, existem os modelos baseados em trajetórias, que ao invés da composição química, fazem usos de trajetórias reversas para identificar possíveis locais das fontes emissoras.

2.4.2 Modelos baseados em trajetórias

Os modelos baseados em trajetórias são utilizados para identificar aonde estão localizadas as fontes emissoras dos poluentes atmosféricos. Diferente dos modelos receptores, que são usados para identificar as fontes.

Identificar as fontes emissoras do MP, sua origem, e como funciona o processo de dispersão do vento é importante para presumir a respeito do comportamento do MP, além de auxiliar na identificação de fontes emissoras desse poluente atmosférico (DIMITRIOU; KASSOMENOS, 2014b; KAVOURAS et al., 2013). Análises estatísticas baseadas no tempo de residência das massa de ar têm sido usadas, dentre elas, podemos citar o PSCF (*Potencial Source Contribution Function*) (HUANG et al., 2012; ZHANG; WONG; LEE, 2015), CWT (*Concentration Weighted Trajectory*) (DIMITRIOU et al., 2015) e ART (Análise de Regressão de Trajetória) (KAVOURAS et al., 2013; DIMITRIOU; KASSOMENOS, 2014a). Nesses modelos, os tempos de residência para cada célula é calculado baseado nas trajetórias das massas de ar, que são trajetórias reversas (ou *backward*) porque são as trajetórias obtidas que chegaram em determinado local.

Em geral as trajetórias das massas de ar são obtidas por meio de modelos matemáticos, sendo um dos mais usados o modelo HYSPLIT (*Hybrid Single-Particle Lagrangian Integrated Trajectory Model*), do laboratório NOAA - ARL (*National Oceanic and Atmospheric Administration - Air Resources Laboratory*) (DRAXLER; HESS, 1998; KULSHRESTHA; KUMAR, 2014; PEREZ et al., 2015; STEIN et al., 2015; Rolph, G.D., 2016). Este modelo lagrangeano calcula os pontos de latitude e longitude por onde passaram as trajetórias até chegarem num determinado local receptor. A equação base para cálculo das trajetórias é dada na Equação 2.1, que mostra a mudança da posição no tempo, dada a advecção pelo vento, que determinará a trajetória que a partícula seguirá (STEIN et al., 2015).

$$P_{mean}(t + \Delta t) = P_{mean}(t) + \frac{1}{2}[V(P_{mean}, t) + V(\{P_{mean}(t) + [V(P_{mean}, t)\Delta t\}, t + \Delta t)]\Delta t \quad (2.1)$$

onde $P_{mean}(t)$ é a posição da trajetória no tempo t . O $V(P_{mean}, t)$ é o vetor de velocidade tri-dimensional, na posição e tempo inicial (DRAXLER; HESS, 1998). As trajetórias são calculadas a cada 1 hora.

O modelo HYSPLIT usa dados meteorológicos para calcular as trajetórias das massas de ar. Há alguns disponíveis, como por exemplo, o *Global Data Assimilation System* (GDAS), *NCEP/NCAR global reanalysis data*, *ETA Data Assimilation System* (EDAS) (STEIN et al., 2015). Estes arquivos meteorológicos contém dados das componentes da velocidade do vento, temperatura, e várias outras variáveis. Além disso, estes arquivos já estão no formato para uso no HYSPLIT.

O modelo PSCF calcula a probabilidade que a fonte esteja localizada em determinada região, combinando as trajetórias reversas das massas de ar com a concentração medida num local. A região do local receptor é dividida em um grid de células. A probabilidade de cada célula é calculada pela razão entre o número de vezes que a concentração foi

alta quando passou pela célula, dividido pelo número de vezes que a trajetória passou pela célula (PEKNEY et al., 2006; HUANG et al., 2012). Este modelo tem por objetivo identificar a localização de possíveis fontes emissoras de poluentes atmosféricos.

PSCF foi usado para identificar fontes emissoras de $MP_{2,5}$ em Beijing, China. Sul de Hebei, Norte de Henan e sudoeste de Shandong foram as localização das potenciais fontes. Tais regiões são onde muitas indústrias estão localizadas e também com alto potencial de fontes de carbono orgânico (ZHANG; WONG; LEE, 2015). Diversas outras aplicações, em diferentes poluentes atmosféricos foram exploradas, como para o O_3 e NO_2 em Seul, Coréia do Sul (VELLINGIRI et al., 2016) e O_3 em Atenas e Ioannina, Grécia (DIMITRIOU; KASSOMENOS, 2015).

O modelo CWT é implementado para localizar as fontes principais de um poluente atmosférico. Brevemente, neste modelo é obtido, para cada célula, a média ponderada da concentração do poluente atmosférico que está associado as trajetórias que passaram por aquela região (DIMITRIOU et al., 2015; KONG et al., 2013). Dimitriou et al. (2015) aplicaram o modelo CWT na concentração do $MP_{2,5}$ na Grécia, em Atenas. A Península dos Balcãs e o mar Mediterrâneo foram caracterizados como regiões de potenciais fontes regionais de emissão nesta cidade, devido a combustão e ao aerossol marinho.

Kong et al. (2013) combinaram os métodos PSCF e uma modificação no CWT, chamada de RTWC (*Residence Time Weighted Concentration*) na análise do MP_{10} nas cidades de Hefei e Chaohu, na China. As províncias de Henan e Shanxi, no nordeste e o delta do rio Yangtze, foram localização das regiões de emissão do MP_{10} . Reconhecidamente, essas regiões tem alta emissão antropogênica, como tráfego.

A Seção 2.5 abordará alguns trabalhos sobre a ART, em que a ideia principal é identificar a localização das possíveis fontes emissoras, baseado na relação da concentração do poluente atmosférico e o tempo de residência da massa de ar.

2.5 Análise de Regressão de Trajetória - ART

A Análise de Regressão é uma ferramenta para definir e quantificar a relação entre uma variável dependente ou resposta (Y) e várias outras variáveis independentes ($X_1, X_2, \dots, X_m$). A estrutura da regressão é dada pela Equação 2.2:

$$Y_j = \sum_{k=1}^m B_k X_{jk} \quad (2.2)$$

Onde Y_j é a variável dependente, B_k são os coeficientes do modelo e X_{jk} são as variáveis independentes.

O foco principal, quando se fala de Análise de Regressão de Trajetória (ART) é em modelos de regressão que combinam as trajetórias de massa de ar com a concentração de poluentes atmosféricos. Ou seja, a ART são modelos que tem por finalidade relacionar a concentração de um poluente em função do tempo de residência da massa de ar sobre uma determinada região, sendo o resultado desse modelo usado para identificar a contribuição que cada região (ou potenciais fontes emissoras) exerce sobre os níveis de concentração do poluente atmosférico.

Assim, na ART, a variável dependente é a concentração do poluente atmosférico (Y_j), enquanto que o tempo de residência da massa de ar sobre a região k é a variável independente (XU et al., 2006; RODOPOULOU et al., 2014; DIMITRIOU; KASSOMENOS, 2014a). Ou ainda pela Equação 2.2, Y_j é a concentração do poluente atmosférico no j -ésimo dia, X_{jk} é o tempo total de residência da massa de ar na região k no dia j e, finalmente, B_k é o fator de contribuição do poluente na região k .

O tempo total de residência é obtido pelos pontos da trajetória da partícula na massa de ar. Ele é calculado pela soma do número de vezes que a trajetória permaneceu pela região pré-determinada, ou seja, soma-se quantos pontos de latitude e longitude estão dentro da região (KAVOURAS et al., 2013). A contribuição da região k na concentração do poluente atmosférico é obtida pela multiplicação de B_k por X_{jk} (Equação 2.2).

As trajetórias das massas de ar, para *input* da ART, podem ser obtidas pelo modelo HYSPLIT, do laboratório NOAA - ARL (STEIN et al., 2015; Rolph, G.D., 2016). Os trabalhos de Xu et al. (2006), Chalbot, McElroy e Kavouras (2013), Kavouras et al. (2013), Dimitriou e Kassomenos (2014a) e Rodopoulou et al. (2014) fizeram uso do modelo HYSPLIT para a Análise de Regressão.

O modelo HYSPLIT usa parâmetros meteorológicos para calcular as trajetórias das massas de ar que alcançam um local receptor. Existem diferentes conjuntos de dados meteorológicos que podem ser usados com o modelo HYSPLIT. Por exemplo, o *Global Data Assimilation System* (GDAS) foi usado por Kavouras et al. (2013), Chalbot, McElroy e Kavouras (2013) e Chalbot et al. (2013). NCEP/NCAR *global reanalysis data* foi usado por Huang et al. (2010). Xu et al. (2006) e Rodopoulou et al. (2014) usaram *ETA Data Assimilation System* (EDAS). Não foi possível identificar qual conjunto de dados meteorológicos foi usado por Dimitriou e Kassomenos (2014a).

Em todos os trabalhos apresentados na Tabela 2.2, as trajetórias reversas foram produzidas para alcançar um local receptor a 500 metros acima do nível do solo. Este limite foi selecionado baseado nas propriedades da camada limite atmosférica, pois é considerado adequado para detecção da influência de diferentes tipos de circulação atmosférica (CHALBOT; MCELROY; KAVOURAS, 2013; DIMITRIOU; KASSOMENOS, 2014a).

A Tabela 2.2 apresenta alguns estudos que usaram da ART. Um estudo, feito em Paris, identificou a contribuição das massas de ar de diversas regiões na concentração diária do MP_{10} e $MP_{2,5}$, usando modelos de regressão linear multivariado. Obteve-se que cerca de 50% da contribuição diária do MP foi de fontes locais (DIMITRIOU; KASSOMENOS, 2014a). Kavouras et al. (2013) usaram ART para identificar as fontes emissoras do $MP_{2,5}$ e $MP_{10-2,5}$ em Amsterdã, Atenas, Birmingham e Helsinki entre Outubro de 2002 e Dezembro de 2004. Encontrou-se que, para Amsterdã e Birmingham, a maior contribuição foram de massas de ar da Europa Central; para Helsinki, da Rússia e Europa Central; e para Atenas, foi do Norte da África e Mar Mediterrâneo.

Baseado em dados coletados entre 2002 e 2004, Chalbot et al. (2013) apontaram as influências locais (tráfego, poeira da estrada, usina elétrica) e fontes regionais (emissões de navios e transporte da Europa Central e Oeste Europeu) na concentração do MP_{10} e $MP_{2,5}$ na cidade de Atenas, Grécia, usando ART.

Outros trabalhos, nos EUA, usaram da ART para determinar os componentes locais e regionais do MP e do ozônio (RODOPOULOU et al., 2014) e também da concentração

Tabela 2.2: Características do mais relevantes artigos sobre Análise de Regressão de Trajetória.

Autor (ano)	Poluente atmosférico	Local	Principais resultados
Xu et al. (2006)	Sulfato de amônia	Oeste dos EUA	Identificou-se a contribuição do transporte marítimo do oceano Pacífico
Huang et al. (2010)	<i>Black Carbon</i>	Ártico	Notou-se que 67% da concentração do Black Carbon veio da Rússia.
Chalbot et al. (2013)	MP ₁₀ e MP _{2,5}	Atenas, Grécia	Foi encontrado que mais que 40% do MP foi originado de regiões distantes mais que 500 Km de Atenas.
Chalbot, McElroy e Kavouras (2013)	MP _{2,5}	Arkansas, EUA	Foi identificado o efeito de queimadas na concentração do MP _{2,5} .
Kavouras et al. (2013)	MP _{2,5} e MP _{10-2,5}	Amsterdã (Holanda), Atenas (Grécia), Birmingham (Inglaterra), e Helsinki (Finlândia)	Encontrou-se que para Amsterdã e Birmingham a maior contribuição foi de massas de ar da Europa Central, para Helsinki, o MP é originado na Rússia e Europa Central e para Atenas, do Norte da África e do Mar mediterrâneo.
Dimitriou e Kassomenos (2014a)	MP ₁₀ e MP _{2,5}	Paris, França	observou-se que aproximadamente 50% da contribuição das massas de ar na concentração do MP ₁₀ foram de fontes locais.
Rodopoulou et al. (2014)	MP ₁₀ , MP _{2,5} e O ₃	Doña Ana County, EUA	Fontes dentro de um raio de 500 km contribuíram para a concentração do MP, e a poeira da estrada foi uma das principais fontes do MP ₁₀ e MP _{2,5} .

do sulfato de amônia (XU et al., 2006). Rodopoulou et al. (2014) encontraram que fontes dentro de 500 km do sul da área de estudo contribuíram na maior parte na concentração do MP_{10} , $MP_{2,5}$ e O_3 , sendo a poeira da estrada de vias parvimentadas e não-pavimentadas a principal fonte dos particulados. Xu et al. (2006) sugeriram que embarcações e outras emissões portuárias da costa do Pacífico são as mais importantes contribuições na concentração do sulfato de amônia sobre grandes áreas do Oeste dos EUA.

Chalbot, McElroy e Kavouras (2013) identificaram o efeito das queimadas na concentração do $MP_{2,5}$ usando a ART, em Arkansas, no vale do Mississipi. Além do efeito das queimadas, o impacto de atividades portuárias do Golfo do México e cidades costeiras foram identificados. Chalbot, McElroy e Kavouras (2013) concluíram que a qualidade do ar de áreas urbanas foi afetada pela emissão de fontes a 1000 km de distância.

No Ártico, Huang et al. (2010) encontraram que 67% da concentração do *black carbon* é proveniente da Rússia, enquanto que 18% e 15% são provenientes da União Europeia e América do Norte, respectivamente.

Os modelos de regressão usados nos estudos da Tabela 2.2, assumem que a variável dependente segue uma distribuição Normal para a realização dos testes estatísticos. Com a suposição de normalidade, os valores possíveis da variável dependente é assumido qualquer valor real, ou seja, pode assumir valores negativos e positivos. Além disso, uma das suposições desse modelo, é em relação à independência dos dados, ou sobre a suposição de não-correlação dos valores da variável dependente. Entretanto, dois pontos podem ser observados. Primeiro, como trata-se de concentração de poluentes, os valores possíveis são sempre positivos. Em segundo, a suposição de independência pode não ser a mais adequada, pois os dados estão dispostos numa série temporal. Os trabalhos selecionados na Tabela 2.2 não abordam essas duas situações.

Kavouras et al. (2013) e Rodopoulou et al. (2014) descreveram algumas limitações do modelo de regressão da ART. Primeiro, os erros na localização das trajetórias podem não ser aleatórios e conduzir a grandes erros no cálculo da contribuição média, principalmente em dias com concentração diária muito alta (ou muito baixa). Em segundo, contribuições médias para dias individuais podem ser altamente instáveis.

A ART é uma técnica muito interessante para identificar fontes emissoras de poluentes atmosféricos. Entretanto, há possibilidades de melhoria nos modelos já utilizados, como incorporar uma estrutura de séries temporais (MORETTIN; TOLOI, 2006), ou ainda utilizar de modelos lineares generalizados (CORDEIRO, 1986; MCCULLAGH; NELDER, 1989; PAULA, 2013), para adequar melhor ao comportamento da concentração do poluente. A aplicação da modelagem de séries temporais ou de modelos lineares generalizados são uma tentativa de superar as limitações mencionadas por Kavouras et al. (2013) e Rodopoulou et al. (2014).

Análise de séries temporais é usada para prever a concentração do poluente em diversos locais. Metodologias de *Box-Jenkins*, como os modelos, ARMA (*Auto-Regressive Moving Average*), ARIMA (*Auto-Regressive Integrated Moving Average*) e SARIMA (*Seasonal Auto-Regressive Integrated Moving Average*) (MORETTIN; TOLOI, 2006; LEITE et al., 2013), e as redes neurais artificiais (HAN; KAMBER, 2006) são alguns exemplos de modelos de análise de séries temporais.

Na cidade de São Carlos, São Paulo, Pozza et al. (2010) usaram os modelos SARIMA e suavização de *Holt-Winters* para a previsão do $MP_{2,5}$ e $MP_{10-2,5}$. Os modelos sazonais SARIMA apresentaram um melhor desempenho. Souza et al. (2015) compararam redes neurais artificiais individuais com *ensembles* de redes neurais artificiais para a previsão da concentração do MP_{10} na cidade de Piracicaba, São Paulo. A abordagem baseada nos *ensembles* conduziram a resultados melhores em relação a previsão. Os modelos SARFIMA (*Seasonal Auto-Regressive Fractional Integrated Moving Average*) com modelagem da variância, usando processos GARCH (Generalized Autorregressive Conditional Heteroscedastic) foram estudados na série temporal do MP_{10} na cidade de Cariacica, Espírito Santo. Estes modelos ajustam séries temporais com volatilidade, em que se considera que a variância do erro se altera de acordo com o tempo (REISEN et al., 2014). Os modelos destes estudos (POZZA et al., 2010; SOUZA et al., 2015; REISEN et al., 2014) consideram no ajuste da série apenas os próprios dados da concentração MP, sem incluir nenhuma variável independente no modelo, ou seja, apenas modelando a série ao longo do tempo.

Goyal, Chan e Jaiswal (2006) utilizaram de modelos de regressão, modelos ARIMA e uma combinação entre os dois para a previsão do MP_{10} na cidades de Delhi e Hong Kong. Encontraram que o modelo combinado das duas técnicas foi melhor, conseguindo explicar mais de 70% da variação total. Liu e Johnson (2002) também usaram os modelos de regressão combinados com modelos ARIMA para prever picos diários de concentração de O_3 na cidade de Milwaukee, EUA. Neste estudo, esse modelo conseguiu explicar 80,9% da variação total. Num estudo na cidade de Vitória, esses modelos combinados foram chamados de ARMAX, e usados para prever a concentração horária de ozônio (MONTE; ALBUQUERQUE; REISEN, 2015). Como variáveis regressoras foram usadas a temperatura, umidade relativa, velocidade do vento e concentração do NO_2 . Pinto, Reisen e Monte (2018) utilizaram uma combinação de modelos SARIMA com variáveis regressoras, chamado de SARIMAX, para a previsão do MP_{10} diário na região da grande Vitória. Dentre as variáveis meteorológicas avaliadas, velocidade do vento e precipitação pluvial foram significativas para explicar a concentração do MP_{10} .

Já os modelos propostos por (GOYAL; CHAN; JAISWAL, 2006; LIU; JOHNSON, 2002; MONTE; ALBUQUERQUE; REISEN, 2015; PINTO; REISEN; MONTE, 2018) consideraram variáveis independentes combinando com modelos ARIMA ou SARIMA. Porém, nenhum deles considerou como variáveis independentes o tempo de residência das massas de ar, ou seja, o uso de um modelo ART.

Modelos lineares generalizados são uma extensão dos modelos lineares de regressão. Tais modelos flexibilizam a distribuição de probabilidade da variável dependente. Nesses modelos a variável resposta pode ter as distribuições normal, gama, binomial, poisson, normal inversa, entre outras (MCCULLAGH; NELDER, 1989; PAULA, 2013; CORDEIRO, 1986). Outra classe de modelos são os modelos aditivos generalizados para locação, escala e forma (GAMLSS, do inglês *Generalized Additive Models for Location, Scale and Shape*), que abrangem outras distribuições de probabilidades, como a Log-normal, Weibull, entre outras (RIGBY; STASINOPOULOS, 2005; STASINOPOULOS; RIGBY, 2007). Agarwal e Nagendra (2016) ajustaram a distribuição log-normal para dados de MP_{10} , $MP_{2,5}$ e MP_1 . Sharma et al. (2013) compararam diversas ditribuições assimétricas para ajuste de

dados de SO_2 , NO_2 , MP_{10} e PTS em Delhi, na Índia. Para o MP_{10} as distribuições Gama e Log-normal se ajustaram melhor.

Huebnerova e Michalek (2014) utilizaram um modelo linear generalizado com distribuição gama para identificar os fatores meteorológicos que afetam a poluição do ar, mais especificamente, o MP_{10} em Brno, na República Checa. As medições ocorreram em duas estações de monitoramento, localizadas em vias de alto tráfego de veículos e numa área industrial. Na cidade de Barreiro, em Portugal, Garcia et al. (2016) aplicaram um modelo linear generalizado com distribuição de Poisson para analisar a relação da concentração do MP_{10} em relação a gases poluentes (CO , NO_2 , NO_x , O_3 , SO_2) e variáveis meteorológicas (temperatura, umidade relativa e velocidade do vento). Um modelo linear generalizado também foi usado em Beijing, na China, para relacionar o $\text{MP}_{2,5}$ com parâmetros meteorológicos (Direção e velocidade do vento, temperatura, pressão, umidade relativa, precipitação). Conseguiu-se explicar aproximadamente 80% da variabilidade total com esse modelo (LIANG et al., 2017). Não foi possível identificar a distribuição usada nesse estudo.

Nos estudos de (AGARWAL; NAGENDRA, 2016; SHARMA et al., 2013; HUEBNEROVA; MICHALEK, 2014; GARCIA et al., 2016; LIANG et al., 2017), distribuições diferentes da normal, foram usadas para ajuste de dados de concentração, mas nenhum deles combinou com dados de trajetórias de massas de ar.

O presente estudo pretende utilizar três abordagens para a ART. A primeira é o modelo usual de regressão (modelo normal) já empregado nos trabalhos anteriores (Tabela 2.2). A segunda é utilizar um modelo linear generalizado ou um modelo aditivo generalizado para locação escala e forma, que considere a estrutura da variável resposta (concentração do MP), isto é, considere que essa variável é um valor sempre positivo. A terceira é um modelo de séries temporais baseado na metodologia de Box-Jenkins. As diferentes abordagens sugeridas para os modelos ART neste trabalho são uma tentativa de corrigir as limitações descritas nos estudos de Kavouras et al. (2013) e Rodopoulou et al. (2014).

Além da ART, existem outras ferramentas estatísticas que podem auxiliar no entendimento das relações existentes entre a concentração do MP e as variáveis meteorológicas. Duas importantes ferramentas estatísticas se destacam nesta aplicação: a Análise de Cluster (AC) e Análise de Componentes Principais (ACP), que serão discutidas na Seção 2.6.

2.6 Uso de AC e ACP no estudo de variáveis meteorológicas e poluentes atmosféricos

As análises multivariadas de AC e ACP têm por objetivo principal encontrar relações multivariadas complexas, que a princípio são difíceis de serem notadas.

A Análise de Cluster (AC) é amplamente aplicada para agrupar vários objetos em grupos homogêneos, sendo que o nível de associação entre objetos no mesmo cluster será mais forte do que o nível de associação entre objetos de diferentes clusters (DOMINICK et al., 2012; PAVÓN-DOMÍNGUEZ; JIMÉNEZ-HORNERO; RAVÉ, 2014).

AC é uma técnica usada para agrupar observações onde, nenhuma suposições são feitas sobre o número de grupos ou estrutura dos grupos. É baseada em similaridades ou

distâncias, como a Euclidiana, máxima, Manhattan, Minkowski, entre outras. A fim de criar grupos na AC, existem métodos hierárquicos e não hierárquicos. Para o primeiro método existem vários algoritmos como ligação simples, ligação completa, Ward, etc. Para o segundo método, o mais conhecido é o *k-means* (JOHNSON; WICHERN, 1988; AUSTIN et al., 2012; DOMINICK et al., 2012).

A Análise de Componentes Principais (ACP) é uma ferramenta estatística que visa obter um número reduzido de combinações lineares não-correlacionadas das variáveis originais, chamadas de componentes principais (CP). As CPs são ordenadas de forma que a 1ª CP explica a maior fração da variabilidade dos dados originais. A 2ª CP explica a segunda maior fração, e assim por diante. A ideia é que as CPs consigam explicar uma grande parte da variabilidade total das variáveis originais, sem muita perda de informação (JOHNSON; WICHERN, 1988; HE; LU, 2012; DOMINICK et al., 2012; KASSOMENOS et al., 2014).

ACP pode ser dada pela Equação 2.3.

$$Z_i = a_{1i}x_1 + a_{2i}x_2 + a_{3i}x_3 + \dots + a_{ni}x_n \quad (2.3)$$

Onde Z é a componente principal, a são os coeficientes da componente principal, x é o valor da variável, i é o número da componente e n é o número de variáveis (DOMINICK et al., 2012).

O uso da AC e ACP podem gerar importante informação para permitir identificar a relação entre os poluentes e variáveis meteorológicas. A Tabela 2.3 apresenta alguns estudos que usaram a AC e/ou ACP.

Nos EUA, um estudo utilizou da AC para agrupar locais de acordo com a composição do $MP_{2,5}$. Os pesquisadores obtiveram 4 clusters: Leste dos EUA, áreas industrializadas, Oeste/Centro dos EUA, e o último, nas áreas da costa (AUSTIN et al., 2013). Usando a AC hierárquica, 8 estações de monitoramento da região metropolitana de Sevilla foram agrupadas em 3 clusters diferentes baseados no padrão de comportamento do O_3 (PAVÓN-DOMÍNGUEZ; JIMÉNEZ-HORNERO; RAVÉ, 2014). Em Xi'an, na China, Huang et al. (2015) estudaram 13 locais com similaridade baseados no perfil do $MP_{2,5}$. Três clusters foram identificados, sendo os clusters 1 e 2 formados por locais de atividades industriais e uma ampla estrutura de produção, porém no cluster 2 foi formado por locais com atividades industriais de alta tecnologia, como elétricas e de equipamentos de petróleo, por exemplo. Por fim, o cluster 3 agrupou, principalmente, locais com intensa atividade agrícola e atividades turísticas.

Austin et al. (2012) classificaram dias em grupos baseados nos perfis de concentração dos poluentes (gases e particulados), na cidade de Boston, entre 2004 e 2009. Eles identificaram e caracterizaram 5 grupos de dias, de acordo com propriedades químicas, físicas e meteorológicas. Os dias do cluster 1 sugeriram maior impacto das emissões de tráfego; dias do cluster 2 sugeriram suspensão e re-suspensão de materiais crustais; O cluster 3 indicou a presença de emissões primárias; O cluster 4 refletiu episódios de sulfato; e o cluster 5 indicou uma forte contribuição de poeira e queima de madeira. Essa classificação foi feita utilizando AC hierárquica com método de *k-means*.

Tabela 2.3: Características de alguns artigos relevantes, associados com AC e ACP.

Autor (ano)	Local	Técnica empregada	Cluster (AC)	Variáveis relacionadas por ACP
Statheropoulos, Vassiliadis e Pappa (1998)	Atenas, Grécia	ACP	–	Poluentes atmosféricos e variáveis meteorológicas
Lu, He e Dong (2011)	Hong Kong, China	AC e ACP	Estações de monitoramento	Estações de monitoramento
Dominick et al. (2012)	Malásia	AC e ACP	Estações de monitoramento	Poluentes atmosféricos e variáveis meteorológicas
Austin et al. (2012)	Boston, EUA	AC	Dias do ano	–
He e Lu (2012)	Hong Kong	ACP	–	Variáveis meteorológicas
Austin et al. (2013)	EUA	AC	Locais de monitoramento	–
Dimitriou e Kassomenos (2014b)	Alemanha	ACP	–	Poluentes atmosféricos
Kassomenos et al. (2014)	Atenas, Grécia, Londres, Inglaterra and Madri, Espanha	ACP	–	Poluentes atmosféricos
Latif et al. (2014)	Malásia	AC e ACP	Horas do dia	Poluentes atmosféricos
Luna et al. (2014)	Rio de Janeiro, Brasil	ACP	–	Poluentes atmosféricos e variáveis meteorológicas
Pavón-Domínguez, Jiménez-Hornero e Ravé (2014)	Sevilha, Espanha	AC	Estações de monitoramento	–
Souza et al. (2014)	Vitória, Brasil	ACP	–	Poluentes atmosféricos
Huang et al. (2015)	China	AC	Locais de monitoramento	–
Kwon et al. (2015)	Seul, Coréia do Sul	ACP	–	Variáveis meteorológicas e MP
Targino e Krecl (2016)	Londrina, Brasil	AC	Dias do ano	–
Binaku e Schmeling (2017)	Chicago, USA	ACP	–	Poluentes atmosféricos e variáveis meteorológicas
Souza et al. (2017)	Campo Grande, Brasil	ACP	–	Poluentes atmosféricos e variáveis meteorológicas

Na cidade de Londrina-PR, Brasil, Targino e Krecl (2016) utilizaram a AC para agrupar dias com perfis similares em relação as variáveis meteorológicas (temperatura, umidade relativa e velocidade do vento), e posteriormente identificaram a relação com o padrão diurno do *black carbon*. Encontraram que o *black carbon* é relacionado com a queima de biomassa e transporte da fumaça de longo alcance. Entretanto, os autores sugerem que o controle de fontes locais, como queima de biomassa e tráfego intenso podem conduzir a diminuição da concentração do *black carbon* nesta cidade.

Na cidade de Atenas, na Grécia, foram comparados os níveis de concentração de poluentes CO, NO, NO₂, O₃ e SO₂, utilizando da ACP. Foram levadas em consideração também as variáveis meteorológicas de umidade relativa, temperatura, duração do tempo de sol e velocidade e direção do vento. Os poluentes foram relacionados à combustão de gasolina e óleo, e interações com o ozônio, enquanto que os principais fatores meteorológicos foram condições secas (no verão) e velocidade do vento (no inverno) (STATHEROPOULOS; VASSILIADIS; PAPPA, 1998).

Em Seul, Coréia do Sul, o MP foi coletado em 6 estações de metrô para estudar sua influência dentro e fora dos trens, e a ACP foi usada para relacioná-lo com variáveis meteorológicas. Obtiveram que a concentração do MP é influenciada pela ventilação, afetando tanto a concentração do MP dentro e fora dos trens (KWON et al., 2015). Entre os anos de 2011 e 2012, Luna et al. (2014) usaram da ACP para relacionar variáveis meteorológicas (velocidade do vento, radiação solar, temperatura e umidade) e poluentes atmosféricos (NO₂, NO, NO_x, CO e O₃) em 3 estações de monitoramento na cidade do Rio de Janeiro. Encontraram que o NO, NO_x e radiação solar tem forte impacto na concentração do O₃.

A ACP foi utilizada para relacionar poluentes atmosféricos (MP₁₀, MP_{2,5}, NO, NO₂, NO_x, SO₂, O₃ e CO) nas cidades de Atenas, Londres e Madri, nos períodos de frio e calor, em duas estações de monitoramento em cada uma das cidades. As Componentes Principais (CP) foram associadas a condições de tráfego, atividades foto-químicas e combustão de combustíveis (KASSOMENOS et al., 2014). Dimitriou e Kassomenos (2014b) utilizaram da ACP para agrupar poluentes atmosféricos com o objetivo de encontrar fontes emissoras. Identificaram 4 fontes emissoras, sendo: emissões veiculares (CP 1); poluição fotoquímica (CP 2); emissões industriais (CP 3) e particulados emitidos de indústrias e tráfego veicular (CP 4).

Souza et al. (2014) utilizaram a ACP como *input* ao modelo aditivo generalizado de Poisson. A ACP agrupou os poluentes atmosféricos (CO, NO₂, O₃, MP₁₀ e SO₂) em 3 CPs, e então foram usadas como variáveis independentes no modelo de Poisson, sendo o número de internações, a variável resposta. Encontraram que um aumento de 10,49 $\mu\text{g}/\text{m}^3$ de MP₁₀, resultou num aumento de 3% no risco relativo de admissões hospitalares para a região de Vitória-ES, Brasil.

Em 2012, He e Lu (2012) utilizaram da ACP para relacionar variáveis meteorológicas, e posteriormente (por meio de um modelo de regressão) relacionar com a concentração do O₃, em estações de monitoramento em Hong Kong. As variáveis meteorológicas explicaram 31% e 34% (nas estações Sha Tin e Kwun Tong, respectivamente) da variação do ozônio, sendo que a mais relevante foi a quantidade de nuvens.

Na cidade de Campo Grande-MS, durante 2014, Souza et al. (2017) usaram ACP para relacionar variáveis meteorológicas (temperatura, umidade relativa, precipitação, velocidade do vento, radiação solar e índice de claridade) com dois poluentes do ar (ozônio e monóxido de carbono). Temperatura, velocidade do vento e radiação solar foram a principal contribuição na concentração do ozônio (38,2%).

Durante o verão (Julho e Agosto) de 2010, 2011 e 2012, em Chicago, nos EUA, dados meteorológicos, aerossóis e gases traços foram obtidos por Binaku e Schmeling (2017). Os autores usaram ACP para identificar a relação entre variáveis meteorológicas e de poluição atmosférica. Quatro CPs explicaram aproximadamente 70% da variabilidade dos dados, em que a CP1 foi associada com a relação do ozônio e nitrato com respeito a temperatura e umidade, enquanto a CP2 foi relacionada com a presença de acetato e formato, resultantes de emissões primárias ou transporte. Finalmente, CP3 e CP4 representou transporte de vento do norte de cloreto, nitrato e ozônio.

A AC e ACP foram comparadas a fim de obter áreas da cidade de Hong Kong com comportamentos de poluição similares. Encontraram que a direção do vento e emissões veiculares foram as principais influências na concentração de poluentes. Além disso, concluíram que as duas ferramentas são ótimos métodos de se gerenciar sistemas de monitoramento da qualidade do ar (LU; HE; DONG, 2011).

Dominick et al. (2012) apresentaram uma avaliação da qualidade do ar na Malásia, usando análise estatística multivariada, como AC, ACP e análise de regressão múltipla. As principais fontes detectadas foram veículos, indústrias e áreas com alta densidade populacional. O MP_{10} foi o poluente que mais contribuiu na poluição do ar, e notou-se que fatores meteorológicos como temperatura do ambiente e umidade influenciaram na concentração do MP_{10} e de outros poluentes. Umidade e temperatura tiveram uma correlação inversa, ou seja, baixa umidade (relacionada a baixos índices de chuva), levaram a altas temperaturas e altas concentrações dos poluentes.

Um outro estudo, também conduzido na Malásia (LATIF et al., 2014), usaram AC e Regressão por Componente Principal (RCP). Esta última análise é uma combinação de ACP com modelos de regressão, em que a CP é usada como variável dependente na regressão. Primeiramente, AC foi usada para avaliar padrões diurnos baseados na concentração média anual de vários poluentes atmosféricos (NO , NO_2 , NO_x , MP_{10} , CO e SO_2). Em segundo lugar, RCP foi usada para avaliar variáveis meteorológicas (temperatura, umidade, velocidade do vento e radiação ultravioleta) mais significativas. Ou seja, a análise de regressão foi realizada pelas CPs, formada pelos poluentes atmosféricos, em função das variáveis meteorológicas. Foi encontrado que fontes de emissão, como veículos automotores, foram os que principalmente contribuíram com a poluição do ar. Além disso, parâmetros meteorológicos foram muito importantes na variabilidade da poluição do ar, influenciando na sua dispersão e reações químicas fotosensíveis.

Como visto, a AC e ACP são análises bastante utilizadas em diversos trabalhos. Por exemplo, é possível, encontrar perfis espaciais e temporais do comportamento de poluentes.

A AC e ACP usadas em conjunto, podem gerar informações importantes na identificação da relação entre as variáveis poluentes e meteorológicas. Poucos estudos utilizam conjuntamente essas duas técnicas, como Austin et al. (2012), Dominick et al. (2012).

Num contexto estadual, nenhum estudo foi feito aqui no estado de São Paulo (pelo conhecimento adquirido), utilizando dessas técnicas conjuntamente em diferentes estações de monitoramento ou mesmo cidades.

Uma possibilidade interessante seria elaborar uma abordagem usando as 3 técnicas em conjunto (ART, AC e ACP). Num primeiro momento, a Análise de Regressão de Trajetória, poderia identificar os locais das fontes emissoras dos poluentes (MP_{10} e $MP_{2,5}$). Em seguida, a AC possibilitaria agrupar os dias, baseado na contribuição de cada região obtida na ART, no perfil meteorológico e do comportamento dos poluentes atmosféricos, obtendo um padrão temporal. Por fim, a ACP quantificaria a relação existente entre as variáveis dentro de cada cluster obtido. Dominick et al. (2012), em seu estudo, considerou o uso conjunto da AC, ACP e modelos de regressão, com objetivo de identificar padrões espaciais da qualidade do ar e também estimar principais fontes da poluição do ar na Malásia. Porém, não foi empregado as trajetórias de massas de ar em nenhuma análise.

Deste modo, este trabalho pretende abordar essas 3 técnicas (ART, AC e ACP), de forma a contribuir na investigação de problemas na poluição do ar em cidades paulistas. Essas análises serão descritas com maiores detalhes ao longo do trabalho, no Capítulo 3.

Capítulo 3

Métodos

Este capítulo é dividido em 6 seções. Na primeira, Seção 3.1, são descritos as duas cidades de estudo. Na Seção 3.2, são apresentados como os dados foram obtidos, da base do QUALAR. Na Seção 3.3 são descritos como foram obtidas as trajetórias das massas de ar do modelo HYSPLIT. Nas Seção 3.4, Seção 3.5 e Seção 3.6 são descritas as ferramentas estatísticas (ART, AC e ACP, respectivamente) utilizadas na análise dos resultados deste trabalho.

3.1 Área de estudo

As cidades estudadas foram: Piracicaba e Santos. Essas cidades possuem monitoramento tanto do MP_{10} , quanto do $MP_{2,5}$. Esse é um dos motivos para escolha das mesmas. Santos possui duas estações automáticas, porém os dados foram obtidos da estação da Ponta da Praia. Esta estação tem como um dos seus objetivos, avaliar os impactos das emissões dos poluentes provenientes do porto (CETESB, 2015a). Com a escolha dessas cidades, há uma cobertura de duas diferentes regiões do Estado de São Paulo, abrangendo o Interior e Litoral. No total serão 2 estações de monitoramento avaliadas, uma para cada cidade. Vale ressaltar que a metodologia descrita neste Capítulo 3, foi aplicada em cada cidade separadamente.

Nas Figuras 3.1 e 3.2 são apresentados os mapas com as manchas urbanas e a malha rodoviária próxima as cidades de Piracicaba (Figura 3.1) e Santos (Figura 3.2). Também são mostrados a localização das estações de monitoramento automáticas da CETESB. Esses mapas foram feitos no sistema DataGEO, da Secretaria de Infraestrutura e Meio Ambiente do Estado de São Paulo (DATAGEO, 2019).

Piracicaba é uma cidade do interior, distribuída em 1378,069 Km², com uma população estimada em 2018 de 400949 habitantes (IBGE, 2019). Piracicaba fica próxima a Limeira, Rio Claro, Americana e a região metropolitana de Campinas (Figura 3.1). As principais rodovias próximas a Piracicaba são SP147 (Deputado Laércio Corte), SP127 (Cornélio Pires), SP304 (Luiz de Queiroz) e SP308 (Rodovia do Açúcar), além de estar bem próxima da rodovia SP348 (Bandeirantes) (Figura 3.1). O clima de Piracicaba é subtropical úmido com inverno seco e verão quente (Cwa), segundo a classificação de Köppen (ALVARES et al., 2013; FRANÇA et al., 2010). Janeiro é o mês mais chuvoso (média histórica de 255,55

mm) e agosto o mês mais seco (média de 28,7 mm), com pluviosidade média anual de 1356,5 mm (CIIAGRO, 2019). A média histórica de temperatura, revela que o mês de Fevereiro é o mais quente (25,0 °C), enquanto Julho é o mês mais frio (18,1 °C) (CIIAGRO, 2019).

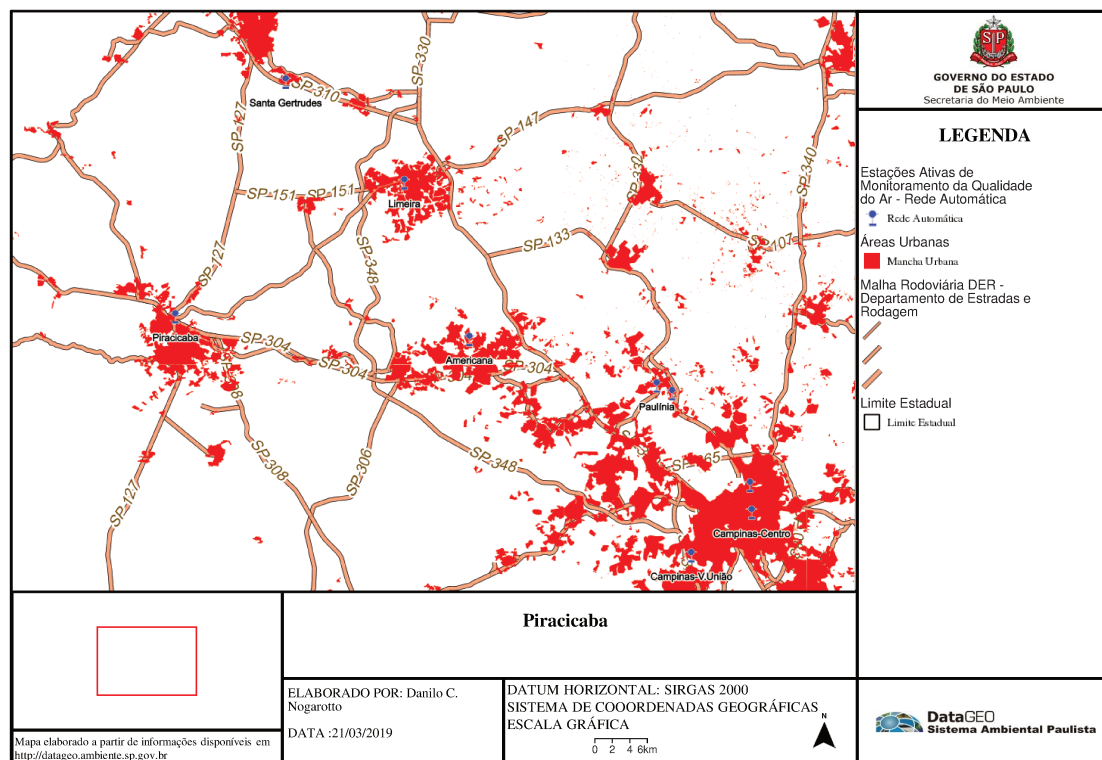


Figura 3.1: Mapa de Piracicaba. Fonte: (DATAGEO, 2019)

Santos é uma cidade litorânea, com população estimada em 2018 de 432957 habitantes, numa área de 281,033 Km² (IBGE, 2019). Fica próxima ao polo industrial de Cubatão, além da região metropolitana de São Paulo (Figura 3.2). As principais rodovias de acesso a cidade são a Via Anchieta (SP150), rodovias Imigrantes, Padre Manuel da Nóbrega, Rio-Santos e Cônego Domêmico Ranganoni. De acordo com a classificação de Köppen, o clima de Santos é Af, clima tropical úmido ou superúmido (ALVARES et al., 2013; COMINI-FROTA et al., 2013), com o mês de janeiro, o mais chuvoso (média histórica de 291,4 mm) e agosto o mês menos chuvoso com média histórica de 74,1 mm, enquanto o total médio anual é de 2050,6 mm. Considerando a média histórica de temperatura, o mês de fevereiro é o mais quente, com 26,9 °C, e julho o mês mais frio, com 19,4 °C (CIIAGRO, 2019).

Os relatórios da CETESB (CETESB, 2014; CETESB, 2015b) apresentam qualitativamente os principais fenômenos meteorológicos atuantes no estado de São Paulo, durante o período de estudo (2014 e 2015). O ano de 2014 foi muito seco, com baixas precipitações pluviométricas. No primeiro trimestre, houve atuação do Anticiclone Subtropical do Atlântico Sul e também de vórtices ciclônicos anômalos na média e alta troposfera. Esses fenômenos dificultaram a ZCAS (Zona de Convergência do Atlântico Sul) e ZCOU (Zona de Convergência de Umidade), que são, em partes, responsáveis pela precipitação.

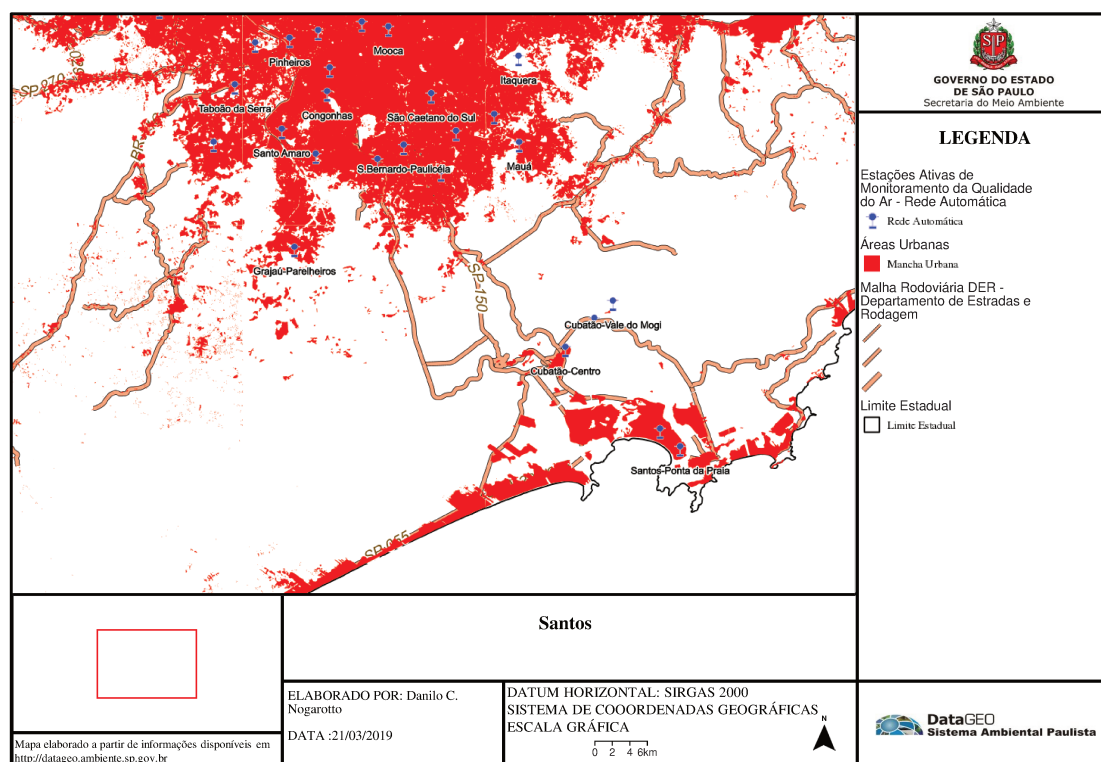


Figura 3.2: Mapa de Santos. Fonte: (DATAGEO, 2019)

No quarto trimestre, também houve atuação do Anticiclone Subtropical do Atlântico Sul, que provocou estabilidade atmosférica. E com isso, inibiu a chegada de frentes frias, que poderiam provocar instabilidade, e consequentemente, precipitações. Durante o ano de 2015, ocorreu o fenômeno global do El Niño. Em Janeiro/2015, houve um bloqueio atmosférico, que ocasionou baixas precipitações neste mês. Entretanto, em Fevereiro e Março de 2015, as precipitações ficaram acima da média, devido aos episódios de ZCAS e ZCOU. Em Setembro/2015, devido ao predomínio de uma massa de ar quente (sistema de alta pressão subtropical), não houve ocorrência de chuvas, pois essas massas de ar quente, dificultaram a entrada de sistemas instáveis, de baixa pressão. No mês de Dezembro/2015, em Santos, as chuvas ficaram acima da média, devido atuação do El Niño.

3.2 Dados

Os dados são referentes às variáveis meteorológicas (Umidade Relativa, Velocidade do Vento e Temperatura) e variáveis poluentes (MP_{10} e $MP_{2,5}$) das cidades de Santos e Piracicaba, SP, e foram obtidos na base QUALAR, disponível no site da CETESB (QUALAR, 2016). Essa base contém os dados históricos das estações de monitoramento de diversas cidades de SP, tanto das estações de monitoramento automática, quanto da manual. Neste estudo, foram selecionadas as informações das estações de monitoramento automático, que dispõe os dados horários de cada variável. O período dos dados foram dos anos de 2014 e 2015.

Considerou-se os dados médios diários nas análises. Com os dados horários, foi calculada a média diária das variáveis. Entretanto, nas bases de dados, retiradas das estações de monitoramento automático, há casos de dados faltantes. Na Tabela 3.1 são apresentados os percentuais de dados faltantes em cada uma das estações de monitoramento para cada variável. Como visto, não há muitos dados faltantes na base de dados, sendo o máximo 3,6% para o MP_{2,5} na cidade de Santos. Esse percentual corresponde a 26 dias sem medição.

A análise de séries temporais, baseada na metodologia de Box-Jenkins, supõe que a série esteja completa para se obter o ajustes desses modelos (MORETTIN; TOLOI, 2006). Deste modo, essas lacunas de dados faltantes foram preenchidas, para que a série histórica fique completa. O preenchimento foi feito pelo método *last and next*, onde os valores faltantes são substituídos pela média do último e do próximo valor da variável (PLAIA; BONDI, 2006).

Tabela 3.1: Região, Cidades e Percentual de dados faltantes por variável

Região	Cidades	MP ₁₀	MP _{2,5}	UR	VV	Temp
Interior	Piracicaba	1,1%	2,7%	1,1%	1,1%	1,1%
Litoral	Santos	3,2%	3,6%	3,2%	3,3%	3,2%

3.3 Trajetórias das massas de ar

O modelo HYSPLIT (*Hybrid Single-Particle Lagrangian Integrated Trajectory*), do laboratório NOAA – ARL (*National Oceanic and Atmospheric Administration – Air Resources Laboratory*) (DRAXLER; HESS, 1998; STEIN et al., 2015; Rolph, G.D., 2016) com conjunto de dados meteorológicos NCEP/NCAR Global Reanalysis (KALNAY et al., 1996; HOINASKI et al., 2014) foi usado para obter as trajetórias de massas de ar que atingem um dado destino (no caso, as coordenadas das estações de monitoramento de cada cidade). Essas trajetórias são chamadas de reversas ou *backward*. Considerou-se três dias de trajetórias reversas (72 h), quatro vezes por dia (a cada seis horas), iniciando-se as 3:00 h (3:00 h, 9:00 h, 15:00 h e 21:00 h UTC (*Universal Time Coordinated*), ou 0:00 h, 6:00 h, 12:00 h e 18:00 h horário local). As trajetórias foram calculadas para alcançar a cidade a 500 m acima do solo. Este limite foi selecionado como adequado para detectar a influência de diferentes tipos de circulação atmosférica nos níveis do MP (DIMITRIOU; KASSOMENOS, 2014a; KAVOURAS et al., 2013; RODOPOULOU et al., 2014). As coordenadas dos pontos de trajetórias são obtidas pelo modelo HYSPLIT, em uma base horária, ou seja, cada trajetória é constituída de 73 pontos de trajetória (3 dias *backward* por trajetória $\times$ 24 h por dia de trajetória + um ponto de chegada).

Por exemplo, a Figura 3.3 mostra uma trajetória reversa para a cidade de Piracicaba. São 73 pontos de latitude e longitude (coordenadas) gerados pelo modelo HYSPLIT. Esta trajetória corresponde ao dia 01/01/2014 as 3:00 h (UTC) ou 0:00 h (local).

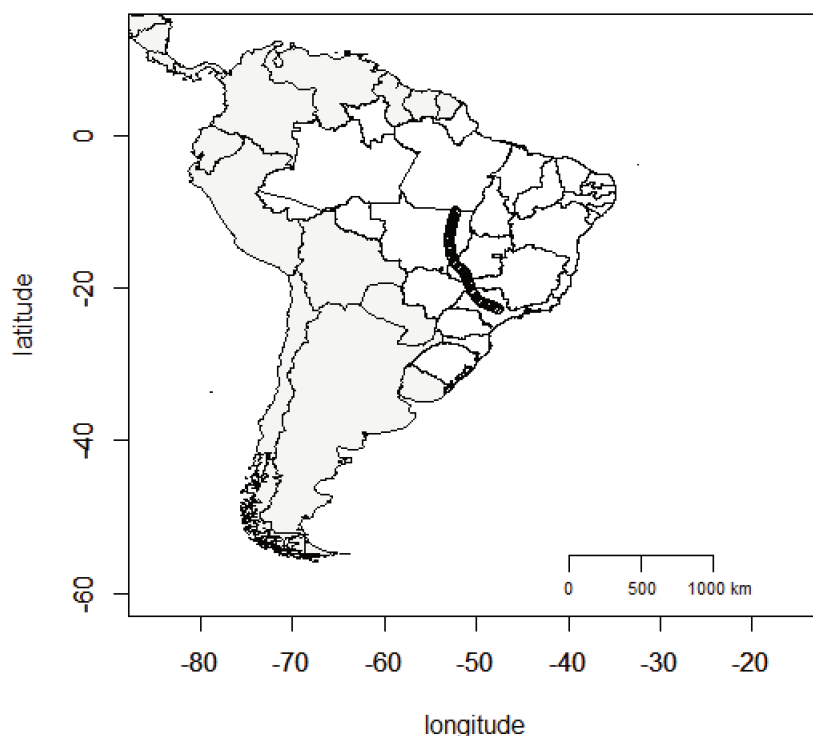


Figura 3.3: Exemplo de uma trajetória reversa para a cidade de Piracicaba

3.4 Análise de Regressão de Trajetória

Após a obtenção das trajetórias de massa de ar no modelo HYSPLIT, é necessário definir regiões, em que são analisadas na ART. As regiões pré-definidas incluem 4 regiões ao redor da cidade (ao Nordeste, Sudeste, Noroeste e Sudoeste) para descrever as contribuições locais. Essas 4 regiões estão contidas num quadrado, cujo lado está distante 5° da coordenada do local da estação automática de monitoramento em cada cidade (Tabela 3.2). Essas regiões serão sempre numeradas de 1 a 4, conforme exemplo para a cidade de Piracicaba (Figura 3.4). As outras regiões representam áreas com características similares. Regiões mais distantes do local receptor, ficaram com áreas maiores em relação as regiões próximas, para assim, considerarmos os maiores erros em pontos de trajetória distantes do local do receptor. A ideia é tentar diferenciar a contribuição de fontes de áreas urbanas de fontes mais distantes (DIMITRIOU; KASSOMENOS, 2014a; KAVOURAS et al., 2013). Consequentemente, os limites de cada região variam para cada cidade.

Tabela 3.2: Coordenadas das estações automáticas em cada cidade

Cidade	Latitude	Longitude
Piracicaba	$-22,7010^\circ$	$-47,6500^\circ$
Santos	$-23,9810^\circ$	$-46,3000^\circ$

Para a cidade de Piracicaba, as regiões ficaram definidas conforme Figura 3.4. As regiões 1 a 4 incluem as contribuições locais, e as demais são regiões com contribuições de fontes mais distantes da cidade. O centro do quadrado formado pelas regiões 1 a 4 é a coordenada da cidade de Piracicaba.

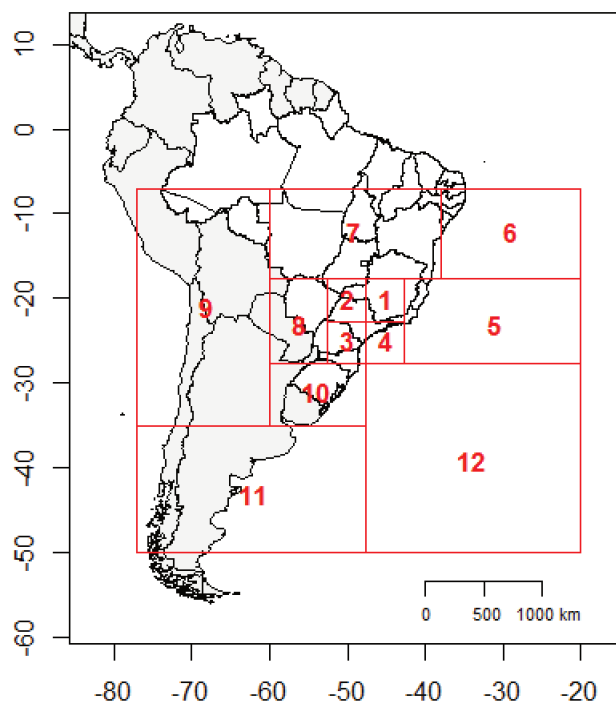


Figura 3.4: Regiões de residência de massa de ar em Piracicaba

O tempo de residência calculado sobre cada célula (ou seja, a quantidade de pontos de trajetória dentro de cada célula) foram somados para cada região pré-definida a fim de determinar o tempo gasto sobre cada região. Assim, para cada dia, foi possível obter quantas vezes a trajetória permaneceu sobre uma das regiões, contando as coordenadas dentro de cada região. Por exemplo, para o dia 01/01/2014, Piracicaba, são 72 pontos de trajetória (excluindo o ponto de chegada) vezes 4 trajetórias por dia, totalizando 288 pontos de trajetórias (coordenadas). Dos 288 pontos de trajetória, 164 pontos ficaram sobre a região 1, 40 sobre a região 2, 16 sobre a região 3 e 68 sobre a região 7, e nenhum ponto sobre as demais regiões. Deste modo, temos o tempo de residência da massa de ar em cada região para cada dia, em todas as cidades.

Após a obtenção, os dados foram analisados a fim de responder os objetivos propostos. Todas as análises e gráficos foram realizadas com o Excel e o software estatístico R, que é um software livre, bastante acessível para diversas análises estatísticas (R Core Team, 2017).

A quantidade de tempo que o ar gasta sobre uma região é linearmente relacionada com a contribuição da região na medição do poluente num local receptor (XU et al., 2006; DIMITRIOU; KASSOMENOS, 2014a; KAVOURAS et al., 2013). Assim, foi utilizada a Análise de Regressão de Trajetória, baseada na estrutura de regressão linear dada como sugere a Equação 2.2, reescrita na Equação 3.1.

$$Y_j = \sum_{k=1}^m B_k X_{jk} \quad (3.1)$$

Onde Y_j é a concentração do poluente atmosférico no j -ésimo dia, X_{jk} é o tempo total de residência da massa de ar na região k no dia j e, finalmente, B_k é o fator de contribuição do poluente na região k . O X_{jk} é obtido pela contagem das coordenadas que permaneceram em cada região em cada dia. O parâmetro B_k é obtido de acordo com o modelo escolhido, que será descrito a seguir.

Neste trabalho considerou-se 5 modelos diferentes para abordagem da ART. Cada um dos modelos, a partir de agora, serão identificados pelo nome como foram sublinhados. Os 5 modelos são os seguintes:

- **Modelo Normal**: é o modelo usual de regressão (WEISBERG, 2005), considerando que a variável resposta (MP_j) pode assumir qualquer valor real, segundo uma distribuição normal. Os trabalhos de Xu et al. (2006), Huang et al. (2010), Chalbot, McElroy e Kavouras (2013), Kavouras et al. (2013), Chalbot et al. (2013), Dimitriou e Kassomenos (2014a) e Rodopoulou et al. (2014), usaram a estrutura do modelo de regressão normal. Neste caso, as estimativas de máxima verossimilhança dos parâmetros B_k são obtidos pelo método dos mínimos quadrados ordinários (WEISBERG, 2005). O modelo Normal é baseado na Equação 3.1.
- **Modelo Normal Inversa**: é uma das possibilidades de ajuste de um modelo linear generalizado (MLG) (CORDEIRO, 1986; MCCULLAGH; NELDER, 1989; PAULA, 2013). Neste caso, considera-se que a variável resposta segue uma distribuição Normal Inversa. Neste modelo, leva-se em consideração que a variável resposta (MP_j) sempre é uma variável positiva (com valores possíveis maiores que zero). Nos MLG as estimativas de máxima verossimilhança dos parâmetros B_k são obtidos pelo método iterativo de Newton-Raphson (PAULA, 2013).
- **Modelo Gama**: similar ao modelo Normal Inversa, porém considerando outra distribuição de probabilidade para a variável resposta. Neste caso a distribuição Gama, que também é uma distribuição que somente assume valores positivos. O método de estimação dos parâmetros B_k é o mesmo do Modelo Normal Inversa.
- **Modelo Log-Normal**: parecido com os modelos Normal Inversa e Gama, entretanto, aqui temos uma distribuição de probabilidade que pertence a classe dos modelos lineares generalizados aditivos para localização, escala e forma ou GAMLSS (*generalized additive model for location, scale and shape*) (RIGBY; STASINOPOULOS, 2005). Nos modelos GAMLSS, as estimativas dos parâmetros B_k foram obtidas pelo algoritmo RS (RIGBY; STASINOPOULOS, 2005; STASINOPOULOS; RIGBY, 2007).

Os modelos Normal Inversa, Gama e Log-normal (os GLM e GAMLSS) tem por característica ajustar uma função da média da distribuição de probabilidade escolhida (no caso, Normal Inversa, Gama ou Log-normal). Neste trabalho, considerou-se ajustar a função identidade, ou seja, a própria média ou valor esperado. Ou seja,

nos modelos GLM e GAMLSS o valor médio da distribuição escolhida (μ_j) é função da estrutura de regressão, baseada na ART, conforme a Equação 3.2.

$$\mu_j = \sum_{k=1}^m B_k X_{jk} \quad (3.2)$$

- **Modelo ARMA(p,q) (Auto-Regressivo de Médias Móveis)**: neste último modelo leva-se em consideração que a variável resposta é uma série temporal (GOYAL; CHAN; JAISWAL, 2006; MORETTIN; TOLOI, 2006). A ideia foi usar a abordagem de Box-Jenkins para modelar a estrutura temporal dos dados, através de um processo ARMA, que não foi levada em consideração nos modelos anteriores. Nos modelos ARMA, as estimativas de máxima verossimilhança dos parâmetros B_k foram obtidas pelo método iterativo BFGS (Broyden, Fletcher, Goldfarb e Shanno) (HONGFANG; RUI, 2010). Em outros termos, este último modelo é uma combinação dos modelos de regressão com um processo ARMA(p,q), ou seja, um modelo auto-regressivo de médias móveis com variáveis explicativas, também chamados de ARMAX (MONTE; ALBUQUERQUE; REISEN, 2015; PINTO; REISEN; MONTE, 2018). O modelo ARMA(p,q), combina a estrutura de regressão da Equação 3.1 com um processo ARMA(p,q), isto é, um processo de ordem p (na parte Auto-Regressiva) e ordem q (na parte de Médias-Móveis), conforme a Equação 3.3.

$$Y_j = \sum_{k=1}^m B_k X_{jk} + ARMA(p, q) \quad (3.3)$$

Onde Y_j é a concentração do poluente atmosférico no j -ésimo dia, X_{jk} é o tempo total de residência da massa de ar na região k no dia j e, finalmente, B_k é o fator de contribuição do poluente na região k . O processo ARMA(p,q) é dado pela Equação 3.4.

$$ARMA(p, q) = \alpha_1 Y_{j-1} + \dots + \alpha_p Y_{j-p} + \epsilon_j + \beta_1 \epsilon_{j-1} + \dots + \beta_q \epsilon_{j-q} \quad (3.4)$$

Detalhes do modelo de regressão combinado com um processo ARMA podem ser encontrados em Goyal, Chan e Jaiswal (2006), Monte, Albuquerque e Reisen (2015), Pinto, Reisen e Monte (2018). O processo ARMA(p,q) deve passar por um processo de identificação. Nesse processo devem ser escolhidos os valores de p e q, usando as funções de autocorrelação e autocorrelação parcial da série diária do MP (MORETTIN; TOLOI, 2006). Também foi considerado o valor do Critério de Informação de Akaike (AIC). Essa medida é utilizada para comparar modelos, sendo que quanto menor seu valor, melhor o modelo. Além do AIC, dado os valores de p e q, os resíduos foram analisados pela função de auto-correlação e o teste de Ljung-Box, a fim de verificar a ausência de auto-correlação dos resíduos (MORETTIN; TOLOI, 2006). Desta forma, foi escolhido uma combinação de valores de p e q, e assim, esse foi o modelo ARMA(p,q) comparado com os demais modelos (Normal, Normal Inversa, Gama e Log-normal).

Com os resultados obtidos, foi possível identificar os locais de origens das fontes de emissão do MP, e também como foi o comportamento do processo de transporte pelas

massas de ar na influência da concentração do MP. Ao se formular as seguintes hipótese nula (H_0) e hipótese alternativa (H_1):

$$\begin{cases} H_0 : B_k = 0 \\ H_1 : B_k \neq 0 \end{cases}$$

é possível testar quais as regiões que são relevantes estatisticamente. Ou seja, caso a H_0 seja rejeitada, podemos afirmar que a região k é significativa estatisticamente, e ela contribui na composição do MP. Caso contrário, afirmamos que a região k não contribui na composição do MP.

Além disso, tendo em vista as cidades escolhidas, poderemos ver a contribuição de fontes locais, regionais, continentais e litorâneas na concentração do MP em diferentes regiões (interior e litoral) do Estado de São Paulo. Tais contribuições calculadas de cada região, refletem a contribuição após a ocorrência de todas as prováveis perdas como deposição úmida, dispersão do vento, diluição química, entre outras (XU et al., 2006; KAVOURAS et al., 2013; DIMITRIOU; KASSOMENOS, 2014a).

Para comparar o desempenho dos 5 modelos, a Correlação de Pearson (CP) e a raiz quadrada do erro quadrático médio (REQM) entre o valor observado e o valor ajustado foram calculados. Também foi considerado o valor do Critério de Informação de Akaike (AIC) e a log-verossimilhança (LOG). Destas medidas, valores menores do REQM e AIC, enquanto que valores maiores de CP e LOG são esperados para os modelos com melhor desempenho.

O desempenho satisfatório de qualquer modelo e sua adequação podem ser avaliados em termos dos resíduos (GOYAL; CHAN; JAISWAL, 2006). A fim de comparação entre os modelos também foi considerada uma análise dos resíduos, considerando os resíduos padronizados (WEISBERG, 2005; GOYAL; CHAN; JAISWAL, 2006; MORETTIN; TOLOI, 2006; PAULA, 2013). Gráfico dos resíduos versus a ordem foram construídos. Se não houver nenhum tipo de padrão nesses resíduos, ou seja, se eles tiverem um comportamento aleatório (gaussiano), podemos concluir que temos um ajuste adequado dos dados.

Outro gráfico dos resíduos utilizados foi o gráfico normal de probabilidade (ou *QQ-plot*), ou ainda, dos resíduos contra os percentis da distribuição Normal Padrão. Nesse gráfico, espera-se que os pontos fiquem próximos a linha de tendência da normal, indicando assim, que não há falta de ajuste no modelo (WEISBERG, 2005; PAULA, 2013).

Além da análise dos resíduos, o modelo da ART foi comparado em termos de previsão. Para isso, a concentração diária e o tempo de residência diário do mês de Janeiro de 2016, foram utilizados para avaliar a previsão do modelo. Desta forma, pode-se analisar o desempenho do modelo, para um conjunto de dados (de Janeiro de 2016) que não foi usado para ajuste do modelo.

Para tentar identificar os locais das possíveis fontes emissoras do MP, foi elaborado o mapa de densidade das trajetórias reversas obtidas no HYSPLIT. Foi usada a densidade Kernel, para visualização da superfície suavizada de todas as coordenadas das trajetórias (MCGOWAN; CLARK, 2008; GE et al., 2016; SALMABADI; SAEEDI, 2018), nas regiões de maior contribuição.

3.5 Análise de Cluster

Em seguida, os dias foram agrupados de acordo com níveis de similaridade das variáveis meteorológicas e contribuição das regiões significativas. A contribuição das regiões foi obtida da ART, a partir do modelo escolhido pelo melhor desempenho.

Foi utilizada a Análise de Cluster Hierárquica Aglomerativa, com o método de Ward, enquanto que a distância euclidiana foi usada para medir similaridade (HAN; KAMBER, 2006; AUSTIN et al., 2012). Nesse tipo de agrupamento, primeiramente cada unidade está em clusters diferentes. Então, as distâncias entre os clusters são calculadas, e o par com a menor distância é agrupado, formando um novo cluster (LU; HE; DONG, 2011; DOMINICK et al., 2012; LATIF et al., 2014). Uma característica deste método é que os resultados de um estágio anterior são sempre aninhados com os resultados de um estágio posterior, apresentando semelhança com a estrutura de uma árvore (JOHNSON; WICHERN, 1988; FAVERO et al., 2009). Essa estrutura é chamada de dendograma, o qual mostra os níveis de similaridade indicado no seu eixo vertical (BASSAB; MIAZAKI; ANDRADE, 1990).

Nessa etapa, a ideia foi identificar (agrupar) dias com comportamentos similares em relação às variáveis do estudo. A AC foi feita baseada nos dados de contribuição de cada região, obtidos pelo modelo escolhido dentre os 5 comparados. As variáveis meteorológicas (temperatura, umidade relativa e velocidade do vento) também foram usadas para obter os grupos. Além disso, foi possível identificar um perfil temporal dos dados do MP, de acordo com os grupos de dias encontrados.

3.6 Análise de Componentes Principais

Depois de agrupados os dias, foi realizada a ACP, dentro de cada grupo (*cluster*). Foi considerada a rotação varimax para produzir o melhor relacionamento entre as CP (Componentes Principais) e as variáveis originais. Neste caso, a rotação garante que cada variável é maximamente correlacionada com apenas uma componente e tem uma associação próxima de zero com as demais. As cargas fatoriais após a rotação são importantes para saber o quanto a variável contribui em particular para a CP e em que medida uma variável é semelhante às outras. Quanto maior o fator de carga da variável, maior é a contribuição dela para a CP em particular (JOHNSON; WICHERN, 1988; STATHEROPOULOS; VASSILIADIS; PAPPA, 1998). Dominick et al. (2012) sugerem que se o fator for maior que 0,75, é considerado uma forte correlação; entre 0,50 e 0,75 classifica-se como moderada; 0,30 a 0,50 como fraca.

Nesta parte da análise, foram quantificadas as relações entre as variáveis meteorológicas e as variáveis poluentes dentro de cada cluster, identificando quais dessas variáveis meteorológicas mais influenciam no MP, e como influenciam.

Capítulo 4

Resultados e Discussão

Este capítulo apresenta os resultados e discussão da análise do MP_{10} e $MP_{2,5}$, divididos em duas seções. A Seção 4.1 mostra os resultados para o MP_{10} e a Seção 4.2 apresenta os resultados para o $MP_{2,5}$. Cada Seção é dividida em duas sub-seções, referentes as duas cidades de estudo, Piracicaba e Santos. Por fim, a Seção 4.3 apresenta uma comparação dos resultados obtidos para as duas cidades.

Alguns gráficos foram inseridos no apêndice, divididos em três capítulos. No Capítulo A estão os gráficos dos resíduos dos modelos Normal, Normal inversa, Gama e Log-normal. Assim como no Capítulo B estão os gráficos das contribuições por região destes mesmos modelos. Por fim, no Capítulo C são apresentados os gráficos dos dendogramas resultantes da AC. Ao longo do texto, as Figuras estão devidamente referenciadas.

4.1 Material Particulado inalável - MP_{10}

4.1.1 Piracicaba

Os resultados são referentes a cidade de Piracicaba, localizada no Interior do Estado de São Paulo, totalizando 730 dias nesse período. De acordo com as trajetórias obtidas pelo modelo HYSPLIT e as características das regiões geográficas, foram definidas as 12 regiões, conforme Figura 4.1 e Tabela 4.1.

Note que as regiões de 1 a 4 são influências locais. A região 4 também tem influência do litoral do Estado de São Paulo. As regiões 5, 6 e 12 são influências litorâneas, sendo respectivamente, do litoral da região Sudeste, da região Nordeste e do Atlântico Sul. Por fim, as regiões 7, 8, 9, 10 e 11 representam as regiões brasileiras e também países sul-americanos.

Com os dados da concentração diária do MP_{10} , em primeiro lugar, escolheu-se os valores de p e q (relativo as ordens do processo AR e MA, respectivamente) da série temporal. Considerou-se diversas combinações, a fim de selecionar apenas uma delas. Conforme Tabela 4.2, o menor valor de AIC foi para o modelo $AR(2)$, ou seja, $p=2$ e $q=0$. Assim, o modelo $AR(2)$ foi escolhido, e comparado com os demais modelos (Normal, Normal Inversa, Gama e Log-normal). Vale ressaltar, que este modelo não apresentou resultados razoáveis para os resíduos, na análise da função de auto-correlação e o teste de Ljung-Box Figura A.1. Porém, dentre os considerados, foi o melhor.

Tabela 4.1: Descrição das regiões de residência de massa de ar em Piracicaba

Região	Descrição	Fonte
1	Interior e sul de Minas Gerais e parte do Rio de Janeiro	Local
2	Interior de São Paulo e Triângulo mineiro	Local
3	Sul de São Paulo, Paraná e Santa Catarina	Local
4	Região metropolitana e Litoral Paulista	Local/Litorânea
5	Espírito Santo e Litoral do Sudeste	Litorânea
6	Litoral do Nordeste	Litorânea
7	Norte/Nordeste/Centro-Oeste	Continental
8	Mato Grosso do Sul e Paraguai	Continental
9	Acre e Rondônia, e Bolívia, Peru, Chile e Argentina	Continental
10	Rio Grande do Sul e Uruguai	Continental
11	Sul da Argentina e litoral argentino	Continental
12	Atlântico Sul	Litorânea

Tabela 4.2: Critério de Seleção para os modelos ARMA(p,q) - MP_{10} - Piracicaba

Modelo	Quantidade de parâmetros	AIC
ARMA(1,0) = AR(1)	1	5805,172
ARMA(2,0) = AR(2)	2	5797,097
ARMA(0,1) = MA(1)	1	6049,080
ARMA(0,2) = MA(2)	2	5916,377
ARMA(1,1)	2	5797,098
ARMA(2,1)	3	5799,071
ARMA(1,2)	3	5799,066
ARMA(2,2)	4	5800,871

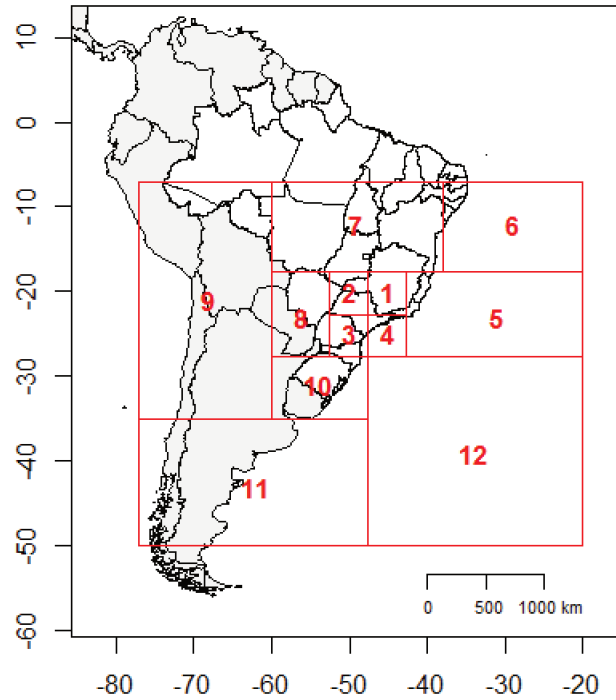


Figura 4.1: Regiões de residência de massa de ar em Piracicaba

Após a escolha do modelo de séries temporais (AR(2)), os cinco modelos foram ajustados e comparados. O ajuste dos modelos levou em consideração a concentração diária do MP_{10} e o tempo de residência por região em cada dia, de acordo com as Equação 3.1, Equação 3.2 e Equação 3.3. A Tabela 4.3 apresenta os valores de CP, REQM, LOG e AIC para todos os modelos. Perceba que o modelo AR(2) apresentou as melhores medidas. O valor da CP (0,779) ficou acima do dobro dos valores obtidos nos demais modelos. O REQM, LOG e AIC também foram melhores no modelo AR(2). Isso demonstra que o modelo AR(2) tem um desempenho de predição eficiente e satisfatório. Ao se comparar somente os modelos Normal, Normal inversa, Gama e Log-normal, percebe-se que em relação ao CP, eles estão muito próximos. Em relação ao REQM, o modelo Log-normal ficou um pouco acima dos demais. Porém, ao se analisar o LOG e AIC, o modelo Normal inversa apresenta os melhores resultados.

As Figuras A.2, A.3, A.4, A.5 e 4.2 mostram os gráficos dos resíduos padronizados ao longo do tempo. Destes gráficos, nota-se que os quatro primeiros parecem ter algum tipo de padrão no seu comportamento, formando como se fossem ondas senoidais. Esse comportamento pode estar relacionado com o período de atuação de fenômenos meteorológicos. Por exemplo, no último trimestre de 2014 (a partir do dia 270 até o dia 365 na série) houve a atuação do anticiclone tropical do Atlântico Sul (CETESB, 2014), que influenciou no comportamento dos resíduos (Figuras A.2, A.3, A.4 e A.5). O gráfico dos resíduos do modelo AR(2) (Figura 4.2) parecem ser mais aleatórios ao redor do zero (comportamento gaussiano), apesar de apresentar alguns valores mais distantes (aberrantes). O modelo AR(2), ao considerar o comportamento da série temporal, também considera a atuação dos fenômenos meteorológicos.

Tabela 4.3: Medidas comparativas - MP_{10} - Piracicaba

Modelo	CP	REQM	LOG	AIC
Normal	0,360	18,683	-3172,964	6371,928
Normal Inversa	0,353	18,733	-3015,714	6057,428
Gama	0,357	18,702	-3040,156	6106,312
Log-Normal	0,356	19,133	-3017,358	6060,715
AR(2)	0,779	12,560	-2883,549	5797,097

Para finalizar a análise dos resíduos, as Figuras A.6, A.7, A.8, A.9 e 4.3 apresentam uma comparação dos resíduos padronizados com a distribuição Normal Padrão, isto é, distribuição normal com média zero e variância 1. Esse gráfico é conhecido como sendo o gráfico normal de probabilidades, e quanto mais próximos da reta, melhor o ajuste do modelo (WEISBERG, 2005; PAULA, 2013). Observe que para o modelo normal Inversa os resíduos ficaram bem próximos da reta, assim como para os modelos Gama e Log-normal. Os resíduos do modelo Normal foram os que apresentaram pior desempenho. Já os resíduos do modelo AR(2) se distanciaram da reta, principalmente nos extremos, ou seja, nos valores mais baixos ou mais altos dos resíduos.

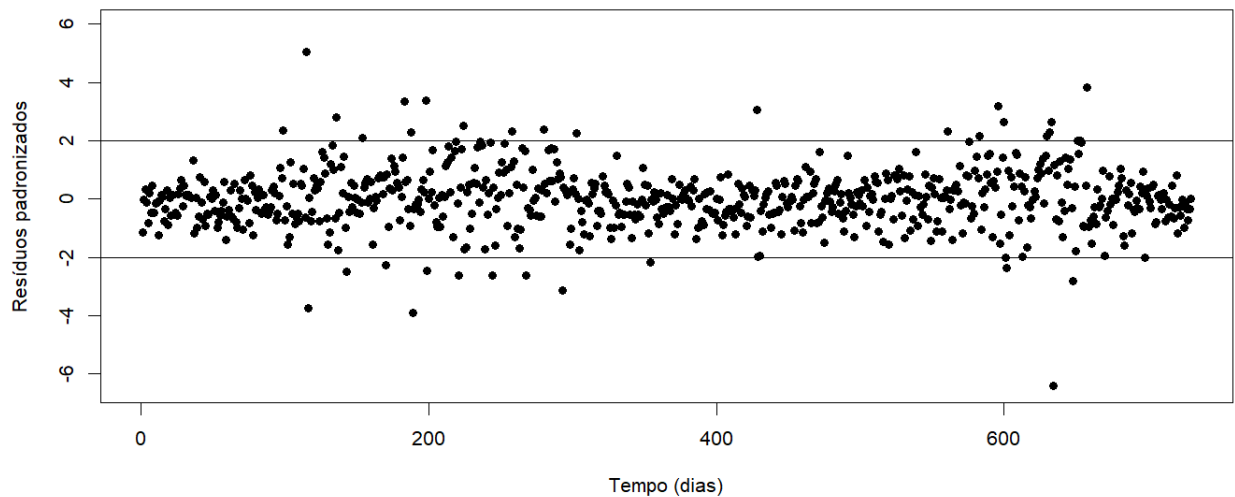


Figura 4.2: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo AR(2) - Piracicaba

Pela análise das medidas comparativas e a análise dos resíduos pode-se notar que existem diferenças entre o ajuste de cada um dos modelos. Aparentemente, o modelo AR(2) consegue capturar um comportamento relevante, que é a dependência temporal dos dados. Os resultados da Tabela 4.3 e Figura 4.2 apontam nesse sentido. Porém, esse modelo não considera a característica da variável resposta, que é um valor sempre positivo, (concentração do MP_{10}). Por considerar essa característica, o modelo Normal inversa apresenta um comportamento mais adequado dos resíduos na Figura A.7.

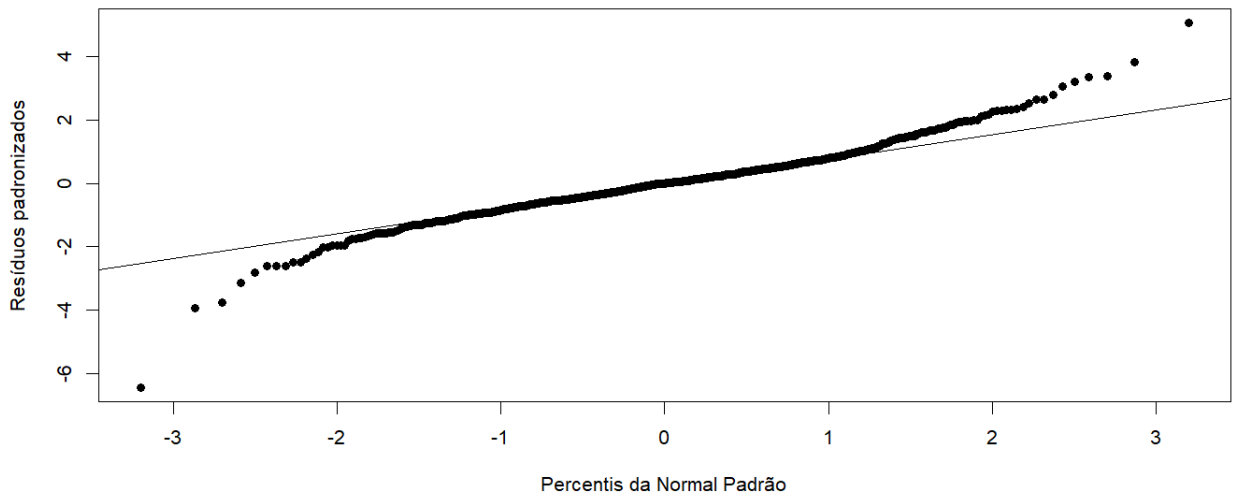


Figura 4.3: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo AR(2) - Piracicaba

Finalmente, as Figuras B.1, B.2, B.3, B.4 e 4.4 mostram a proporção de contribuição de cada região pré-determinada anteriormente para os modelos Normal, Normal inversa, Gama, Log-normal e AR(2), respectivamente. Em relação à proporção de contribuição obtida em cada um dos modelos houve algumas pequenas diferenças nos valores estimados. Uma diferença notada está nas regiões que são significativas estatisticamente. Observa-se que as cores da barra indicam se a região é significativa ou não a um nível de 5%. Se a região é significativa, tem-se que essa região influencia na composição da concentração diária do MP_{10} , caso contrário, não há contribuição significativa da região. Assim, nota-se que no ajuste dos três primeiros modelos, as regiões 9 e 11 não foram significativas, enquanto que para o ajuste do modelo AR(2), as regiões 9, 10 e 11 não foram significativas. No modelo Log-normal, todas as regiões foram significativas estatisticamente. Essa diferença no resultado do ajuste dos modelos pode ser mais um indicativo das diferenças entre as abordagens adotadas.

Como visto até agora, os modelos Normal inversa e AR(2) foram os que se mostraram melhores. O modelo Normal inversa é interessante por considerar a característica da concentração do MP_{10} como sendo sempre um valor maior que zero, e o modelo AR(2) é interessante por considerar a dependência temporal das observações. Vamos analisar a contribuição média estimada por região nesses dois modelos, olhando as Figuras B.2, 4.4 e a Figura 4.5. Na Figura 4.5 são comparados os dois modelos escolhidos, sendo que, se a região tem uma cor mais quente, maior a contribuição, e se a cor é mais fria, menor é a contribuição.

Considerando as 3 regiões de menor contribuição (9, 10 e 11) elas representam apenas 1,3% e 0,5% para o ajuste dos modelos Normal inversa e AR(2), respectivamente. Além disso, no ajuste do modelo Normal inversa, a região 10 (que é significativa) representa apenas 1% de contribuição na composição diária do MP_{10} em Piracicaba. Assim, os Estados do Rio Grande do Sul, Acre e Rondônia, além dos países sul-americanos Chile,

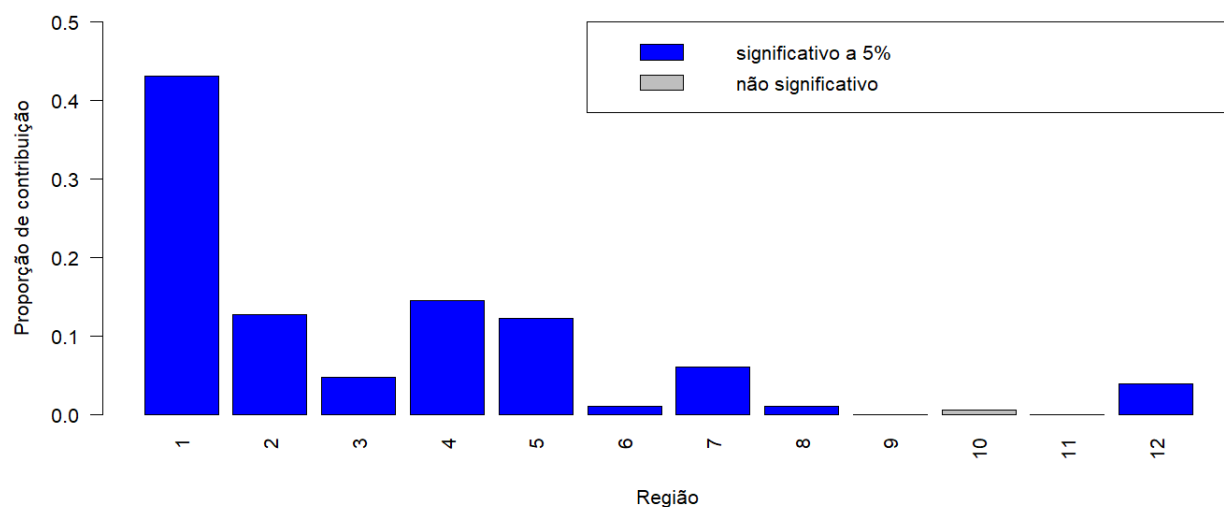


Figura 4.4: Contribuição média do MP_{10} por área - Modelo AR(2) - Piracicaba

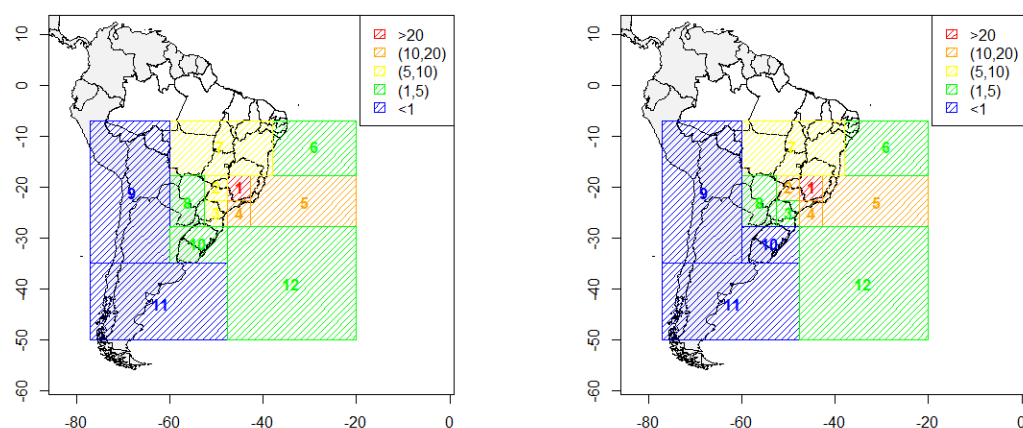


Figura 4.5: Mapa de contribuição das regiões no modelo Normal inversa e AR(2), respectivamente - Piracicaba - MP_{10}

Peru, Uruguai e Argentina pouco contribuem na concentração do MP_{10} para a cidade de Piracicaba. Ressaltando que essa região foi não significativa para o modelo AR(2).

As maiores contribuições ocorreram das regiões locais (Regiões 1 a 4), com 75% e 75,7%, respectivamente para o modelo Normal inversa e AR(2). Sendo que a Região 1 (Interior e Sul de Minas Gerais e parte do Rio de Janeiro) é a de maior contribuição (47% para o modelo Normal inversa e 43% para o modelo AR(2)). Dentre as fontes locais, a região 3 foi a de menor contribuição, com 5,6% para o modelo Normal inversa e 4,7% para o modelo AR(2).

As regiões 1 a 4 representam grande parcela da região sudeste e sul do Brasil, conhecidas como as regiões mais desenvolvidas do país, com um grande número de indústrias e alto

tráfego de veículos automotores, que contribuem significativamente na concentração do MP_{10} . As regiões 5, 6 e 12 são formadas principalmente pelo litoral. Juntas representam 17,5% e 21,3% da contribuição do MP_{10} para os modelos Normal inversa e AR(2), respectivamente. Dessas proporções, a região 5 tem a maior representatividade, com 10,4% e 12,2%, respectivamente em cada modelo. Essa região é composta, principalmente, pelo litoral do Rio de Janeiro e Espírito Santo.

Por fim, as regiões 7 e 8 representam as contribuições de fontes regionais e continentais, totalizando 8,0% e 7,1% da contribuição do MP_{10} para os ajustes dos modelos Normal inversa e AR(2), respectivamente. Sozinha, a região 7 representa 5,1% da contribuição no modelo Normal inversa e 6,0% no modelo AR(2).

Resumindo, as fontes locais foram as que, principalmente, contribuíram na concentração diária do MP_{10} , em seguida das fontes litorâneas e por fim das fontes regionais e continentais.

A Figura 4.6 apresenta o mapa de densidade das trajetórias das três regiões de maior contribuição, regiões 1, 2 e 4. Neste mapa, as potenciais rotas das trajetórias que passaram nessas regiões foram calculadas usando uma densidade kernel (MCGOWAN; CLARK, 2008; GE et al., 2016; SALMABADI; SAEEDI, 2018). Com isso, obteve-se uma superfície suavizada dos pontos de coordenadas das trajetórias de massa de ar que chegaram em Piracicaba, onde o vermelho indica uma alta densidade, e o azul, uma baixa densidade. Quanto maior a densidade, mais tempo a massa de ar permaneceu naquela área. Note que, as trajetórias parecem predominantemente chegar da direção nordeste de Piracicaba, desde Belo Horizonte, passando por Sul de Minas, e com as mais altas densidade na região de Campinas, Araras, Araraquara e Jundiaí (Figura 4.6). Essas regiões se caracterizam por intensa atividade industrial e alto desenvolvimento agrícola. Há também várias áreas em que são realizadas queima de palha de cana-de-açúcar (GONÇALVES et al., 2017; CETESB, 2018). Além disso, importantes rodovias, como Anhanguera e Bandeirantes, estão localizadas nessas áreas de maior densidade, onde o transporte de carga, de veículos pesados a Diesel, correspondem a maior parte da emissão do MP (MMA, 2014). Por fim, esta região de alta densidade possui as cidades de Paulínia (com enorme parque industrial petroquímico) e Santa Gertrudes e Rio Claro, formando importante polo cerâmico (CETESB, 2018).

Ainda comparando os dois modelos (Normal Inversa e AR(2)), temos as Figuras 4.7 e 4.8, que mostram os valores observados diariamente da concentração do MP_{10} nos anos de 2014 e 2015, respectivamente. Nestas Figuras, são comparadas os valores reais (observados) com os ajustes diários do modelo Normal inversa e AR(2). Observa-se que o ajuste do modelo Normal inversa, não consegue acompanhar a série, permanecendo próximo a uma média geral dos dados. Diferente do ajuste do modelo AR(2), em que nota-se um acompanhamento bem próximo da série real. Os valores de picos, ou valores muito baixos da concentração só são observados no modelo AR(2). Perceba que apenas o modelo AR(2) aponta a maioria dos dias acima do limite estabelecidos pela OMS.

Quando se considera a contribuição média geral, os modelos ficaram bem próximos, conforme visto nas Figuras B.1, B.2, B.3, B.4 e 4.4. Porém, quando se observa diariamente, o modelo AR(2) é melhor que os demais. Apesar disso, percebe-se que há uma defasagem

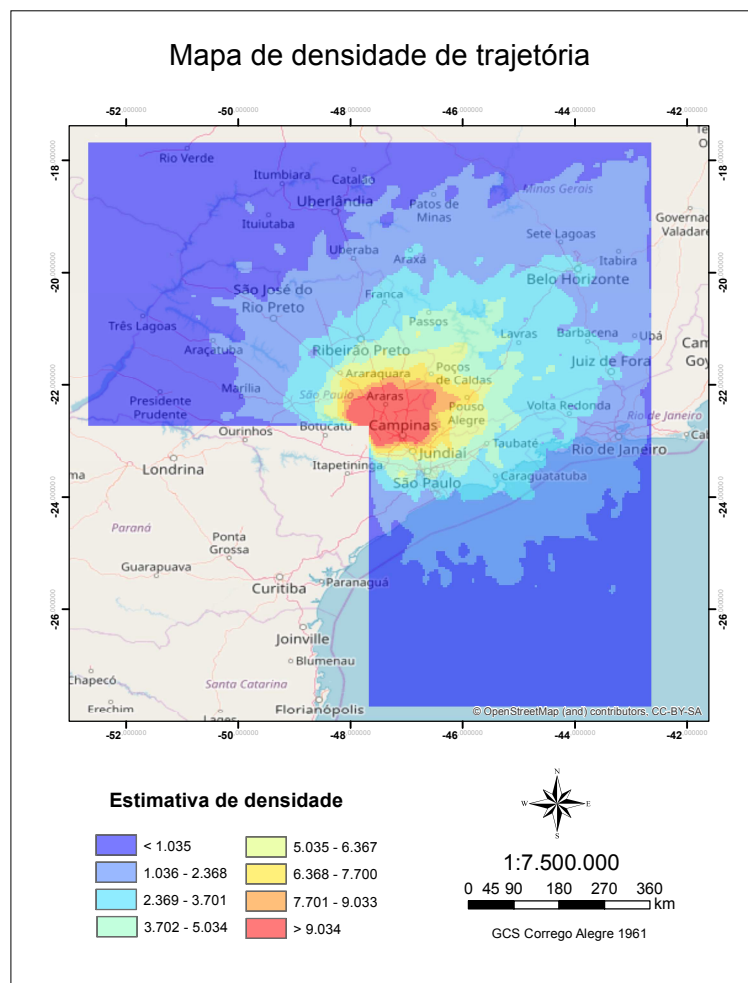


Figura 4.6: Mapa de densidade das trajetórias *backward* das regiões 1, 2 e 4 - Piracicaba

no ajuste deste modelo. Aparentemente, o pico visto em determinado dia, somente é ajustado no dia seguinte pelo modelo, indicando uma defasagem no ajuste.

Os dados de concentração do MP_{10} e os tempos de residência de cada região do mês de Janeiro de 2016 foram usados para avaliar o poder de previsão do modelo AR(2). Note que na Figura 4.9 são mostrados os valores previstos pelo modelo, e os valores reais da concentração do MP_{10} . De modo geral, a previsão superestimou a concentração ao longo do mês.

Deste modo, conforme o discutido, o modelo AR(2) é a melhor escolha dentre os cinco modelos comparados. Com o ajuste deste modelo, prosseguiremos com a AC e a ACP.

A AC foi realizada considerando as variáveis meteorológicas (Temperatura, Umidade Relativa e Velocidade do Vento) e a proporção de contribuição diária do MP_{10} por região (somente as regiões significativas 1-8 e 12). Todos os 730 dias do período (anos 2014 e 2015) de medição foram divididos em 4 grupos, baseados na análise do dendograma

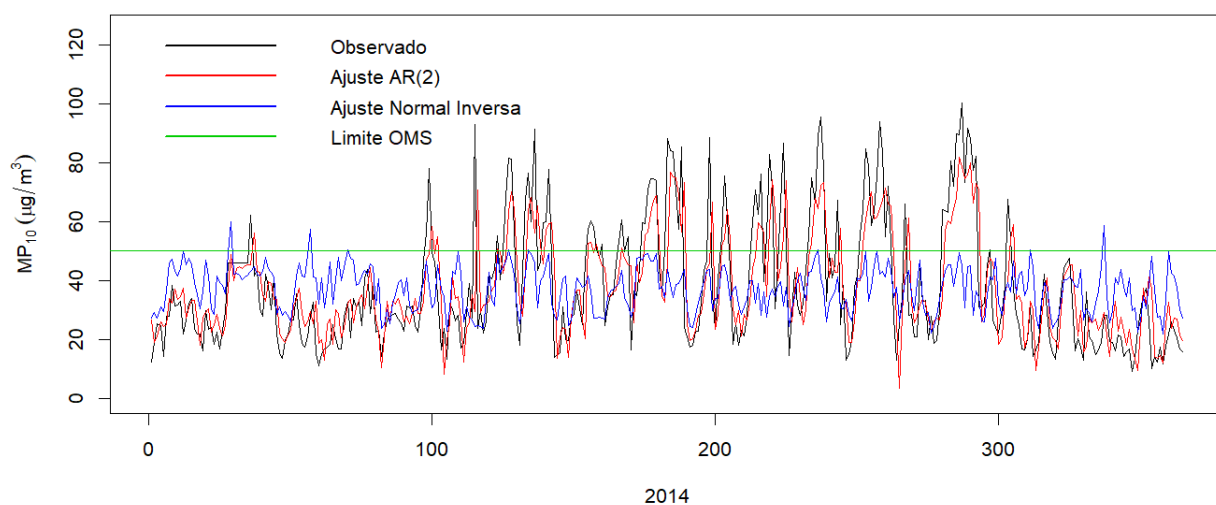


Figura 4.7: Ajuste da concentração diária do MP_{10} em 2014 em comparação com valores observados - Piracicaba

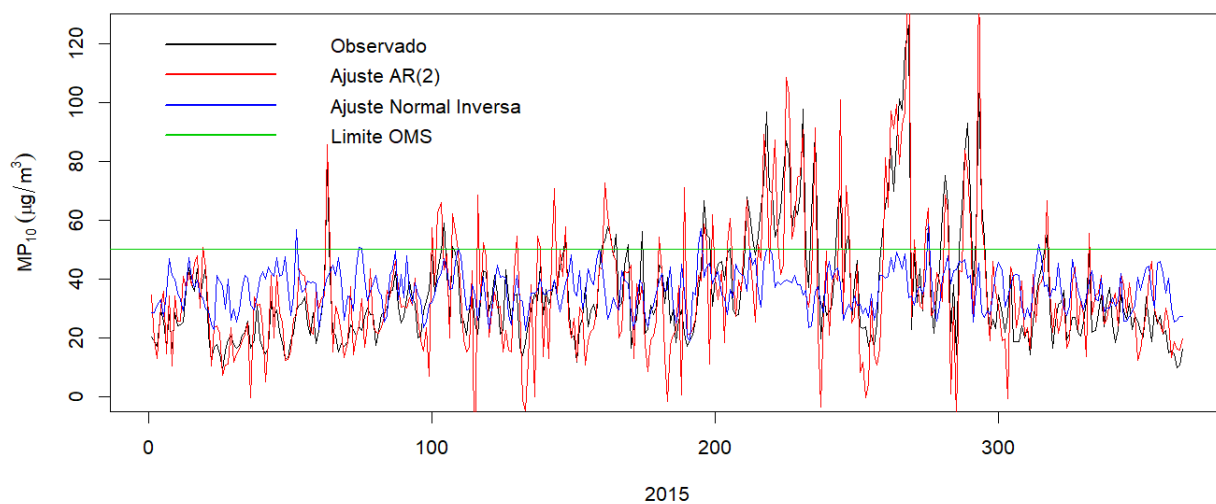


Figura 4.8: Ajuste da concentração diária do MP_{10} em 2015 em comparação com valores observados - Piracicaba

(Figura C.1). O grupo 1 foi formado por 111 dias, o grupo 2 por 228 dias, o grupo 3 por 219 dias e, por fim, o grupo 4 foi formado por 172 dias.

A Tabela 4.4 mostra os percentuais de dias dentro de cada estação do ano em cada um dos quatro grupos formados na AC. Na divisão em cada estação do ano, considerou-se os meses de Dezembro, Janeiro e Fevereiro, como a estação de verão. Os meses de Março, Abril e Maio, como outono. Os meses de Junho, Julho e Agosto como inverno. E por fim, os meses de Setembro, Outubro e Novembro como primavera (IAG, 2018).

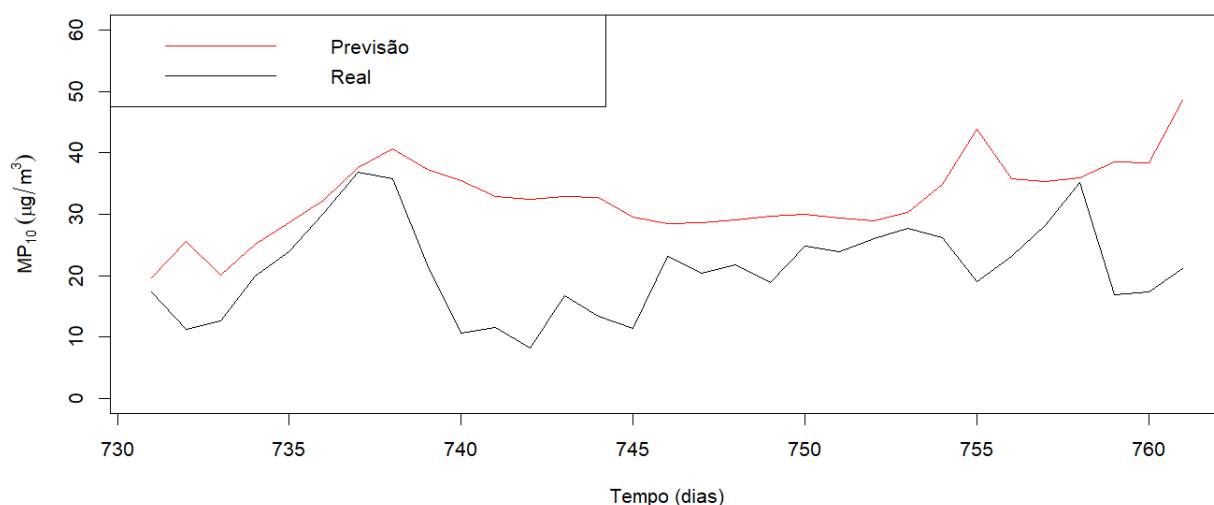


Figura 4.9: Previsão do MP_{10} e valores reais para o mês de Janeiro de 2016 - Piracicaba

A Tabela 4.5 apresenta os percentuais de cada grupo em cada uma das regiões que contribuíram na composição dos níveis do MP_{10} em Piracicaba.

Tabela 4.4: Percentual de dias dentro de cada estação do ano por grupo (MP_{10} - Piracicaba)

	Primavera	Verão	Outono	Inverno
Grupo 1	26,1%	33,3%	23,4%	17,1%
Grupo 2	27,2%	19,7%	22,8%	30,3%
Grupo 3	24,2%	23,3%	23,7%	28,8%
Grupo 4	22,1%	27,3%	31,4%	19,2%

As Figuras 4.10, 4.11, 4.12 e 4.13 mostram a distribuição dos dados, por meio dos box-plots, da concentração do MP_{10} , e das variáveis meteorológicas (Temperatura, Umidade relativa e Velocidade do vento), respectivamente, divididos nos 4 grupos.

A Tabela 4.6 mostra a proporção da variância explicada por cada CP (Componente Principal) na ACP. Esta análise foi feita dentro de cada grupo obtido na AC, considerando as variáveis meteorológicas e a concentração do MP_{10} . Perceba que, considerando apenas 2 CPs, mais de 70% da variância é explicada em todos os grupos. Sendo que, nos grupos 3 e 4, mais de 78% da variabilidade total é explicada por 2 componentes. As Figuras 4.14, 4.15, 4.16 e 4.17 apresentam os coeficientes da ACP após a rotação varimax. Quanto maior seu valor absoluto, maior sua importância dentro da CP.

Nas Figuras 4.14, 4.15, 4.16 e 4.17 as variáveis foram nomeadas de outra forma para facilitar a visualização no gráfico de barras. Assim, Temp, UR e VV, correspondem a Temperatura, Umidade relativa e Velocidade do Vento respectivamente.

Tabela 4.5: Percentual de contribuição do grupo dentro de cada região de residência da massa de ar (MP_{10} - Piracicaba)

	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Região 1	4,4%	10,1%	43,4%	42,2%
Região 2	59,4%	14,6%	4,3%	21,6%
Região 3	25,1%	62,7%	2,5%	9,7%
Região 4	3,8%	68,5%	13,1%	14,6%
Região 5	1,1%	15,0%	74,9%	9,0%
Região 6	7,5%	23,1%	42,0%	27,4%
Região 7	38,8%	18,6%	15,6%	27,1%
Região 8	46,6%	52,5%	0,2%	0,6%
Região 12	0,2%	74,4%	24,2%	1,2%

Tabela 4.6: Resumo da ACP - MP_{10} - Piracicaba

Grupo 1	CP1	CP2	CP3	CP4
Proporção da variabilidade	0,429	0,299	0,199	0,073
Proporção acumulada	0,429	0,728	0,927	1,000
Grupo 2				
Proporção da variabilidade	0,404	0,309	0,209	0,078
Proporção acumulada	0,404	0,713	0,922	1,000
Grupo 3				
Proporção da variabilidade	0,522	0,265	0,154	0,059
Proporção acumulada	0,522	0,787	0,941	1,000
Grupo 4				
Proporção da variabilidade	0,414	0,366	0,160	0,059
Proporção acumulada	0,414	0,781	0,941	1,000

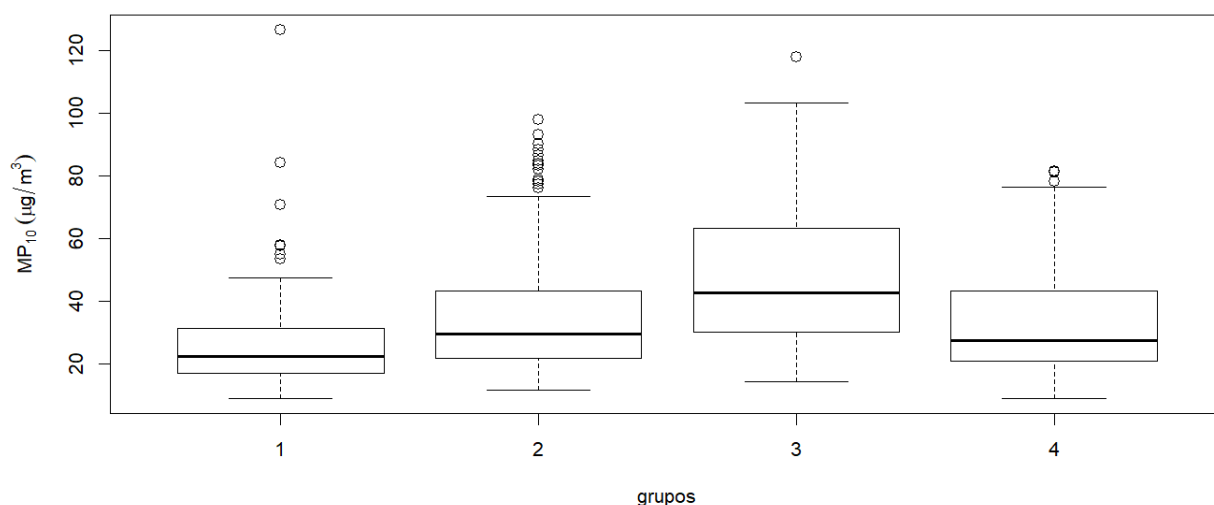


Figura 4.10: Box-plot do MP_{10} ($\mu\text{g}/\text{m}^3$) por grupo - MP_{10} - Piracicaba

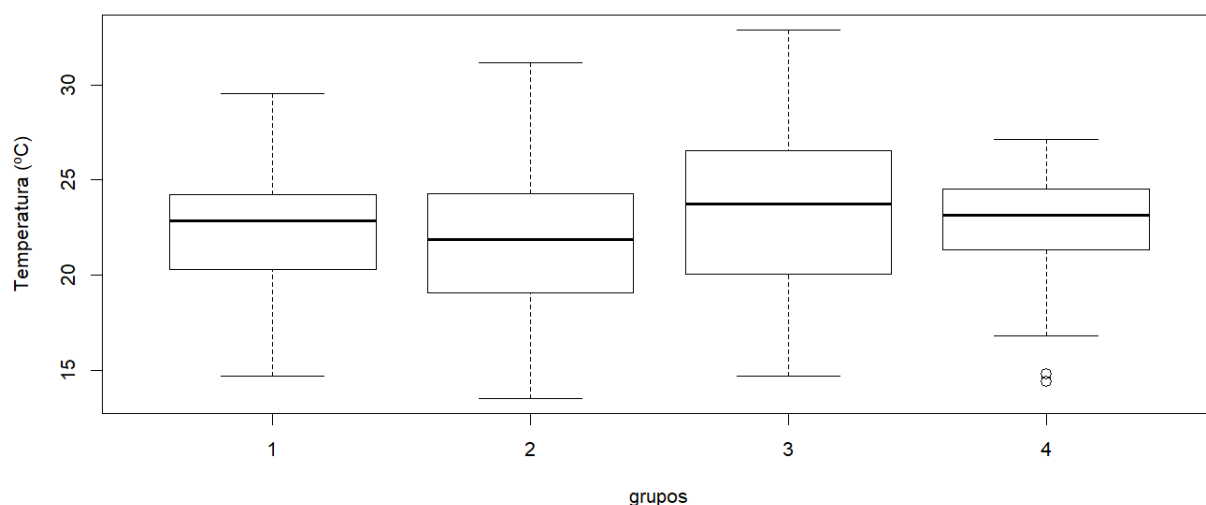


Figura 4.11: Box-plot da Temperatura ($^{\circ}\text{C}$) por grupo - MP_{10} - Piracicaba

O grupo 1 é o menor grupo, formado por 111 dias. Neste grupo, os dias são predominantemente da estação do verão (Tabela 4.4) e ocorre os maiores valores de Umidade Relativa, o que facilita a “limpeza” da atmosfera, pois o nível mediano de MP_{10} é menor dentre todos os grupos (Figuras 4.10 e 4.12). Note que o percentual de umidade relativa é próximo a 80% neste grupo. Essa relação da Umidade Relativa e MP_{10} é inversamente proporcional, como mostrada na Figura 4.14, na CP1. Esta componente explica 42,9% da variabilidade total do grupo 1 (Tabela 4.6). A CP2 explica um percentual de 29,9%, onde a relação predominante é da Velocidade do Vento com a Temperatura. As regiões 2, 7 e 8 são as que tem as maiores contribuições nesse grupo (Tabela 4.5), portanto com contribuições desde o Oeste ao Norte de Piracicaba, ou mais detalhado, do Interior de

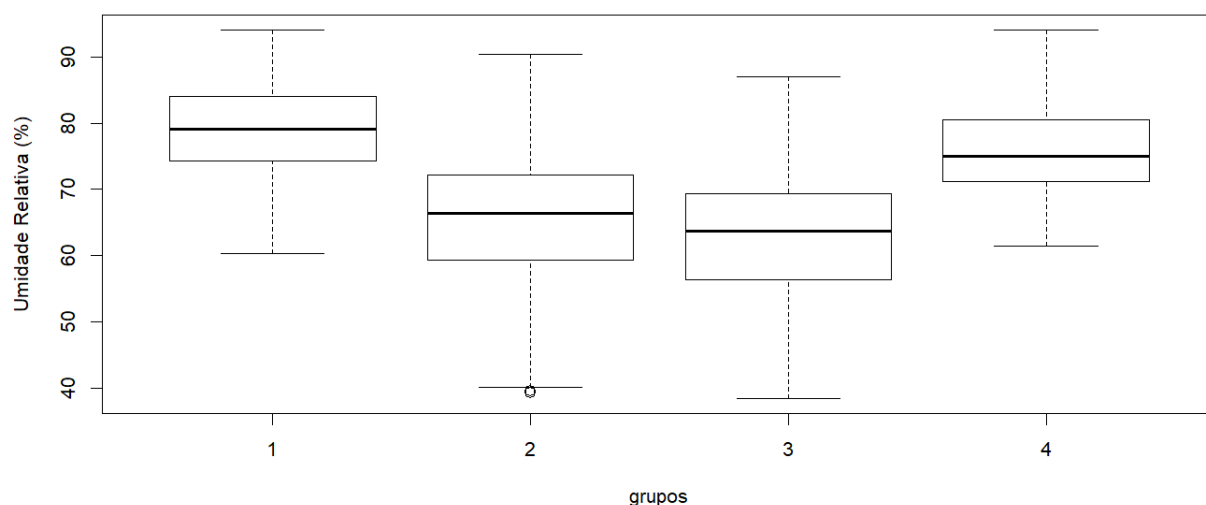


Figura 4.12: Box-plot da Umidade Relativa (%) por grupo - MP_{10} - Piracicaba

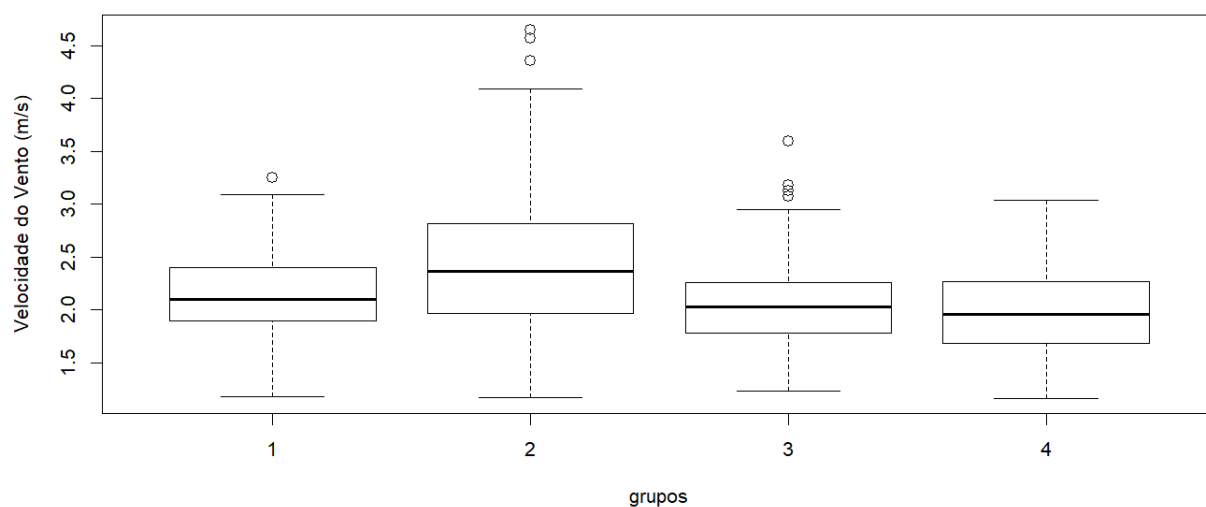


Figura 4.13: Box-plot da Velocidade do Vento (m/s) por grupo - MP_{10} - Piracicaba

São Paulo, Triângulo mineiro, Norte, Nordeste e Centro-oeste do Brasil, além do Mato Grosso do Sul e Paraguai (Figura 4.1 e Tabela 4.1).

O grupo 2 é o maior grupo, com 228 dias. O segundo maior valor mediano da concentração do MP_{10} ocorre neste grupo (Figura 4.10). Este grupo tem predominância dos dias no inverno (Tabela 4.4), com os menores valores na temperatura e maiores velocidade do vento (Figuras 4.11 e 4.13, respectivamente). A relação mais forte (vista na CP1) é da Umidade Relativa com a Temperatura (inversamente proporcional), representando 40,4% (Tabela 4.6 e Figura 4.15). Enquanto que, na CP2, a velocidade do vento atua dispersando o MP, ou seja, conforme aumenta-se a velocidade do vento, diminui a concentração do MP_{10} (Figura 4.15). As regiões 3, 4, 8 e 12 foram as de maiores contribuições nesse

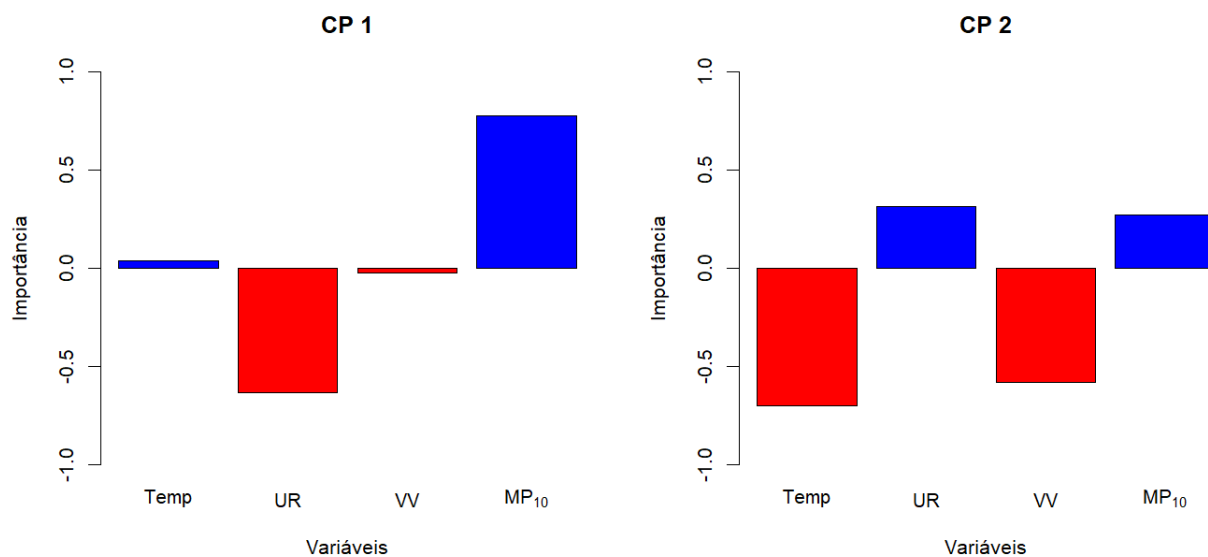


Figura 4.14: Importância das CP 1 e CP 2 no grupo 1 - MP_{10} - Piracicaba

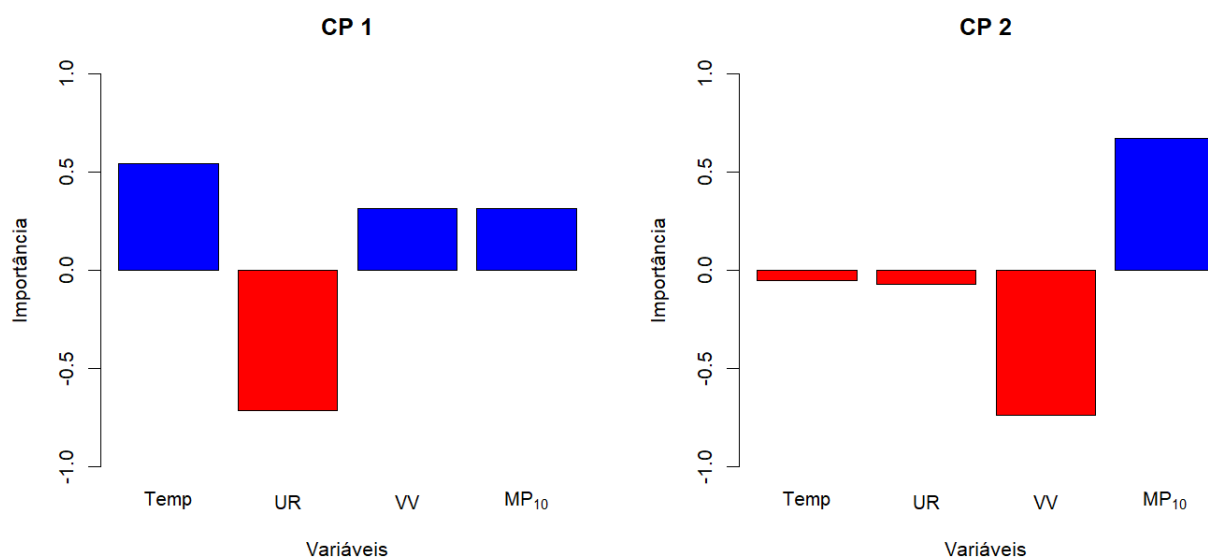


Figura 4.15: Importância das CP 1 e CP 2 no grupo 2 - MP_{10} - Piracicaba

grupo (Tabela 4.5). Essas regiões estão localizadas mais ao sul de Piracicaba, incluindo contribuição litorânea, do Atlântico Sul.

Note que, em relação ao nível mediano de MP_{10} , o grupo 3 foi o que apresentou maior valor (Figura 4.10). Este grupo, formado por 219 dias, sendo a maioria no período de inverno (Tabela 4.4), o que aumenta a concentração do MP_{10} , devido ao tempo seco, isto é, com baixa umidade relativa, próximas a 60% (Figura 4.12). A ACP identificou a Umidade Relativa como o principal fator para os altos valores do MP_{10} neste grupo (Figura 4.16, CP1) com 52,2% da variabilidade explicada por esta relação. As regiões 1, 5 e 6 mais contribuem, formadas pelo interior de Minas Gerais e parte do Rio de Janeiro,

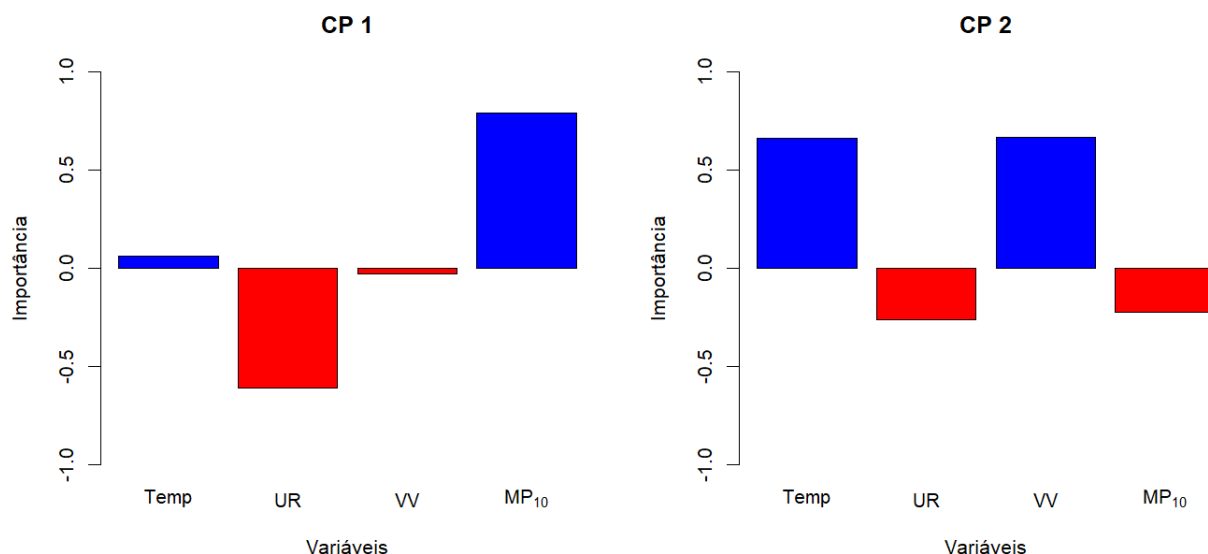


Figura 4.16: Importância das CP 1 e CP 2 no grupo 3 - MP_{10} - Piracicaba

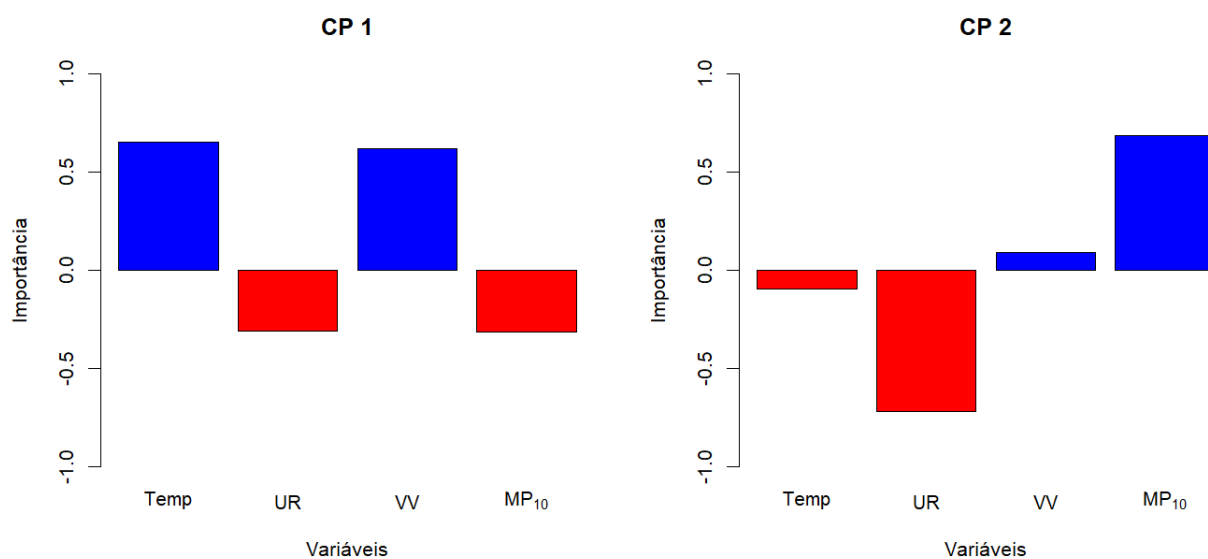


Figura 4.17: Importância das CP 1 e CP 2 no grupo 4 - MP_{10} - Piracicaba

Espírito Santo, além do litoral das regiões Sudeste e Nordeste. São regiões desde o Norte, até o Leste de Piracicaba.

Por fim, o grupo 4 é formado por 172 dias. A estação do ano predominante é o outono, seguida do verão (Tabela 4.4), com umidade relativa alta (Figura 4.12) e níveis do MP_{10} não tão altos (Figura 4.10). Como no verão ocorrem as maiores temperaturas, esta variável foi a mais importante neste grupo, juntamente com a velocidade do vento, explicando 41,4% da variância (Tabela 4.6 e Figura 4.17). 36,6% da variabilidade é explicada pela relação inversamente proporcional entre MP_{10} e Umidade Relativa. A região que prevalece é a 1, relativa ao interior de Minas Gerais e parte do Rio de Janeiro, com contribuições de fontes locais. Esta foi a região de maior contribuição (Tabela 4.1).

De modo geral, pode-se perceber a AC agrupou regiões próximas dentro de um mesmo grupo. Esse fato indica que a correlação espacial dos dados é um fator relevante nos níveis do MP_{10} . Além disso, as variáveis meteorológicas tiveram comportamento diferente nos grupos (bem evidenciado nos resultados da ACP), também indicando um fator importante na concentração do MP_{10} . A ACP mostrou que variável meteorológica UR foi a de maior importância na relação com MP_{10} .

4.1.2 Santos

Nesta seção, os resultados são da cidade de Santos, localizada no Litoral do Estado de São Paulo. Doze regiões (Figura 4.18 e Tabela 4.7) foram definidas, baseadas no alcance das trajetórias obtidas no modelo HYSPLIT, e também nas características das regiões, similarmente ao ocorrido em Piracicaba.

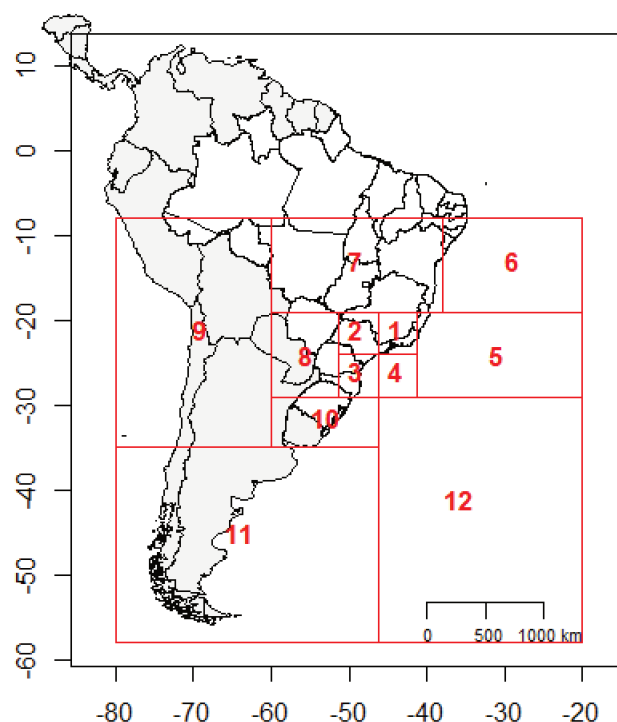


Figura 4.18: Regiões de residência de massa de ar em Santos

Destas doze regiões, as regiões 1 a 4 ficam ao redor da cidade de Santos, sendo cada uma numa área $5^\circ \times 5^\circ$. As regiões 1 e 2 são consideradas locais, e as regiões 3 e 4 são locais, mas também tem razoável parte formada pelo litoral. As regiões 5, 6, 11 e 12 são, na sua maioria, formadas pelo litoral. As regiões 7, 8, 9 e 10 são, principalmente, regiões continentais.

Os modelos Normal, Normal inversa, Gama e Log-normal já tem os parâmetros definidos, baseados nas Equação 3.1 e Equação 3.2. Entretanto, o modelo ARMA(p,q), precisa num primeiro momento, definir as ordens do processo auto-regressivo (p) e do processo de médias-móveis (q). A Tabela 4.8 apresenta os valores do critério de informação AIC para algumas possibilidades do modelo ARMA. Note que o menor valor do AIC é para

Tabela 4.7: Descrição das regiões de residência de massa de ar em Santos

Região	Descrição	Fonte
1	Litoral Norte paulista, Sul de Minas Gerais e Rio de Janeiro	Local
2	Estado de São Paulo (interior e capital)	Local
3	Litoral Sul de São Paulo, Leste e Litoral do Paraná e Santa Catarina	Local/Litorânea
4	Litoral Paulista	Local/Litorânea
5	Espírito Santo e Litoral do Sudeste	Litorânea
6	Litoral do Nordeste	Litorânea
7	Norte/Nordeste/Centro-Oeste	Continental
8	Leste do Paraná e Santa Catarina, Mato Grosso do Sul e Paraguai	Continental
9	Acre e Rondônia, e Bolívia, Peru, Chile e Argentina	Continental
10	Rio Grande do Sul, litoral gaúcho e Uruguai	Continental
11	Argentina e litoral argentino	Continental
12	Atlântico Sul	Litorânea

o modelo ARMA(2,2). Além disso, esse modelo apresentou os melhores resultados em relação a função de auto-correlação e o teste de Ljung-Box dos resíduos Figura A.10.

Tabela 4.8: Critério de Seleção para os modelos ARMA(p,q) - MP_{10} - Santos

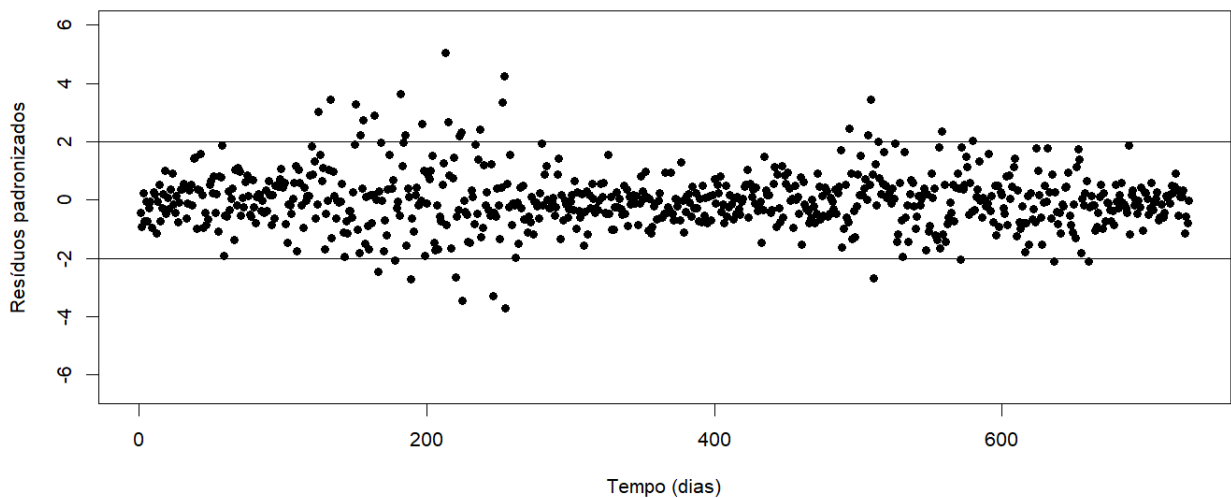
Modelo	Quantidade de parâmetros	AIC
ARMA(1,0) = AR(1)	1	6359,584
ARMA(2,0) = AR(2)	2	6361,580
ARMA(0,1) = MA(1)	1	6418,892
ARMA(0,2) = MA(2)	2	6382,230
ARMA(1,1)	2	6361,579
ARMA(2,1)	3	6362,647
ARMA(1,2)	3	6334,571
ARMA(2,2)	4	6326,696

Deste modo, o modelo ARMA(2,2) foi escolhido, dentre os modelos de séries temporais, e foi comparado com os modelos Normal, Normal Inversa, Gama e Log-normal. A Tabela 4.9 mostra os valores de CP, REQM, LOG e AIC para os cinco modelos para ajuste diário da concentração do MP_{10} em Santos. O modelo ARMA(2,2) apresentou os melhores resultados dentre as quatro medidas comparativas, ou seja, maiores valores para CP e LOG, e menores valores de REQM e AIC. Ao comparar os outros quatro modelos, nota-se que o modelo Normal é melhor em relação ao CP e REQM, enquanto que o modelo Log-normal é melhor em relação ao LOG e AIC. A diferença entre os dois modelos é que o segundo, considera o comportamento da variável resposta (concentração do MP) ser um valor sempre positivo, diferente do modelo Normal.

Tabela 4.9: Medidas comparativas - MP_{10} - Santos

Modelo	CP	REQM	LOG	AIC
Normal	0,441	22,299	-3302,132	6630,264
Normal Inversa	0,427	22,487	-3166,214	6358,427
Gama	0,436	22,362	-3154,478	6334,956
Log-Normal	0,430	22,901	-3145,307	6316,614
ARMA(2,2)	0,689	18,002	-3146,348	6326,696

As Figuras A.11, A.12, A.13, A.14 e 4.19 mostram os resíduos padronizados pela ordem das observações. Espera-se um comportamento aleatório dos resíduos ao longo do tempo. Os resíduos dos modelos Normal (Figura A.11), Normal inversa (Figura A.12), Gama (Figura A.13) e Log-normal (Figura A.14) parecem apresentar uma tendência em forma de “ondas”, diferente do comportamento dos resíduos na Figura 4.19, do modelo ARMA, que se apresentam de forma mais aleatória (gaussiana) dentro dos limites -2 a 2. Esta característica, indica um melhor ajuste dos dados observados com os dados modelados.

Figura 4.19: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo ARMA(2,2) - Santos

Além dos gráficos dos resíduos pela ordem, também são apresentados os gráficos normal de probabilidade para os resíduos, nas Figuras A.15, A.16, A.17, A.18 e 4.20. Nestes gráficos é feita a comparação dos resíduos padronizados com os percentis da distribuição normal padrão. Os resíduos dos modelos Gama, Normal Inversa e Log-normal ficaram bem próximos da reta, indicando uma adequação da distribuição proposta para a variável resposta (MP_{10}). Enquanto os resíduos dos modelos Normal e ARMA(2,2) ficaram mais dispersos em relação a reta, principalmente nos extremos.

Mais uma evidência de que há diferença nas abordagens entre os cinco modelos foram as contribuições médias de cada um, mostrados nas Figuras B.5, B.6, B.7, B.8 e 4.21, res-

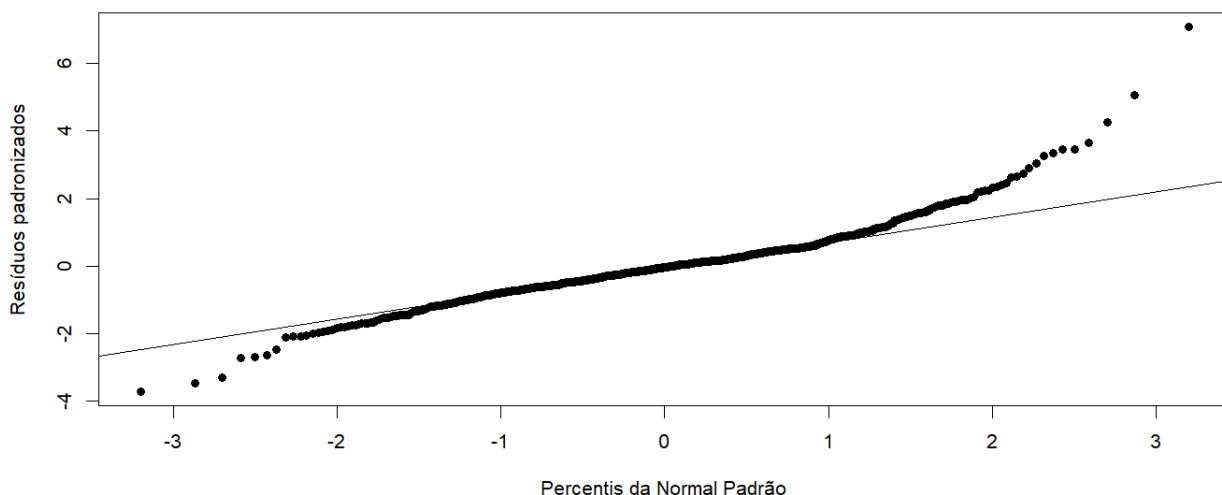


Figura 4.20: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo ARMA(2,2) - Santos

pectivamente para os modelos Normal, Normal inversa, Gama, Log-normal e ARMA(2,2). As proporções de contribuição de cada região e as regiões significativas estatisticamente foram diferentes entre os modelos. No modelo Normal e Log-normal, todas as regiões foram significativas. Nos modelos Normal inversa e Gama, a região 9 não foi significativa. E por fim, no modelo ARMA(2,2), a região 11 não foi significativa a 5% de significância. Neste caso, os modelos lineares generalizados (Normal inversa e Gama), que tem suposições bem similares (exceto pela distribuição de probabilidade), tiveram resultados iguais em termos de regiões significativas.

Baseado na análise dos resíduos, e nas medidas comparativas (Tabela 4.9), o modelo ARMA(2,2) apresentou melhores resultados. Na sequência, o modelo Log-normal foi o que apresentou medidas comparativas melhores (LOG e AIC), e resíduos comparado com percentis da normal padrão (Figura A.18) mais adequados. Desta forma, os modelos ARMA(2,2) e Log-normal foram comparados na contribuição da concentração média do MP_{10} dos anos de 2014 e 2015 para a cidade de Santos. Na Figura 4.22 são comparados os dois modelos escolhidos (Log-normal e ARMA(2,2)). Se a região está preenchida com a cor vermelha, há maior contribuição, e se a região está preenchida com a cor azul, menor é a contribuição da região nos níveis do MP_{10} .

A região de maior contribuição nos dois modelos foi a região 1, com 35,3% (modelo Log-normal) e 40,9% (modelo ARMA(2,2)). Essa região é referente ao Sul de Minas Gerais, Litoral Norte de São Paulo e o Estado do Rio de Janeiro. A segunda região de maior contribuição foi a 4, com 15,5% (modelo Log-normal) e a região 2, com 15,4% (modelo ARMA(2,2)). A região 4 contribuiu com 10,5% no modelo ARMA(2,2), e a região 2 contribuiu com 12,3% no modelo Log-normal. A região 2 corresponde a influência de praticamente todo o Estado de São Paulo, interior e região metropolitana. A região 4 evidencia exclusivamente influência litorânea na concentração do MP_{10} (Tabela 4.7 e Figura 4.18). Dentre as quatro regiões de influência local, a região 3 foi a de menor

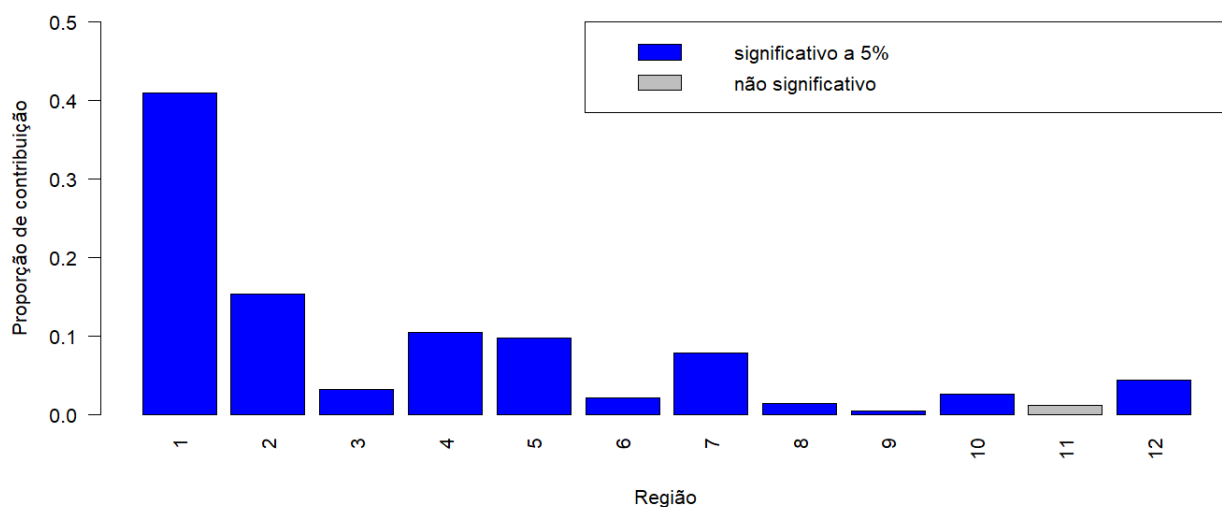


Figura 4.21: Contribuição média do MP_{10} por área - Modelo ARMA(2,2) - Santos

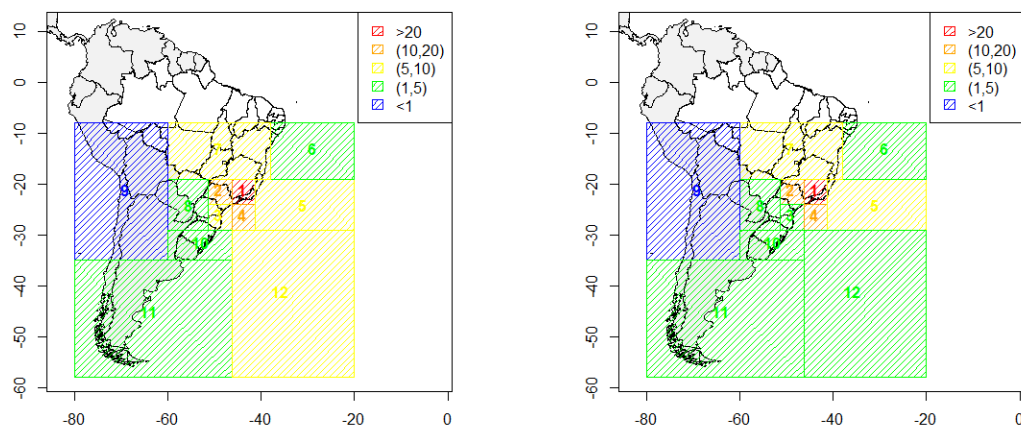


Figura 4.22: Mapa de contribuição das regiões no modelo Log-Normal e ARMA(2,2), respectivamente - Santos - MP_{10}

percentual. Essa região reflete influência de massa de ar dos estados do Paraná e Santa Catarina, e seus litorais, além do Sul do Estado de São Paulo. O percentual de contribuição foi de 5,5% e 3,3% para os modelos Log-normal e ARMA(2,2), respectivamente. De modo geral, as regiões com fontes locais corresponderam a 68,5% de contribuição no modelo Log-normal, e 70,1% no modelo ARMA(2,2). Dentro dessa contribuição total, vale destacar o porto de Santos, que tem importante relevância nas atividades da região da baixada santista, além da intensa movimentação de caminhões no local (CETESB, 2015a).

As regiões 5, 6 e 12 tem em suas áreas praticamente correspondentes ao litoral do Atlântico. No modelo Log-normal, a contribuição total dessas áreas foi de 16,4%, enquanto

que no modelo ARMA(2,2) foi de 16,3%. O litoral da região Sudeste do Brasil (área 5) foi a de maior contribuição, com 9,0% (modelo Log-normal) e 9,7% (modelo ARMA). Em seguida, a região 12 (litoral do Atlântico Sul) corresponde a pouco mais de 4% no modelo ARMA(2,2) e quase 6% no modelo Log-normal. E o Litoral Nordestino, área 6, foi a de menor contribuição dentre essas regiões, com 1,4% e 2,1% nos modelos Log-normal e ARMA(2,2), respectivamente.

A região de menor contribuição foi a 9 (referente aos estados do Acre e Rondônia, além da Bolívia, norte da Argentina e Chile, e sul do Peru. Sua contribuição foi menor que 1% nos dois modelos. Na sequência, a região 11 também teve pouca contribuição, sendo que não foi significativa no modelo ARMA(2,2). No modelo Log-normal, sua contribuição foi de 1,9%.

A estimativa de contribuição na concentração média diária da região 10 (Estado do Rio Grande do Sul e Litoral, Uruguai e o litoral) foi de 3,0% e 2,6% nos modelos Log-normal e ARMA(2,2), respectivamente.

Finalmente, as regiões 7 e 8 correspondem a maior parte por influências de massas de ar continentais, tendo contribuído com 10% e 9,3% nos modelos Log-normal e ARMA(2,2). Dessas duas, a região 7 teve relevante contribuição, com 7,4% (modelo Log-normal) e 7,8% (modelo ARMA(2,2)). A região 7 corresponde, principalmente, as áreas dos Estados da Bahia, Goiás, Distrito Federal, Mato Grosso, Tocantins (maior parte, do centro ao Sul do Estado) e Norte de Minas Gerais. A região 8 (Paraguai, Mato Grosso do Sul, e oeste do Paraná e Santa Catarina) teve uma menor contribuição, sendo a estimativa no modelo Log-normal de 2,6% e no modelo ARMA(2,2) foi de 1,5%.

A Figura 4.23 apresenta o mapa de densidade das trajetórias das duas regiões de maior contribuição, regiões 1 e 2. O mapa foi feito com os pontos de coordenadas das trajetórias de massa de ar usando uma densidade kernel (MCGOWAN; CLARK, 2008; GE et al., 2016; SALMABADI; SAEEDI, 2018). O resultado da densidade é uma superfície suavizada das coordenadas das trajetórias *backward* em Santos.

Note que, as trajetórias parecem predominantemente vir da direção nordeste de Santos, desde as cidades mineiras de Juiz de Fora e Varginha, passando por Sul do Rio de Janeiro, incluindo contribuição da Região metropolitana de Campinas. As mais altas densidades se localizam no leste e sul da região metropolitana de São Paulo, e também região de São José dos Campos, e grande parte da baixada Santista (Figura 4.23). Essas regiões se caracterizam por intensa atividade industrial, principalmente por incluir o polo industrial de Cubatão, com empresas do setor siderúrgico, petroquímico e de fertilizantes (CETESB, 2015a). A circulação de caminhões em direção ao polo industrial de Cubatão e ao porto de Santos, provocam ressuspensão de material pela ação dos ventos, e pela própria movimentação dos caminhões, que também emitem particulados devido a queima do diesel (CETESB, 2016b; CETESB, 2018). Há também emissões provenientes de movimentação de grãos (milho e soja principalmente) devido as atividades portuárias (CETESB, 2015a; CETESB, 2018). Além disso, emissões veiculares também impactam nos níveis do MP_{10} em Santos, devido ao intenso tráfego veicular (CETESB, 2015a). Por fim, e dada a proximidade com o mar, o aerossol marinho também contribui para a composição do MP em Santos (CETESB, 2015a).

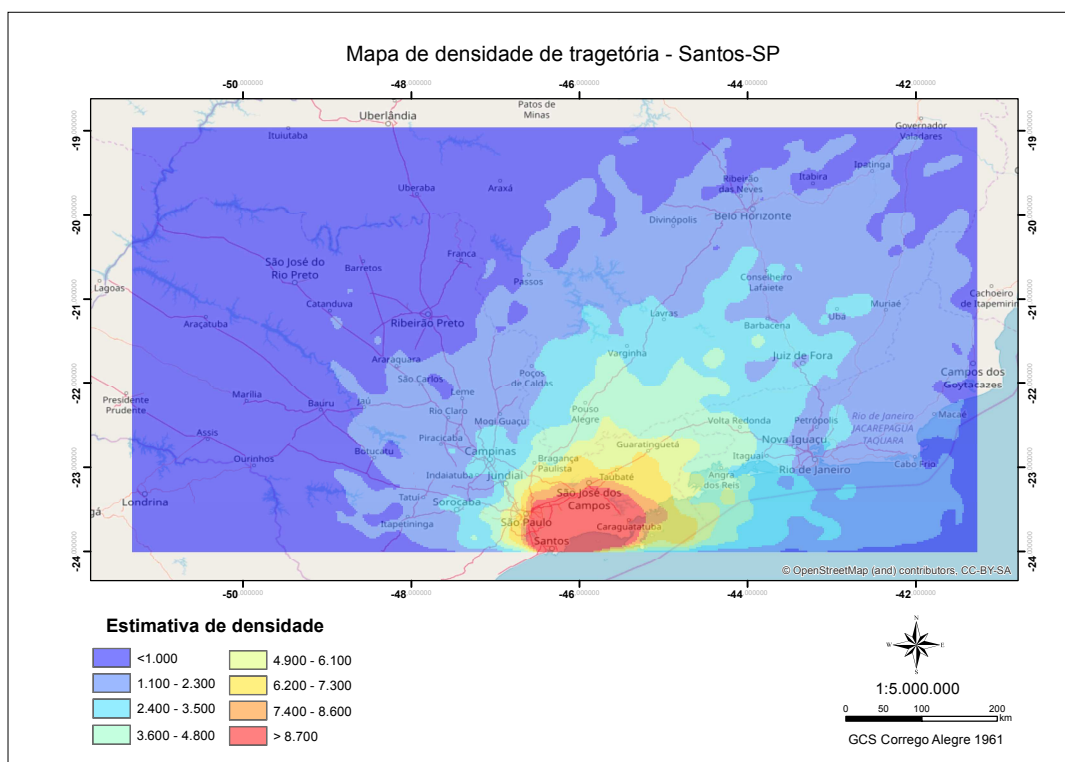


Figura 4.23: Mapa de densidade das trajetórias *backward* das regiões 1 e 2 - Santos

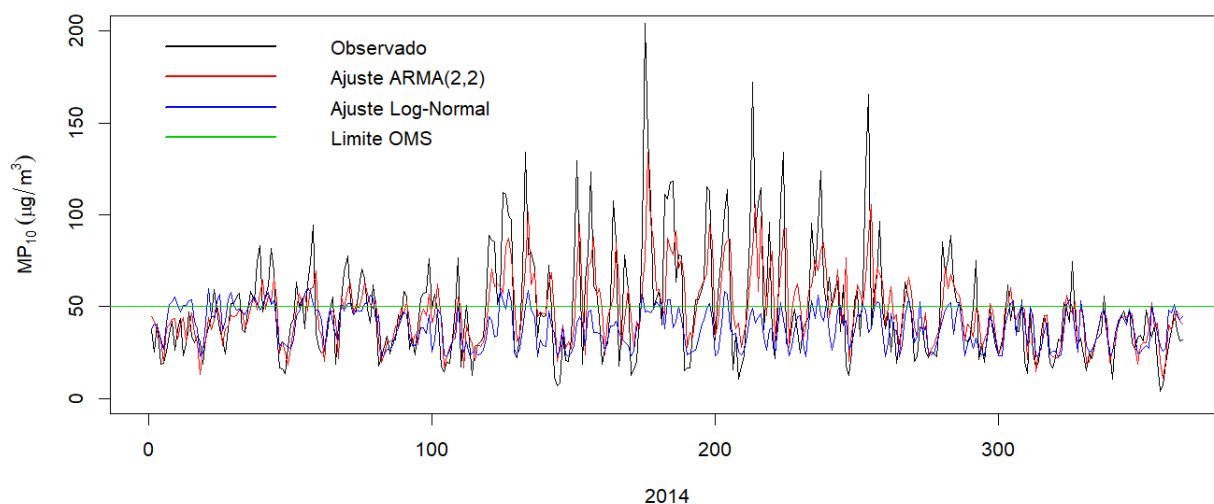


Figura 4.24: Ajuste da concentração diária do MP_{10} em 2014 em comparação com valores observados - Santos

Com os dados do MP_{10} e os tempos de residência de cada região do mês de Janeiro de 2016, em Santos, o poder de previsão do modelo ARMA(2,2) foi avaliado. Na Figura 4.26 são mostrados os valores previstos pelo modelo, e os valores reais da concentração do MP_{10} em Santos. Neste caso, as concentrações previstas ficaram bem próximas as concentrações reais, indicando uma previsão satisfatória.

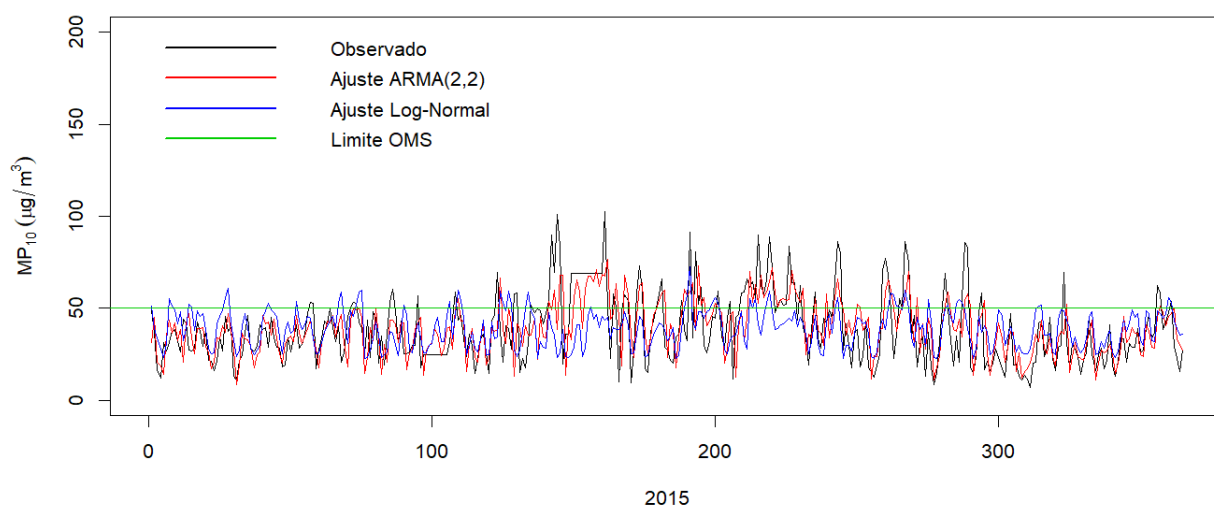


Figura 4.25: Ajuste da concentração diária do MP_{10} em 2015 em comparação com valores observados - Santos

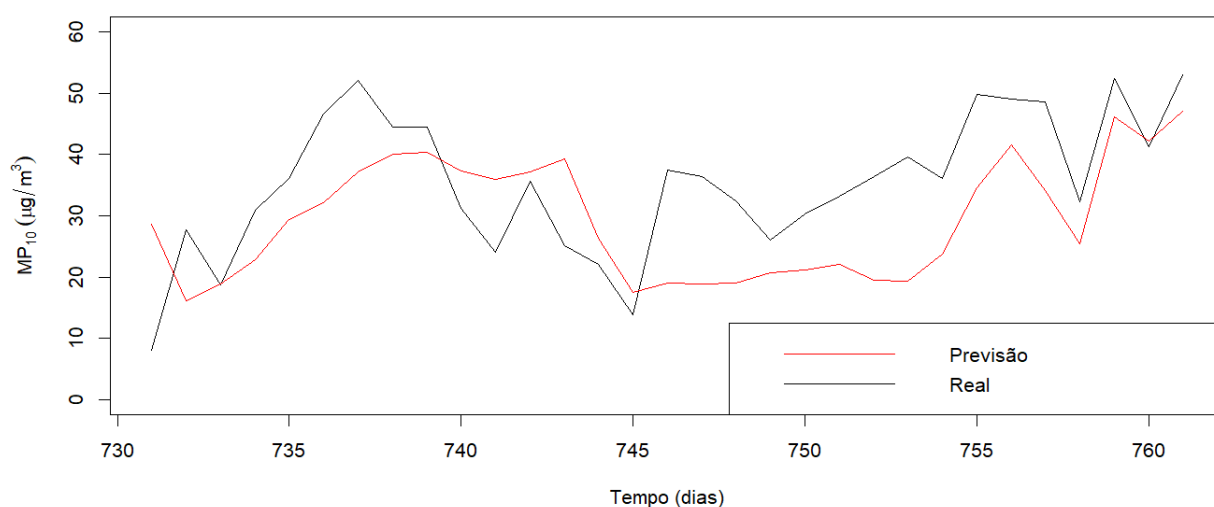


Figura 4.26: Previsão do MP_{10} e valores reais para o mês de Janeiro de 2016 - Santos

A contribuição média durante o período foi bem similar na comparação dos dois modelos. As maiores diferenças foram nas regiões 1, 2 e 4. Nas demais, as estimativas do percentual de contribuição nos níveis do MP_{10} foi bem pequeno. Conclui-se que, na média, os modelos não apresentam diferenças nas estimativas. Porém, as Figuras 4.24 e 4.25 apresentam o ajuste diário da concentração do MP_{10} nos modelos Log-normal e ARMA(2,2) em comparação com os valores diários observados nos anos de 2014 e 2015. Os valores ajustados pelo modelo ARMA(2,2), conseguem capturar a dependência temporal dos dados, acompanhando a série dos valores observados. Diferente do modelo Log-normal que apenas se ajusta em torno de uma média geral dos dados. Ou seja, quando há elevadas ou

baixas concentrações diárias do MP_{10} , apenas o modelo ARMA identifica esses pontos. A linha verde representa os valores do Limite estabelecido da OMS para a média de 24 horas consecutiva. Nota-se que, em alguns dias no início de cada ano, o modelo Log-normal estima uma concentração acima do limite, quando na verdade a concentração do MP_{10} está abaixo. Deste modo, dados as evidências apresentadas, o modelo ARMA(2,2) é o melhor dentre os cinco modelos propostos. A AC e ACP foi feita com os dados ajustados deste modelo.

Na AC, os valores das estimativas diárias da proporção de contribuição de cada região (exceto a região 11, que não foi significativa), e os valores observados das variáveis meteorológicas, temperatura, umidade relativa e velocidade do vento foram usadas para obter os grupos. Os dias foram divididos em quatro grupos distintos (Figura C.3), sendo o grupo 1 formado por 115 dias, o grupo 2 por 257 dias, o grupo 3 por 97 dias, e o grupo 4 por 261 dias.

A Tabela 4.10 e a Tabela 4.11 apresentam o percentual de dias dentro de cada grupo, em cada estação do ano, e o percentual de cada grupo em cada uma das regiões que contribuíram na composição da concentração do MP_{10} em Santos, respectivamente.

Tabela 4.10: Percentual de dias dentro de cada estação do ano por grupo (MP_{10} - Santos)

	Primavera	Verão	Outono	Inverno
Grupo 1	26,1%	24,3%	16,5%	33,0%
Grupo 2	28,8%	16,3%	29,2%	25,7%
Grupo 3	21,6%	20,6%	30,9%	26,8%
Grupo 4	21,8%	34,5%	23,0%	20,7%

Tabela 4.11: Percentual de contribuição do grupo dentro de cada região de residência da massa de ar (MP_{10} - Santos)

	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Região 1	5,4%	11,1%	12,9%	70,6%
Região 2	67,4%	9,5%	0,8%	22,3%
Região 3	21,7%	67,4%	1,3%	9,6%
Região 4	2,1%	71,9%	10,6%	15,5%
Região 5	0,8%	10,7%	62,7%	25,8%
Região 6	4,2%	0,1%	6,3%	89,4%
Região 7	42,4%	5,5%	0,8%	51,3%
Região 8	38,6%	59,7%	0,0%	1,7%
Região 9	3,4%	96,6%	0,0%	0,0%
Região 10	3,3%	93,4%	0,4%	2,9%
Região 12	0,3%	80,4%	16,9%	2,4%

Nas Figuras 4.27, 4.28, 4.29 e 4.30 estão elaborados os gráficos de box-plot por grupo, para cada uma das variáveis meteorológicas (temperatura, umidade relativa e velocidade do vento) mais a concentração do MP_{10} .

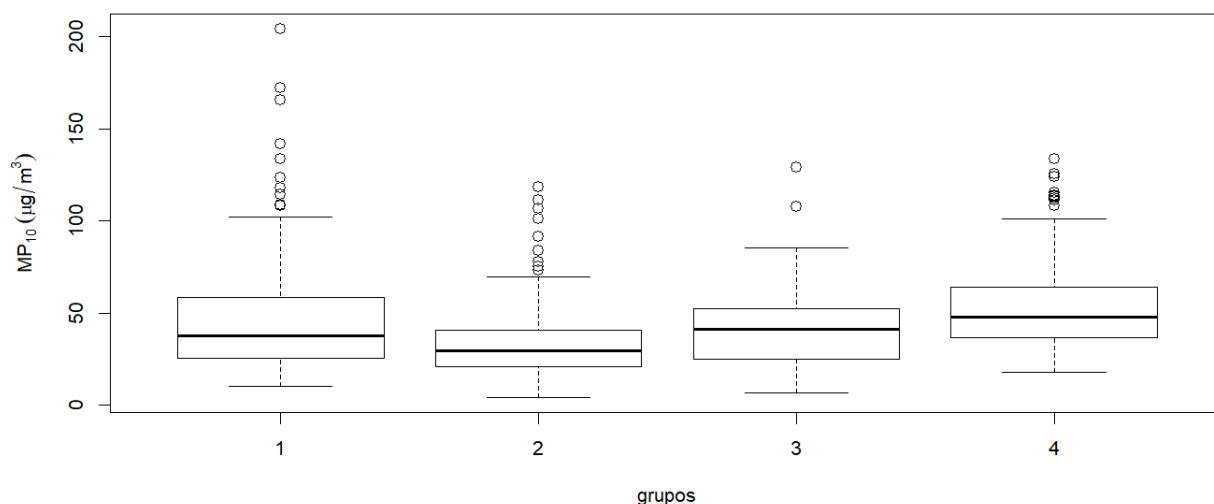


Figura 4.27: Box-plot do MP_{10} ($\mu g/m^3$) por grupo - MP_{10} - Santos

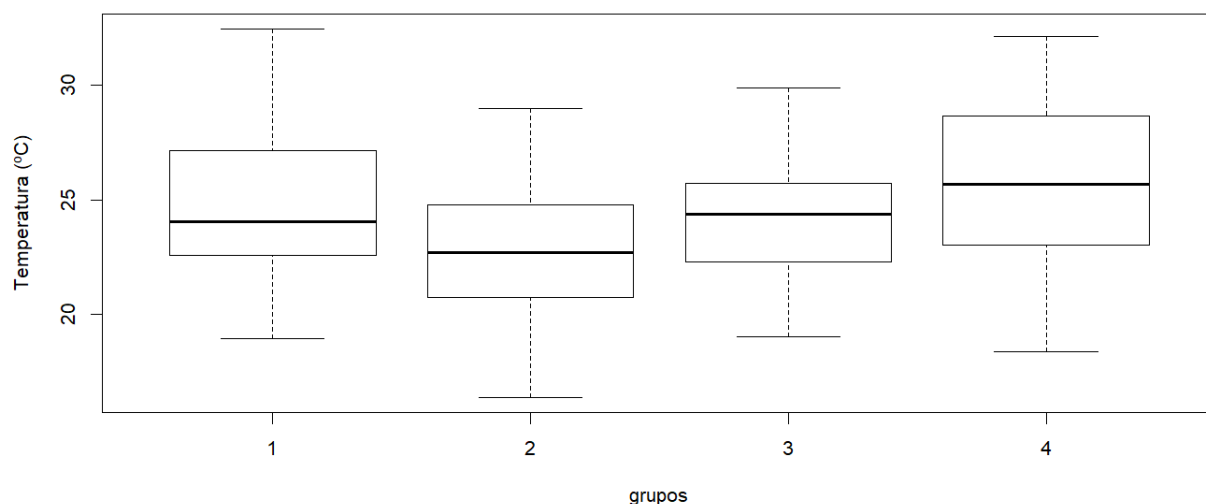


Figura 4.28: Box-plot da Temperatura ($^{\circ}C$) por grupo - MP_{10} - Santos

A Tabela 4.12 exibe a proporção da variabilidade explicada na ACP, para cada CP, e também a proporção acumulada, em cada grupo de dias, obtidos da AC. Na ACP foram considerados a concentração do MP_{10} e as medições diárias das variáveis meteorológicas (3 no total), por isso no máximo são quatro CPs. Nos grupos 1 e 4, escolhendo 2 CPs, temos um acúmulo de mais de 0,80 da variabilidade explicada. Nos grupos 2 e 3, as duas primeiras CPs retêm mais de 0,70 da proporção da variabilidade. Assim, a interpretação

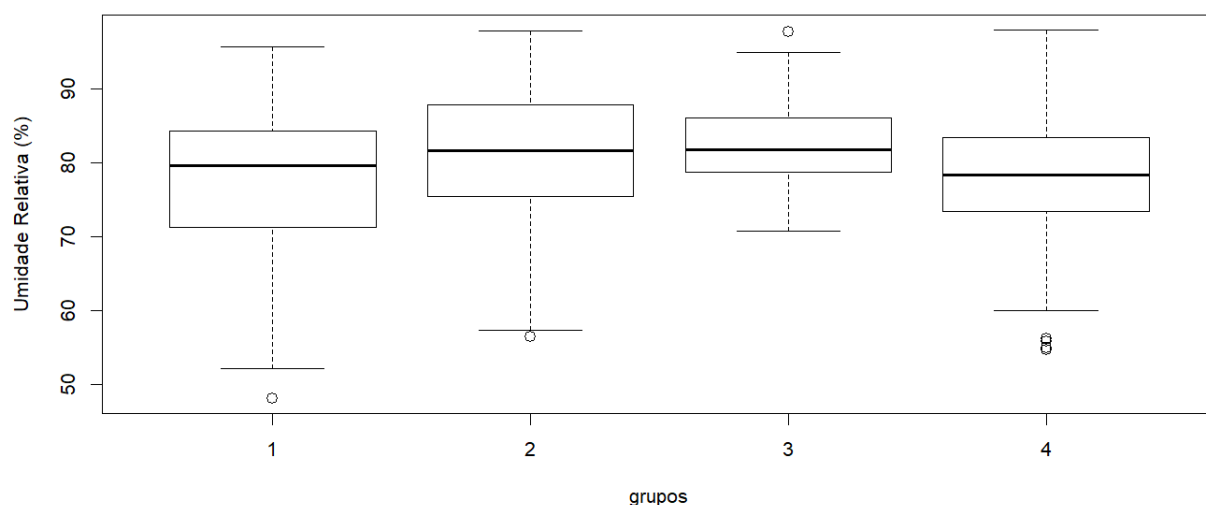


Figura 4.29: Box-plot da Umidade Relativa (%) por grupo - MP_{10} - Santos

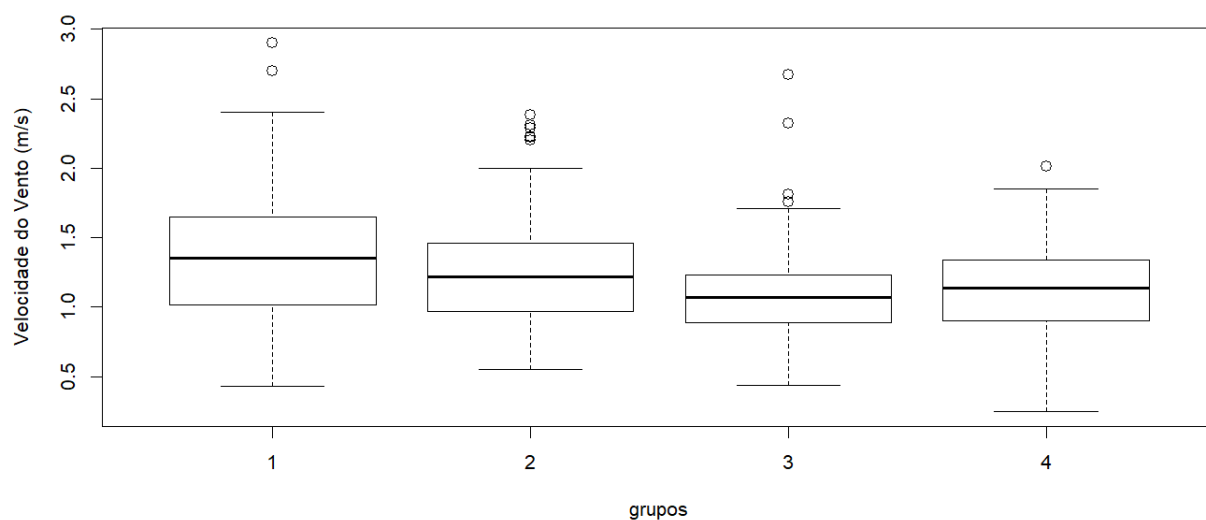


Figura 4.30: Box-plot da Velocidade do Vento (m/s) por grupo - MP_{10} - Santos

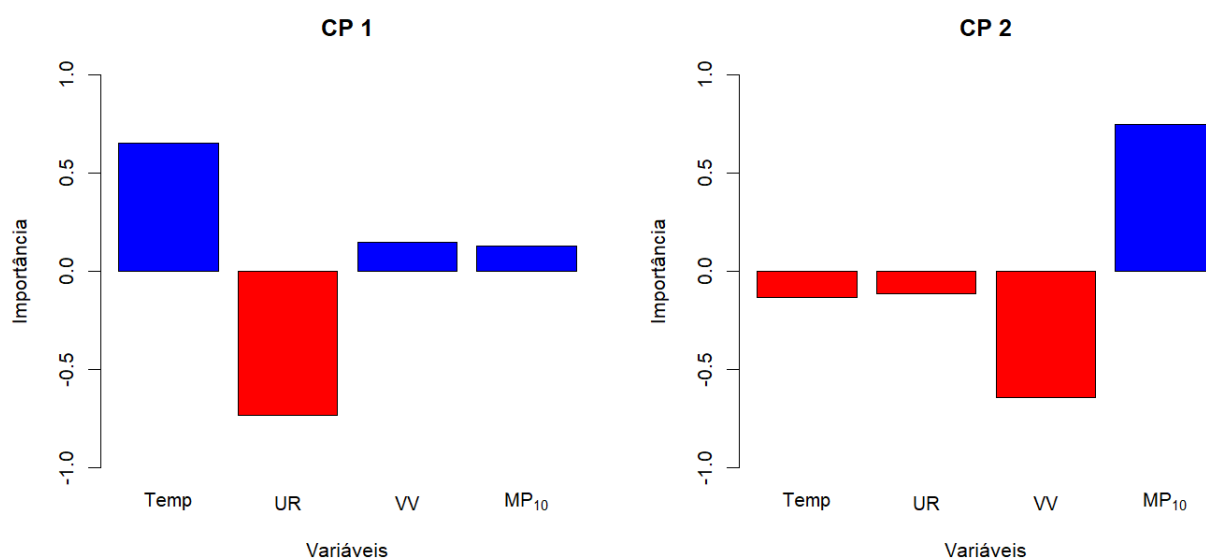
da relação existente entre essas variáveis foi feita considerando apenas as duas primeiras CPs por grupo.

O grupo 1 foi formado por 115 dias, predominantemente no inverno, com 33% dos dias nesta estação (Tabela 4.10). As massas de ar das regiões 2, 7 e 8 foram de maior influência nesses dias (Tabela 4.11), ou seja, influências de fontes provenientes do norte até o oeste de Santos (Figura 4.18, Tabela 4.7). São influências mais continentais, ou seja, de massas de ar mais secas. A relação destacada na CP1 foi da temperatura com a umidade relativa (Figura 4.31). A CP2 mostra a variável mais relacionada com os níveis do MP_{10} , que é a velocidade do vento (Figura 4.31). Este grupo apresenta os maiores

Tabela 4.12: Resumo da ACP - MP_{10} - Santos

Grupo 1	CP1	CP2	CP3	CP4
Proporção da variabilidade	0,497	0,329	0,122	0,052
Proporção acumulada	0,497	0,826	0,948	1,000
Grupo 2				
Proporção da variabilidade	0,377	0,354	0,185	0,084
Proporção acumulada	0,377	0,731	0,916	1,000
Grupo 3				
Proporção da variabilidade	0,449	0,337	0,127	0,087
Proporção acumulada	0,449	0,786	0,913	1,000
Grupo 4				
Proporção da variabilidade	0,533	0,300	0,096	0,071
Proporção acumulada	0,533	0,833	0,929	1,000

valores medianos da velocidade do vento (Figura 4.30), além das máximas concentrações do MP_{10} (Figura 4.27).

Figura 4.31: Importância das CP 1 e CP 2 no grupo 1 - MP_{10} - Santos

O grupo 2, agrupado com 257 dias, tem a maioria dos dias no outono (29,2%) e a minoria dos dias no verão (16,3%), conforme Tabela 4.10. Neste grupo, ficaram os dias com menor valor mediano da concentração do MP_{10} (Figura 4.27) e menores temperaturas (Figura 4.28). Assim como no grupo 1, a relação mais importante foi da temperatura com a umidade relativa, com 37,7% da variabilidade (Tabela 4.12). A segunda CP neste grupo, teve uma proporção da variabilidade explicada muito próxima, com 35,4%, explicitando a relação entre velocidade do vento e MP_{10} (Figura 4.32). As regiões de massa de ar mais predominante no grupo 2 foram as 3, 4, 8, 9, 10 e 12 (Tabela 4.11). Deste modo, as influências foram desde fontes marinhas (regiões 4 e 12, principalmente), até de fontes

continentais (regiões 3, 8, 9 e 10), como observado na Figura 4.18 e Tabela 4.7. São massas de ar provenientes do sul até o oeste da cidade de Santos.

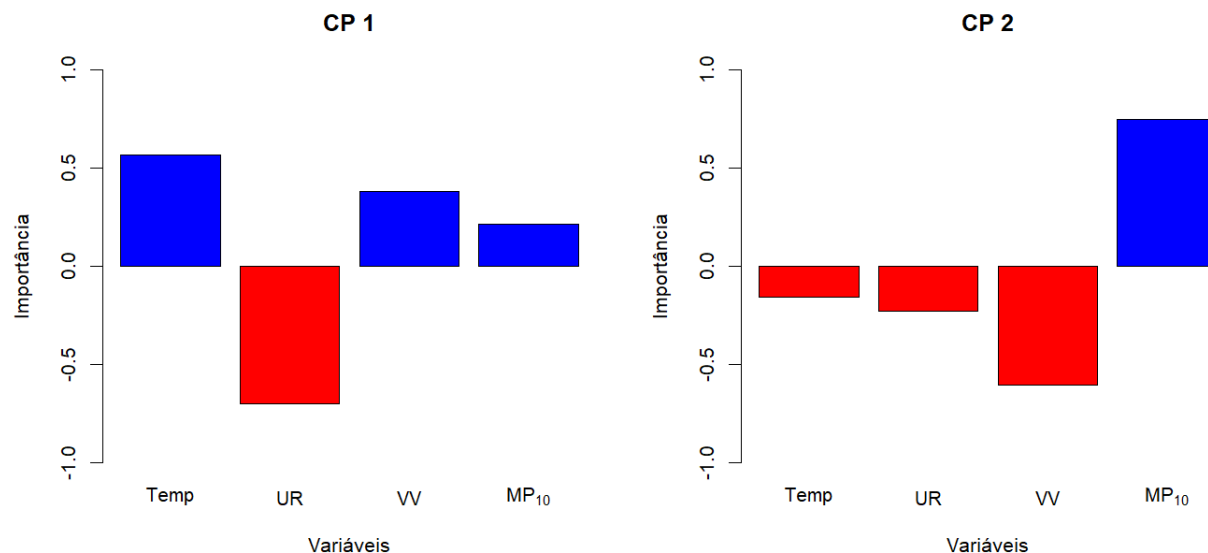


Figura 4.32: Importância das CP 1 e CP 2 no grupo 2 - MP_{10} - Santos

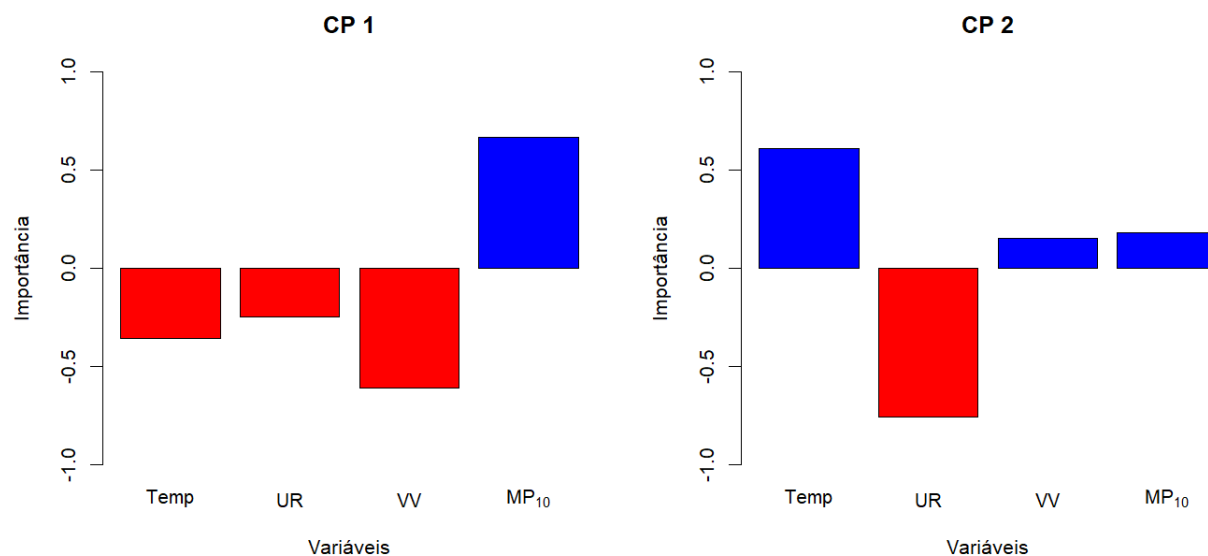


Figura 4.33: Importância das CP 1 e CP 2 no grupo 3 - MP_{10} - Santos

O grupo 3, com 97 dias, é o menor dos quatro grupos obtidos na AC. Nestes dias, as massas de ar foram basicamente da região 5 (Tabela 4.11), que é quase exclusivamente litorânea, o litoral do sudeste, ao leste de Santos (Figura 4.18 e Tabela 4.7). A estação predominante nestes dias foi o outono, com 30,9% (Tabela 4.10). A umidade relativa foi a maior (na mediana) neste grupo em relação ao demais grupos (Figura 4.29), enquanto o menor nível mediano da velocidade do vento ocorreu neste grupo (Figura 4.30). A ACP

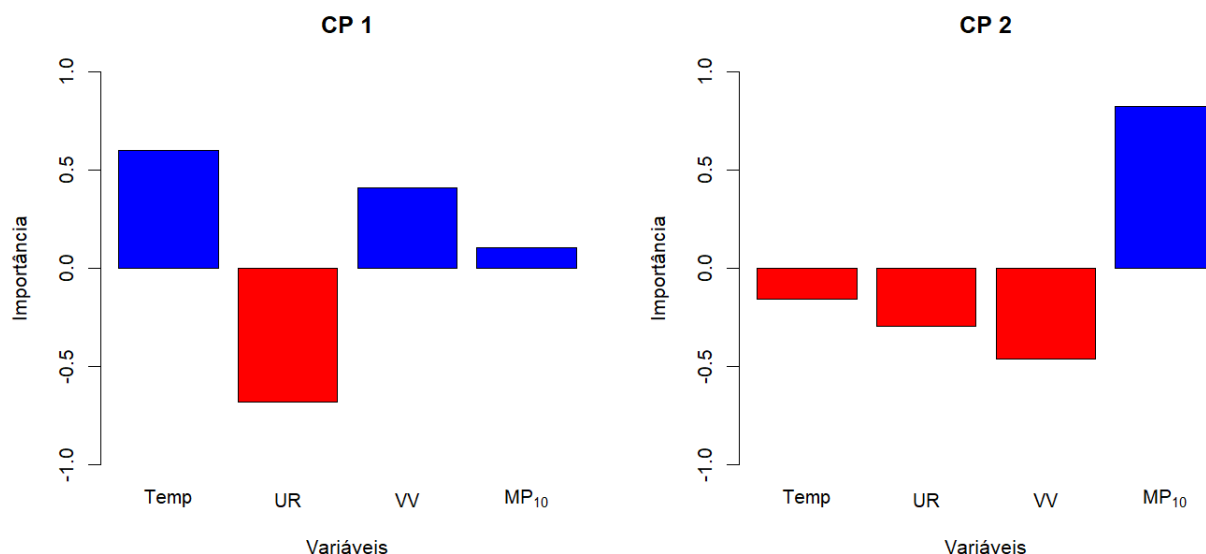


Figura 4.34: Importância das CP 1 e CP 2 no grupo 4 - MP_{10} - Santos

indicou que a relação mais relevante no grupo 3 foi a velocidade do vento com a concentração do MP_{10} , conforme Figura 4.33, sendo 44,9% da variabilidade total (Tabela 4.12). A segunda relação ficou com temperatura e umidade relativa, inversamente proporcionais.

Por último, temos o grupo 4, com o maior número de dias (261). A estação predominante nestes dias foi o verão (Tabela 4.10, com 34,5%, e com maior nível mediano de temperatura (Figura 4.28) e menor nível mediano de umidade relativa (Figura 4.29), e maior nível mediano de MP_{10} (Figura 4.27). As regiões de massa de ar principais foram a 1, 6 e 7, ou seja, provenientes do litoral nordestino (região 6), litoral norte paulista, Rio de Janeiro e Sul de Minas Gerais (região 1) e do sul das regiões Norte e Nordeste do Brasil, e norte de Minas Gerais (região 7). São regiões do leste ao norte de Santos. Na CP1, novamente a temperatura e umidade relativa são as variáveis relacionadas inversamente (Figura 4.34), assim como nos grupos 1 e 2. Na CP2, a relação com a concentração do MP_{10} é com a velocidade do vento, também inversamente proporcional (Figura 4.34).

Note que todos os quatro grupos destacaram a relação do MP_{10} com a velocidade do vento, inversamente proporcional, ou seja, o vento atua dispersando as partículas (YNOUE et al., 2017). Além disso, regiões próximas ficaram agrupadas dentro de um mesmo grupo, indicando um agrupamento adequado. Vale ressaltar que Santos é uma cidade litorânea, que se localiza próximo a serra que interfere diretamente nas condições climáticas da região, e por consequência na qualidade do ar.

4.2 Material Particulado inalável fino - $MP_{2,5}$

As subseções 4.2.1 e 4.2.2 apresentam os resultados da análise dos dados de $MP_{2,5}$. Para isso, as regiões de residência da massa de ar consideradas aqui foram conforme a Figura 4.1, na cidade de Piracicaba, e para Santos, as regiões foram conforme apresentadas na Figura 4.18.

4.2.1 Piracicaba

Nesta seção tem-se a análise do $MP_{2,5}$ ou Material Particulado fino para a cidade de Piracicaba. A Tabela 4.13 apresenta os valores de AIC para diversas escolhas de p e q . Novamente foi escolhido apenas uma combinação. Os modelos ARMA(2,1) e ARMA(2,2) tiveram valores muito próximos do AIC. Neste caso, o modelo ARMA(2,1) foi escolhido, por apresentar um menor número de parâmetros e menor AIC, e então comparado com os outros quatro modelos. Este modelo apresentou bons resultados para a função de auto-correlação e o teste de Ljung-Box, indicando que os resíduos não apresentam auto-correlação Figura A.19.

Tabela 4.13: Critério de Seleção para os modelos ARMA(p,q) - $MP_{2,5}$ - Piracicaba

Modelo	Quantidade de parâmetros	AIC
ARMA(1,0) = AR(1)	1	4322,456
ARMA(2,0) = AR(2)	2	4314,134
ARMA(0,1) = MA(1)	1	4619,320
ARMA(0,2) = MA(2)	2	4461,669
ARMA(1,1)	2	4312,559
ARMA(2,1)	3	4298,204
ARMA(1,2)	3	4310,630
ARMA(2,2)	4	4298,604

A Tabela 4.14 mostra os valores de CP, REQM, LOG e AIC para os cinco modelos. Novamente o modelo de séries temporais (ARMA) apresentou os melhores valores. A CP é maior que 0,8, enquanto os demais modelos (Normal, Normal Inversa, Gama e Log-normal) ficaram próximos de 0,36. O mesmo ocorre com o REQM, em que o modelo ARMA(2,1) ficou abaixo de 5, e os demais maiores do que 7. LOG e AIC também foram melhores no modelo ARMA. Comparando apenas os modelos Normal, Normal inversa, Gama e Log-normal, CP e REQM ficaram muito próximos. Entretanto, em relação ao LOG e AIC, houve uma ligeira vantagem para o modelo Log-normal.

Tabela 4.14: Medidas comparativas - $MP_{2,5}$ - Piracicaba

Modelo	CP	REQM	LOG	AIC
Normal	0,365	7,143	-2471,110	4968,219
Normal Inversa	0,361	7,154	-2337,054	4700,107
Gama	0,363	7,148	-2350,426	4726,852
Log-Normal	0,361	7,316	-2335,321	4696,641
ARMA(2,1)	0,811	4,492	-2133,102	4298,204

Partindo para a análise dos resíduos, as Figuras A.20, A.21, A.22, A.23 e 4.35 mostram os gráficos dos resíduos padronizados pela ordem das observações. Os resíduos do modelo de séries temporais (ARMA(2,1)) parecem mais aleatórios em torno do zero (gaussianos),

se comparado com os demais. Sendo que os resíduos dos modelos Normal, Normal inversa, Gama e Log-normal formaram como se fossem ciclos. Esses ciclos podem estar relacionados com fenômenos meteorológicos que atuaram ao longo de 2014 e 2015. No 1º trimestre de 2015 (entre os dias 366 e 450), houve atuação dos fenômenos ZCAS e ZCOU, o que provocou aumento das chuvas na região sudeste (CETESB, 2015b), e coincide com um ciclo em que os resíduos ficaram abaixo de zero (Figuras A.20, A.21, A.22 e A.23).

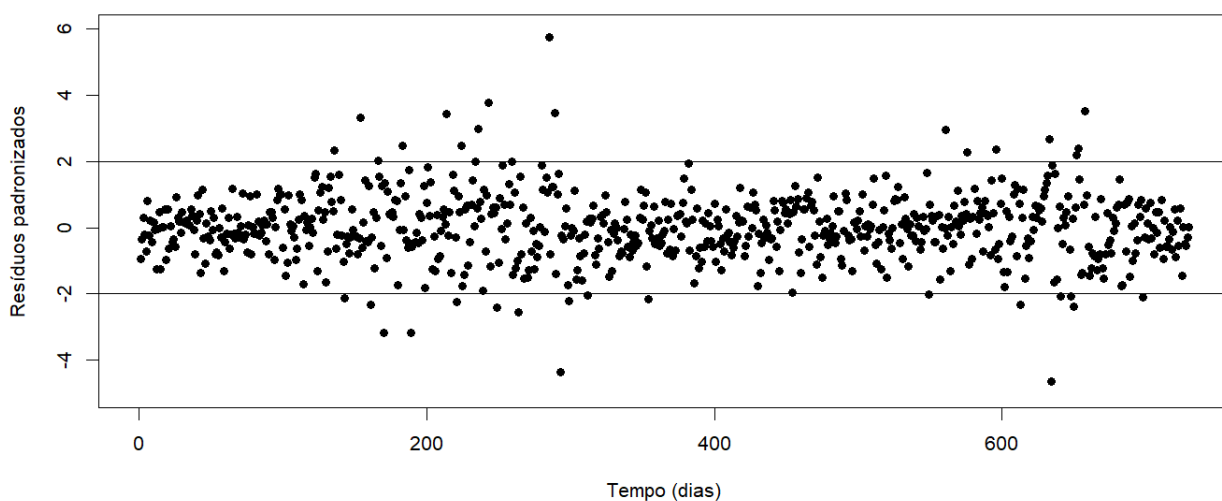


Figura 4.35: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo ARMA(2,1) - Piracicaba

Os próximos gráficos dos resíduos são os *qq-plot*, conforme as Figuras A.24, A.25, A.26, A.27 e 4.36. O melhor ajuste, neste caso, são dos modelos Normal inversa e Log-normal, seguido pelo modelo Gama. Os resíduos de ambos os ajustes estão próximos a reta. Os resíduos do modelo ARMA(2,1) se distanciaram da reta, principalmente nos extremos (Figura 4.36). Das cinco Figuras, os resíduos do modelo Normal, foram os que apresentaram pior resultado.

Mais uma vez é possível notar que há diferenças entre os ajustes dos cinco modelos. O modelo ARMA(2,1) consegue capturar a dependência temporal dos dados, conforme os resultados da Tabela 4.14 e Figura 4.35. Por outro lado, os modelos Log-normal e Normal inversa (que são distribuições assimétricas e positivas) se mostraram mais adequado para ajustar outra característica importante dos dados, que diz respeito ao fato da concentração do MP ser um valor sempre positivo. Por isso os resíduos das Figuras A.25 e A.27 foram melhores que os demais.

As cinco Figuras, as B.9, B.10, B.11, B.12 e 4.37 apresentam a contribuição (em proporção) de cada região definida anteriormente para os modelos Normal, Normal inversa, Gama, Log-normal e ARMA(2,1), respectivamente. Em cada modelo a estimativa é um pouco diferente, porém, em média, elas ficaram bem próximas. No ajuste do $MP_{2,5}$, as regiões significativas estatisticamente (a 5%) foram as mesmas nos modelos Normal, Normal inversa, Gama e ARMA(2,1). Em todos eles, as regiões 9, 10 e 11 não contribuíram

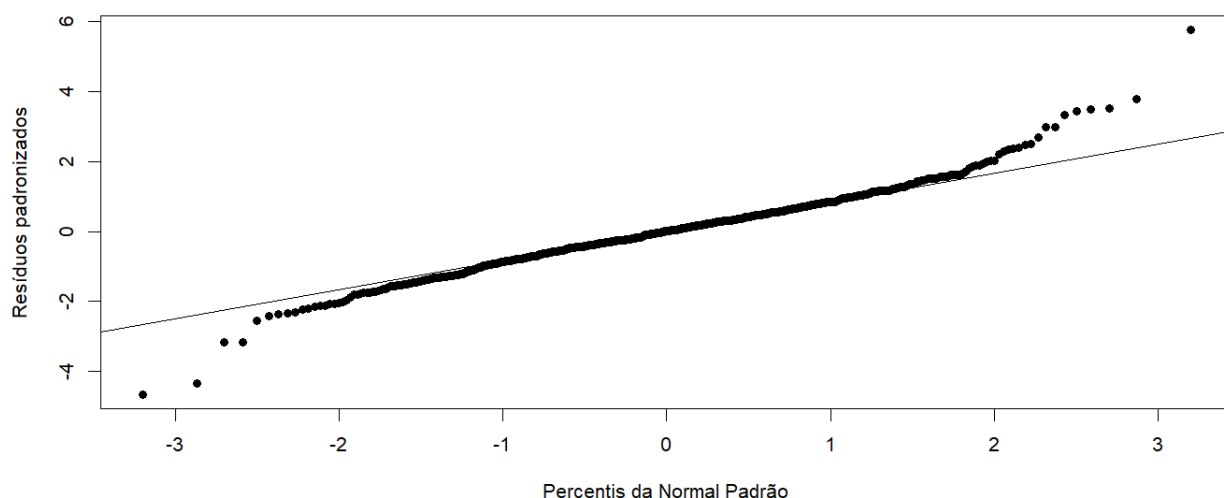


Figura 4.36: Gráfico normal de probabilidades - *qq-plot* - ($MP_{2,5}$) - Modelo ARMA(2,1) - Piracicaba

na concentração do material particulado fino. Entretanto, no modelo Log-normal, apenas a região 9 não foi significativa, ou seja, não contribuiu na concentração do $MP_{2,5}$.

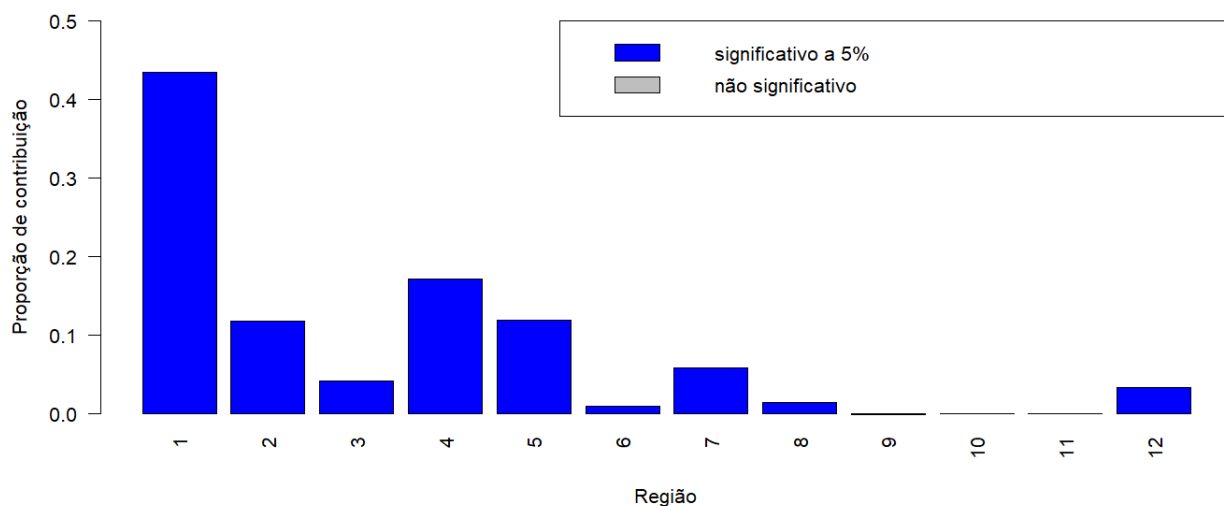


Figura 4.37: Contribuição média do $MP_{2,5}$ por área - Modelo ARMA(2,1) - Piracicaba

Nesta seção, foram comparados os modelos Log-normal e ARMA(2,1), por apresentarem os melhores resultados discutidos até agora. O modelo Log-normal foi escolhido ao se observar os resultados na Tabela 4.14. Deste modo, foram analisadas a contribuição média estimada por região nesses dois modelos, olhando as Figuras B.12 e 4.37. Na Figura 4.38 são comparados os dois modelos Log-normal e ARMA(2,1). As cores indicam

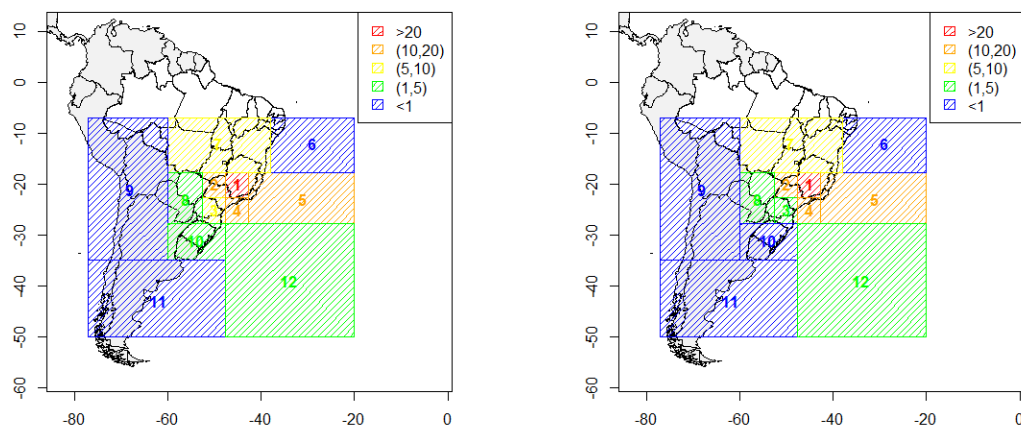


Figura 4.38: Mapa de contribuição das regiões no modelo Log-Normal e ARMA(2,1), respectivamente - Piracicaba - $MP_{2,5}$

diferentes percentuais de contribuição, sendo o vermelho, para indicar regiões de maior contribuição, e até alcançar a cor azul, que representa uma menor contribuição.

Pelos resultados do modelo ARMA(2,1), as regiões 9, 10 e 11 não foram significativas estatisticamente, ou seja, elas não influenciam na composição da concentração diária do $MP_{2,5}$ em Piracicaba. Ou ainda, os Estados do Rio Grande do Sul, Acre e Rondônia, além dos países sul-americanos Chile, Peru, Uruguai e Argentina não contribuem na concentração do $MP_{2,5}$. Diferente do modelo Log-normal, em que apenas a região 9 (Acre, Rondônia, Peru, Bolívia, Chile e Argentina) não contribuiu para os níveis do $MP_{2,5}$. Neste modelo, os resultados apontaram para uma contribuição de quase 2% para as regiões 10 e 11.

As regiões 1 a 4 (fontes locais) contribuíram com 73,3% e 76,5%, respectivamente para o modelo Log-normal e ARMA(2,1). Assim como na contribuição do MP_{10} , a Região 1 (Interior e Sul de Minas Gerais e parte do Rio de Janeiro) é a mais relevante na contribuição do $MP_{2,5}$, sendo 38,9% para o modelo Normal inversa e 43,4% para o modelo ARMA(2,1). Em sequência decrescente, contribuíram as Regiões 4 (17,5%), 2 (11,5%) e 3 (5,5%) para o modelo Log-normal, e as Regiões 4 (17,2%), 2 (11,8%) e 3 (4,1%) para o modelo ARMA(2,1).

A contribuição significativa na concentração diária média do $MP_{2,5}$ dessas regiões (1 a 4), representando grande parte das regiões sul e sudeste do Brasil, são devido ao seu desenvolvimento, sendo constituída por elevado número de indústrias e também pelo enorme tráfego de veículos.

Em relação a contribuição do litoral, as regiões 5, 6 e 12, representam 17,3% e 16,2% da contribuição diária média do $MP_{2,5}$ para os modelos Log-normal e ARMA(2,1), respectivamente. O litoral da região sudeste (Região 5) representa a maior parcela com 12,7% (modelo Log-normal) e 11,9% (modelo ARMA(2,1)).

As regiões 7 e 8 contribuíram conjuntamente um percentual muito próximo em cada um dos modelos, sendo 7,4% para o modelo Log-normal e 7,2% para o modelo ARMA(2,1).

Porém, houve uma pequena diferença na divisão entre as duas regiões. No modelo Normal inversa, a Região 7 representa 5,1%, enquanto que no modelo ARMA(2,1), a Região 7 representa 5,8%. Lembrando que essas duas regiões representam as influências de fontes regionais e continentais.

Comparando-se os dois modelos, nota-se pequenas diferenças na proporção de contribuição de cada região. Essa diferença pode ser explicada pelas diferentes características das duas abordagens. Interessante reforçar que a análise dos resíduos apresentada indica que o modelo ARMA(2,1) é o mais indicado nessa situação.

Assim como na concentração do MP_{10} , as principais fontes de contribuição da concentração diária média do $MP_{2,5}$ foram as fontes locais. Como visto na Figura 4.6, as fontes de processos de combustão, como emissões veicular e industrial aparentemente são as principais para a região de Piracicaba. Além disso, o $MP_{2,5}$ pode ter sua fonte de aerossóis secundários, como nitratos e sulfatos (CETESB, 2018).

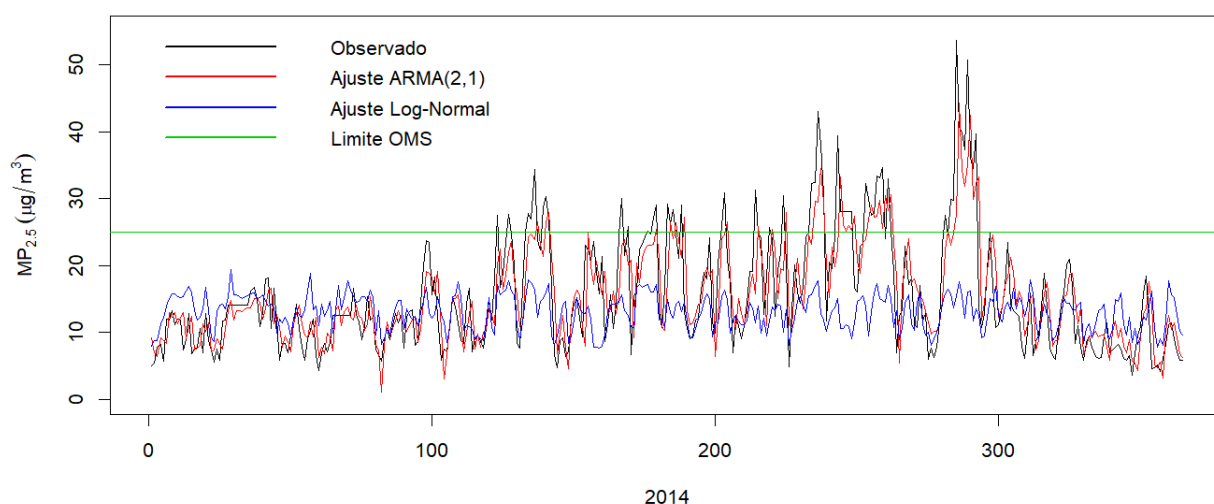


Figura 4.39: Ajuste da concentração diária do $MP_{2,5}$ em 2014 em comparação com valores observados - Piracicaba

Com os dados do mês de Janeiro de 2016 de concentração do $MP_{2,5}$ e os tempos de residência de cada região, o poder de previsão do modelo ARMA(2,1) foi analisado. Na Figura 4.41 são mostrados os valores reais da concentração do $MP_{2,5}$ em Piracicaba comparados com os valores previstos pelo modelo. Observe que as concentrações previstas conseguiram acompanhar as concentrações reais ao longo do mês, indicando uma bom ajuste.

Na comparação do ajuste diário dos dois modelos (Log-normal e ARMA(2,1)), temos as Figuras 4.39 e 4.40, apresentando os valores observados e as estimativas diárias da concentração do $MP_{2,5}$ nos anos de 2014 e 2015, respectivamente. Note mais uma vez que o ajuste do modelo Log-normal não consegue capturar a tendência temporal dos dados. Na análise da média geral de contribuição todos os modelos são muito parecidos. Entretanto, no ajuste diário, apenas os modelos de séries temporais conseguem capturar a tendência, seja no ajuste do MP_{10} , ou no ajuste do $MP_{2,5}$. Quando se compara com

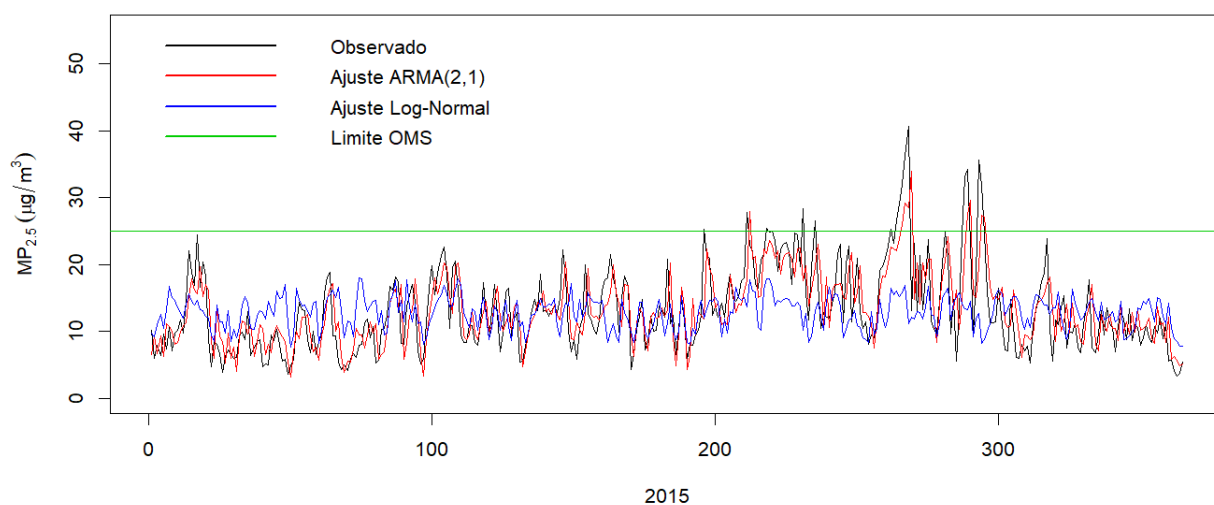


Figura 4.40: Ajuste da concentração diária do $MP_{2,5}$ em 2015 em comparação com valores observados - Piracicaba

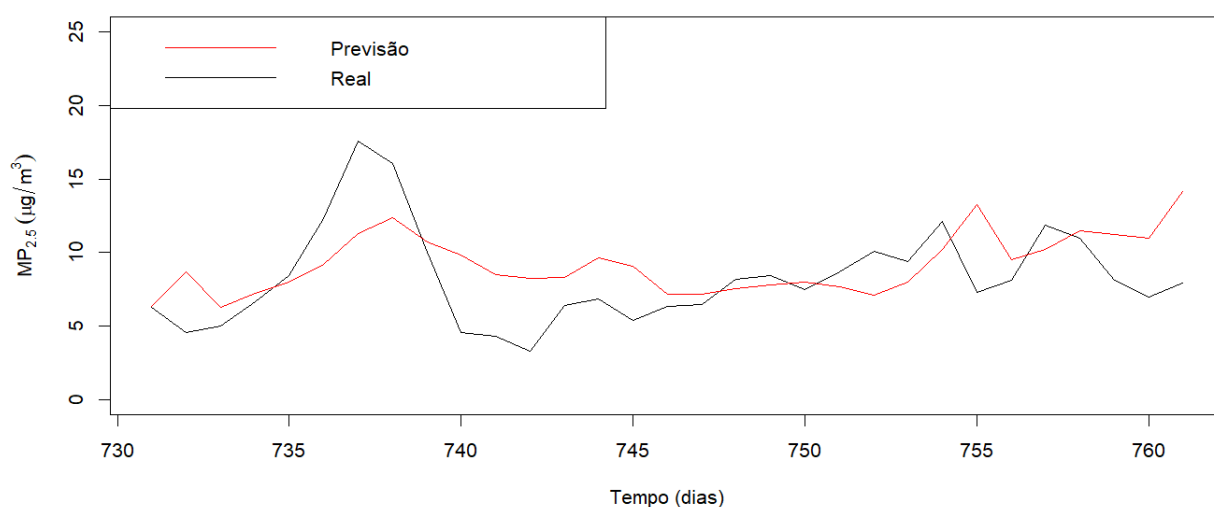


Figura 4.41: Previsão do $MP_{2,5}$ e valores reais para o mês de Janeiro de 2016 - Piracicaba

os limites da OMS, apenas o modelo ARMA(2,1) consegue indicar dias que ocorreram concentrações acima do limite. Os demais modelos, não conseguem acompanhar a série observada, e por consequência não estima os dias acima do limite.

Assim, para o ajuste do $MP_{2,5}$ o modelo ARMA(2,1) é melhor que os demais. A desvantagem é que também nota-se que há uma defasagem no ajuste. O pico observado em determinado dia, somente é ajustado no dia seguinte, pelo modelo. Como medida corretiva, neste caso, há a possibilidade de ajuste dos modelos ARFIMA (*Auto-Regressive Fractionally Integrated Moving Average*). Reisen et al. (2014) utilizaram os modelos ARFIMA

sazonais, para modelar e prever a concentração diária do MP_{10} na cidade de Cariacica (ES).

Tabela 4.15: Percentual de dias dentro de cada estação do ano por grupo ($MP_{2,5}$ - Piracicaba)

	Primavera	Verão	Outono	Inverno
Grupo 1	20,4%	25,1%	34,6%	19,9%
Grupo 2	20,6%	22,3%	27,4%	29,7%
Grupo 3	17,8%	21,8%	33,0%	27,4%
Grupo 4	43,1%	29,9%	3,0%	24,0%

Com o modelo ARMA(2,1), continuaremos com a AC e ACP. Mais uma vez, o modelo de séries temporais (modelos ARMA) mostram-se como a melhor alternativa, dentre os cinco modelos considerados.

De acordo com as variáveis meteorológicas e a proporção de contribuição diária do $MP_{2,5}$ por região (somente as regiões significativas 1-8 e 12) foi realizada a AC. Todos os dias do período foram divididos em 4 grupos (Figura C.2). O grupo 1 foi composto por 191 dias e o grupo 2 foi composto por 175 dias. O grupo 3 e 4 foram formados por 197 e 167 dias cada um, respectivamente.

A Tabela 4.15 apresenta os percentuais de dias dentro de cada estação do ano em cada um dos quatro grupos formados na AC. A divisão em cada estação do ano foi mantida a mesma usada anteriormente, conforme IAG (2018).

A Tabela 4.16 mostra os percentuais de cada grupo em cada uma das regiões que foram estatisticamente significativas no ajuste do $MP_{2,5}$ em Piracicaba.

Tabela 4.16: Percentual de contribuição do grupo dentro de cada região de residência da massa de ar ($MP_{2,5}$ - Piracicaba)

	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Região 1	23,6%	6,2%	42,4%	27,8%
Região 2	42,8%	31,8%	14,5%	10,9%
Região 3	25,2%	49,0%	8,0%	17,8%
Região 4	20,6%	43,5%	15,1%	20,8%
Região 5	22,3%	8,4%	40,5%	28,8%
Região 6	11,8%	5,0%	32,4%	50,8%
Região 7	41,2%	18,5%	18,6%	21,7%
Região 8	41,1%	39,0%	0,6%	19,2%
Região 12	13,1%	51,1%	17,2%	18,6%

Foram feitos gráficos do boxplot da concentração do $MP_{2,5}$ (Figura 4.42), da temperatura (Figura 4.43), da umidade relativa (Figura 4.44) e da velocidade do vento (Figura 4.45) dividido nos quatro grupos obtidos na AC.

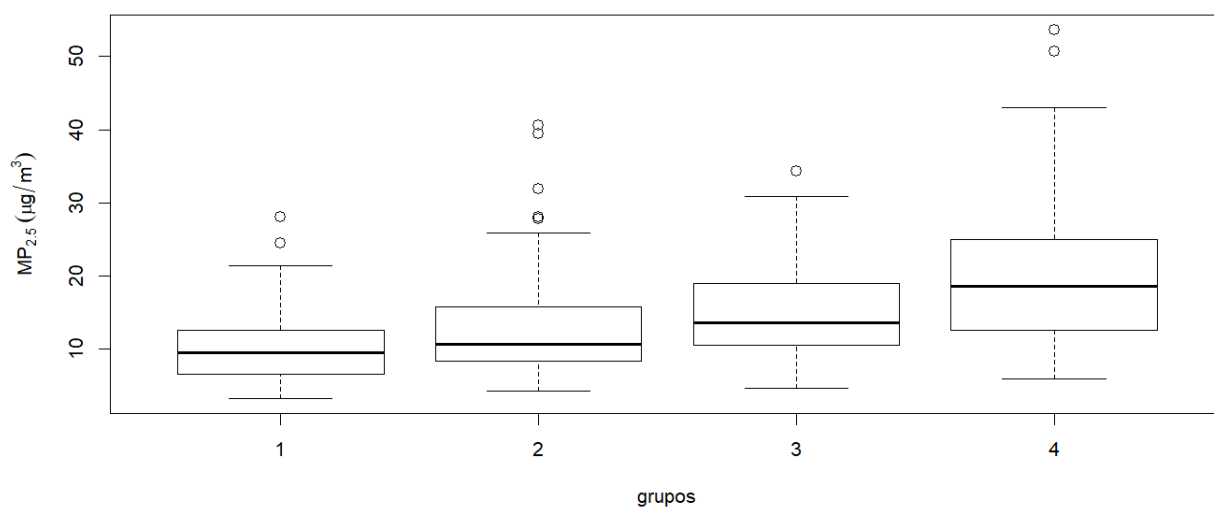


Figura 4.42: Box-plot do $MP_{2,5}$ ($\mu g/m^3$) por grupo - $MP_{2,5}$ - Piracicaba

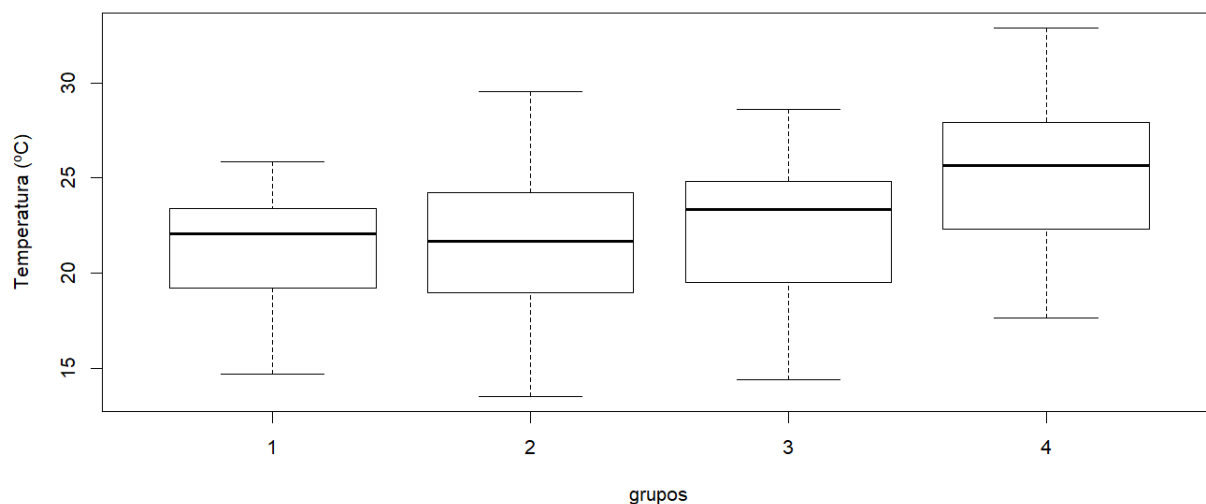


Figura 4.43: Box-plot da Temperatura ($^{\circ}C$) por grupo - $MP_{2,5}$ - Piracicaba

A Tabela 4.17 apresenta o resumo da ACP, com a proporção da variância explicada por cada CP, nos quatro grupos. A ACP foi realizada considerando as variáveis meteorológicas e a concentração do $MP_{2,5}$. Todos os grupos tiveram mais de 64% da variabilidade total sendo explicada pelas duas primeiras CP. O grupo 3, as duas primeiras CP acumulam 74,5%. Assim, em todos os grupos foram consideradas apenas as duas primeiras CPs na discussão dos resultados.

As Figuras 4.46, 4.47, 4.48 e 4.49 mostram os coeficientes da ACP, para as CP1 e CP2 depois da rotação varimax, em todos os quatro grupos. De novo, as variáveis Temperatura, Umidade Relativa e Velocidade do Vento foram abreviadas usando Temp, UR e VV.

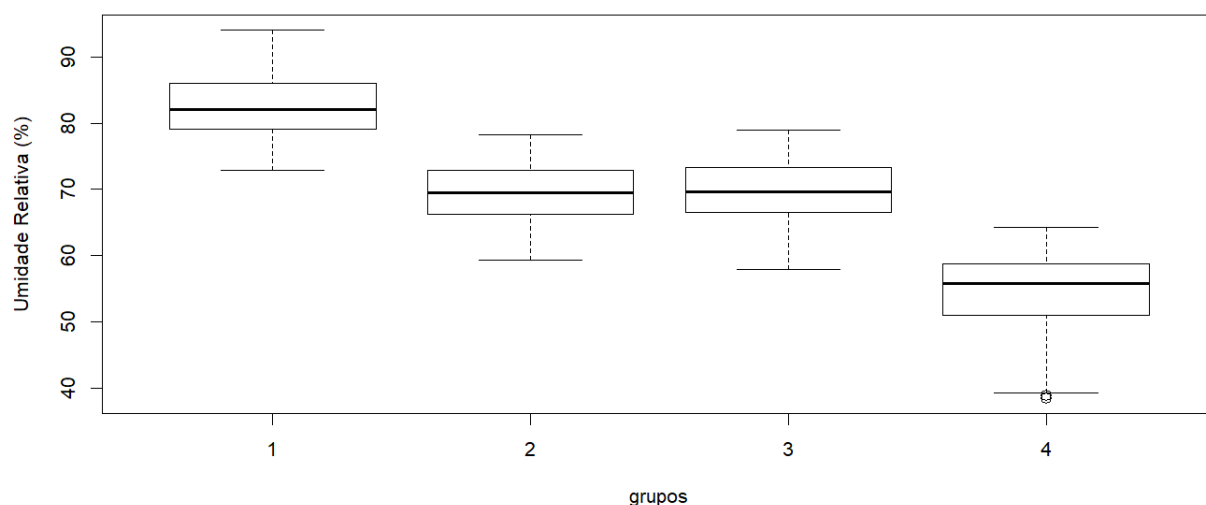


Figura 4.44: Box-plot da Umidade Relativa (%) por grupo - $MP_{2,5}$ - Piracicaba

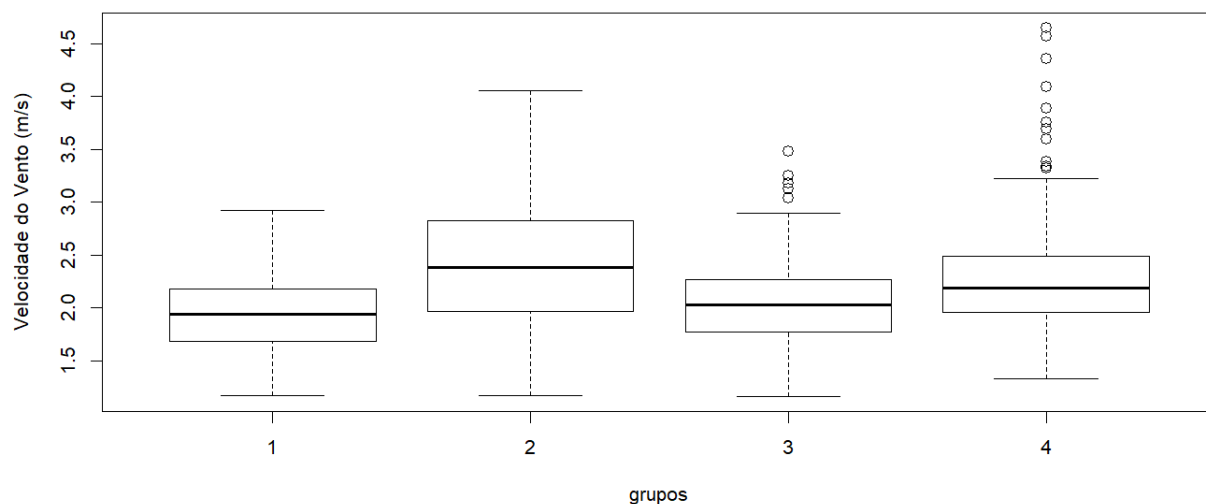


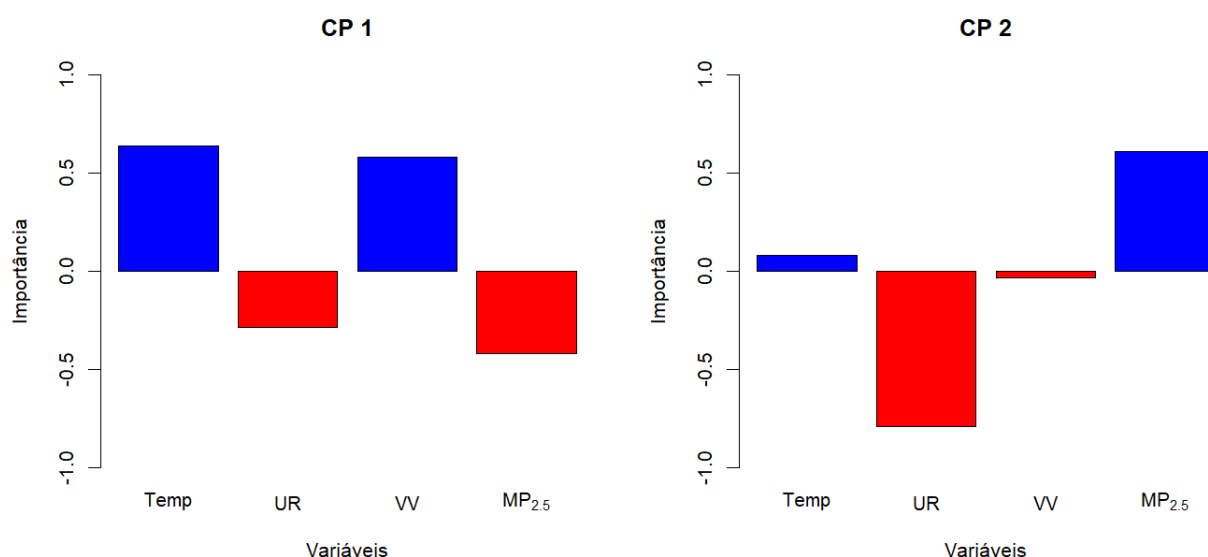
Figura 4.45: Box-plot da Velocidade do Vento (m/s) por grupo - $MP_{2,5}$ - Piracicaba

O grupo 1 é agrupado com 191 dias, com predomínio dos dias na estação do outono, com 34,6% (Tabela 4.15). Neste grupo, tem-se os menores níveis mediano de $MP_{2,5}$ e maiores valores para a Umidade Relativa (Figuras 4.42 e 4.44), mostrando uma relação inversamente proporcional. A relação entre UR e $MP_{2,5}$ equivale a 30,3% da variabilidade total dos dados neste grupo, conforme mostrado na Tabela 4.17, porém, esta relação é segunda mais importante, pois está na CP2 (Figura 4.46). A relação da CP1 (com 37,6%) é da temperatura e velocidade do vento. Dentre as regiões definidas na Tabela 4.1, destacam-se as 2, 7 e 8 (Tabela 4.16).

O grupo 2 é formado por 175 dias, com predominância do inverno e outono, com menor nível mediano de temperatura (Figura 4.43) e altas velocidade do vento (Figura 4.45). As

Tabela 4.17: Resumo da ACP - $MP_{2,5}$ - Piracicaba

Grupo 1	CP1	CP2	CP3	CP4
Proporção da variabilidade	0,376	0,303	0,197	0,124
Proporção acumulada	0,376	0,679	0,876	1,000
Grupo 2				
Proporção da variabilidade	0,347	0,294	0,227	0,132
Proporção acumulada	0,347	0,641	0,868	1,000
Grupo 3				
Proporção da variabilidade	0,468	0,277	0,141	0,114
Proporção acumulada	0,468	0,745	0,886	1,000
Grupo 4				
Proporção da variabilidade	0,399	0,291	0,202	0,108
Proporção acumulada	0,399	0,690	0,892	1,000

Figura 4.46: Importância das CP 1 e CP 2 no grupo 1 - $MP_{2,5}$ - Piracicaba

regiões de maior contribuição neste grupo foram a 3, 4 e 12 (Tabela 4.16). São influências de massas de ar vindas do estado do Paraná e Santa Catarina, além de massas do oceano Atlântico, mais ao sul. Devido as altas velocidade do vento, o nível mediano de $MP_{2,5}$ neste grupo é o segundo menor, pois o vento auxilia na dispersão do MP. A ACP, mostra a principal relação deste grupo com a velocidade do vento e a temperatura, com 34,7% da variabilidade total. A CP2 evidencia a relação inversamente proporcional entre umidade relativa e $MP_{2,5}$.

O maior grupo é o 3, com 197 dias, predominantemente do outono (33,0%, Tabela 4.15). Apresenta o segundo maior nível mediano de $MP_{2,5}$ (Figura 4.42). A relação da CP1 foi da temperatura e velocidade do vento inversamente proporcionais a concentração do MP, com uma variabilidade de quase 50% (Tabela 4.17). Na CP2, apenas a

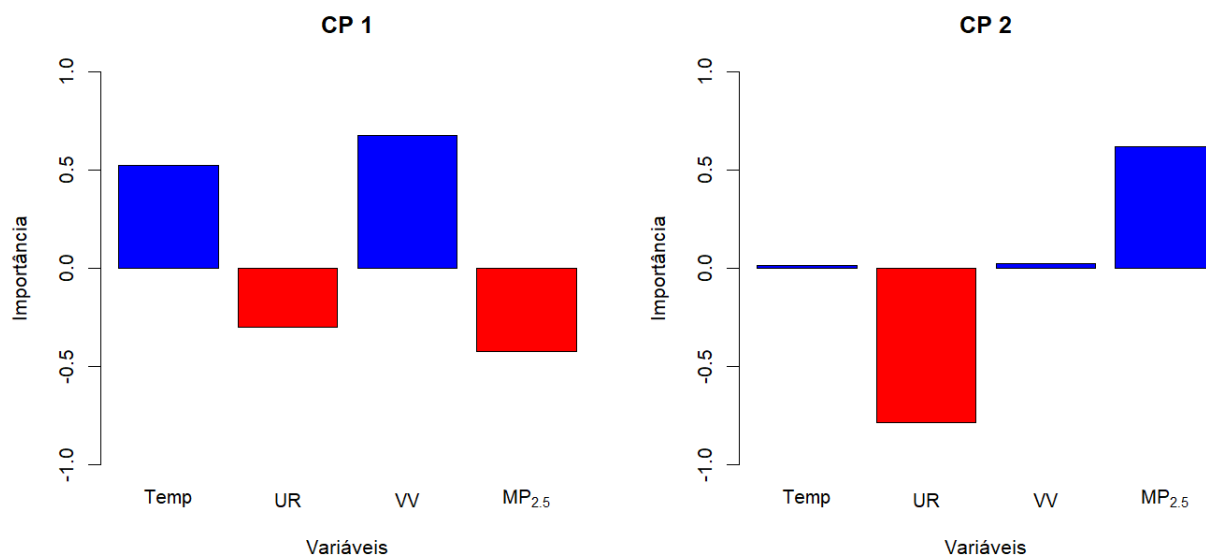


Figura 4.47: Importância das CP 1 e CP 2 no grupo 2 - $MP_{2,5}$ - Piracicaba

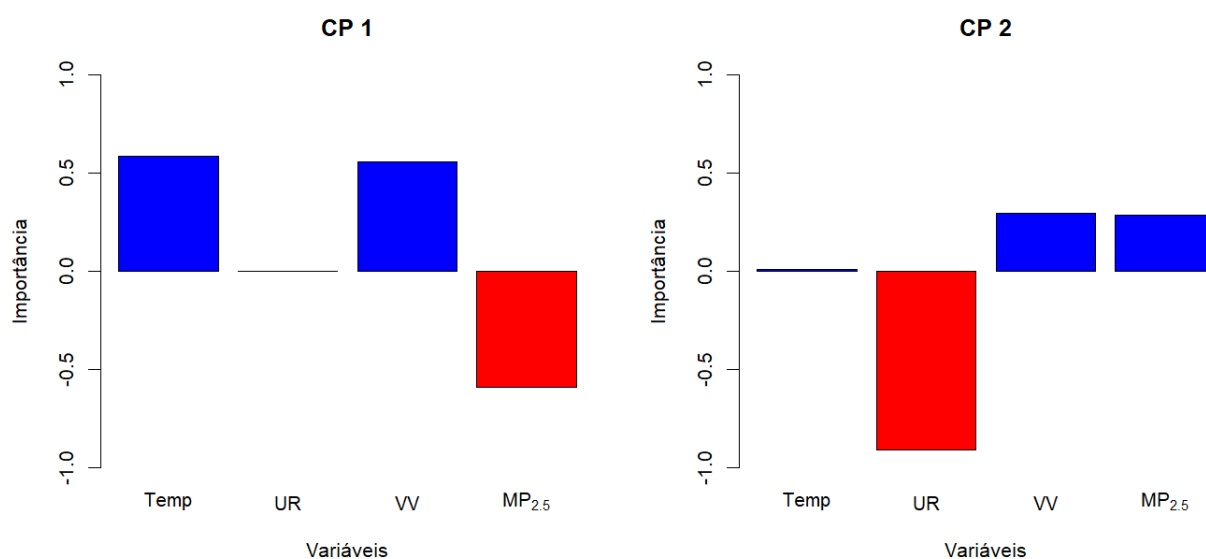


Figura 4.48: Importância das CP 1 e CP 2 no grupo 3 - $MP_{2,5}$ - Piracicaba

variável umidade relativa foi destacada, com 27,7% da variabilidade total. As regiões 1 e 5 mais contribuem neste grupo. São regiões do litoral do sudeste, interior de Minas Gerais, Espírito Santo e parte do Rio de Janeiro (Tabela 4.16).

O menor grupo é o 4, com 167 dias, e predominância da estação da primavera (43,1%) e também do verão (29,9%), de acordo com a Tabela 4.15. Neste grupo 4 percebe-se que quase não há dias na estação de outono, com apenas 3% (Tabela 4.15). Como há o predomínio das estações quentes, nota-se maior valor mediano de temperatura neste grupo (Figura 4.43), porém apresentou menor valor de umidade relativa (Figura 4.44), e o maior nível na concentração do material particulado fino. Quase 40% é explicada pela primeira CP, que indica que a principal relação é entre umidade relativa e $MP_{2,5}$, inver-

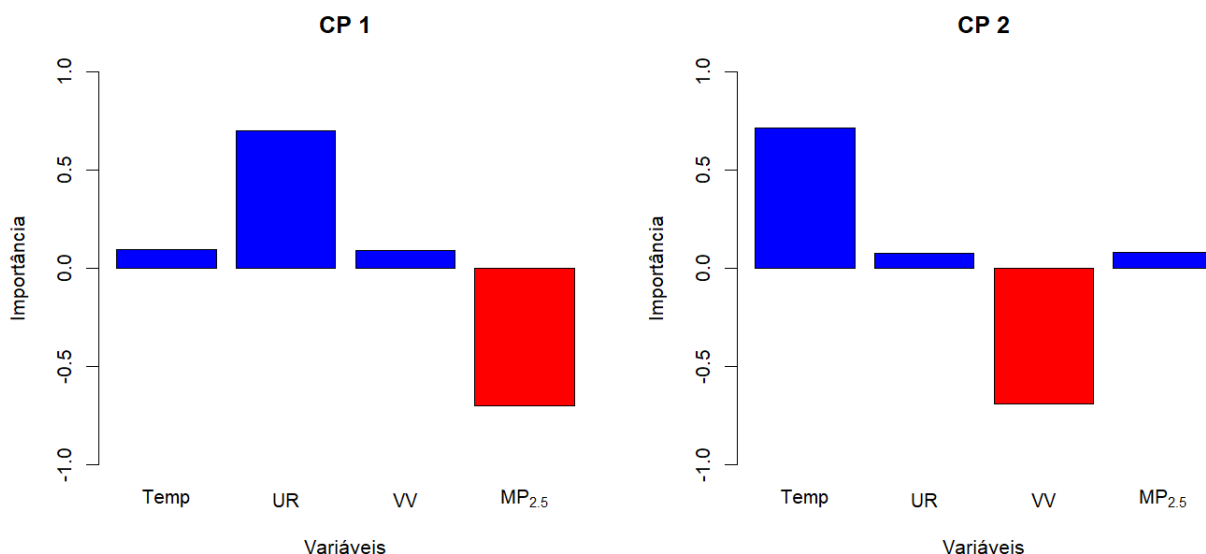


Figura 4.49: Importância das CP 1 e CP 2 no grupo 4 - $MP_{2,5}$ - Piracicaba

samente proporcional. Neste grupo também há uma relação inversamente proporcional entre temperatura e velocidade do vento, apresentada na CP2 (Figura 4.47). A região 6, formada pelo litoral nordestino, é a que mais contribui no grupo 4, conforme Tabela 4.16.

Assim como observado na AC do MP_{10} , as regiões próximas ficaram dentro de um mesmo grupo. Regiões 2, 7 e 8 no grupo 1, regiões 3, 4 e 12 no grupo 2, regiões 1 e 5 no grupo 3, e região 6 no grupo 4.

Dos 4 grupos, três deles mostraram temperatura e velocidade do vento como a principal relação obtida da ACP. Ou seja, há indícios que essas variáveis podem influenciar de modo mais intenso nos níveis do $MP_{2,5}$. Diferente do observado na análise do MP_{10} , em que a variável umidade relativa ficou na CP1 em três grupos, indicando essa variável como mais relevante. Apesar disso, em 3 grupos, o $MP_{2,5}$ ficou numa mesma CP com umidade relativa.

4.2.2 Santos

Esta seção apresenta os resultados obtidos da análise do $MP_{2,5}$ para a cidade de Santos. Primeiramente, é preciso estimar as ordens do processo auto-regressivo (p) e do processo de médias-móveis (q) do modelo ARMA(p,q). A Tabela 4.18 mostra os valores do critério de informação AIC para alguns modelo ARMA pré-escolhidos. O modelo ARMA(2,1) apresenta o menor valor do AIC, além dos melhores resultados em relação a função de auto-correlação e o teste de Ljung-Box dos resíduos, ou seja, os resíduos obtidos desse modelo são não-auto-correlacionados (Figura A.28).

Na Tabela 4.19 tem-se os valores de CP, REQM, LOG e AIC para os cinco modelos. O modelo ARMA(2,1) apresentou melhores resultados em todas as medidas comparativas, isto é, maiores valores de CP e LOG, e menores valores de REQM e AIC. Quando se observa apenas os resultados dos modelos Normal, Normal inversa, Gama e Log-normal, em relação ao CP e REQM, o modelo Normal teve os melhores valores, enquanto que

Tabela 4.18: Critério de Seleção para os modelos ARMA(p,q) - $MP_{2,5}$ - Santos

Modelo	Quantidade de parâmetros	AIC
ARMA(1,0) = AR(1)	1	4718,202
ARMA(2,0) = AR(2)	2	4714,754
ARMA(0,1) = MA(1)	1	4789,960
ARMA(0,2) = MA(2)	2	4749,274
ARMA(1,1)	2	4713,128
ARMA(2,1)	3	4696,025
ARMA(1,2)	3	4705,915
ARMA(2,2)	4	4697,844

o modelo Gama mostrou-se melhor que os demais na comparação do AIC e LOG. Os resultados do modelo Log-normal ficaram bem próximos aos resultados do modelo Gama.

Tabela 4.19: Medidas comparativas - $MP_{2,5}$ - Santos

Modelo	CP	REQM	LOG	AIC
Normal	0,538	7,149	-2471,669	4969,338
Normal Inversa	0,526	7,224	-2430,749	4887,497
Gama	0,534	7,172	-2401,231	4828,462
Log-Normal	0,532	7,311	-2403,369	4832,738
ARMA(2,1)	0,718	5,901	-2332,012	4696,025

Os gráficos dos resíduos padronizados pela ordem das observações estão nas Figuras A.29, A.30, A.31, A.32 e 4.50. Os resíduos padronizados do modelo ARMA(2,1) se dispõe de modo aleatório em torno do zero, em sua maioria dentro dos limites de -2 a 2 (comportamento gaussiano). Mais uma vez, os resíduos dos demais modelos (Normal, Normal inversa, Gama e Log-normal), se comportaram formando como se fosse uma onda senoidal. Isso indica, que por um certo período de tempo, as concentrações diárias do $MP_{2,5}$ foram subestimadas, e num próximo período foram superestimadas, e assim sucessivamente. Mais uma vez, há evidência da atuação dos fenômenos meteorológicos. Em Setembro/2015 (a partir do dia 609) não ocorreram muitas chuvas devido a uma massa de ar quente (sistema de alta pressão subtropical) atuando no estado, o que impediu a entrada de um sistema instáveis (sistema de baixa pressão) (CETESB, 2015b).

As Figuras A.33, A.34, A.35, A.36 e 4.51 mostram os resíduos padronizados comparados com o percentil da distribuição normal padrão (gráfico *qq-plot*). Os modelos GLM e GAMLSS (Normal inversa, Gama e Log-normal) tiveram comportamento muito similar, com os pontos ficando bem próximos da reta. Na Figura 4.51 (modelo ARMA), os resíduos ficaram próximos também da reta, apenas se distanciando no extremo positivo. Já na Figura A.33 (modelo normal), os resíduos apresentaram o pior desempenho, permanecendo distante da reta nos dois extremos.

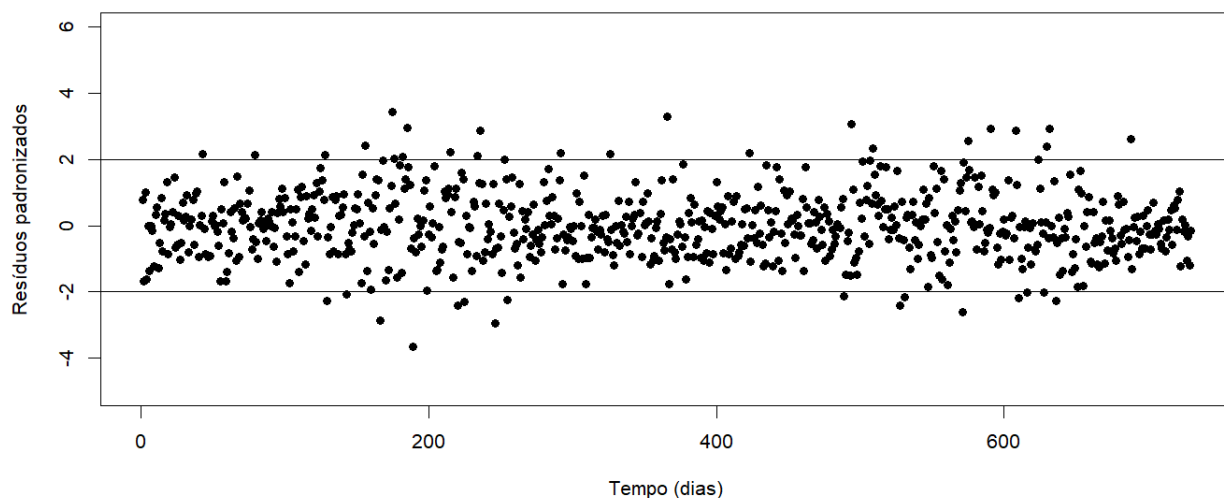


Figura 4.50: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo ARMA(2,1) - Santos

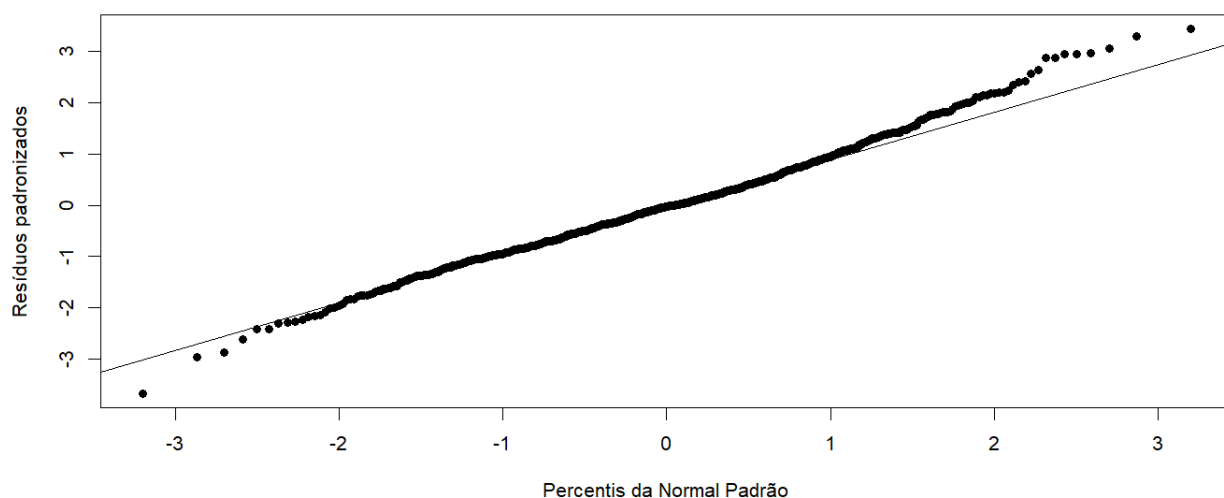


Figura 4.51: Gráfico normal de probabilidades - *qq-plot* - ($MP_{2,5}$) - Modelo ARMA(2,1) - Santos

Os resultados até aqui apresentam diferenças entre os modelos. Os resultados da Tabela 4.19 e Figura 4.50 vão ao encontro do modelo ARMA(2,1), por ajustar a estrutura temporal dos dados. Por outro lado, os resultados do *qq-plot* dos modelos lineares generalizados (Figuras A.34, A.35 e A.36) se mostraram mais ajustados.

A contribuição proporcional média de cada região estão nas Figuras B.13, B.14, B.15, B.16 e 4.52, para os modelos Normal, Normal inversa, Gama, Log-normal e ARMA(2,1), respectivamente. A região 9 não foi significativa a 5% nos modelos Normal, Normal inversa, Gama e ARMA(2,1). A região 9, corresponde aos estados do Acre e Rondônia,

e os países Bolívia, Peru, Chile e Argentina (Figura 4.18 e Tabela 4.7). No modelo Log-normal, todas as regiões foram significativas.

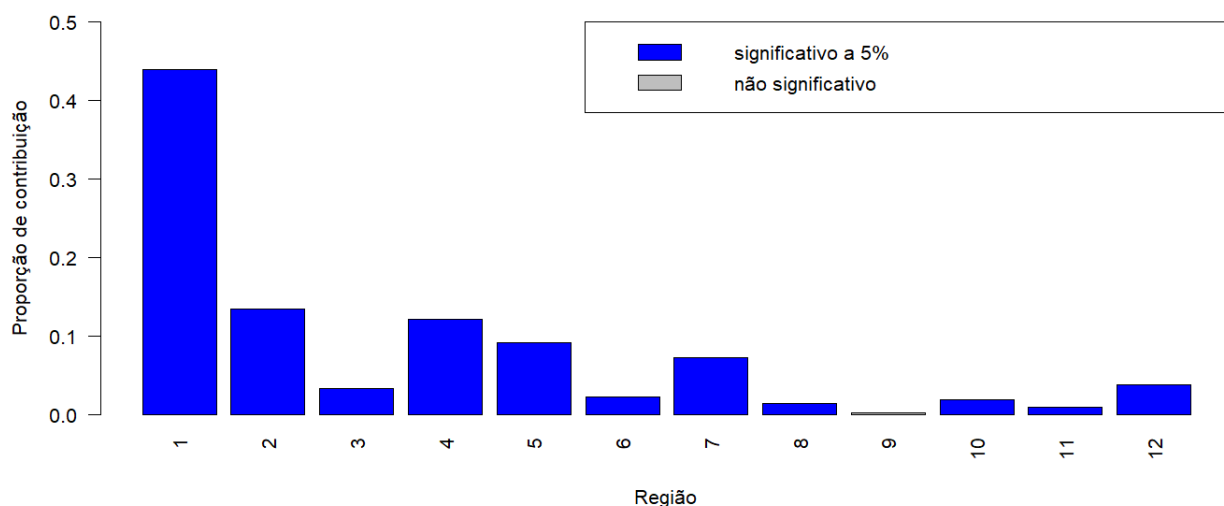


Figura 4.52: Contribuição média do $MP_{2,5}$ por área - Modelo ARMA(2,1) - Santos

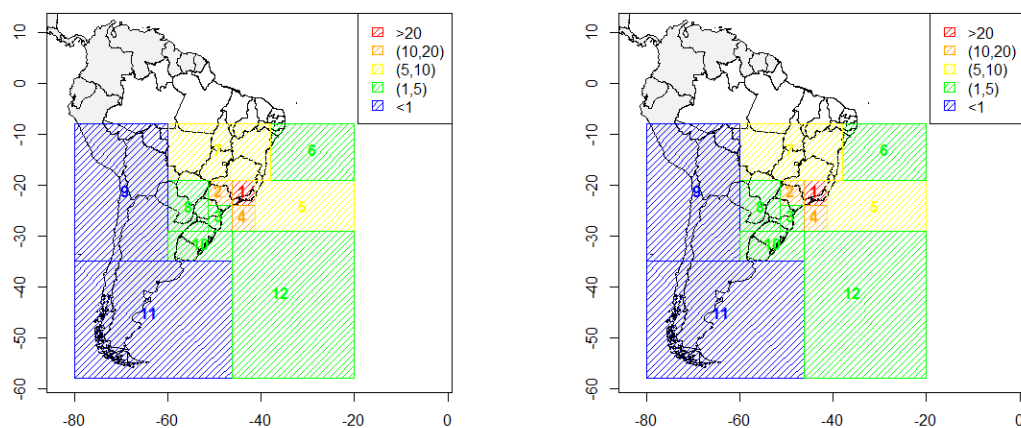


Figura 4.53: Mapa de contribuição das regiões no modelo Gama e ARMA(2,1), respectivamente - Santos - $MP_{2,5}$

Baseado nos resultados da análise de resíduos e das medidas comparativas entre os modelos (Tabela 4.19), os modelos Gama e ARMA(2,1) foram comparados em relação a contribuição média de cada região (Figuras B.15 e 4.52). A Figura 4.53 mostra os percentuais de contribuição por região nos modelos Gama e ARMA(2,1), para o $MP_{2,5}$ na cidade de Santos. As cores mais quentes indicam percentuais de contribuição maior, enquanto as cores mais frias representam uma menor contribuição da região nos níveis do $MP_{2,5}$.

As fontes locais (regiões 1 a 4) contribuem com mais de 70% nos dois modelos, sendo 76,0% no modelo Gama e 72,8% no modelo ARMA. As regiões 1 e 2 foram as de maior contribuição, e correspondem a praticamente massas de ar do Estado de São Paulo inteiro, além do Sul de Minas, e o Estado do Rio de Janeiro. No modelo Gama, os percentuais correspondem a 47,9% e 12,7% nas regiões 1 e 2, respectivamente. Enquanto que, no modelo ARMA(2,1), as estimativas de contribuição das regiões 1 e 2 foram de 43,9% e 13,4%. A região 3 reflete, na maior parte, influência de massa de ar dos estados do Paraná e Santa Catarina, e seus litorais. A contribuição desta região foi de 4,4%, no modelo Gama, e 3,3% no modelo ARMA(2,1). A região 4, com 11,0% (modelo Gama) e 12,2% (modelo ARMA), indica a contribuição média do Litoral paulista, neste caso, correspondem a massas de ar úmidas, provenientes do oceano.

Além da região 4, as regiões 5, 6 e 12 também estimam a contribuição de massas de ar vindas do litoral (Figura 4.18 e Tabela 4.7). A contribuição no modelo Gama, destas regiões foi de 11,9%, e no modelo ARMA foi de 15,2%. A região 5, litoral da região sudeste, teve o maior valor estimado, com 6,6% (Gama) e 9,2% (ARMA). A região 12 teve aproximadamente 3,5% de contribuição nos dois modelos, ao passo que, a região 6 teve aproximadamente 2%.

As regiões 10 e 11, de massas provenientes principalmente de fontes continentais, apresentaram estimativa de 2,9% no modelo Gama, e 3,0% no modelo ARMA(2,1). Deste total das duas regiões, o maior percentual é da região 10, com 1,9% e 2,0% nos modelos Gama e ARMA, respectivamente.

Por fim, restam as regiões 7 e 8. A região 7 (Estados da Bahia, Goiás, Distrito Federal, Mato Grosso, Tocantins) corresponde a uma estimativa de contribuição de 6,1% no modelo Gama, e 7,3% no modelo ARMA. A região 8 (Paraguai, Mato Grosso do Sul, e oeste do Paraná e Santa Catarina) apresentou estimativa de 3,0% no modelo Gama, e 1,5% no modelo ARMA(2,1). Com isso, o percentual estimado das duas regiões foi de 9,0% e 8,8% nos modelos Gama e ARMA(2,1), respectivamente.

Estes percentuais, representam uma contribuição média do período de 2014 a 2015. Tanto para o MP_{10} , quanto para o $MP_{2,5}$, nota-se que as regiões 1 e 2 foram a de maior contribuição. A Figura 4.23 mostra o mapa de densidade kernel dos pontos de coordenadas das trajetórias. Emissões do porto, assim como movimentação de caminhões em direção a Cubatão e Santos, tem influência no material particulado medido em Santos (CETESB, 2015a; CETESB, 2018). Processos de combustão, como queima de combustíveis nos veículos tem influência na composição do $MP_{2,5}$ em Santos (CETESB, 2015a).

Os dados do mês de Janeiro de 2016 da concentração do $MP_{2,5}$ e os tempos de residência de cada região, na cidade de Santos, foram utilizados para avaliar a previsão do modelo ARMA(2,1). Na Figura 4.56 são mostrados os valores reais da concentração do $MP_{2,5}$ comparados com os valores previstos pelo modelo. Note que as concentrações previstas conseguiram acompanhar as concentrações reais ao longo do mês, indicando um bom poder de previsão.

Novamente, a contribuição média do período de 2014 e 2015 foi bem similar entre os dois modelos, ou seja, há pequenas diferenças nas estimativas das contribuições, conforme visto nas Figuras B.15 e 4.52. Todavia, ao comparar as estimativas diárias da concentração do $MP_{2,5}$ nos dois modelos com o valor real da concentração, observa-se que o modelo

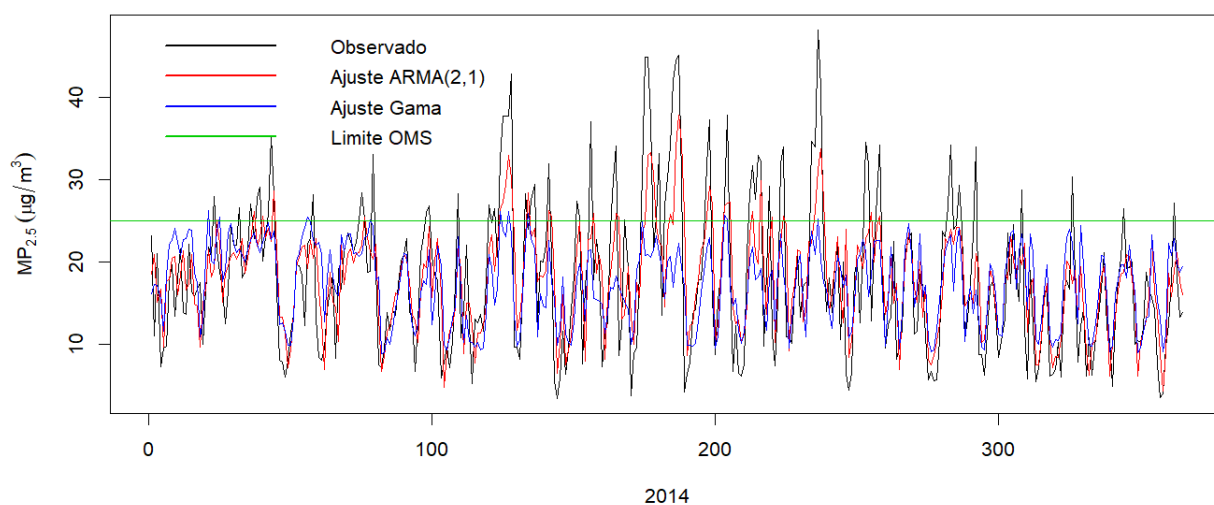


Figura 4.54: Ajuste da concentração diária do $MP_{2,5}$ em 2014 em comparação com valores observados - Santos

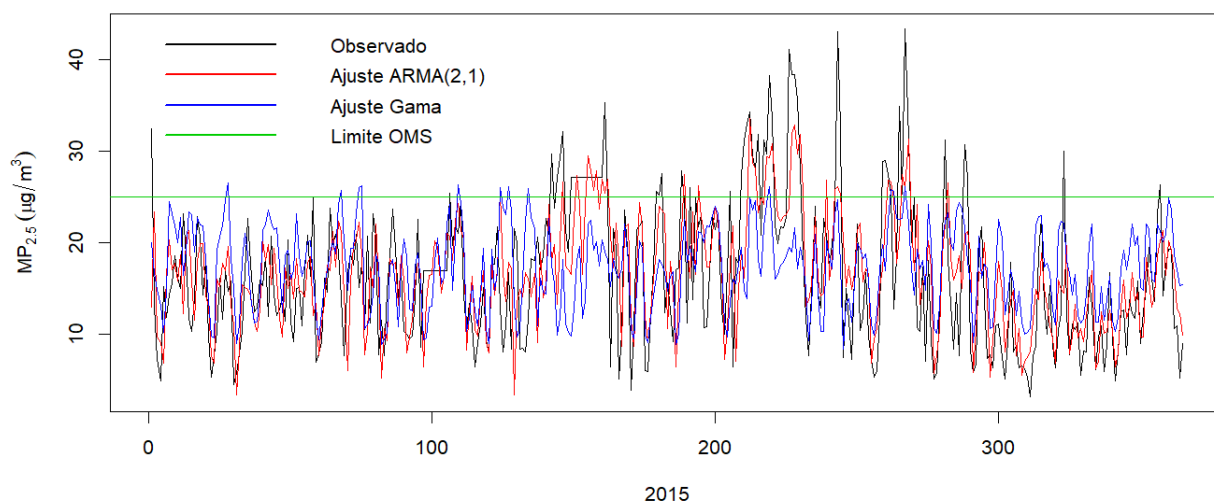


Figura 4.55: Ajuste da concentração diária do $MP_{2,5}$ em 2015 em comparação com valores observados - Santos

ARMA(2,1) é melhor que o modelo Gama. As Figuras 4.54 e 4.55 apresenta os ajustes diários dos dois modelos e o valor observado nos anos de 2014 e 2015, respectivamente. Além disso, nas duas figuras o limite estabelecido pela OMS é traçado. Apenas o modelo ARMA(2,1) consegue acompanhar a série real, inclusive detectando dias com valores bem acima dos limites. Por outro lado, o ajuste do modelo Gama, apenas acompanha a série em torno de uma média geral. Por isso, que na média os dois modelos são bem parecidos, mas ao se perceber diariamente, há ganhos razoáveis em se utilizar o modelo ARMA. Assim, a AC e ACP foi realizada baseada nos resultados obtidos por esse modelo.

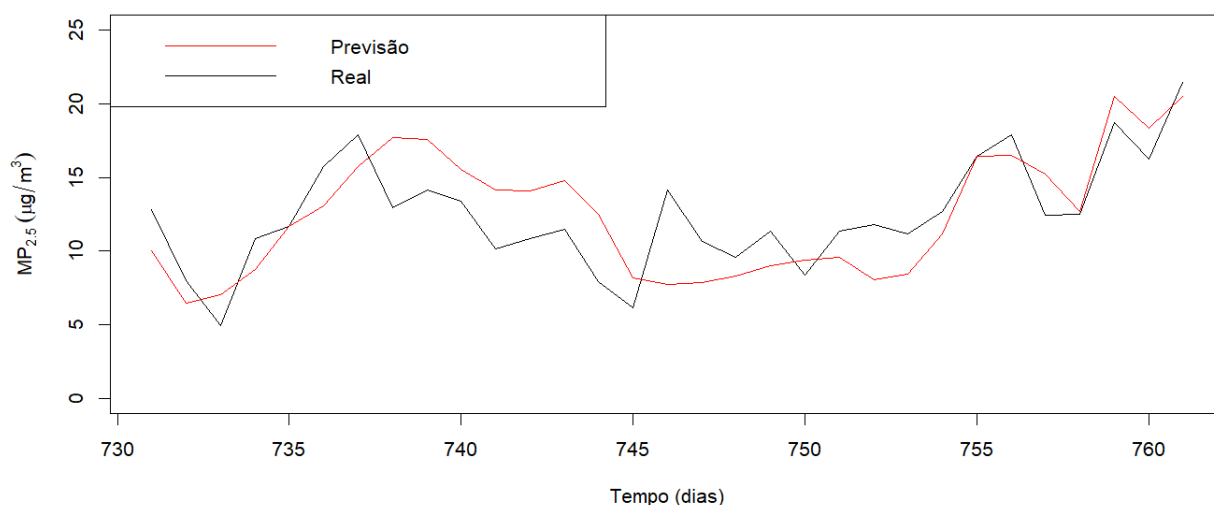


Figura 4.56: Previsão do $MP_{2,5}$ e valores reais para o mês de Janeiro de 2016 - Santos

Com as estimativas diárias obtidas do modelo ARMA(2,1) para cada uma das regiões significativas a 5% (exceção a região 9), e também com as médias diárias das três variáveis meteorológicas foi feita a AC, a fim de obter os grupos. Dessas variáveis, a AC obteve quatro grupos de dias, escolhidos de acordo com o dendograma (Figura C.4). O grupo 1 foi formado por 142 dias, o grupo 2 por 261 dias, o grupo 3 por 203 dias, e o grupo 4 foi formado por 124 dias.

A Tabela 4.20 apresenta o percentual de dias dentro de cada grupo, em cada estação do ano, e a Tabela 4.21 mostra o percentual de cada grupo em cada uma das regiões que contribuíram na composição da concentração do $MP_{2,5}$ em Santos durante os anos de 2014 e 2015.

Tabela 4.20: Percentual de dias dentro de cada estação do ano por grupo ($MP_{2,5}$ - Santos)

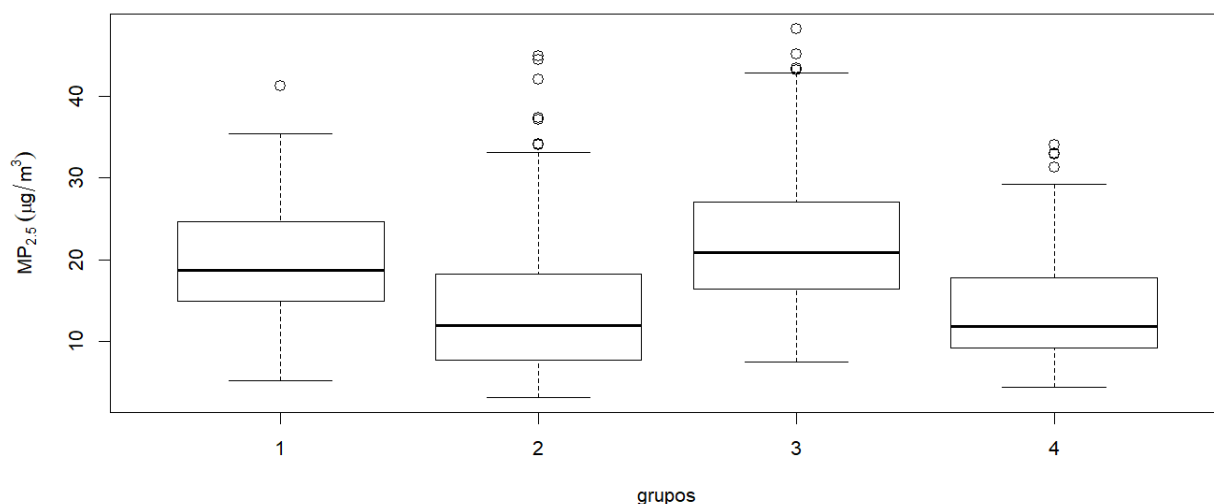
	Primavera	Verão	Outono	Inverno
Grupo 1	12,0%	62,7%	12,7%	12,7%
Grupo 2	27,2%	15,7%	25,3%	31,8%
Grupo 3	26,6%	14,8%	31,0%	27,6%
Grupo 4	32,3%	16,1%	29,8%	21,8%

Nas Figuras 4.57, 4.58, 4.59 e 4.60 tem-se a distribuição das medições de $MP_{2,5}$, temperatura, umidade relativa e velocidade do vento, respectivamente, por meio dos gráficos de box-plot para cada um dos 4 grupos obtidos na AC.

O resumo da ACP, por grupo, é mostrado na Tabela 4.22, em que se tem a proporção da variância explicada por cada CP. Na ACP, considerou-se as variáveis meteorológicas medidas diariamente e a concentração diária do $MP_{2,5}$. Em todos os grupos foram analisadas as duas primeiras CP, pois estas acumularam mais de 70% da variabilidade total

Tabela 4.21: Percentual de contribuição do grupo dentro de cada região de residência da massa de ar ($MP_{2,5}$ - Santos)

	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Região 1	28,0%	14,1%	52,2%	5,7%
Região 2	28,7%	56,1%	11,8%	3,3%
Região 3	4,1%	51,5%	6,1%	38,3%
Região 4	6,4%	53,3%	11,1%	29,2%
Região 5	14,3%	24,4%	54,5%	6,8%
Região 6	67,0%	7,1%	25,8%	0,0%
Região 7	45,6%	31,3%	21,4%	1,8%
Região 8	9,8%	46,5%	0,2%	43,5%
Região 10	0,2%	54,1%	1,3%	44,3%
Região 11	0,0%	57,2%	0,0%	42,8%
Região 12	0,7%	44,6%	9,7%	45,1%

Figura 4.57: Box-plot do $MP_{2,5}$ ($\mu\text{g}/\text{m}^3$) por grupo - $MP_{2,5}$ - Santos

dos dados. O grupo 2 acumulou 75,2% com as duas primeiras CPs, e o grupo 4, acumulou 70,8% com as duas primeiras CPs. Estas foram a maior e menor proporção acumulada entre os grupos (Tabela 4.22).

Nas Figuras 4.61, 4.62, 4.63 e 4.64 as variáveis Temperatura, Umidade Relativa e Velocidade do Vento foram abreviadas usando Temp, UR e VV. Estas Figuras destacam os coeficientes obtidos na ACP, para a CP1 e CP2 após a rotação varimax, dentro de cada um dos 4 grupos.

O primeiro grupo, grupo 1, é formado por 142 dias, sendo predominantemente na estação do verão, com 62,7% (Tabela 4.20), e por causa disto, há os maiores valores de temperatura (Figura 4.58). Os dias deste grupo apresentaram o segundo maior valor da

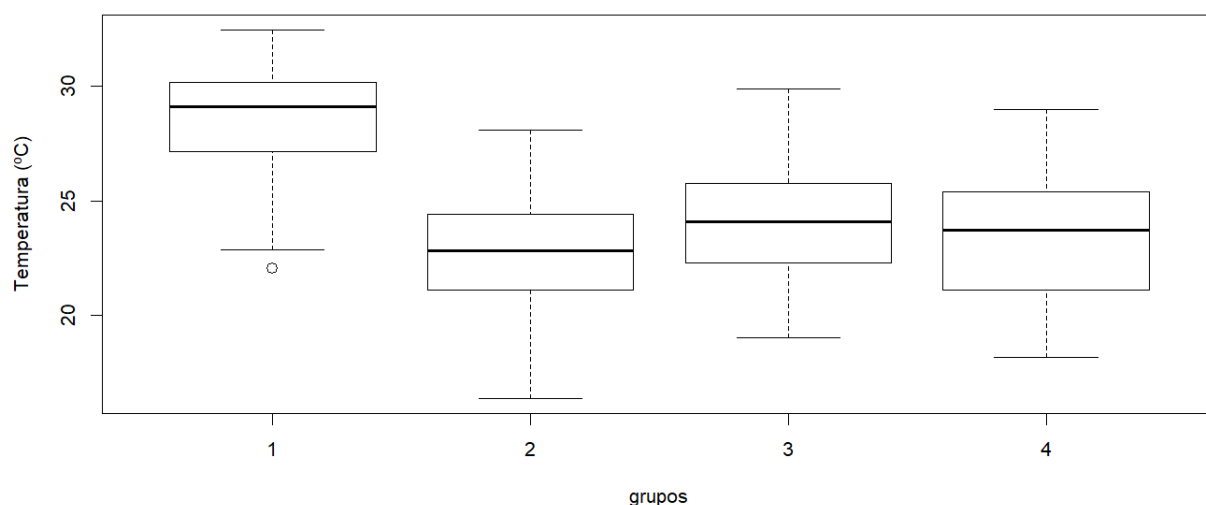


Figura 4.58: Box-plot do Temperatura (°C) por grupo - $MP_{2,5}$ - Santos

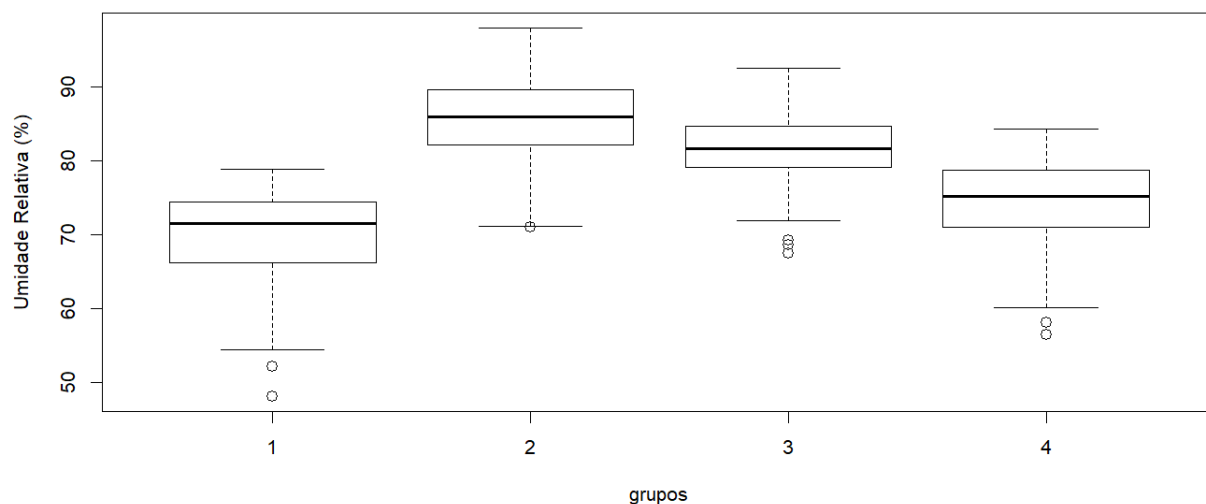
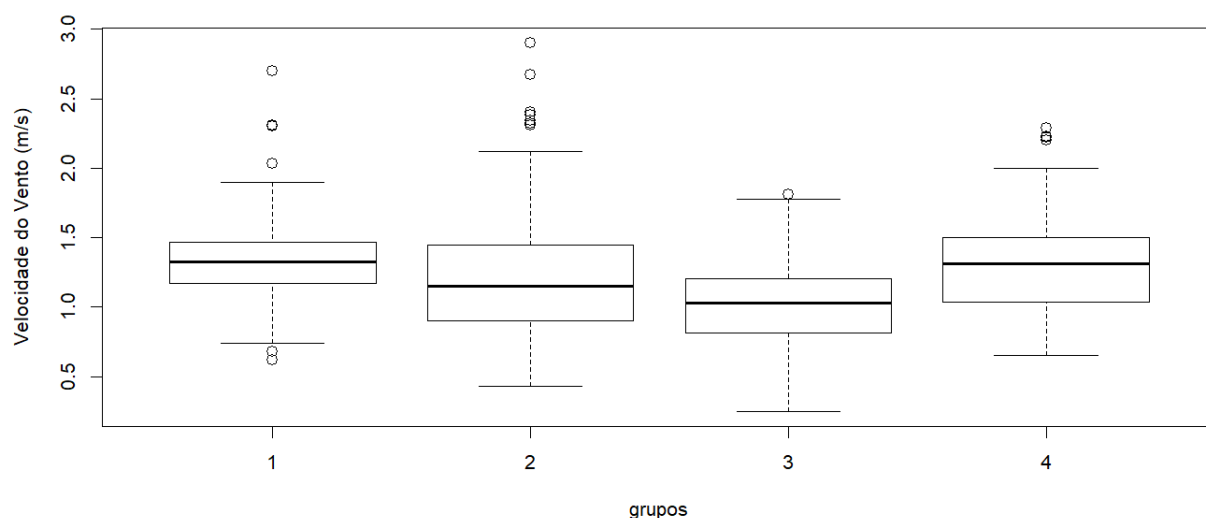


Figura 4.59: Box-plot da Umidade Relativa (%) por grupo - $MP_{2,5}$ - Santos

mediana do MP (Figura 4.57), provavelmente influenciado por valores baixos de UR (Figura 4.59), mesmo com o segundo nível mediano de velocidade do vento (Figura 4.60). A CP1 indicou a maior relação da temperatura e velocidade do vento inversamente proporcionais com a concentração do $MP_{2,5}$ (Figura 4.61), com 49% da variabilidade explicada (Tabela 4.22). A CP2 mostrou destacada apenas a umidade relativa, com 24,6% da variabilidade explicada (Tabela 4.22). As regiões de maior contribuição foram as 6 e 7 (Tabela 4.21), correspondentes, principalmente, pelo nordeste e seu litoral (Figura 4.18 e Tabela 4.7).

O grupo 2, formado por 261 dias, é o maior dentre os quatro grupos, sendo que os dias foram predominantemente do inverno e primavera (Tabela 4.20). A umidade relativa foi

Figura 4.60: Box-plot da Velocidade do Vento (m/s) por grupo - $MP_{2,5}$ - SantosTabela 4.22: Resumo da ACP - $MP_{2,5}$ - Santos

Grupo 1	CP1	CP2	CP3	CP4
Proporção da variabilidade	0,490	0,246	0,180	0,084
Proporção acumulada	0,490	0,736	0,916	1,000
Grupo 2				
Proporção da variabilidade	0,394	0,357	0,156	0,093
Proporção acumulada	0,394	0,752	0,907	1,000
Grupo 3				
Proporção da variabilidade	0,444	0,270	0,171	0,115
Proporção acumulada	0,444	0,714	0,885	1,000
Grupo 4				
Proporção da variabilidade	0,444	0,264	0,204	0,088
Proporção acumulada	0,444	0,708	0,912	1,000

a maior dos quatro grupos (Figura 4.59), o que auxiliou no menor nível mediano de $MP_{2,5}$ (Figura 4.57). A velocidade do vento também apresentou altos valores (Figura 4.60), que auxilia na dispersão do MP. As CP1 e CP2 tiveram percentuais da variabilidade explicada muito próximos, com 39,4% e 35,7% (Tabela 4.22). A CP1 apresentou a relação inversa entre o $MP_{2,5}$ e velocidade do vento, enquanto que a CP2 mostrou a relação inversa da temperatura e umidade relativa (Figura 4.62). Várias regiões tiveram contribuição na composição do $MP_{2,5}$ neste grupo, como as regiões de fontes locais e litorâneas (2, 3 e 4), fontes continentais e litorâneas (8, 10, 11 e 12), conforme observado na Tabela 4.21. Essas regiões correspondem a massas de ar vindas desde o oeste até o sudeste de Santos (Figura 4.18).

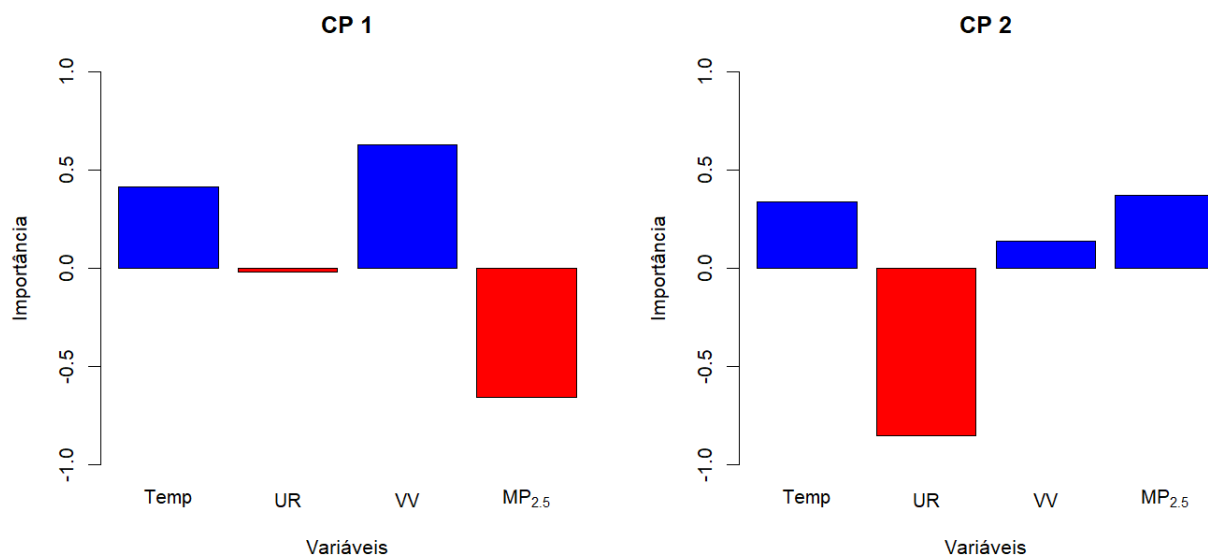


Figura 4.61: Importância das CP 1 e CP 2 no grupo 1 - $MP_{2,5}$ - Santos

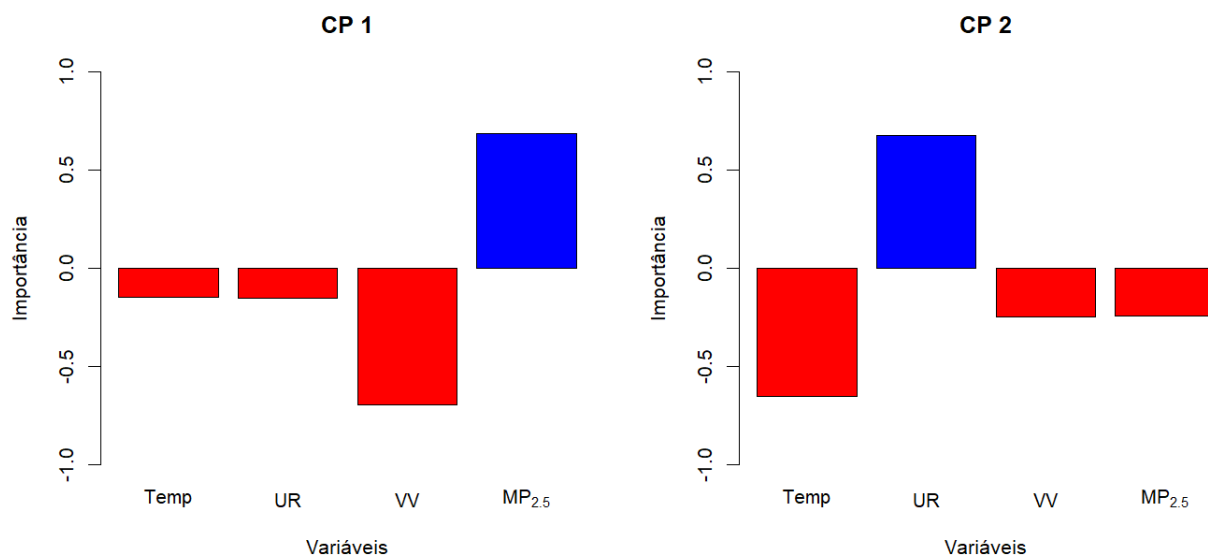


Figura 4.62: Importância das CP 1 e CP 2 no grupo 2 - $MP_{2,5}$ - Santos

O grupo 3 foi formado por 203 dias, com predominância dos dias nas estações do outono e inverno (Tabela 4.20). Devido a isso, provavelmente, tem-se os maiores níveis de concentração do $MP_{2,5}$ neste grupo (Figura 4.57), e os menores valores de temperatura (Figura 4.58). A CP1 indicou a relação inversa do $MP_{2,5}$ com temperatura e velocidade do vento (Figura 4.63), com 44,4% (Tabela 4.22). A umidade relativa foi o fator destacado na CP2, e ficou inversamente proporcional as demais variáveis (Figura 4.63). As regiões 1 e 5 mais contribuem neste grupo (Tabela 4.21), com massas de ar provenientes do litoral do sudeste, interior de Minas Gerais, Espírito Santo e parte do Rio de Janeiro (Figura 4.18).

Por fim, o grupo 4 foi formado por 124 dias, o menor grupo. As estações predominantes foram primavera e outono (Tabela 4.20). As regiões 8, 10, 11 e 12 foram a de maior contri-

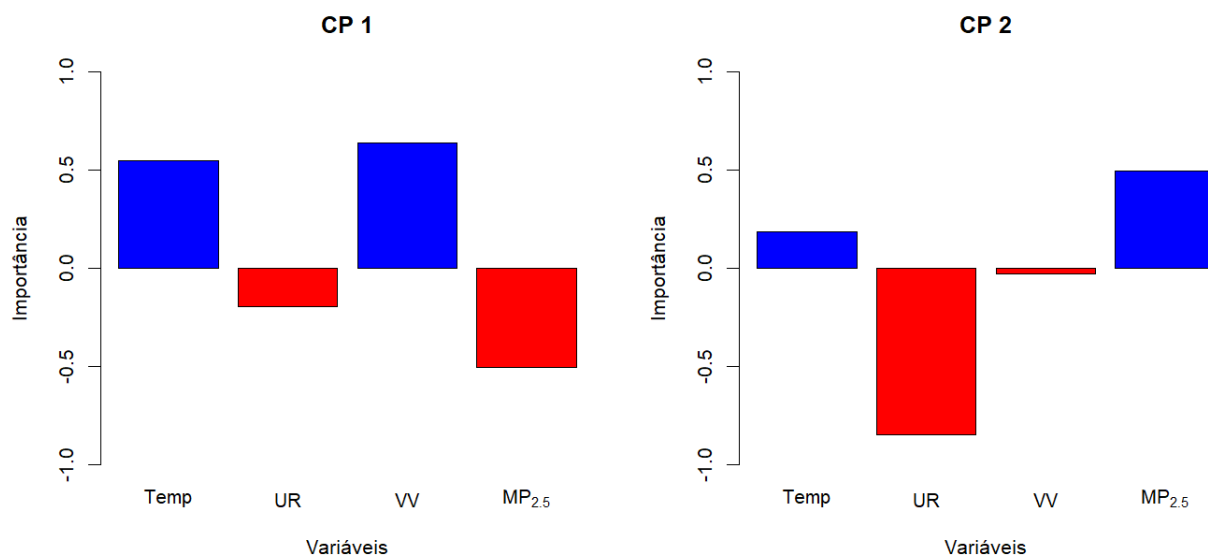


Figura 4.63: Importância das CP 1 e CP 2 no grupo 3 - $MP_{2,5}$ - Santos

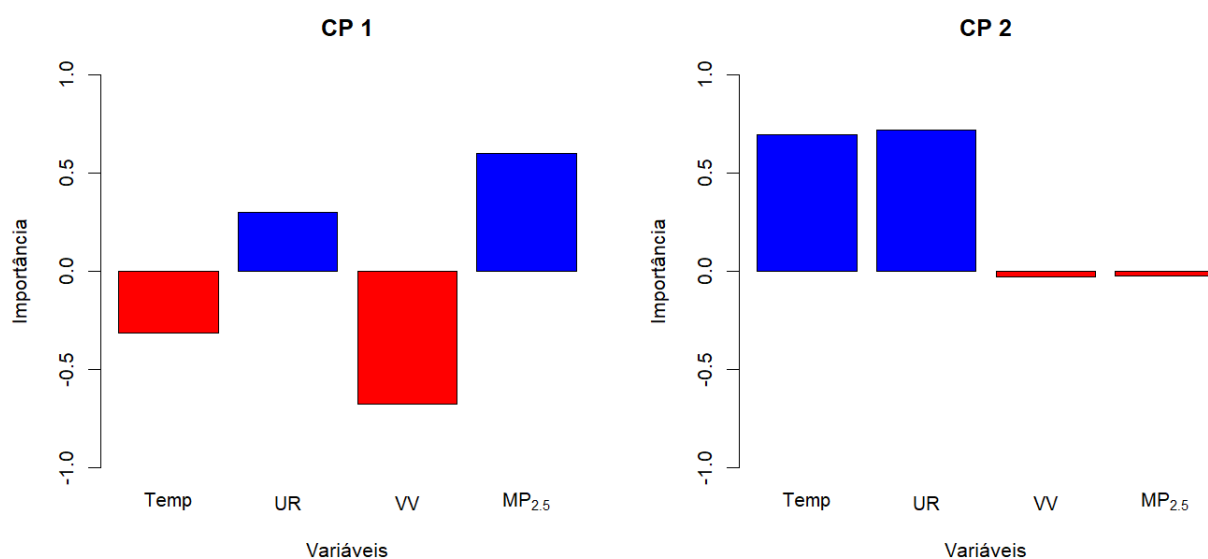


Figura 4.64: Importância das CP 1 e CP 2 no grupo 4 - $MP_{2,5}$ - Santos

buição (Tabela 4.21), com influências continentais e litorâneas, do oeste ao sul de Santos (Figura 4.18 e Tabela 4.7). Este grupo apresenta os menores níveis de $MP_{2,5}$ (Figura 4.57) e a segunda mediana da velocidade do vento, entre os quatro grupos (Figura 4.60). A CP1 também aponta esta relação, inversamente proporcional entre a velocidade do vento e $MP_{2,5}$ (Figura 4.64), com 44,4% da variabilidade explicada por ela (Tabela 4.22). Mais uma vez, nota-se que o vento atua dispersando o $MP_{2,5}$. A CP2, com 26,4% (Tabela 4.22), indicou temperatura e umidade relativa proporcionais, provavelmente devido a estação da primavera, mais quente e úmida.

De modo geral, a AC conseguiu, novamente, agrupar regiões próximas num mesmo grupo, indicando que a estrutura espacial dos dados é relevante para análise do $MP_{2,5}$.

Para o $MP_{2,5}$, a velocidade do vento é a variável mais relevante, pois ficou numa mesma CP, com importância alta, em todos os quatro grupos. Isto pode ser observado porque partículas de $MP_{2,5}$ são mais leves e sofrem maior influência dos ventos, podendo percorrer grandes distâncias até se depositarem (YNOUE et al., 2017).

4.3 Comparação dos resultados de Piracicaba e Santos

As regiões locais foram as de maior contribuição nos níveis dos dois MPs, nas duas cidades. Em Piracicaba, as fontes locais contribuíram 75% para o MP₁₀ e 76,5% para o MP_{2,5}. Em Santos, as fontes locais contribuíram 70% para o MP₁₀ e 72,8% para o MP_{2,5}. Fontes de regiões continentais e litorâneas também contribuem, porém numa menor proporção. As massas de ar predominantemente tem origem na direção nordeste tanto de Piracicaba (Figura 4.6), quanto de Santos (Figura 4.23). Os resultados das possíveis fontes foram diferentes para as duas cidades. Piracicaba, localizada no interior, sofre maior influência da queima de cana-de-açúcar, além de emissões industriais e veiculares, principalmente os movidos a diesel (CETESB, 2018). Cidades como Rio Claro e Santa Gertrudes (polo cerâmico), Paulínia (polo petroquímico) também influenciam nos níveis do MP em Piracicaba (CETESB, 2018). Santos, no litoral, sofre com emissões vindas do porto, e também de emissões industriais provenientes do polo de Cubatão. Processos de combustão, de emissões veiculares, também tem importante influência na composição do MP, e, dada a proximidade com o mar, aerossóis marinhos também é uma das fontes do MP de Santos (CETESB, 2015a).

Em Piracicaba, a umidade relativa interfere mais diretamente no MP, de modo que, conforme maior a umidade, menor os níveis do MP. Em Santos, a velocidade do vento interfere mais diretamente no MP, também inversamente, ou seja, conforme aumenta-se a velocidade do vento, diminui-se os níveis do MP. Esse fato pode estar ligado as diferentes condições climáticas de cada região. Nakazato et al. (2018) apresentou a descrição dos climas de Cubatão (similar a Santos) e Campinas (similar a Piracicaba). Cubatão caracteriza-se por umidade relativa altas durante o ano inteiro, e ausência de período de seca, e altas precipitações, enquanto que a Região de Campinas, apresenta períodos de frio e seca (entre abril e setembro) e quente e úmido (entre outubro e março) bem definidos.

Note os gráficos dos boxplot, para a umidade relativa em Piracicaba (Figuras 4.12 e 4.44) e em Santos (Figuras 4.29 e 4.59). Piracicaba tem valores medianos variando de 50 a 80, enquanto Santos varia entre 70 e 80. Santos praticamente apresenta umidade relativa altas o ano inteiro, diferente de Piracicaba, que apresenta períodos de seca.

Em relação a velocidade do vento, Santos apresenta menores valores se comparado com Piracicaba. Esses períodos de calmaria, dificultam a dispersão dos poluentes (CETESB, 2015a). Os boxplots da velocidade do vento para Piracicaba (Figuras 4.13 e 4.45) mostra valores medianos em torno de 2 m/s ou maior. Já em Santos (Figuras 4.30 e 4.60), os valores medianos ficam próximos ou abaixo de 1,5 m/s. Além disso, em Santos, para cada dia, foi calculado a classe de estabilidade de acordo com as classes de Pasquill (SEINFELD; PANDIS, 2006; MOREIRA; TIRABASSI; MORAES, 2008). O cálculo das classes de Pasquill foi baseado em dados de radiação solar (MOREIRA; TIRABASSI; MORAES, 2008) e cobertura de nuvens (ZHAO; BLACK; BALDWIN, 1997), para os períodos diurno e noturno, respectivamente. Considerando o período diurno, dos 730 dias da amostra, 516 dias foram considerados como instáveis ou levemente instáveis, condições que podem dificultar a dispersão do MP. Já para o período noturno, todos os dias foram classificados como neutro ou estáveis.

Capítulo 5

Conclusões

Os modelos ajustados neste trabalho puderam identificar as contribuições de cada região na concentração do MP_{10} e $MP_{2,5}$ nas cidades de Piracicaba e Santos com uso da Análise de Regressão de Trajetória. Tais contribuições calculadas de cada região, refletem a contribuição após a ocorrência de todas as prováveis perdas como deposição úmida, dispersão do vento, diluição química, entre outras.

Nota-se que os modelos lineares generalizados (normal inversa e gama) e modelos generalizados de locação, escala e forma (log-normal) obtiveram um melhor ajuste se comparado ao modelo de regressão múltipla normal. Entretanto, os modelos de séries temporais, baseados na metodologia de Box-Jenkins, os modelos ARMA, foram melhores que os modelos GLM e GAMLSS. Todos os trabalhos anteriores utilizaram do modelo normal na ART. A principal contribuição deste trabalho pretende abordar a opção dos modelos ARMA para a ART. Os resultados apresentados das medidas comparativas e da análise de resíduos indicaram essa opção como satisfatória.

As regiões locais foram as que mais contribuíram na composição dos dois MPs estudados, tanto em Piracicaba, quanto em Santos. Fontes de regiões continentais e litorâneas também contribuíram, entretanto, num menor percentual. As trajetórias das massas de ar indicam origem na direção nordeste nas duas cidades. As possíveis fontes foram diferentes para as duas cidades. Piracicaba, no interior, sofre maior influência da queima de cana-de-açúcar, emissões industriais e veiculares, principalmente os movidos a diesel. Os polos cerâmicos das cidades de Rio Claro e Santa Gertrudes, e o polo industrial petroquímico de Paulínia também influenciam nos níveis do MP em Piracicaba. Santos, no litoral, por sua vez, sofre com emissões provenientes do porto e emissões industriais vindas do polo de Cubatão. Os processos de combustão de emissões veiculares, também têm influência nos níveis do MP, e, pela localização de Santos, próxima ao mar, aerossóis marinhos também é uma das possíveis fontes do MP.

A AC foi interessante para identificar grupos de dias com comportamentos similares em relação as variáveis meteorológicas (temperatura, umidade relativa e velocidade do vento), e a contribuição do MP de cada região pré-determinada. A ACP identificou a relação do MP com as variáveis meteorológicas mais relevantes dentro de cada grupo obtido.

De modo geral, a AC agrupou dias com contribuições maiores provenientes de regiões próximas. Além disso, cada grupo foi caracterizado por alguma estação do ano predominante. Em Piracicaba, a umidade relativa interfere mais diretamente no MP, de modo

que, conforme maior a umidade relativa, menores os níveis do MP. Já em Santos, a velocidade do vento interfere mais diretamente no MP, de maneira inversa, ou seja, conforme aumenta a velocidade do vento, diminui a concentração do MP. Esse fato pode estar ligado as diferentes condições climáticas das duas cidades. Santos praticamente apresenta umidade relativa altas o ano inteiro, diferente de Piracicaba, que apresenta períodos de seca. Em relação a velocidade do vento, Santos apresenta menores valores se comparado com Piracicaba. Esses períodos de baixa velocidade do vento, desfavorecem a dispersão dos poluentes.

Neste trabalho, modelos lineares generalizados (GLM e GAMLSS) e modelos ARMA foram abordados separadamente. Uma das sugestões de trabalhos futuros é o uso de modelos lineares generalizados com séries temporais, num único modelo (BENJAMIN; RIGBY; STASINOPOULOS, 2003). Cordeiro e Andrade (2009) apresentam este tipo de modelo, chamado de ARMA generalizado (ou GARMA), com aplicações em dados reais seguindo distribuições Gama e Normal Inversa.

A ART também poderia ser abordada separadamente por estações do ano. Neste trabalho, as estações foram identificadas apenas após a AC, com os grupos obtidos. Uma sugestão seria separar os dias dentro das estações do ano, e depois aplicar a ART nesses dias. Uma desvantagem, é que não se teria mais uma série temporal (considerando-se que seriam dados de vários anos), e neste caso, a metodologia de Box-Jenkins não seria mais aplicável. Porém, os modelos lineares generalizados continuariam adequados, assim como o modelo normal.

Outra sugestão de trabalhos futuros é na abrangência de outras cidades, além de Piracicaba e Santos. Por exemplo, a CETESB possui diversas estações de monitoramento automáticas de MP na cidade de São Paulo, além de outras no interior, ou mesmo Cubatão, cidade próxima de Santos conhecida por graves problemas de poluição (ANDRADE et al., 2017).

Capítulo 6

Atividades Desenvolvidas e Perspectivas Futuras

Até o momento foram desenvolvidas algumas atividades científicas durante o período do doutorado, como publicação de um artigo, participação em eventos científicos e auxílio na orientação de três alunas de Iniciação Científica (IC) no ano de 2016, uma aluna no ano de 2017 e uma aluna no ano de 2018.

O artigo “Chemical Composition of Rainwater in an Urban Area of the Southeast of Brazil” foi publicado no Journal “Atmospheric Pollution Research” no ano de 2018 (<<https://doi.org/10.1016/j.apr.2018.10.003>>). Além disso, o artigo “Um novo método de preenchimento de dados faltantes aplicado a séries temporais de concentração de MP₁₀” foi aceito para publicação na “Revista Tecnologia e Sociedade”, também em 2018.

No XXI Congresso Brasileiro de Química (COBEQ-2016) realizado em Fortaleza-CE, foi apresentado o trabalho: “Estudo de relação entre variáveis meteorológicas e poluentes atmosféricos usando técnicas multivariadas”. Em Vitória-ES, foi apresentado em forma de poster, o trabalho: Study of diurnal pattern of air quality using Cluster Analysis, no Air Quality Conference Brazil 3rd CMAS (Community Modeling and Analysis System) South America-2017.

Em Maringá-PR, no Congresso Brasileiro de Sistemas Particulados (XXXVIII ENEMP-2017) foi apresentado o trabalho: “Estudo da Contribuição Regional nos Níveis de MP₁₀ em Piracicaba usando Análise de Regressão de Trajetória”. Ainda neste evento, o trabalho intitulado “Modelagem matemática da dispersão de Material Particulado (MP₁₀) oriundo da indústria cerâmica” foi apresentado em forma de poster.

No XIV SIBESA (Simpósio Ítalo Brasileiro de Engenharia Sanitária e Ambiental), realizado em Junho de 2018 na cidade de Foz do Iguaçu, foi apresentado o trabalho, em formato de poster, intitulado: “Estudo da contribuição regional nos níveis do MP₁₀ em Piracicaba usando análise multivariada”.

Do auxílio na orientação da IC, geraram-se 4 trabalhos. Dois foram publicados no XIV Encontro Nacional de Estudantes de Engenharia Ambiental (ENEEAmb-2016), em Brasília-DF. Outro foi publicado na XIII Semana da Engenharia Ambiental (SEA-2016), em São Carlos-SP, e o terceiro foi publicado no Congresso Brasileiro de Sistemas Particulados (XXXVIII ENEMP-2017), em Maringá-PR. Além disso, dois trabalhos da IC foram apresentados no XXIV Congresso de Iniciação Científica da UNICAMP, em 2016, e um

foi apresentado no XXV Congresso de Iniciação Científica da Unicamp, em 2017, sendo os dois eventos realizado em Campinas-SP.

Por fim, pretende-se publicar os resultados desta tese em periódico indexado, além de participar de outros eventos científicos para divulgação dos mesmos.

6.1 Artigos publicados

MARTINS, E. H.; **NOGAROTTO, D. C.**; MORTATTI, J.; POZZA, S. A. Chemical Composition of Rainwater in an Urban Area of the Southeast of Brazil. *Atmospheric Pollution Research, Fator de Impacto* (2017 JCR): 2,152, 2019, v. 10, p. 520-530. DOI: <<https://doi.org/10.1016/j.apr.2018.10.003>>.

6.2 Artigos aceitos para publicação

NOGAROTTO, D. C.; RISSI, N.M.; POZZA, S. A. Um novo método de preenchimento de dados faltantes aplicado a séries temporais de concentração de MP10. *Revista Tecnologia e Sociedade*, 2018. (ISSN: 1984-3526).

6.3 Trabalhos publicados em anais de eventos (completo)

NOGAROTTO, D. C.; POZZA, S. A. Estudo da contribuição regional nos níveis do MP10 em Piracicaba usando análise multivariada In: XIV - Simpósio Ítalo Brasileiro de Engenharia Sanitária e Ambiental, 2018, Foz do Iguaçu-PR.

NOGAROTTO, D. C.; POZZA, S. A. Estudo da contribuição regional nos níveis do MP10 em Piracicaba usando Análise de Regressão de Trajetória In: ENEMP 2017, 2017. Maringá-PR.

SOUSA, J. B.; ROSSINI, A. J.; **NOGAROTTO, D. C.**; POZZA, S. A. Modelagem matemática na dispersão de Material Particulado (MP10) oriundo da indústria cerâmica In ENEMP 2017, 2017, Maringá-PR.

NOGAROTTO, D. C.; POZZA, S. A. Study of diurnal pattern of air quality using Cluster Analysis In: Air Quality Conference Brazil - 3RD CMAS SOUTH AMERICA, 2017, Vitória-ES-Brasil.

LIMA, M. R. G.; **NOGAROTTO, D. C.**; POZZA, S. A. Uso de Análise de Correlação Canônica: Poluentes Atmosféricos x Variáveis Meteorológicas In: ENEMP 2017, 2017, Maringá-PR.

NOGAROTTO, D.C.; POZZA, S. A. Estudo de relação entre variáveis meteorológicas e poluentes atmosféricos usando técnicas multivariadas. In: XXI Congresso Brasileiro de Engenharia Química (COBEQ), 2016, Fortaleza-CE.

Referências Bibliográficas

AGARWAL, N.; NAGENDRA, S. M. S. Modelling of particulate matter distribution inside the multilevel urban classrooms in tropical climate for exposure assessment. *Building and Environment*, v. 102, p. 73–82, 2016. <<http://dx.doi.org/10.1016/j.buildenv.2016.03.015>>.

ALVARES, C. A. et al. Koppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, v. 22, n. 6, p. 711–728, 2013. <<https://doi.org/10.1127/0941-2948/2013/0507>>.

ANDRADE, M. et al. Ozone sounding in the Metropolitan Area of São Paulo, Brazil: Wet and dry season campaigns of 2006. *Atmospheric Environment*, v. 61, p. 627–640, 2012a. <<https://doi.org/10.1016/j.atmosenv.2012.07.083>>.

ANDRADE, M. et al. Air quality in the megacity of São Paulo: Evolution over the last 30 years and future perspectives. *Atmospheric Environment*, v. 159, p. 66–82, 2017. <<http://dx.doi.org/10.1016/j.atmosenv.2017.03.051>>.

ANDRADE, M. F. et al. Vehicle emissions and PM_{2.5} mass concentrations in six Brazilian cities. *Air Quality, Atmosphere and Health*, v. 5, p. 79–88, 2012b. <<https://doi.org/10.1007/s11869-010-0104-5>>.

AUSTIN, E. et al. A framework for identifying distinct multipollutant profiles in air pollution data. *Environment International*, v. 45, p. 112–121, 2012. <<https://doi.org/10.1016/j.envint.2012.04.003>>.

AUSTIN, E. et al. A framework to spatially cluster air pollution monitoring sites in US based on the PM_{2.5} composition. *Environment International*, v. 59, p. 244–254, 2013. <<https://doi.org/10.1016/j.envint.2013.06.003>>.

BANERJEE, T.; SINGH, S. B.; SRIVASTAVA, R. K. Development and performance evaluation of statistical models correlating air pollutants and meteorological variables at Pantnagar, India. *Atmospheric Research*, v. 99, p. 505–517, 2011. <<https://doi.org/10.1016/j.atmosres.2010.12.003>>.

BASSAB, W.; MIAZAKI, E.; ANDRADE, D. *Introdução a Análise de Agrupamentos*. São Paulo, Brasil: 9 Simpósio Brasileiro de Probabilidade e Estatística, 1990.

BENJAMIN, M. A.; RIGBY, R. A.; STASINOPOULOS, D. M. Generalized Autoregressive Moving Average Models. *Journal of the American Statistical Association*, v. 98, n. 461, p. 214–223, 2003. <<http://dx.doi.org/10.1198/016214503388619238>>.

- BINAKU, K.; SCHMELING, M. Multivariate statistical analyses of air pollutants and meteorology in Chicago during summers 2010-2012. *Air Quality, Atmosphere & Health*, v. 10, p. 1227–1236, 2017. <<http://dx.doi.org/10.1007/s11869-017-0507-7>>.
- BRASIL. *Conselho Nacional do Meio Ambiente (CONAMA) - Resolução 491*. Brasília: Ministério do Meio Ambiente, 2018. Disponível em: <<<http://www2.mma.gov.br/port/conama/legiabre.cfm?codlegi=740>>>.
- BREITNER, S. et al. Short-term effects of air temperature on mortality and effect modification by air pollution in three cities of Bavaria, Germany: A time-series analysis. *Science of the Total Environment*, v. 485-486, p. 49–61, 2014. <<https://doi.org/10.1016/j.scitotenv.2014.03.048>>.
- CANTERAS, F.; MOREIRA, S.; FARIA, B. D. Study of main and signature sources of particulate matter pollutants in Limeira City (Brazil) using SR-TXRF. *X-Ray Spectrometry*, v. 42, p. 290–298, 2013. <<https://doi.org/10.1002/xrs.2468>>.
- CARMO, C. et al. Associação entre material particulado de queimadas e doenças respiratórias na região sul da Amazônia brasileira. *Revista Panamericana de Salud Pública*, v. 27, p. 6–10, 2010. <<https://doi.org/10.1590/s1020-49892010000100002>>.
- CETESB. *Qualidade de ar no Estado de São Paulo*. 2014. Acesso em 14 de Março de 2019. Disponível em: <<https://cetesb.sp.gov.br/ar/publicacoes-relatorios/>>.
- CETESB. *Avaliação da qualidade do ar no município de Santos 2011 a 2014*. São Paulo, 2015. Disponível em: <<https://cetesb.sp.gov.br/ar/publicacoes-relatorios/>>.
- CETESB. *Qualidade de ar no Estado de São Paulo*. 2015. Acesso em 14 de Março de 2019. Disponível em: <<https://cetesb.sp.gov.br/ar/publicacoes-relatorios/>>.
- CETESB. *Companhia de Tecnologia de Saneamento Ambiental*. 2016. Acesso em 11 de Janeiro de 2016. Disponível em: <<<http://www.cetesb.sp.gov.br/ar/qualidade-do-ar/32-qualar>>>.
- CETESB. *Emissões veiculares no estado de São Paulo 2015*. São Paulo, 2016. Disponível em: <<https://cetesb.sp.gov.br/veicular/relatorios-e-publicacoes/>>.
- CETESB. *Qualidade do ar no Estado de São Paulo 2017*. São Paulo, 2018. Disponível em: <<https://cetesb.sp.gov.br/ar/wp-content/uploads/sites/28/2018/05/relatorio-qualidade-ar-2017.pdf>>.
- CHALBOT, M. C. et al. Spatial attribution of sulfate and dust aerosol sources in an urban area using receptor modeling coupled with Lagrangian trajectories. *Atmospheric Pollution Research*, v. 4, p. 346–353, 2013. <<https://doi.org/10.5094/apr.2013.039>>.
- CHALBOT, M. C.; MCELROY, B.; KAVOURAS, I. G. Sources, trends and regional impacts of fine particulate matter in southern Mississippi valley: significance of emissions from sources in the Gulf of Mexico coast. *Atmospheric Chemistry and Physics*, v. 13, p. 3721–3732, 2013. <<https://doi.org/10.5194/acp-13-3721-2013>>.
- CIAGRO. *Centro Integrado de Informações Agrometeorológicas. Resenha Agrometeorológica*. 2019. Acesso em 21 de Março de 2019. Disponível em: <<http://www.ciiagro.sp.gov.br/ciiagroonline/menuresenha.htm>>.

COMINI-FROTA, E. R. et al. Frequency of reported European ancestry among multiple sclerosis patients from four cities in the southern and southeastern regions of Brazil. *Clinical Neurology and Neurosurgery*, v. 115, p. 1642–1646, 2013. <<http://dx.doi.org/10.1016/j.clineuro.2013.02.024>>.

CORDEIRO, G. M. *Modelos Lineares Generalizados*. 1. ed. [S.l.]: Unicamp, 1986.

CORDEIRO, G. M.; ANDRADE, M. G. Transformed generalized linear models. *Journal of Statistical Planning and Inference*, v. 139, p. 2970–2987, 2009. <<http://dx.doi.org/10.1016/j.jspi.2009.02.002>>.

DATAGEO. *Sistema Ambiental Paulista*. 2019. Acesso em 21 de Março de 2019. Disponível em: <<http://datageo.ambiente.sp.gov.br/app/?ctx=DATAGEO#>>.

DIMITRIOU, K. Upgrading the Estimation of Daily PM₁₀ Concentrations Utilizing Prediction Variables Reflecting Atmospheric Process. *Aerosol and Air Quality Research*, v. 16, p. 2245–2254, 2016. <<https://doi.org/10.4209/aaqr.2016.05.0214>>.

DIMITRIOU, K.; KASSOMENOS, P. A study on the reconstitution of daily PM₁₀ and PM_{2.5} levels in Paris with a multivariate linear regression model. *Atmospheric Environment*, v. 98, p. 648–654, 2014a. <<https://doi.org/10.1016/j.atmosenv.2014.09.047>>.

DIMITRIOU, K.; KASSOMENOS, P. Decomposing the profile of PM in two low polluted German cities – Mapping of air mass residence time, focusing on potencial long range transport impacts. *Environment Pollution*, v. 190, p. 91–110, 2014b. <<https://doi.org/10.1016/j.envpol.2014.03.019>>.

DIMITRIOU, K.; KASSOMENOS, P. Three year study of tropospheric ozone with back trajectories at a metropolitan and a medium scale urban area in Greece. *Science of Total Environment*, v. 502, p. 493–501, 2015. <<http://dx.doi.org/10.1016/j.scitotenv.2014.09.072>>.

DIMITRIOU, K. et al. Spatial distribution of source areas of PM_{2.5} by Concentration Weighted Trajectory (CWT) model applied in PM_{2.5} concentration and composition data. *Atmospheric Environment*, v. 116, p. 138–145, 2015. <<http://dx.doi.org/10.1016/j.atmosenv.2015.06.021>>.

DOMINICK, D. et al. Spatial assessment of air quality patterns in Malaysia using multivariate analysis. *Atmospheric Environment*, v. 60, p. 172–181, 2012. <<https://doi.org/10.1016/j.atmosenv.2012.06.021>>.

DRAXLER, R. R.; HESS, G. D. An overview of the HYSPLIT4 modelling system of trajectories, dispersion and deposition. *Australian Meteorological Magazine*, v. 47, p. 295–308, 1998.

EEA. *European Environment Agency*. 2016. Acesso em 27 de Janeiro de 2016. Disponível em: <<<http://www.eea.europa.eu>>>.

FAVERO, L. et al. *Análise de dados: Modelagem multivariada para tomada de decisões*. 2. ed. [S.l.]: Elsevier, 2009.

- FRANÇA, E. J. et al. Characterizing suspended sediments from the Piracicaba River Basin by means of k_0 -INAA. *Nuclear Instruments and Methods in Physics Research A*, v. 622, p. 479–483, 2010. <<http://dx.doi.org/10.1016/j.nima.2010.02.055>>.
- FREITAS, C. U. et al. Air pollution and its impacts on health in Vitoria, Espirito Santo, Brazil. *Revista de Saude Publica*, scielo, v. 50, 2016. <<https://doi.org/10.1590/S1518-8787.2016050005909>>.
- GARCIA, J. M. et al. Developing a methodology to predict PM_{10} concentrations in urban areas using generalized linear models. *Environmental Technology*, v. 37, p. 2316–2325, 2016. <<http://dx.doi.org/10.1080/09593330.2016.1149228>>.
- GE, Y. et al. Potencial transport pathways of dust emanating from the playa of Ebinur Lake, Xinjiang, in arid northwest China. *Atmospheric Research*, v. 178–179, p. 196–206, 2016. <<http://dx.doi.org/10.1016/j.atmosres.2016.04.002>>.
- GIOIA, S. M. C. L. et al. Insights into the dynamics and sources of atmospheric lead and particulate matter in São Paulo, Brazil, from high temporal resolution sampling. *Atmospheric Research*, v. 98, p. 478–485, 2010. <<https://doi.org/10.1016/j.atmosres.2010.08.016>>.
- GONÇALVES, C. et al. Size-segregated aerosol chemical composition from an agro-industrial region of São Paulo state, Brazil. *Air Quality, Atmosphere & Health*, v. 10, p. 483–496, 2017. <<http://dx.doi.org/10.1007/s11869-016-0441-0>>.
- GONÇALVES, F. L. T. et al. The effects of air pollution and meteorological parameters on respiratory morbidity during summer in São Paulo City. *Environment International*, v. 31, p. 343–349, 2005. <<https://doi.org/10.1016/j.envint.2004.08.004>>.
- GOYAL, P.; CHAN, A.; JAISWAL, N. Statistical models for the prediction of respirable suspended particulate matter in urban cities. *Atmospheric Environment*, v. 40, p. 2068–2077, 2006. <<https://doi.org/10.1016/j.atmosenv.2005.11.041>>.
- HAN, J.; KAMBER, M. *Data Mining: Concepts and Techniques*. 2. ed. [S.l.]: Morgan Kaufmann, 2006.
- HE, H.; LU, W. Decomposition of pollution contributors to urban ozone levels concerning regional and local scales. *Building and Environment*, v. 49, p. 97–103, 2012. <<https://doi.org/10.1016/j.buildenv.2011.09.019>>.
- HINDS, W. *Aerosol Technology*. 2. ed. [S.l.]: Wiley, 1998.
- HOINASKI, L. et al. Investigation of rainwater contamination sources in the southern part of Brazil. *Environmental Technology*, v. 35, n. 7, p. 868–881, 2014. <<http://dx.doi.org/10.1080/09593330.2013.854412>>.
- HOINASKI, L. et al. Investigation of PM_{10} sources in Santa Catarina, Brazil through graphical interpretation analysis combined with receptor modelling. *Environmental Technology*, v. 34, p. 2453–2463, 2013. <<http://dx.doi.org/10.1080/21622515.2013.772659>>.

- HONGFANG, C.; RUI, S. ARMA model parameter optimized estimate method. In: *2010 First ACIS International Symposium on Cryptography, and Network Security, Data Mining and Knowledge Discovery, E-Commerce and Its Applications, and Embedded Systems*. IEEE, 2010. Disponível em: <<https://doi.org/10.1109/cdee.2010.13>>.
- HUANG, K. et al. Impact of anthropogenic emission on air quality over a megacity - revealed from an intensive atmospheric campaign during the Chinese Spring Festival. *Atmospheric Chemistry and Physics*, v. 12, p. 11631–11645, 2012. <<https://doi.org/10.5194/acp-12-11631-2012>>.
- HUANG, L. et al. A trajectory analysis of atmospheric transport of black carbon aerosols to Canadian high Arctic in winter and spring (1990–2005). *Atmospheric Chemistry and Physics*, v. 10, p. 5065–5073, 2010. <<https://doi.org/10.5194/acp-10-5065-2010>>.
- HUANG, P. et al. Spatial and Temporal Distribution of PM_{2.5} Pollution in Xián City, China. *International Journal of Environmental Research and Public Health*, v. 12, p. 6608–6625, 2015. <<https://doi.org/10.3390/ijerph120606608>>.
- HUEBNEROVA, Z.; MICHALEK, J. Analysis of daily average PM₁₀ predictions by generalized linear models in Brno, Czech Republic. *Atmospheric Pollution Research*, v. 5, p. 471–476, 2014. <<https://doi.org/10.5094/APR.2014.055>>.
- IAG. *Estação meteorológica do IAG/USP*. 2018. Acesso em 24 de Maio de 2018. Disponível em: <<<http://www.estacao.iag.usp.br/seasons/index.php>>>.
- IBGE. *Instituto Brasileiro de Geografia e Estatística. Canal cidades*. 2019. Acesso em 21 de Março de 2019. Disponível em: <<https://cidades.ibge.gov.br/>>.
- JAIN, S. et al. Source apportionment of PM₁₀ in Delhi, India using PCA/APCS, UNMIX and PMF. *Particuology*, v. 37, p. 107–118, 2018. <<https://doi.org/10.1016/j.partic.2017.05.009>>.
- JOHNSON, R. A.; WICHERN, D. W. (Ed.). *Applied Multivariate Statistical Analysis*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1988. ISBN 0-130-41146-9.
- KALNAY, E. et al. The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, v. 77, n. 3, p. 437–472, 1996.
- KASSOMENOS, P. et al. Study of PM₁₀ and PM_{2.5} levels in three European cities: Analysis of intra and inter urban variations. *Atmospheric Environment*, v. 87, p. 153–163, 2014. <<https://doi.org/10.1016/j.atmosenv.2014.01.004>>.
- KAVOURAS, I. G. et al. Quantitative determination of regional contributions to fine and coarse particle mass in urban receptor sites. *Environmental Pollution*, v. 176, p. 1–9, 2013. <<https://doi.org/10.1016/j.envpol.2013.01.004>>.
- KONG, X. et al. Comparison of transport pathways and potencial sources of PM₁₀ in two cities around a large Chinese lake using the modified trajectories analysis. *Atmospheric Research*, v. 122, p. 284–297, 2013. <<http://dx.doi.org/10.1016/j.atmosres.2012.10.012>>.
- KULSHRESTHA, U.; KUMAR, B. Airmass Trajectories and Long Range Transport of Pollutants: Review of Wet Deposition Scenario in South Asia. *Advances in Meteorology*, p. 1–14, 2014. <<https://doi.org/10.1155/2014/596041>>.

- KWON, S. et al. A multivariate study for characterizing particulate matter (PM₁₀, PM_{2.5}, and PM₁) in Seoul metropolitan subway stations, Korea. *Journal of Hazardous Materials*, v. 297, p. 295–303, 2015. <<https://doi.org/10.1016/j.jhazmat.2015.05.015>>.
- LATIF, M. T. et al. Long term assessment of air quality from a background station on the Malaysian Peninsula. *Science of the Total Environment*, v. 482–483, p. 336–348, 2014. <<https://doi.org/10.1016/j.scitotenv.2014.02.132>>.
- LEITE, D. A. N. O. et al. Avaliação dos parâmetros do índice de qualidade de água segundo o modelo estatístico ARIMA. *HOLOS*, v. 13, n. 1, p. 24, 2013. <<http://dx.doi.org/10.14295/holos.v13i1.4641>>.
- LIANG, P. et al. The role of meteorological conditions and pollution control strategies in reducing air pollution in Beijing during APEC 2014 and Victory Parade 2015. *Atmospheric Chemistry and Physics*, v. 17, p. 13921–13940, 2017. <<https://doi.org/10.5194/acp-17-13921-2017>>.
- LIU, P. G.; JOHNSON, R. Forecasting Peak Daily Ozone Levels - I. A Regression with Time Series Errors Model Having a Principal Component Trigger to Fit 1991 Ozone Levels. *Journal of the Air & Waste Management Association*, v. 52, p. 1064–1074, 2002. <<http://dx.doi.org/10.1080/10473289.2002.10470841>>.
- LU, W. Z.; HE, H. D.; DONG, L. Y. Performance assessment of air quality monitoring networks using principal component analysis and cluster analysis. *Building and Environment*, v. 46, p. 577–583, 2011. <<https://doi.org/10.1016/j.buildenv.2010.09.004>>.
- LUNA, A. et al. Prediction of ozone concentration in tropospheric levels using artificial neural networks and support vector machine at Rio de Janeiro, Brazil. *Atmospheric Environment*, v. 98, p. 98–104, 2014. <<https://doi.org/10.1016/j.atmosenv.2014.08.060>>.
- LYRA, G. B.; SOUZA, M. O.; VIOLA, D. N. Modelos lineares aplicados à estimativa da concentração do Material Particulado (PM₁₀) na cidade do Rio de Janeiro, RJ. *Revista Brasileira de Meteorologia*, v. 26, p. 392–400, 2011. <<http://dx.doi.org/10.1590/S0102-77862011000300006>>.
- MATYASOVSKY, I. et al. Multivariate Analysis respiratory problems and their connection with meteorological parameters and the main biological and chemical air pollutants. *Atmospheric Environment*, v. 45, p. 4152–4159, 2011. <<https://doi.org/10.1016/j.atmosenv.2011.05.024>>.
- MCCULLAGH, P.; NELDER, J. A. *Generalized Linear Models*. 2. ed. [S.l.]: Chapman and Hall, 1989.
- MCGOWAN, H.; CLARK, A. Identification of dust transport pathways from Lake Eyre, Australia using Hysplit. *Atmospheric Environment*, v. 42, p. 6915–6925, 2008. <<http://dx.doi.org/10.1016/j.atmosenv.2008.05.053>>.
- MIRANDA, R. M. et al. Urban air pollution: a representative survey of PM_{2.5} mass concentrations in six Brazilian cities. *Air Quality, Atmosphere and Health*, v. 5, p. 63–77, 2012. <<https://doi.org/10.1007/s11869-010-0124-1>>.

- MMA. *Inventário Nacional de emissões atmosféricas por veículos automotores rodoviários 2013: Ano-base 2012*. [S.l.], 2014. Disponível em: <http://www.antt.gov.br/backend/galeria/arquivos/inventario_de_emissoes_por_veiculos_rodoviarios_2013.pdf>.
- MONTE, E. Z.; ALBUQUERQUE, T. T. A.; REISEN, V. A. Previsão da concentração de ozônio na região da grande Vitória, Espírito Santo, Brasil, utilizando o modelo ARMAX-GARCH. *Revista Brasileira de Meteorologia*, v. 30, n. 3, p. 285–294, 2015. <<http://dx.doi.org/10.1590/0102-778620140060>>.
- MONTE, E. Z.; ALBUQUERQUE, T. T. A.; REISEN, V. A. Impactos das Variáveis Meteorológicas na Qualidade do Ar da Região da Grande Vitória, Espírito Santo, Brasil. *Revista Brasileira de Meteorologia*, v. 31, p. 546–554, 2016. <<http://dx.doi.org/10.1590/0102-7786312314b20150100>>.
- MOREIRA, D. M.; TIRABASSI, T.; MORAES, M. R. Meteorologia e Poluição atmosférica. *Ambiente e Sociedade*, v. 11, n. 1, p. 1–13, 2008.
- MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Séries Temporais*. 2. ed. [S.l.]: Blucher, 2006.
- NAKAZATO, R. K. et al. Efficiency of biomonitoring methods applying tropical bioindicator plants for assessing the phytotoxicity of the air pollutants in SE, Brazil. *Environmental Science and Pollution research*, v. 25, p. 19323–19337, 2018. <<http://dx.doi.org/10.1007/s11356-018-2294-6>>.
- PAULA, G. A. *Modelos de Regressão com apoio computacional*. 2013. Acesso em 09 de Novembro de 2016. Disponível em: <<<https://www.ime.usp.br/~giapaula/>>>.
- PAVÓN-DOMÍNGUEZ, P.; JIMÉNEZ-HORNERO, F.; RAVÉ, E. G. D. Proposal for estimating ground-level ozone concentrations at urban areas based on multivariate statistical methods. *Atmospheric Environment*, v. 90, p. 59–70, 2014. <<https://doi.org/10.1016/j.atmosenv.2014.03.032>>.
- PEKNEY, N. J. et al. Application of PSCF and CPF to PMF-Modeled Sources of PM_{2.5} in Pittsburgh. *Aerosol Science and Technology*, v. 2015, p. 1–20, 2006. <<https://doi.org/10.1155/2015/284213>>.
- PEREZ, I. A. et al. Applications of Air Mass Trajectories. *Advances in Meteorology*, v. 40, p. 952–961, 2015. <<https://doi.org/10.1080/02786820500543324>>.
- PINTO, W. P.; REISEN, V. A.; MONTE, E. Z. Previsão da concentração de material particulado inalável, na Região da Grande Vitória, ES, Brasil, utilizando o modelo SARIMAX. *Engenharia Sanitária e Ambiental*, v. 23, n. 2, p. 307–318, 2018. <<http://dx.doi.org/10.1590/S1413-41522018168758>>.
- PLAIA, A.; BONDI, A. L. Single imputation method of missing values in environmental pollution data sets. *Atmospheric Environment*, v. 40, p. 7316–7330, 2006. <<https://doi.org/10.1016/j.atmosenv.2006.06.040>>.
- POZZA, S. et al. Time series analysis of PM_{2.5} and PM_{10-2.5} mass concentration in the city of São Carlos, Brazil. *International Journal of Environment and Pollution*, v. 41, p. 90–108, 2010. <<https://doi.org/10.1504/IJEP.2010.032247>>.

- QUALAR. *Qualidade do Ar, CETESB, Dados meteorológicos*. 2016. Acesso em 16 de Janeiro de 2016. Disponível em: <<<http://qualar.cetesb.sp.gov.br/>>>.
- QUEROL, X. et al. Speciation and origin of PM₁₀ and PM_{2.5} in Spain. *Aerosol Science*, v. 35, p. 1151–1172, 2004. <<https://doi.org/10.1016/j.jaerosci.2004.04.002>>.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>.
- REISEN, V. A. et al. Modeling and forecasting daily average PM₁₀ concentrations by a seasonal long-memory model with volatility. *Environmental Modelling & Software*, v. 51, p. 286–295, 2014. <<http://dx.doi.org/10.1016/j.envsoft.2013.09.027>>.
- REVUELTA, M. A. et al. Partitioning of magnetic particles in PM₁₀, PM_{2.5} and PM₁ aerosols in the urban atmospheric of Barcelona (Spain). *Environmental Pollution*, v. 188, p. 109–117, 2014. <<https://doi.org/10.1016/j.envpol.2014.01.025>>.
- RIGBY, R. A.; STASINOPOULOS, D. M. Generalized Additive Models for Location, Scale and Shape. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, v. 54, n. 3, p. 507–554, 2005. Retrieved from <<http://www.jstor.org/stable/3592732>>.
- RODOPOULOU, S. et al. Air Pollution and hospital emergency room and admissions for cardiovascular and respiratory diseases in Doña Ana County, New Mexico. *Environmental Research*, v. 129, p. 39–46, 2014. <<https://doi.org/10.1016/j.envres.2013.12.006>>.
- RODOPOULOU, S. et al. Air pollution and cardiovascular and respiratory emergency visits in Central Arkansas: A time-series analysis. *Science of the Total Environment*, v. 536, p. 872–879, 2015. <<https://doi.org/10.1016/j.scitotenv.2015.06.056>>.
- Rolph, G.D. *Real-time Environmental Applications and Display sYstem (READY)*. College Park, MD, 2016. Disponível em: <<http://www.ready.noaa.gov>>.
- ROSA, J. D. D. L. et al. Using PM₁₀ geochemical maps for defining the origin of atmospheric pollution in Andalusia (Southern Spain). *Atmospheric Environment*, v. 44, p. 4595–4605, 2010. <<https://doi.org/10.1016/j.atmosenv.2010.08.009>>.
- São Paulo. *Decreto Estadual N 59.113 de 23 de Abril de 2013*. São Paulo: DOE, 2013. v. 123.
- SALMABADI, H.; SAEEDI, M. Determination of the transport routes of the areas potentially affected by SO₂ emanating from Khatoonabad Copper Smelter (KCS), Kernan province, Iran using HYSPLIT. *Atmospheric Pollution Research*, 2018. <<http://dx.doi.org/10.1016/j.apr.2018.08.008>>.
- SEINFELD, J. H.; PANDIS, S. N. *Atmospheric Chemistry and Physics: from air pollution to climate change*. 2. ed. [S.l.]: Jonh Wiley & Sons, 2006.
- SHARMA, P. et al. An integrated statistical approach for evaluating the exceedence of criteria pollutants in the ambient air of megacity Delhi. *Atmospheric Environment*, v. 70, p. 7–17, 2013. <<http://dx.doi.org/10.1016/j.atmosenv.2013.01.004>>.
- SOUZA, A. et al. Impacto de Fatores meteorológicos sobre as concentrações de ozônio modelados por análise de séries temporais e métodos estatísticos multivariados. *HOLOS*, v. 5, p. 2, 2017. <<https://doi.org/10.15628/holos.2017.5033>>.

- SOUZA, J. B. et al. Principal components and generalized linear modeling in the correlation between hospital admissions and air pollution. *Revista Saúde Pública*, v. 48(3), p. 451–458, 2014. <<https://doi.org/10.1590/s0034-8910.2014048005078>>.
- SOUZA, R. et al. Using ensembles of artificial neural network to improve PM10 forecast. *Chemical Engineering Transactions*, v. 43, p. 2161–2166, 2015. <<https://doi.org/10.3303/CET1543361>>.
- SPIEZIA, L. et al. Short-term exposure to high levels of air pollution as a risk factor for acute isolated pulmonary embolism. *Thrombosis Research*, v. 134, p. 259–263, 2014. <<https://doi.org/10.1016/j.thromres.2014.05.011>>.
- STASINOPOULOS, D. M.; RIGBY, R. A. Generalized Additive Models for Location, Scale and Shape (GAMLSS) in R. *Journal of Statistical Software*, v. 23, n. 7, p. 1–46, 2007.
- STATHEROPOULOS, M.; VASSILIADIS, N.; PAPPA, A. Principal Component and Canonical Correlation analysis for examining air pollution and meteorological data. *Atmospheric Environment*, v. 32, p. 1087–1095, 1998. <[https://doi.org/10.1016/s1352-2310\(97\)00377-4](https://doi.org/10.1016/s1352-2310(97)00377-4)>.
- STEIN, A. et al. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bulletin of the American Meteorological Society*, v. 96, p. 2059–2077, 2015. <<https://doi.org/10.1175/bams-d-14-00110.1>>.
- TARGINO, A.; KRECL, P. Local and Regional Contributions to Black Carbon Aerosols in a Mid-Sized City in Southern Brazil. *Aerosol and Air Quality Research*, v. 16, p. 125–137, 2016. <<https://doi.org/10.4209/aaqr.2015.06.0388>>.
- TIWARI, S. et al. Diurnal and seasonal variations of black carbon and PM2.5 over New Delhi, India: Influence of meteorology. *Atmospheric Research*, v. 125–126, p. 50–62, 2013. <<https://doi.org/10.1016/j.atmosres.2013.01.011>>.
- TOH, Y.; LIM, S.; GLASOW, R. The influence of meteorological factors and biomass burning on surface ozone concentrations at Tanah Rata, Malaysia. *Atmospheric Environment*, v. 70, p. 435–446, 2013. <<https://doi.org/10.1016/j.atmosenv.2013.01.018>>.
- US-EPA. *United States Environmental Protection Agency*. 2017. Acesso em 09 de Novembro de 2017. Disponível em: <<<http://www3.epa.gov>>>.
- VALLERO, D. *Fundamentals of Air Pollution*. 4. ed. [S.l.]: Academic Press, 2008.
- VARDOULAKIS, S.; KASSOMENOS, P. Sources and factors affecting PM10 levels in two European cities: Implications for local air quality management. *Atmospheric Environment*, v. 42, p. 3949–3963, 2008. <<https://doi.org/10.1016/j.atmosenv.2006.12.021>>.
- VELLINGIRI, K. et al. Identification of nitrogen dioxide and ozone source regions for an urban area in Korea using back trajectory analysis. *Atmospheric Research*, v. 176–177, p. 212–221, 2016. <<http://dx.doi.org/10.1016/j.atmosres.2016.02.022>>.
- WEISBERG, S. *Applied Linear Regression*. 3. ed. [S.l.]: John Wiley & Sons, Inc, 2005.

WHO – World Health Organization. *Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide – Global update 2005 – Summary of risk assessment*. Swiss, 2006.

XU, J. et al. Attribution of sulfate aerosols in Federal Class I areas of the western United States based on trajectory regression analysis. *Atmospheric Environment*, v. 40, p. 3433–3447, 2006. <<https://doi.org/10.1016/j.atmosenv.2006.02.009>>.

YNOUE, R. Y. et al. *Meteorologia: noções básicas*. 1. ed. [S.l.]: Oficina de textos, 2017.

ZHANG, Z. Y.; WONG, M. S.; LEE, K. H. Estimation of potencial source regions of Pm_{2,5} in Beijing using backward trajectories. *Atmospheric Pollution Research*, v. 6, p. 173–177, 2015. <<http://dx.doi.org/10.5094/APR.2015.020>>.

ZHAO, Q.; BLACK, T. L.; BALDWIN, M. E. Implementation of the Cloud Prediction Scheme in the Eta Model at NCEP. *American Meteorological Society*, v. 12, p. 697–712, 1997.

ZHOU, B. et al. Daily variations of size-segregated ambient particulate matter in Beijing. *Environmental Pollution*, v. 197, p. 36–42, 2015. <<http://dx.doi.org/10.1016/j.envpol.2014.11.029>>.

ZONG, Z. et al. PMF and PSCF based source apportionment of PM_{2.5} at a regional background site in North China. *Atmospheric Research*, v. 203, p. 207–215, 2018. <<https://doi.org/10.1016/j.atmosres.2017.12.013>>.

Apêndice A

Análise dos Resíduos

A.1 MP_{10} - Piracicaba

A Figura A.1 traz os gráficos de diagnóstico dos resíduos do modelo AR(2) para os dados de de MP_{10} na cidade de Piracicaba.

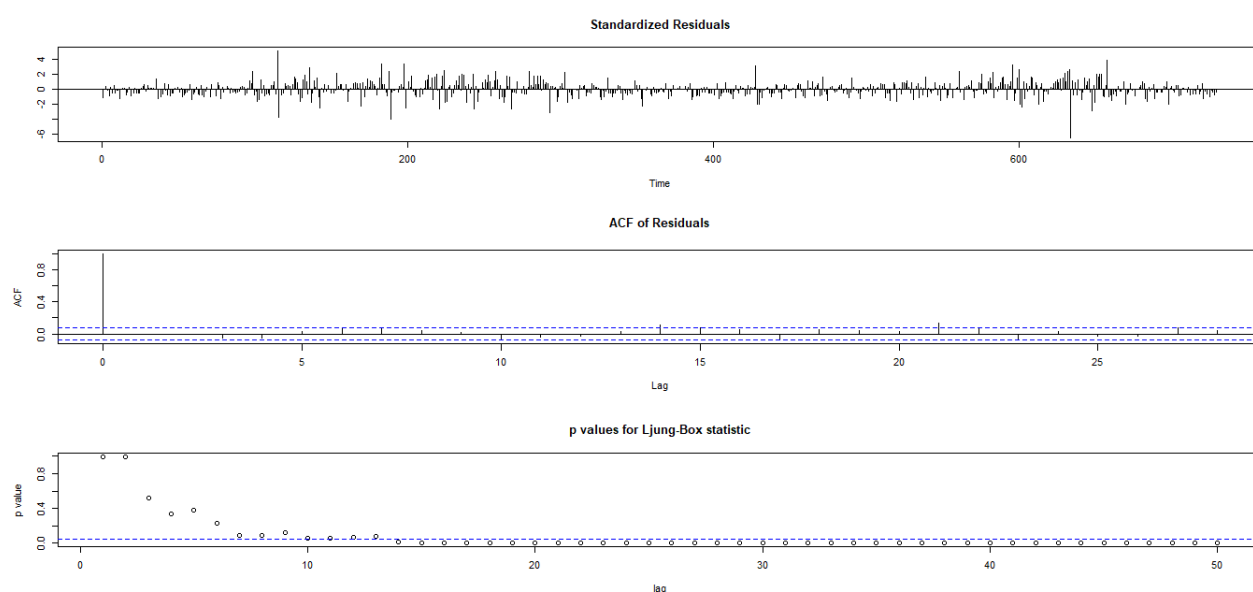


Figura A.1: Diagnóstico dos Resíduos padronizados (MP_{10}) - Modelo AR(2) - Piracicaba

As Figuras A.2, A.3, A.4 e A.5 são referentes aos resíduos padronizados ao longo do tempo para os modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, referente aos dados de MP_{10} em Piracicaba.

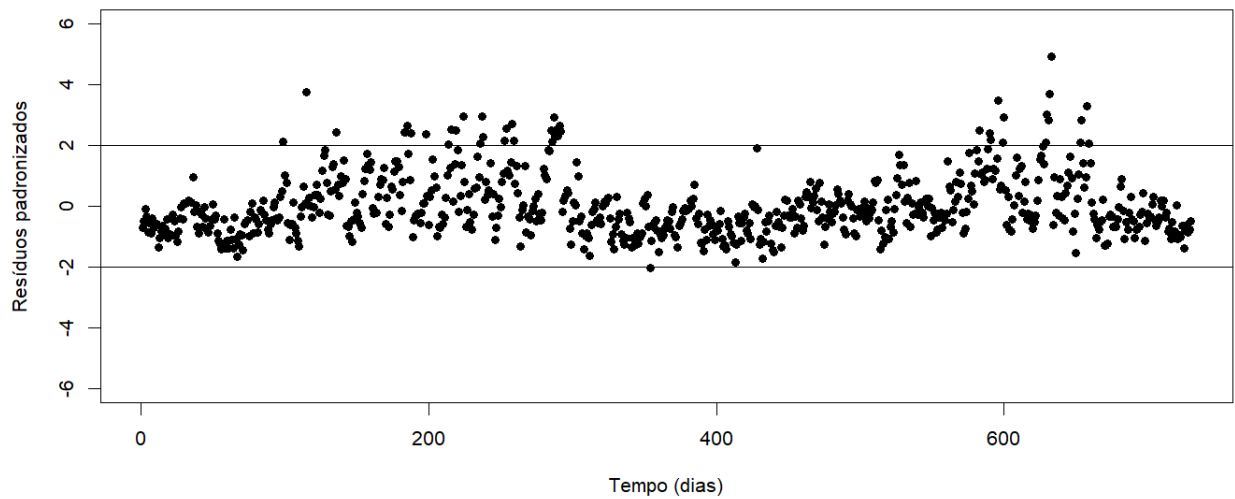


Figura A.2: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo Normal - Piracicaba

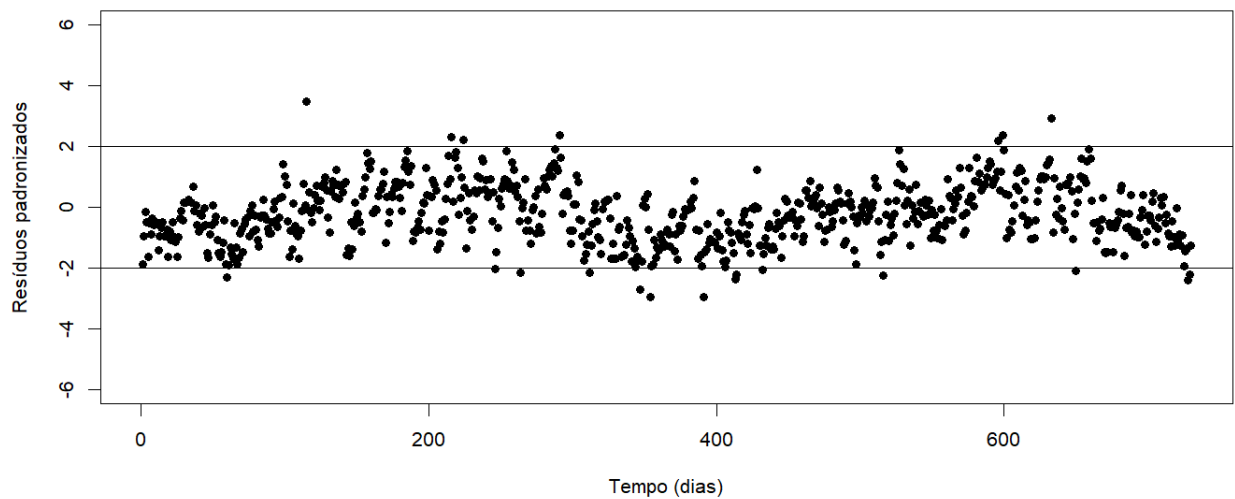


Figura A.3: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo Normal Inversa - Piracicaba

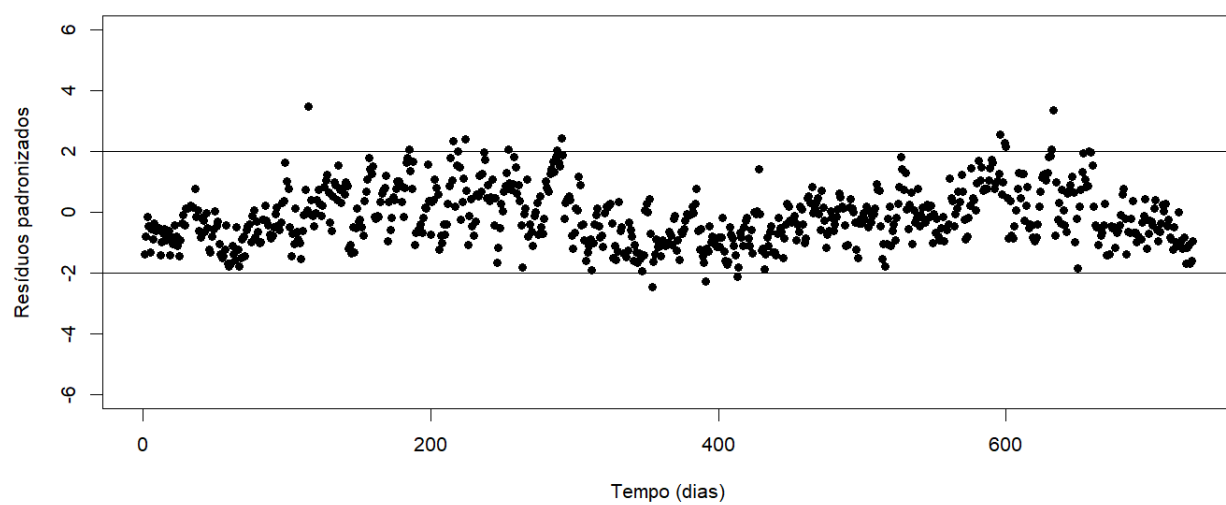


Figura A.4: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo Gama - Piracicaba

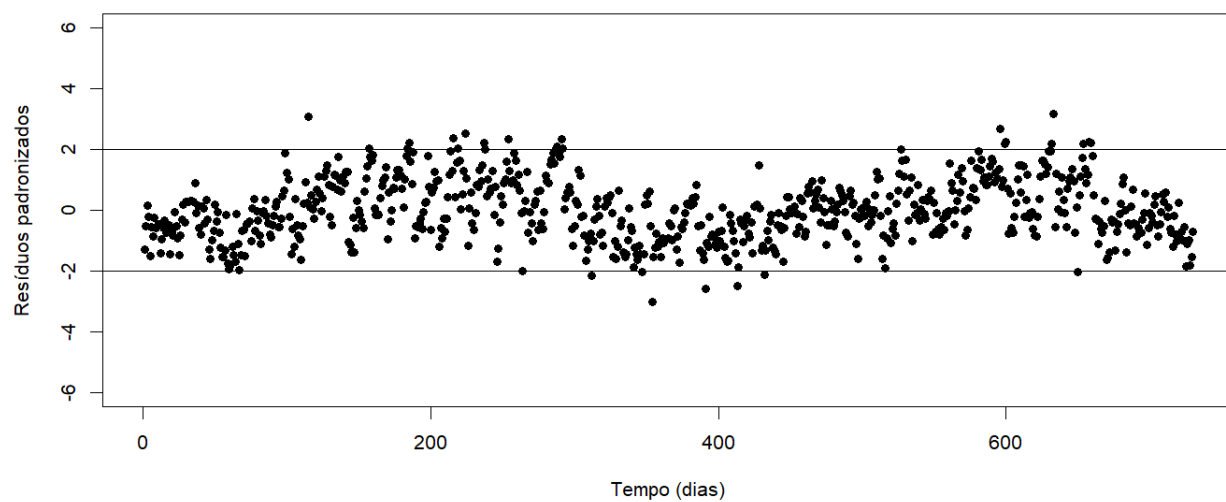


Figura A.5: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo Log-Normal - Piracicaba

As Figuras A.6, A.7, A.8 e A.9 são referentes aos resíduos padronizados comparado com os percentis da distribuição normal padrão para os modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, referente aos dados de MP_{10} em Piracicaba.

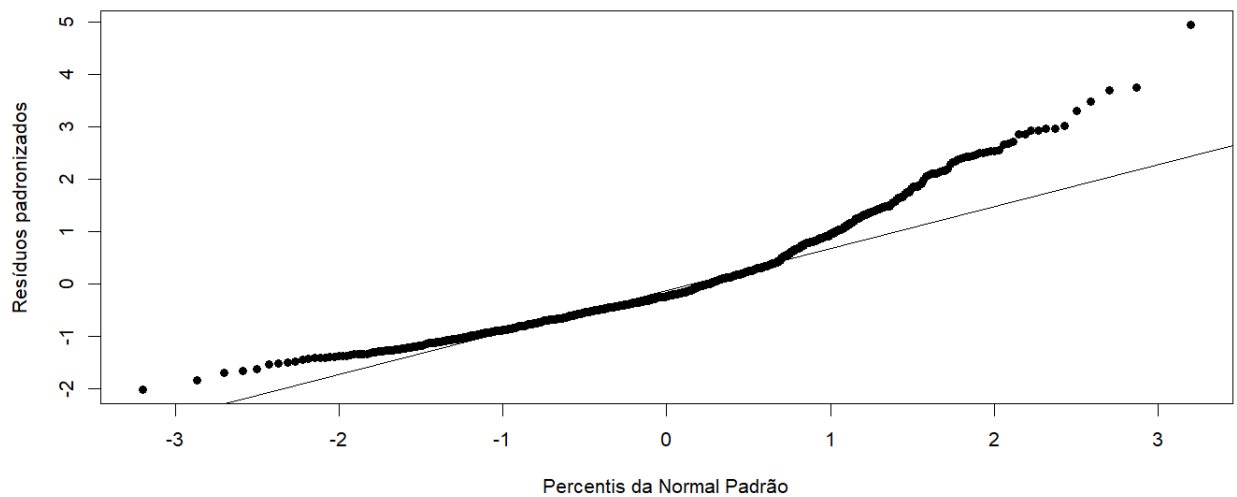


Figura A.6: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo Normal - Piracicaba

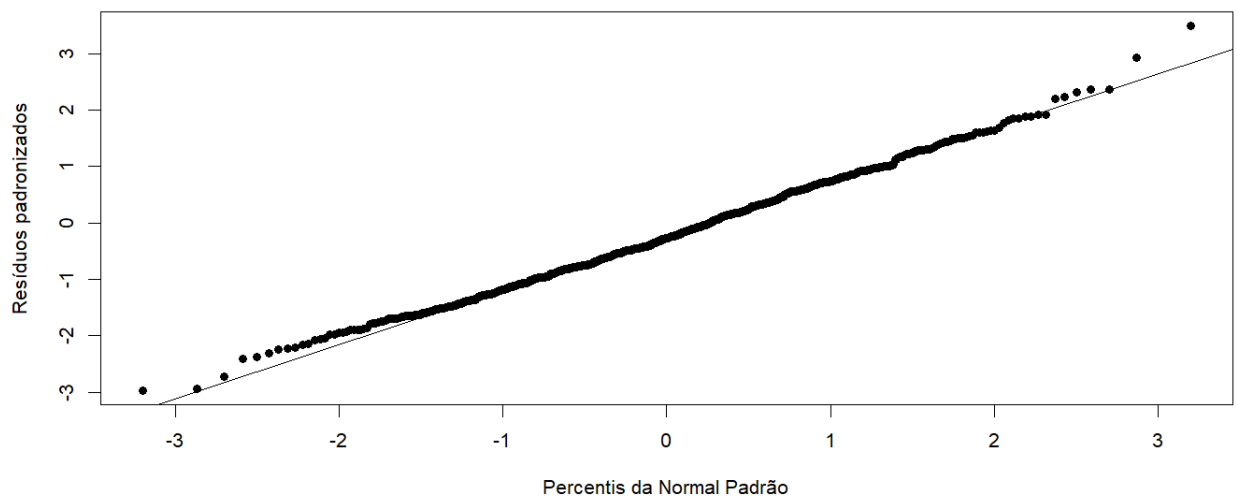


Figura A.7: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo Normal Inversa - Piracicaba

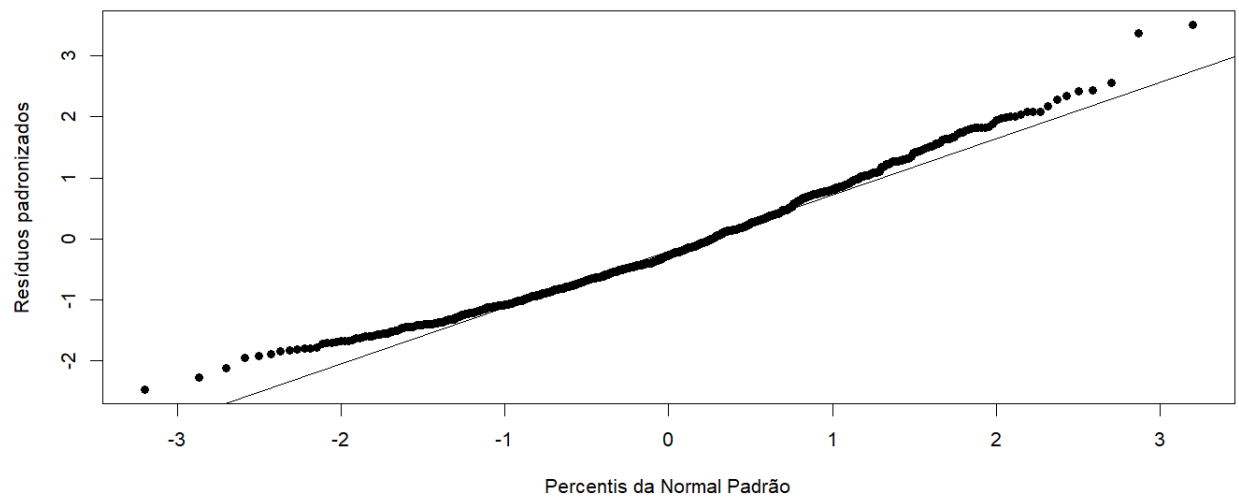


Figura A.8: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo Gama - Piracicaba

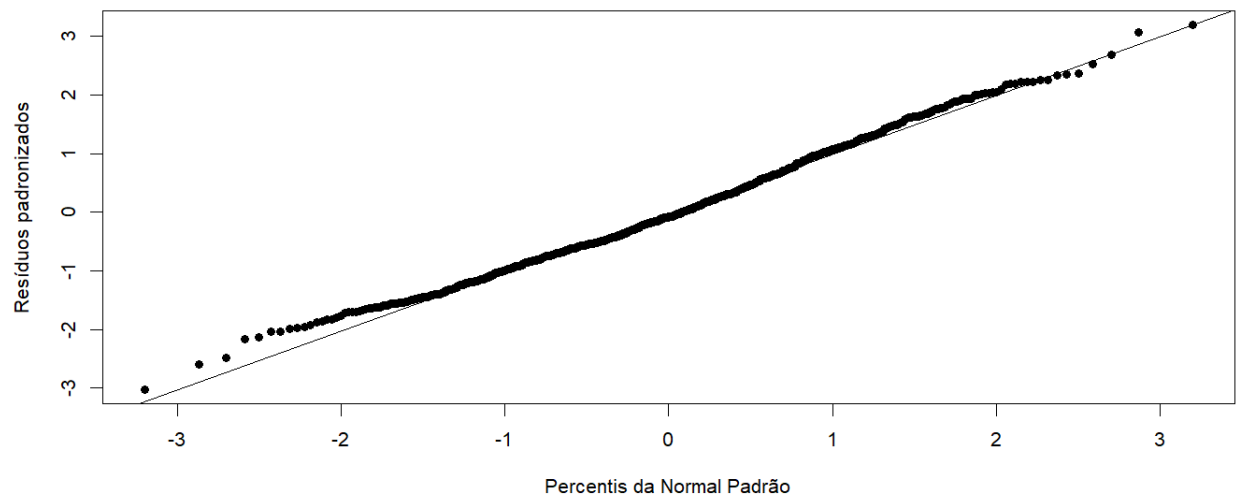


Figura A.9: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo Log-Normal - Piracicaba

A.2 MP_{10} - Santos

A Figura A.10 traz os gráficos de diagnóstico dos resíduos do modelo ARMA(2,2) para os dados de MP_{10} na cidade de Santos.

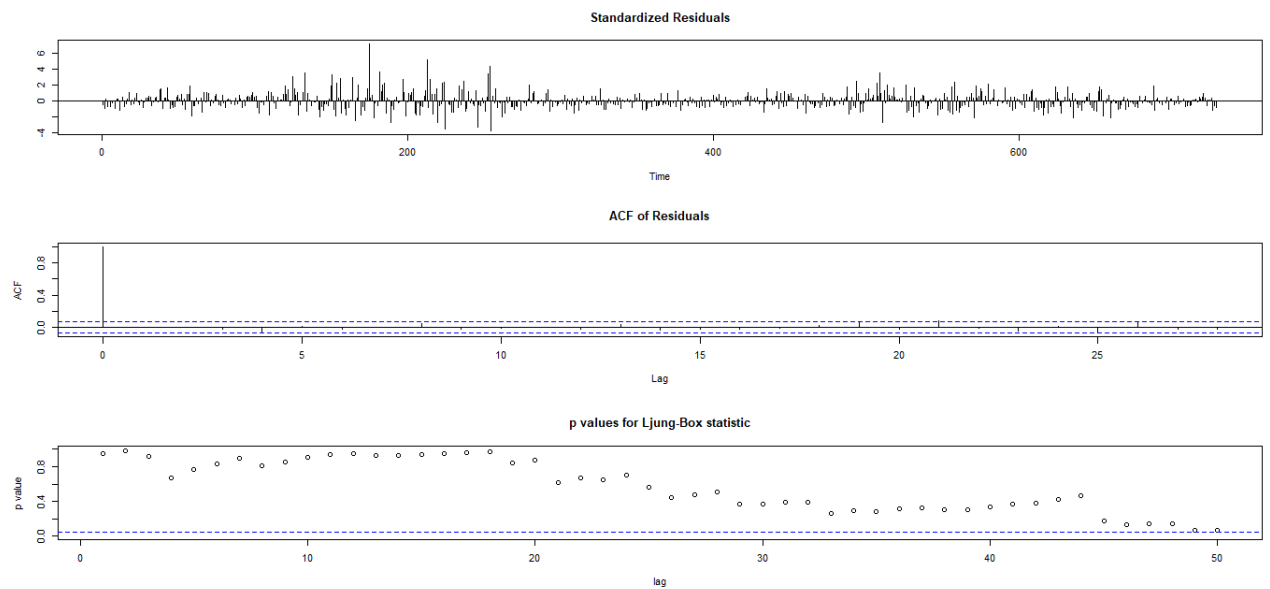


Figura A.10: Diagnóstico dos Resíduos padronizados (MP_{10}) - Modelo ARMA(2,2) - Santos

As Figuras A.11, A.12, A.13 e A.14 são referentes aos resíduos padronizados ao longo do tempo para os modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, referente aos dados de MP_{10} em Santos.

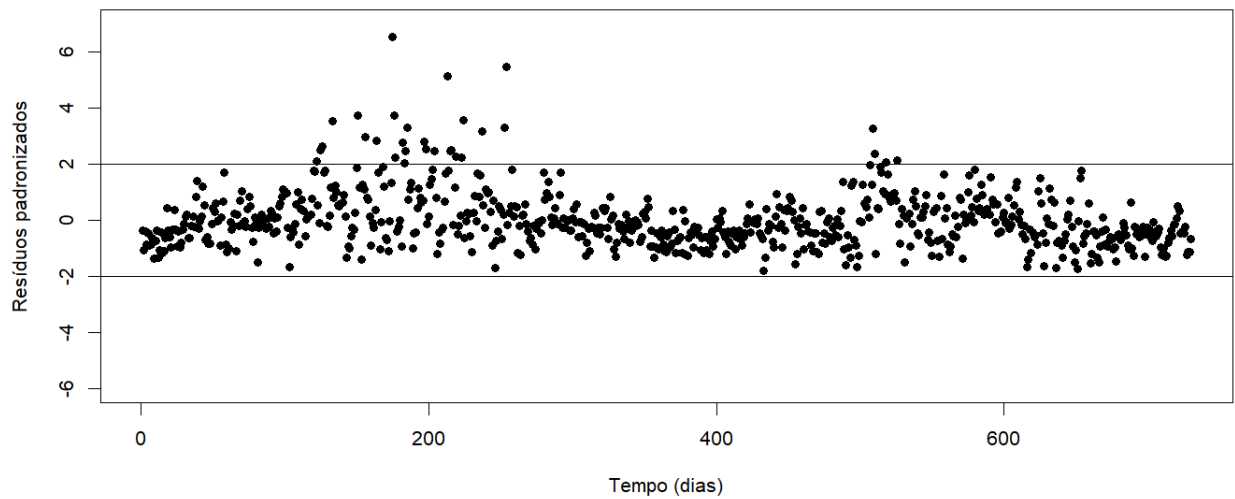


Figura A.11: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo Normal - Santos

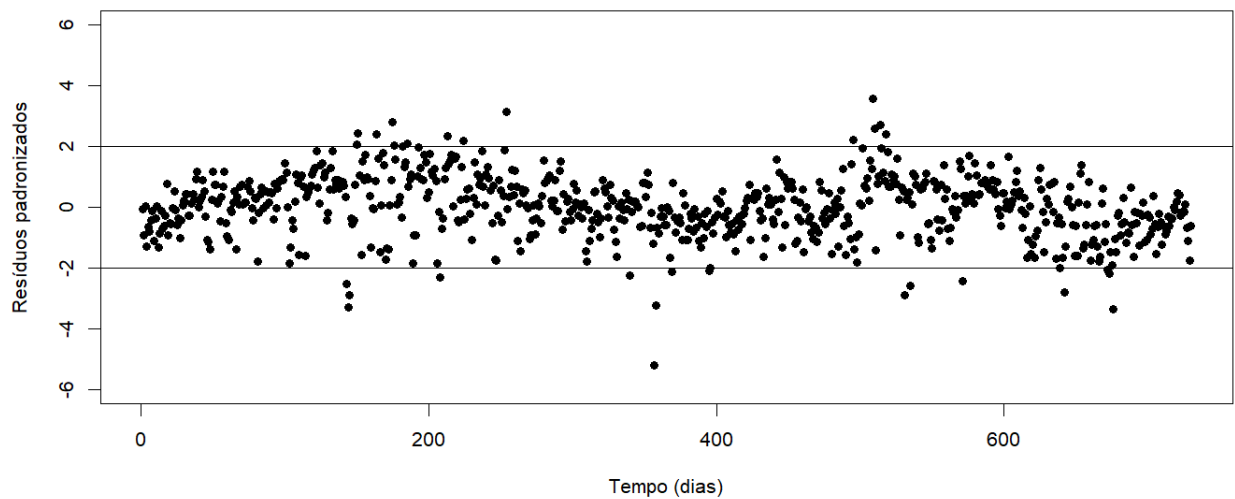


Figura A.12: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo Normal Inversa - Santos

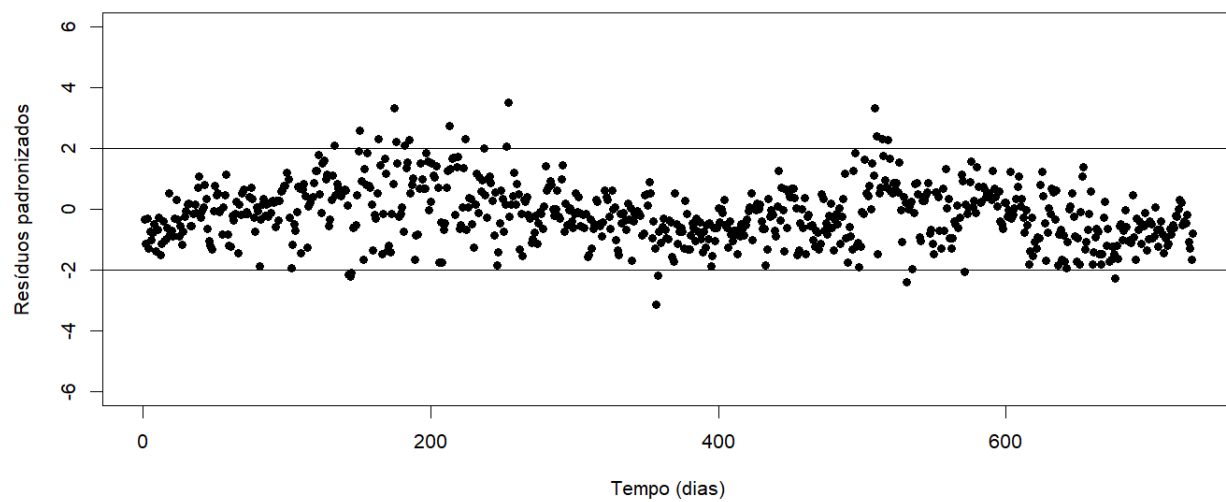


Figura A.13: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo Gama - Santos

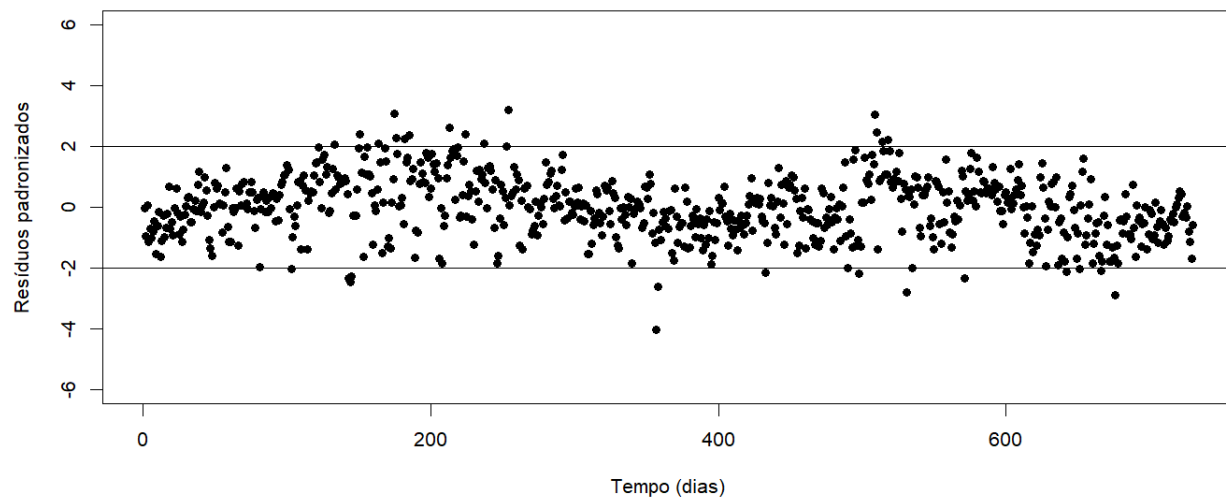


Figura A.14: Resíduos padronizados pela ordem das observações (MP_{10}) - Modelo Log-Normal - Santos

As Figuras A.15, A.16, A.17 e A.18 são referentes aos resíduos padronizados comparado com os percentis da distribuição normal padrão para os modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, referente aos dados de MP_{10} em Santos.

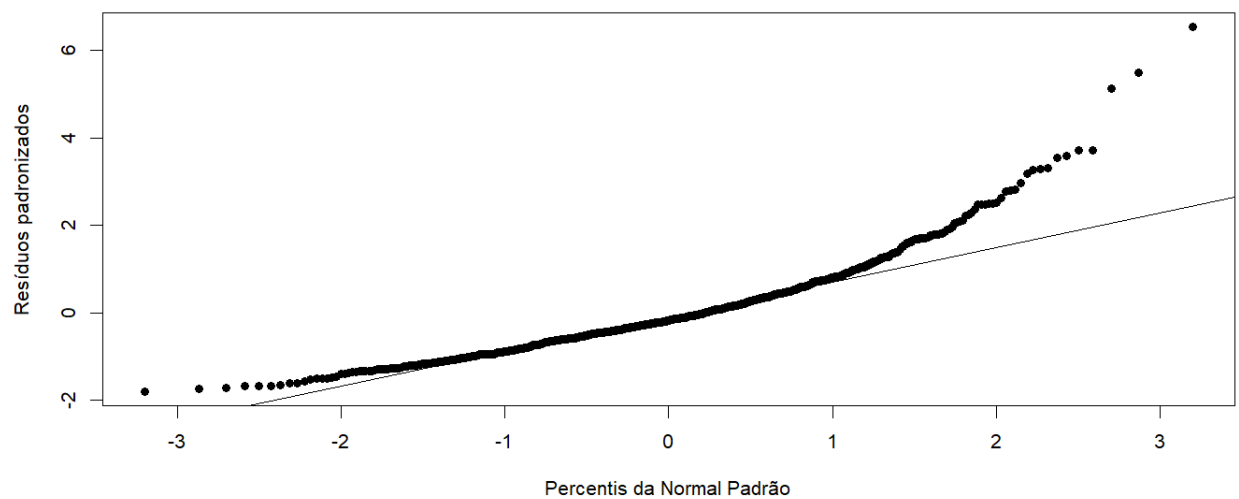


Figura A.15: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo Normal - Santos

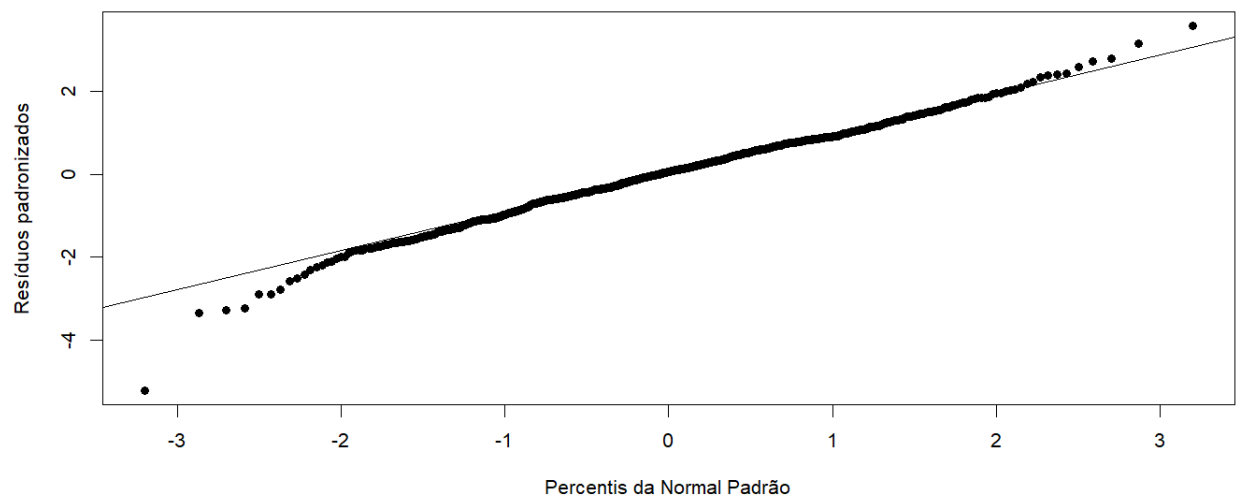
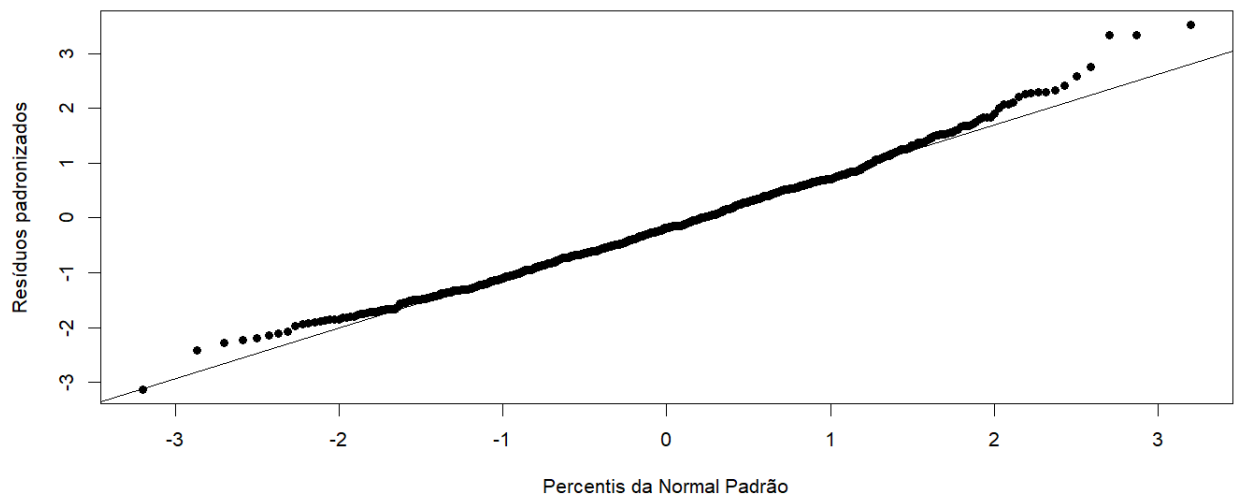
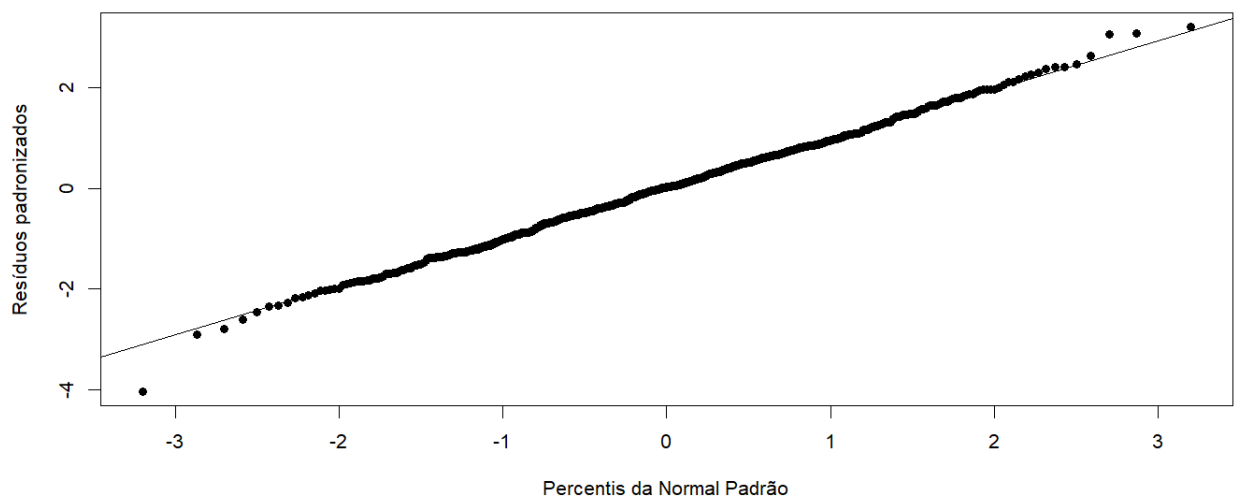


Figura A.16: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo Normal Inversa - Santos

Figura A.17: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo Gama - SantosFigura A.18: Gráfico normal de probabilidades - *qq-plot* - (MP_{10}) - Modelo Log-Normal - Santos

A.3 $MP_{2,5}$ - Piracicaba

A Figura A.19 traz os gráficos de diagnóstico dos resíduos do modelo ARMA(2,1) para os dados de $MP_{2,5}$ na cidade de Piracicaba.

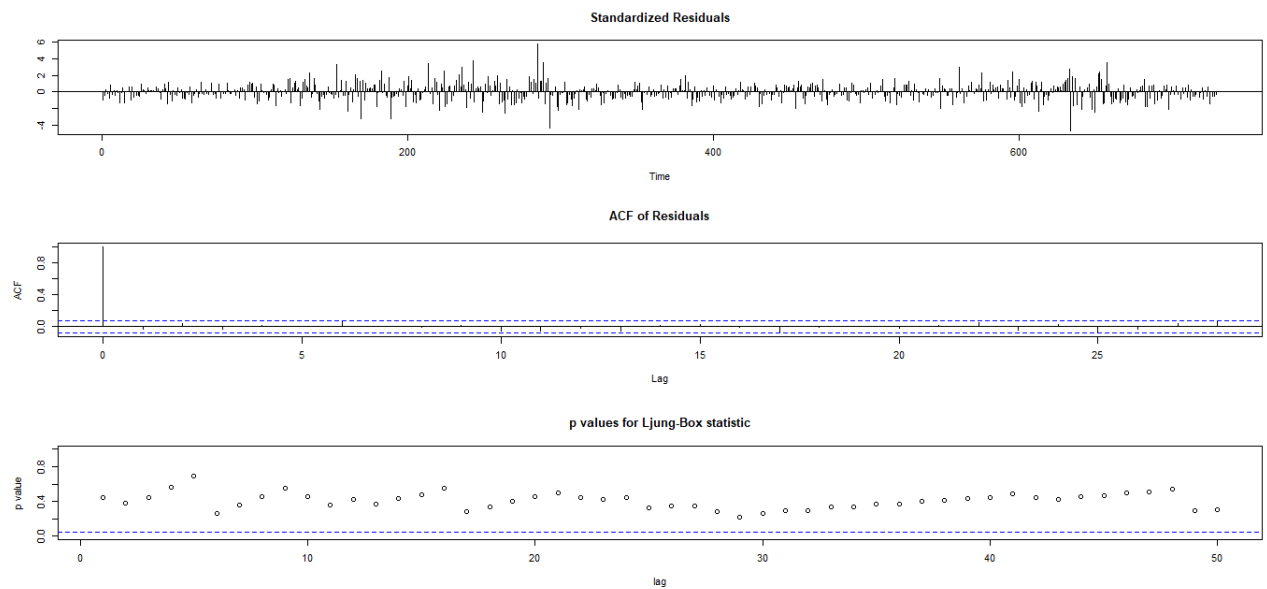


Figura A.19: Diagnóstico dos Resíduos padronizados ($MP_{2,5}$) - Modelo ARMA(2,1) - Piracicaba

As Figuras A.20, A.21, A.22 e A.23 são referentes aos resíduos padronizados ao longo do tempo para os modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, referente aos dados de $MP_{2,5}$ em Piracicaba.

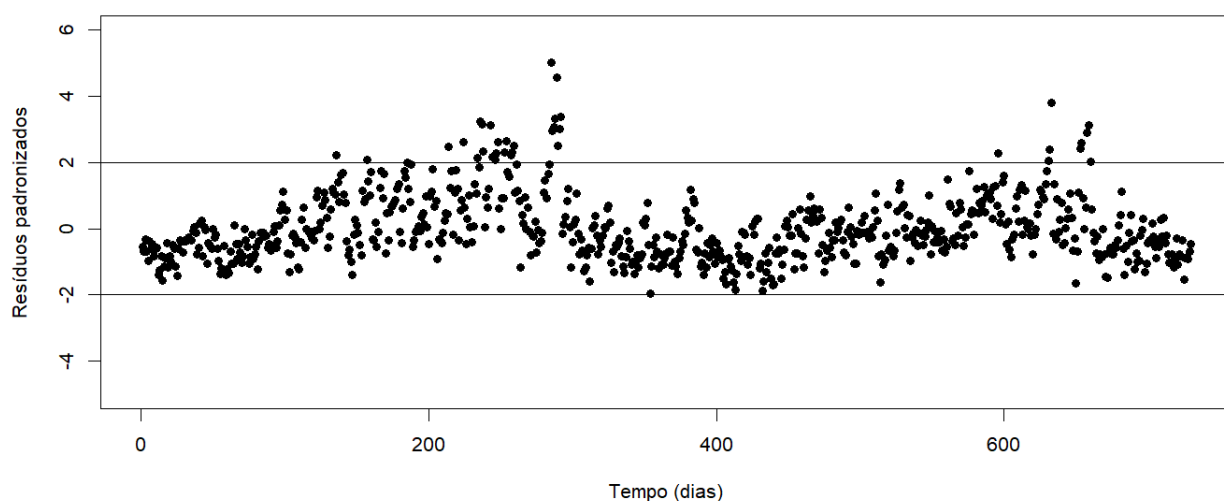


Figura A.20: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Normal - Piracicaba

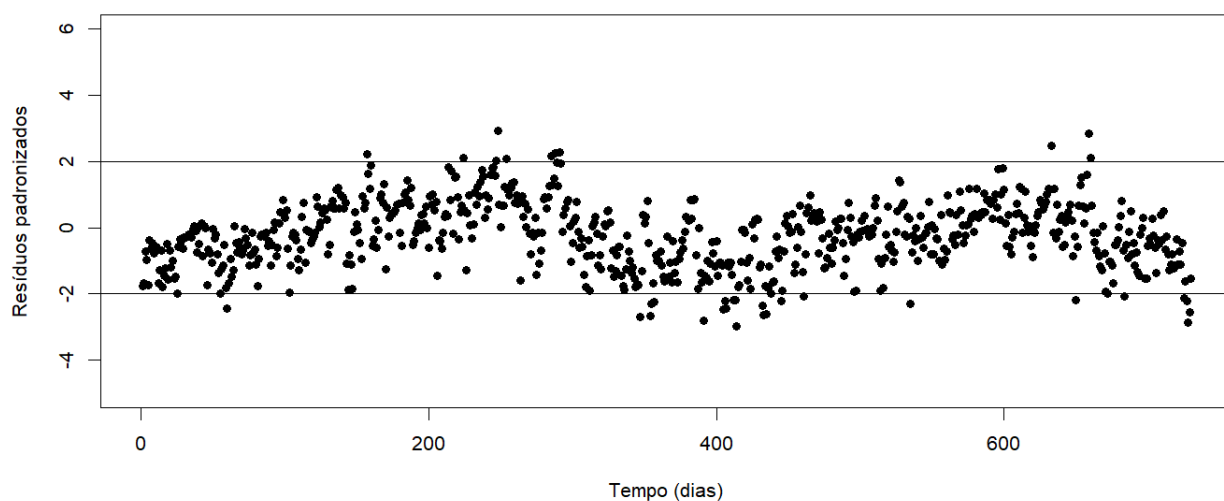


Figura A.21: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Normal Inversa - Piracicaba

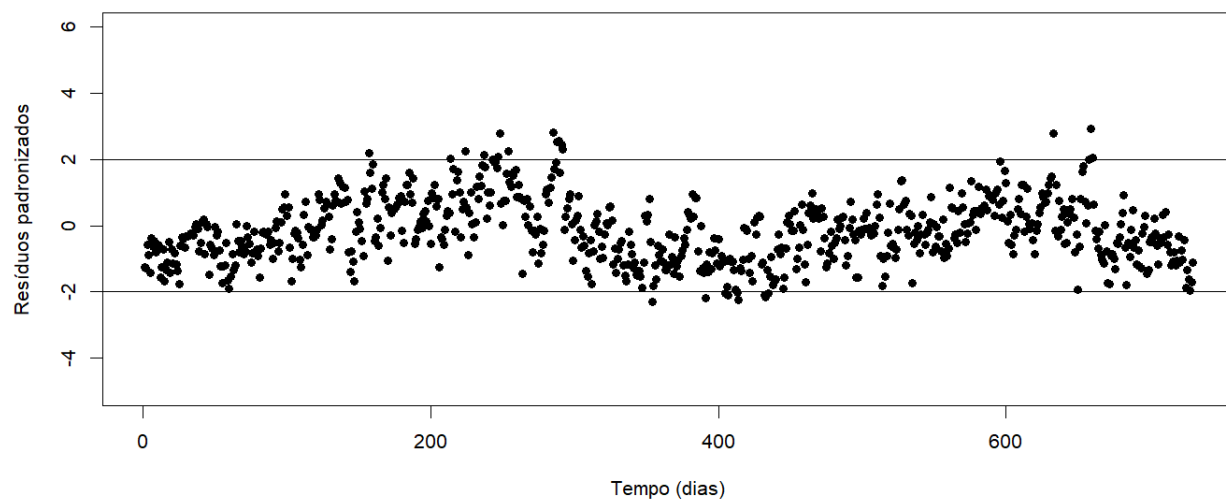


Figura A.22: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Gama - Piracicaba

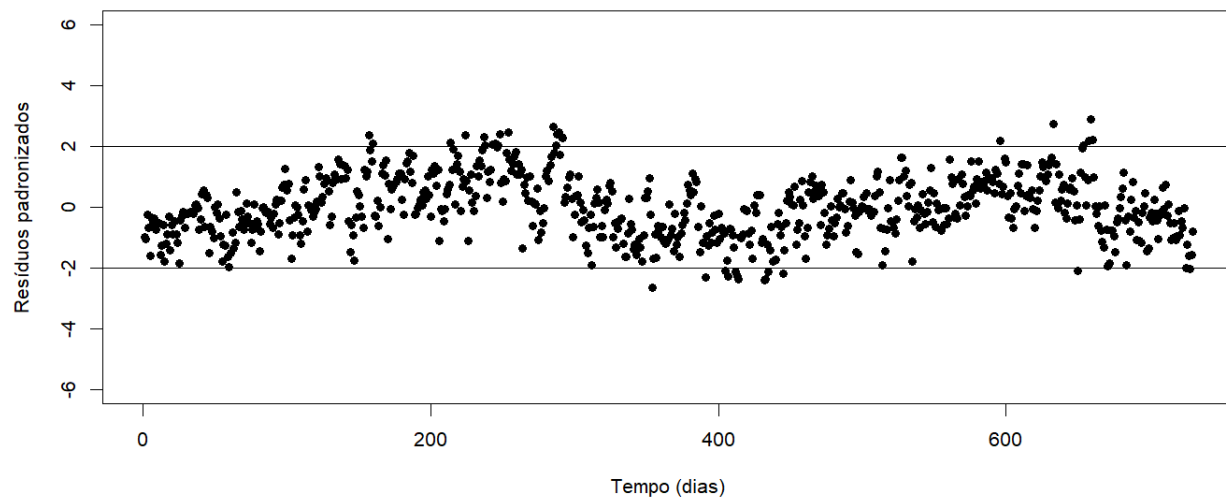


Figura A.23: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Log-Normal - Piracicaba

As Figuras A.24, A.25, A.26 e A.27 são referentes aos resíduos padronizados comparado com os percentis da distribuição normal padrão para os modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, referente aos dados de $MP_{2,5}$ em Piracicaba.

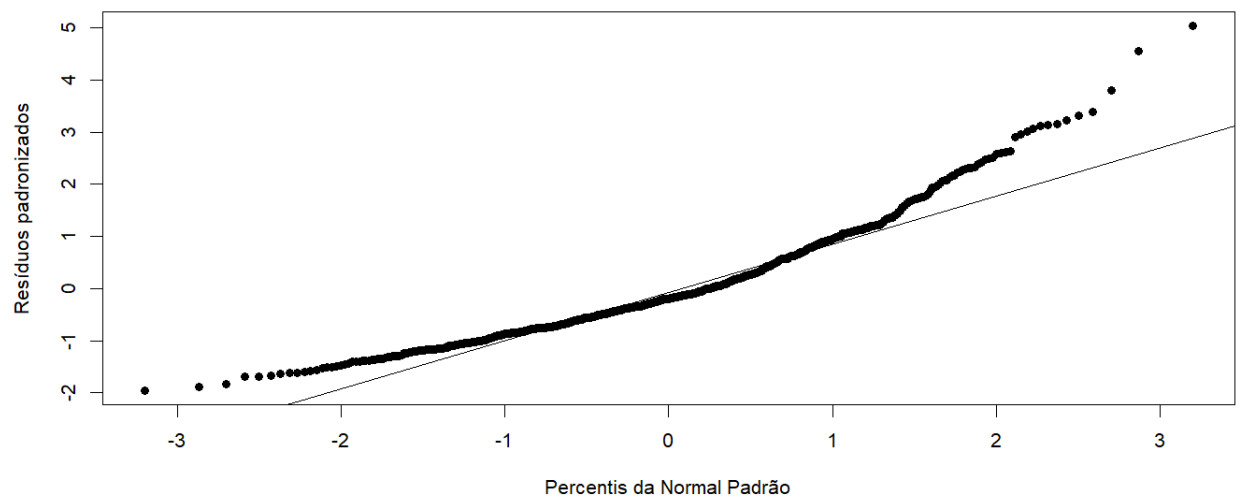


Figura A.24: Gráfico normal de probabilidades - *qq-plot* - ($MP_{2,5}$) - Modelo Normal - Piracicaba

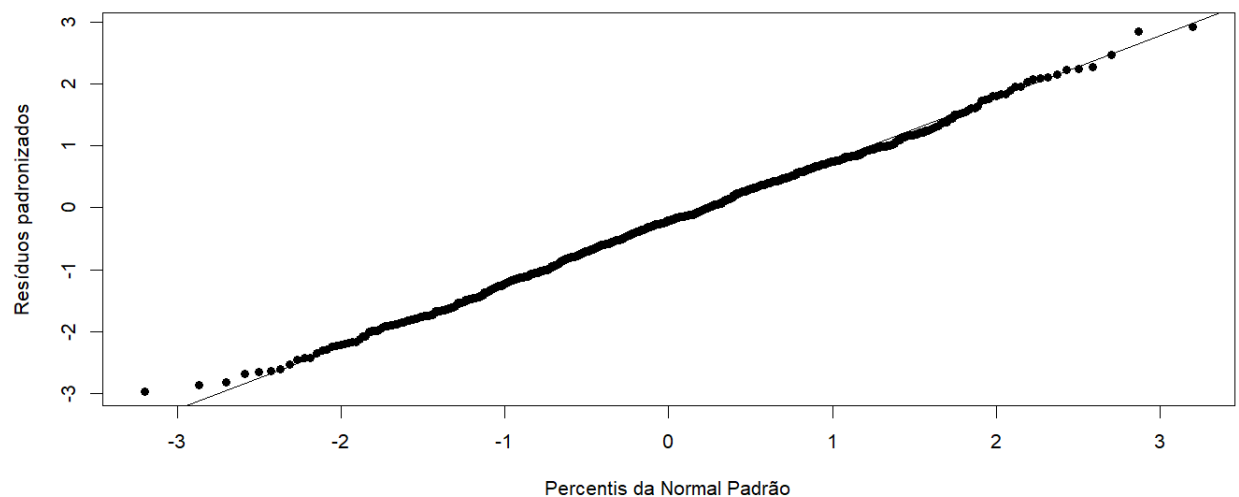


Figura A.25: Gráfico normal de probabilidades - *qq-plot* - ($MP_{2,5}$) - Modelo Normal Inversa - Piracicaba

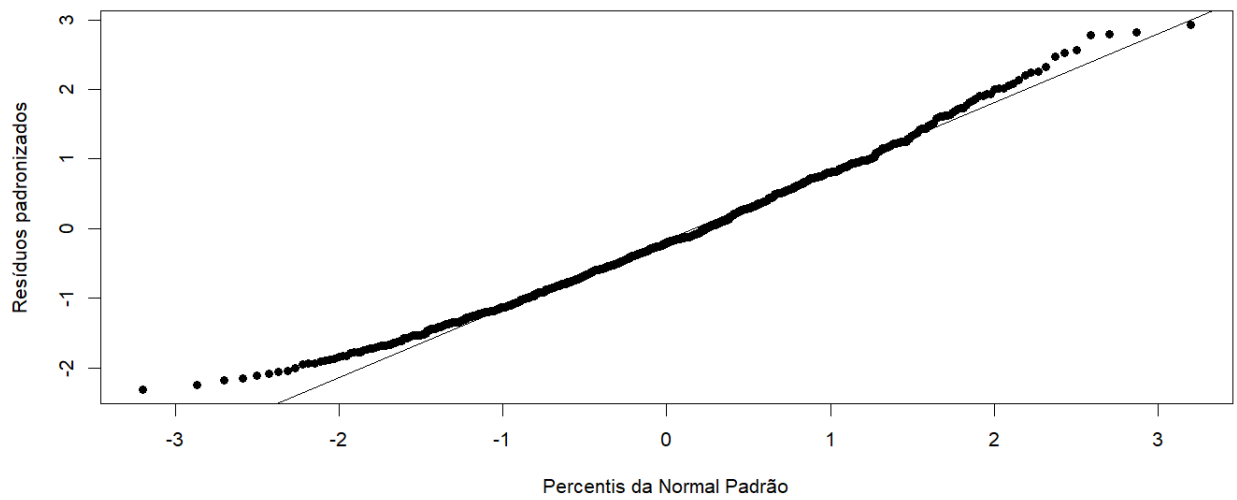


Figura A.26: Gráfico normal de probabilidades - *qq-plot* - ($MP_{2,5}$) - Modelo Gama - Piracicaba

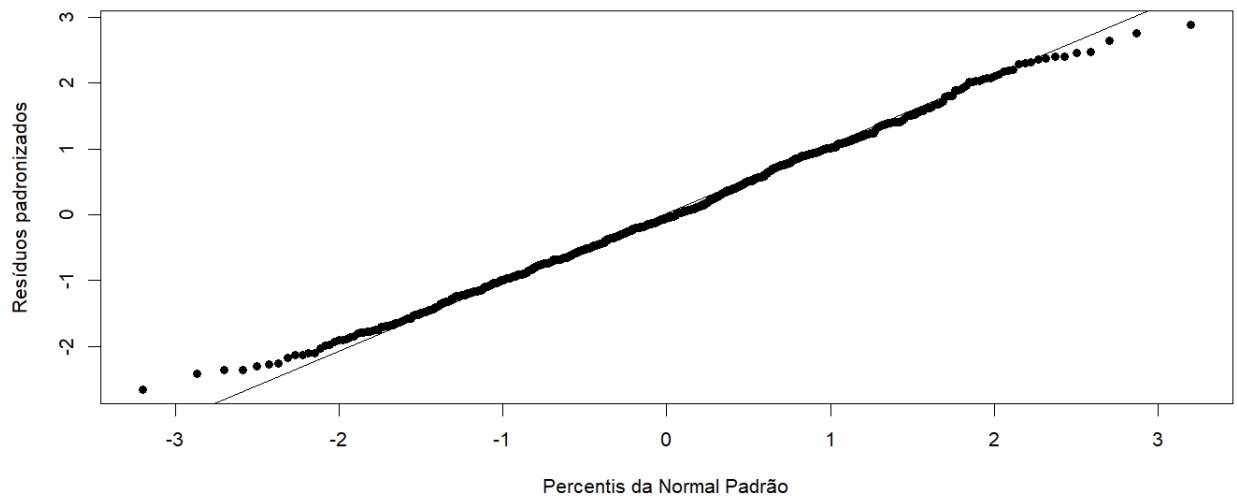


Figura A.27: Gráfico normal de probabilidades - *qq-plot* - ($MP_{2,5}$) - Modelo Log-Normal - Piracicaba

A.4 $MP_{2,5}$ - Santos

A Figura A.28 traz os gráficos de diagnóstico dos resíduos do modelo ARMA(2,1) para os dados de $MP_{2,5}$ na cidade de Santos.

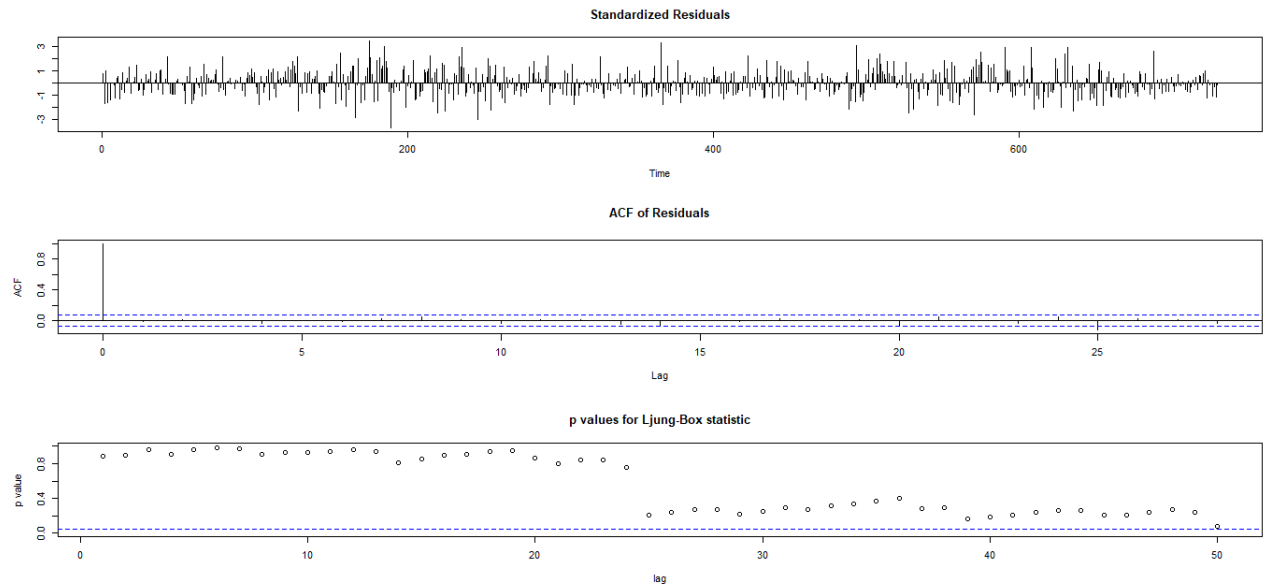


Figura A.28: Diagnóstico dos Resíduos padronizados ($MP_{2,5}$) - Modelo ARMA(2,1) - Santos

As Figuras A.29, A.30, A.31 e A.32 são referentes aos resíduos padronizados ao longo do tempo para os modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, referente aos dados de $MP_{2,5}$ em Santos.

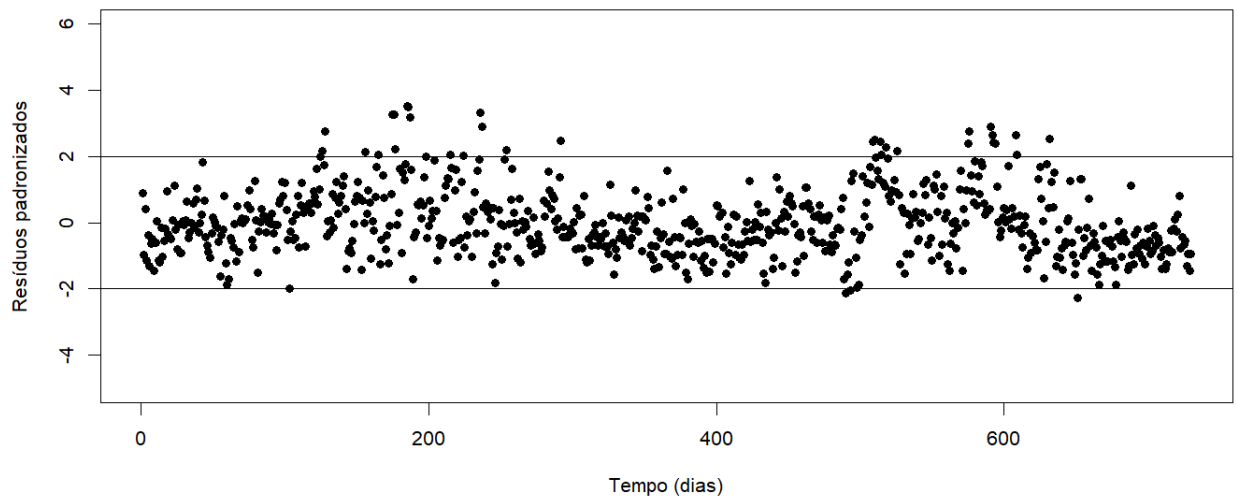


Figura A.29: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Normal - Santos

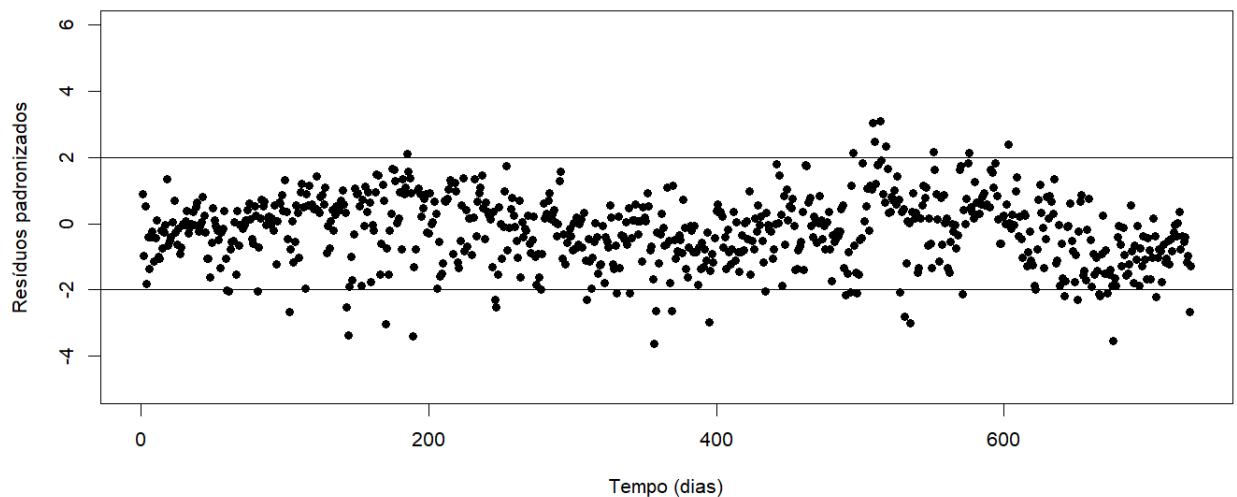


Figura A.30: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Normal Inversa - Santos

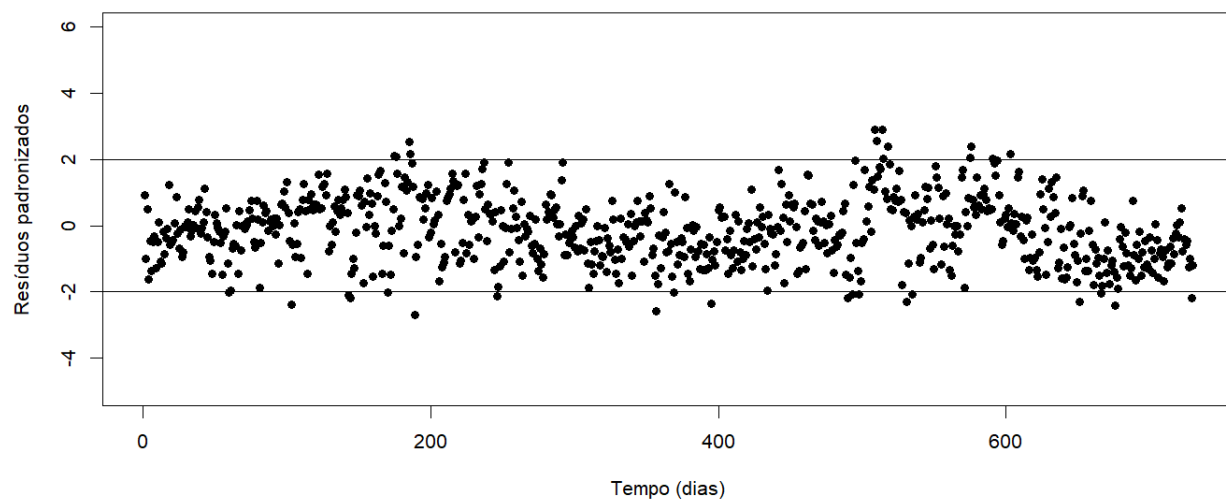


Figura A.31: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Gama - Santos

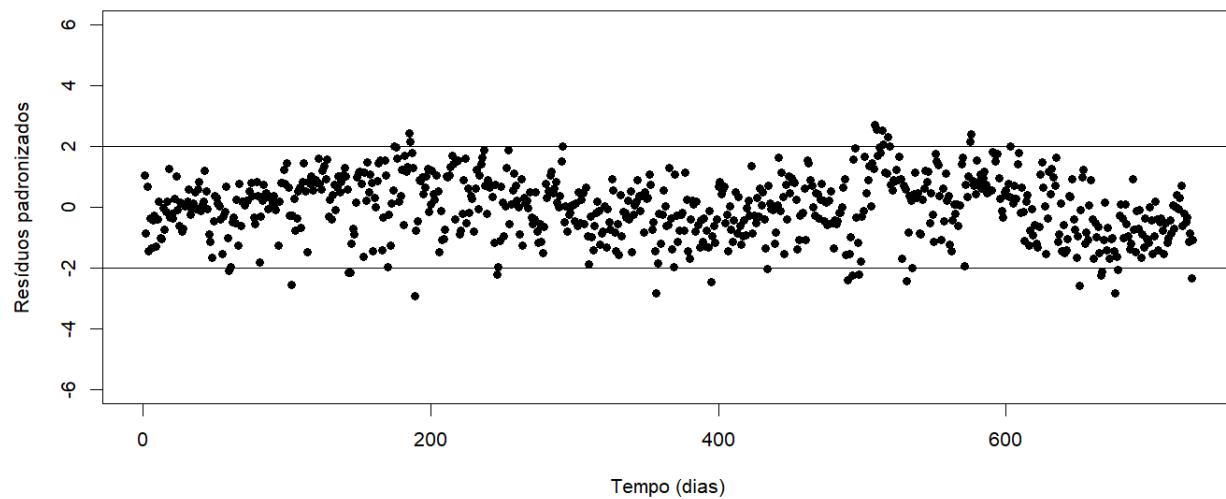


Figura A.32: Resíduos padronizados pela ordem das observações ($MP_{2,5}$) - Modelo Log-Normal - Santos

As Figuras A.33, A.34, A.35 e A.36 são referentes aos resíduos padronizados comparado com os percentis da distribuição normal padrão para os modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, referente aos dados de $MP_{2,5}$ em Santos.

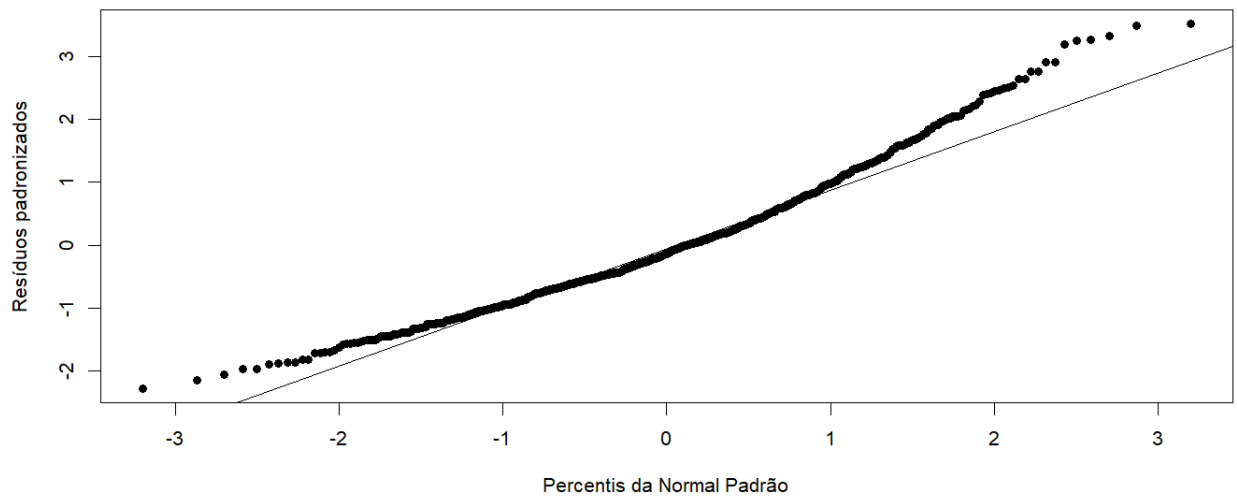


Figura A.33: Gráfico normal de probabilidades - *qq-plot* - ($MP_{2,5}$) - Modelo Normal - Santos

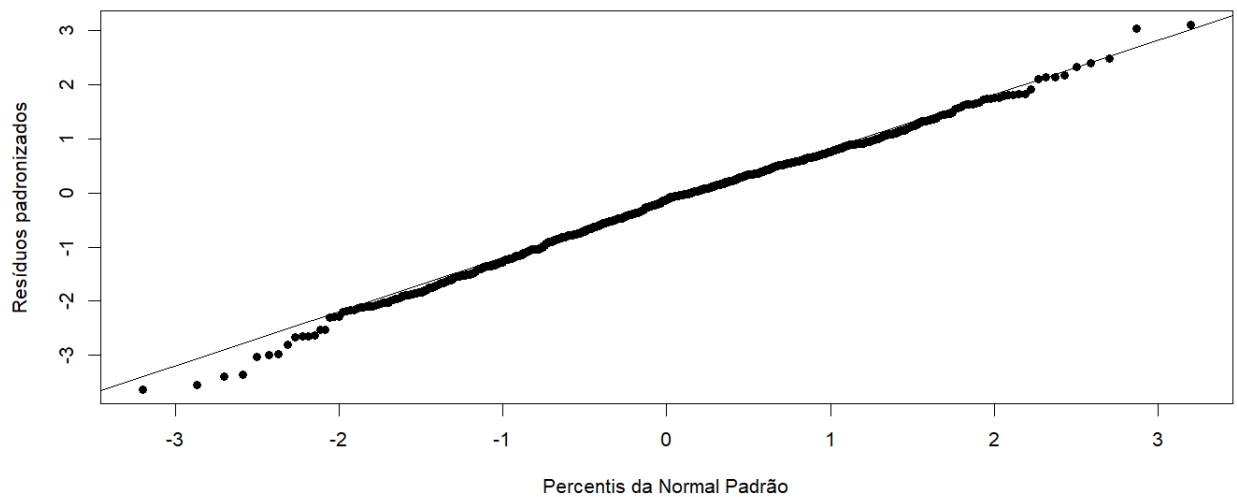
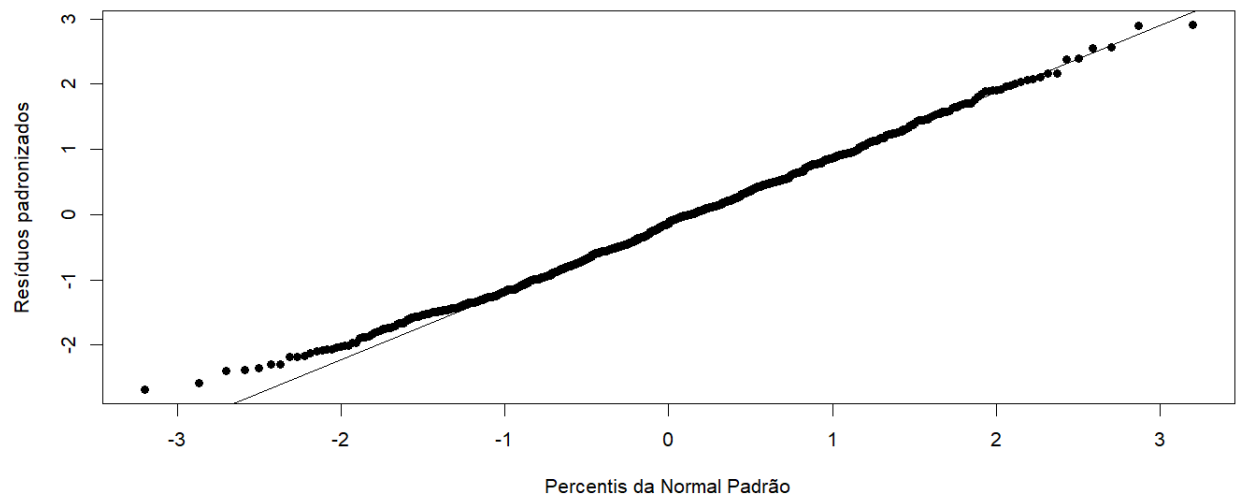
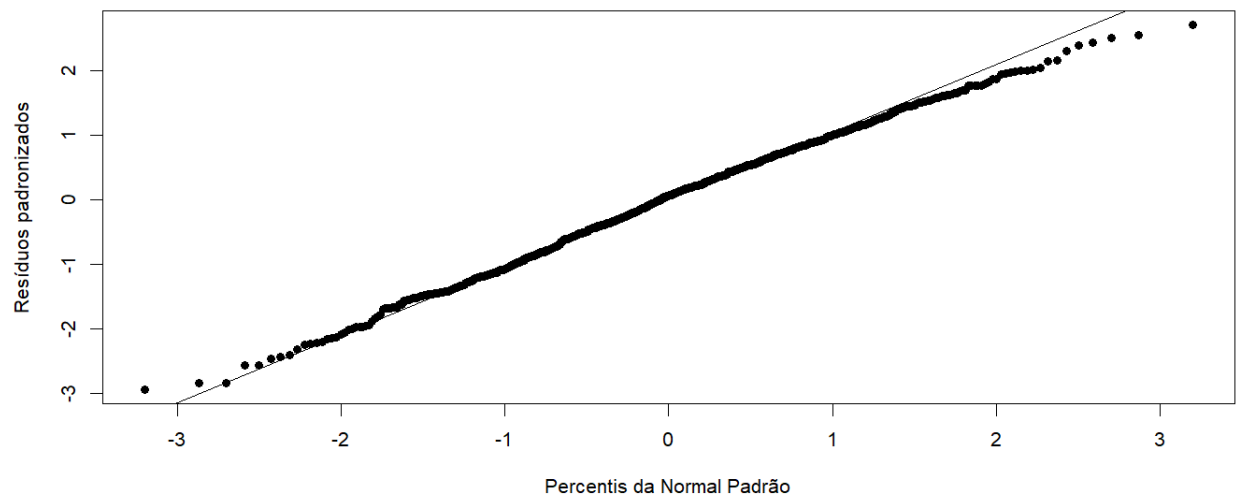


Figura A.34: Gráfico normal de probabilidades - *qq-plot* - ($MP_{2,5}$) - Modelo Normal Inversa - Santos

Figura A.35: Gráfico normal de probabilidades - *qq-plot* - $(MP_{2,5})$ - Modelo Gama - SantosFigura A.36: Gráfico normal de probabilidades - *qq-plot* - $(MP_{2,5})$ - Modelo Log-Normal - Santos

Apêndice B

Contribuição média por região

B.1 MP_{10} - Piracicaba

As Figuras B.1, B.2, B.3 e B.4 mostram as contribuições médias de cada região nos modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, para os dados de MP_{10} em Piracicaba.

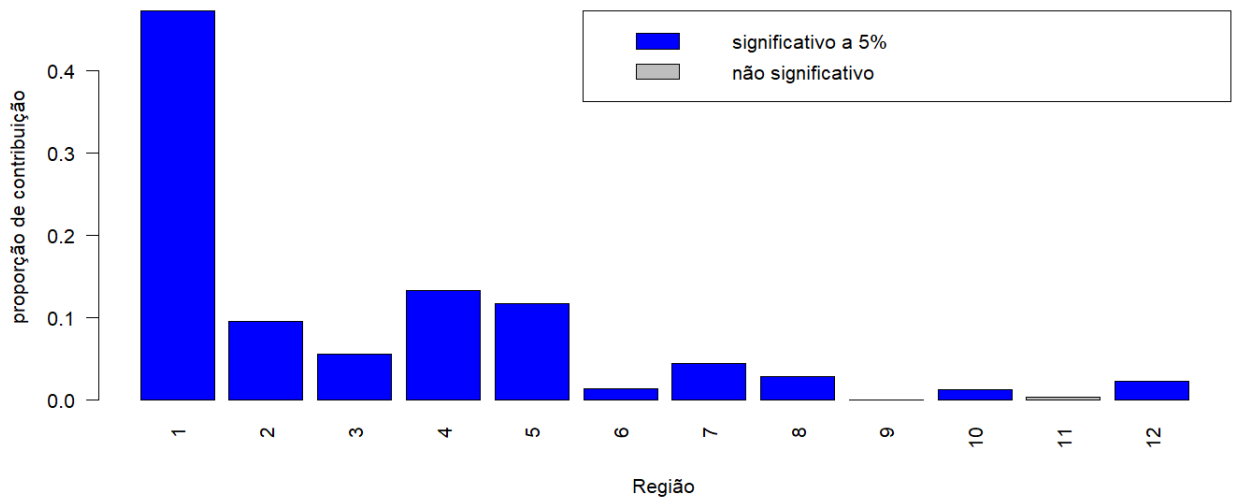
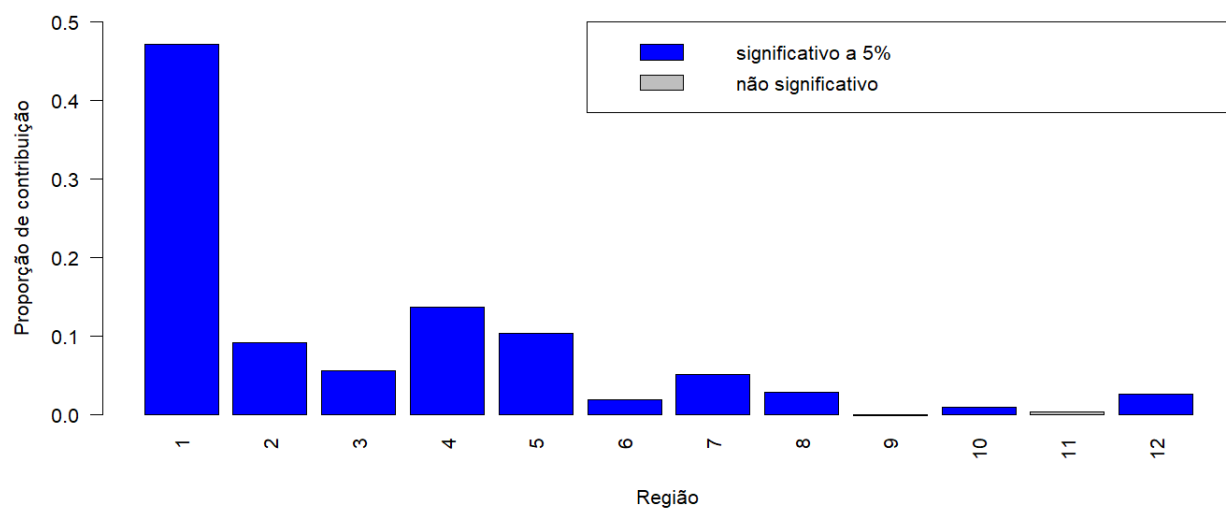
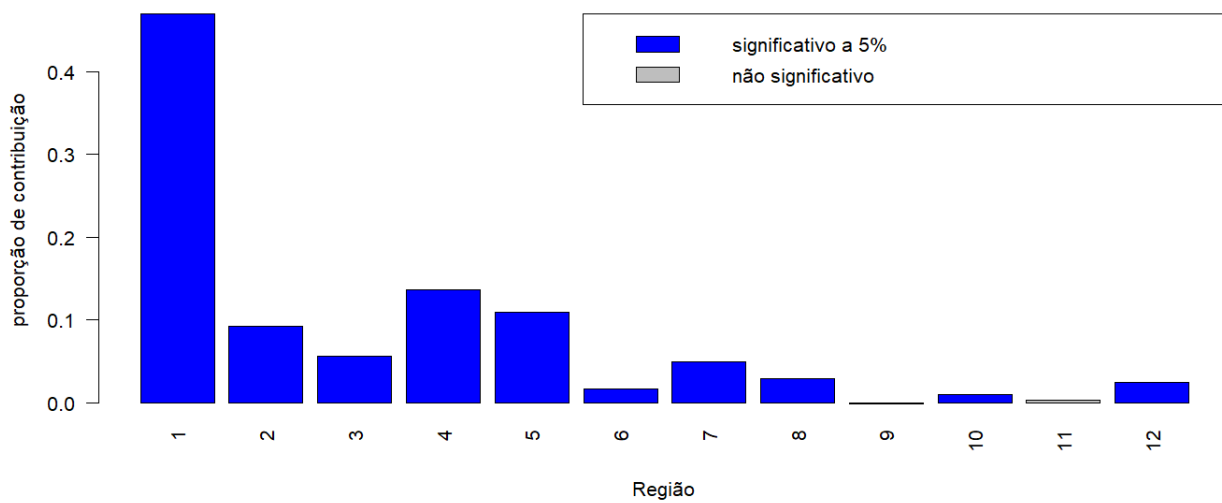
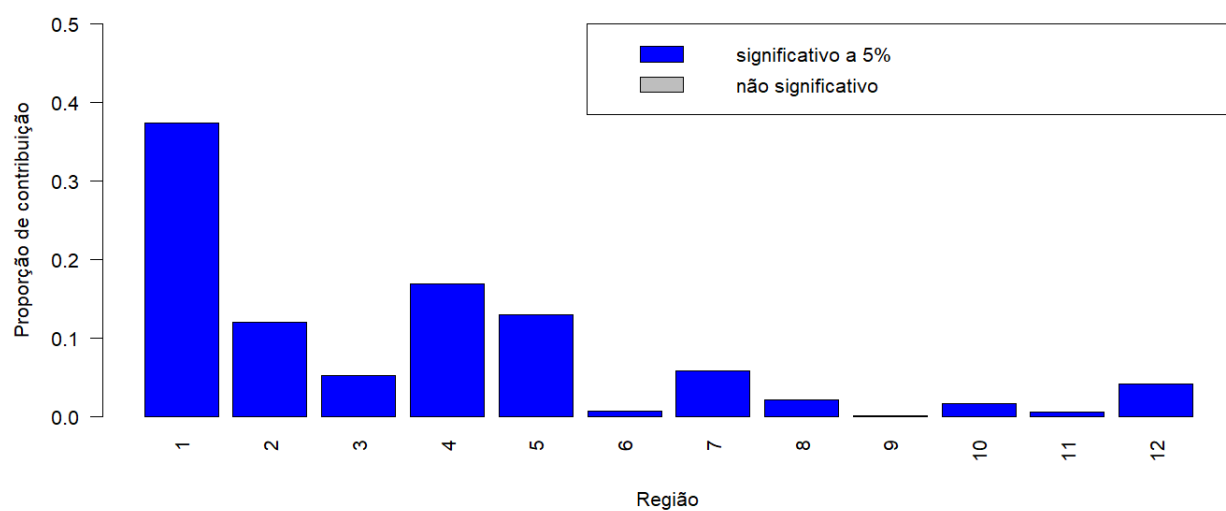


Figura B.1: Contribuição média do MP_{10} por área - Modelo Normal - Piracicaba

Figura B.2: Contribuição média do MP_{10} por área - Modelo Normal Inversa - PiracicabaFigura B.3: Contribuição média do MP_{10} por área - Modelo Gama - Piracicaba

Figura B.4: Contribuição média do MP_{10} por área - Modelo Log-Normal - Piracicaba

B.2 MP_{10} - Santos

As Figuras B.5, B.6, B.7 e B.8 mostram as contribuições médias de cada região nos modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, para os dados de MP_{10} em Santos.

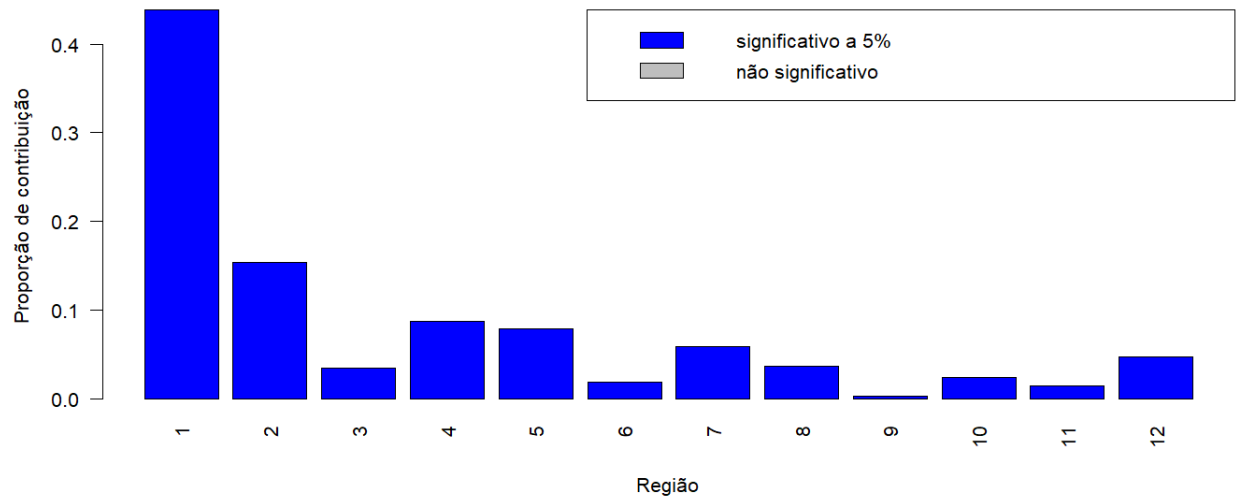


Figura B.5: Contribuição média do MP_{10} por área - Modelo Normal - Santos

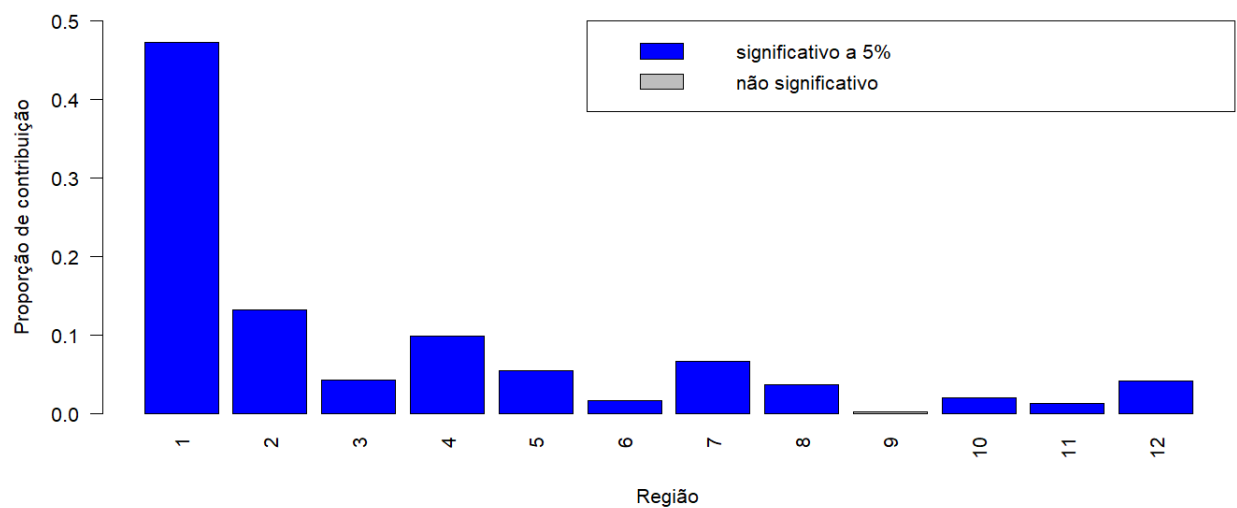
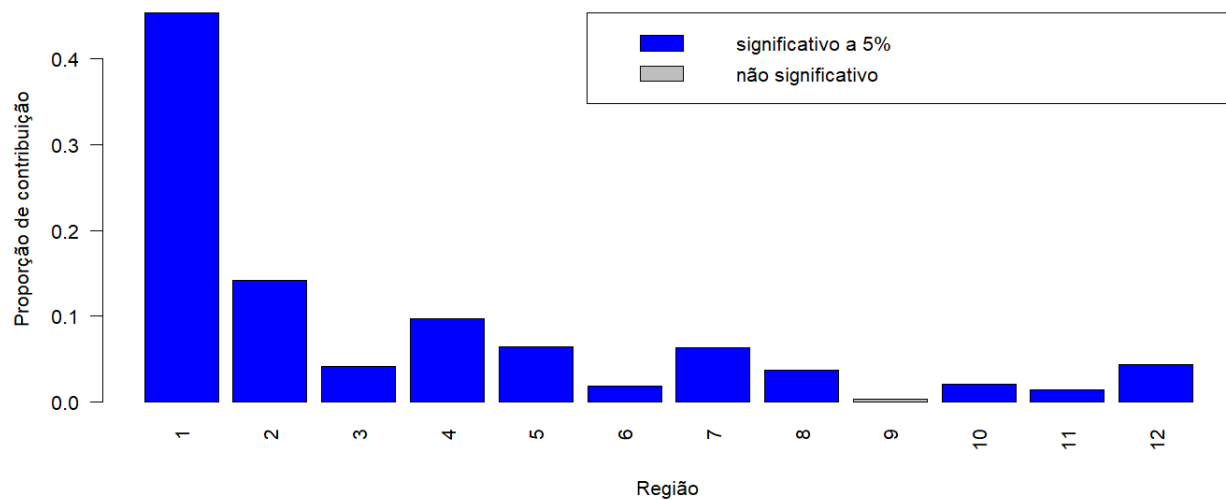
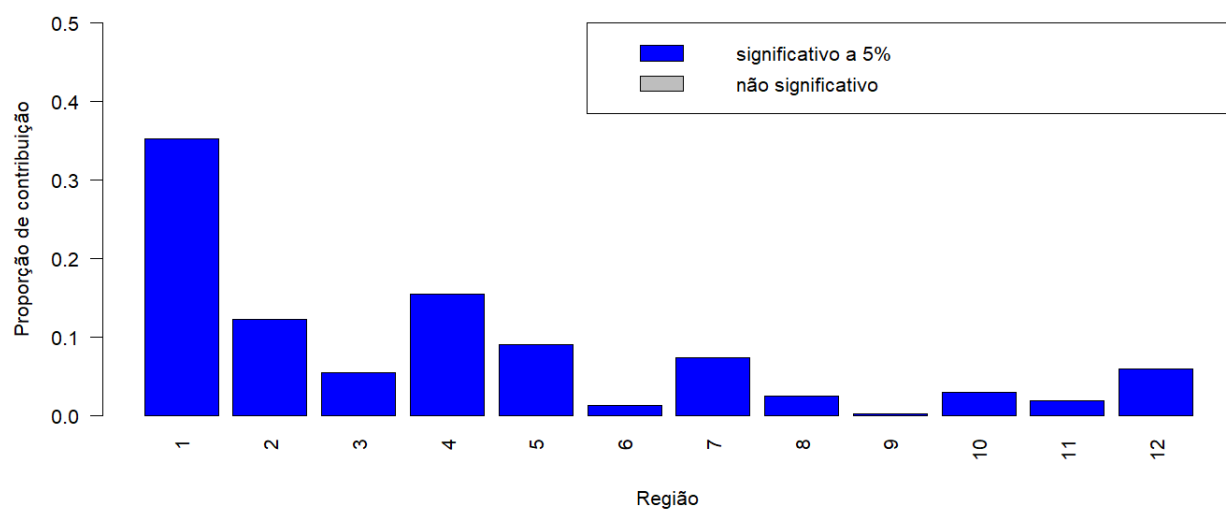


Figura B.6: Contribuição média do MP_{10} por área - Modelo Normal Inversa - Santos

Figura B.7: Contribuição média do MP_{10} por área - Modelo Gama - SantosFigura B.8: Contribuição média do MP_{10} por área - Modelo Log-Normal - Santos

B.3 $MP_{2,5}$ - Piracicaba

As Figuras B.9, B.10, B.11 e B.12 mostram as contribuições médias de cada região nos modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, para os dados de $MP_{2,5}$ em Piracicaba.

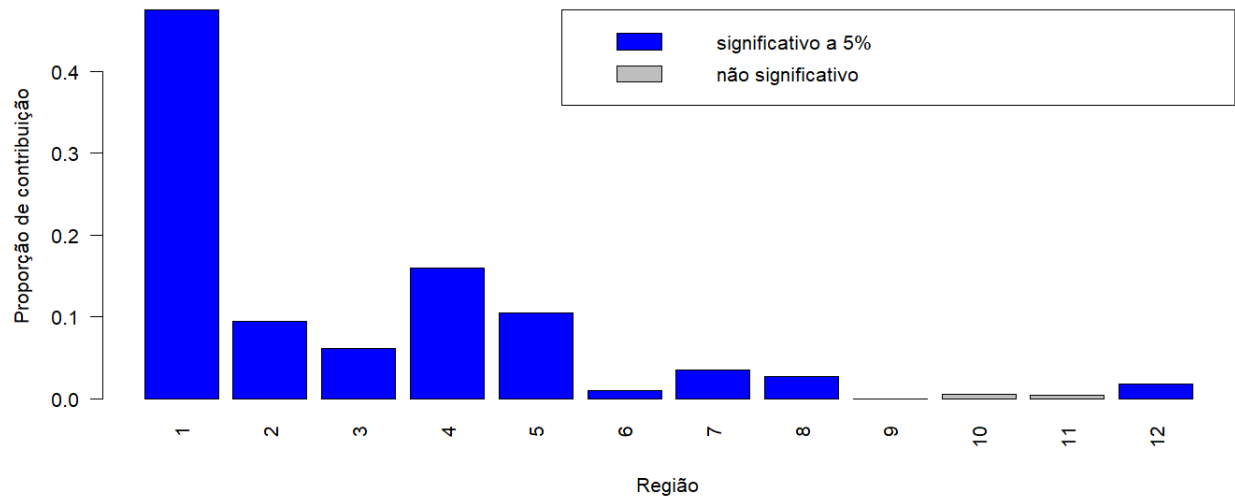


Figura B.9: Contribuição média do $MP_{2,5}$ por área - Modelo Normal - Piracicaba

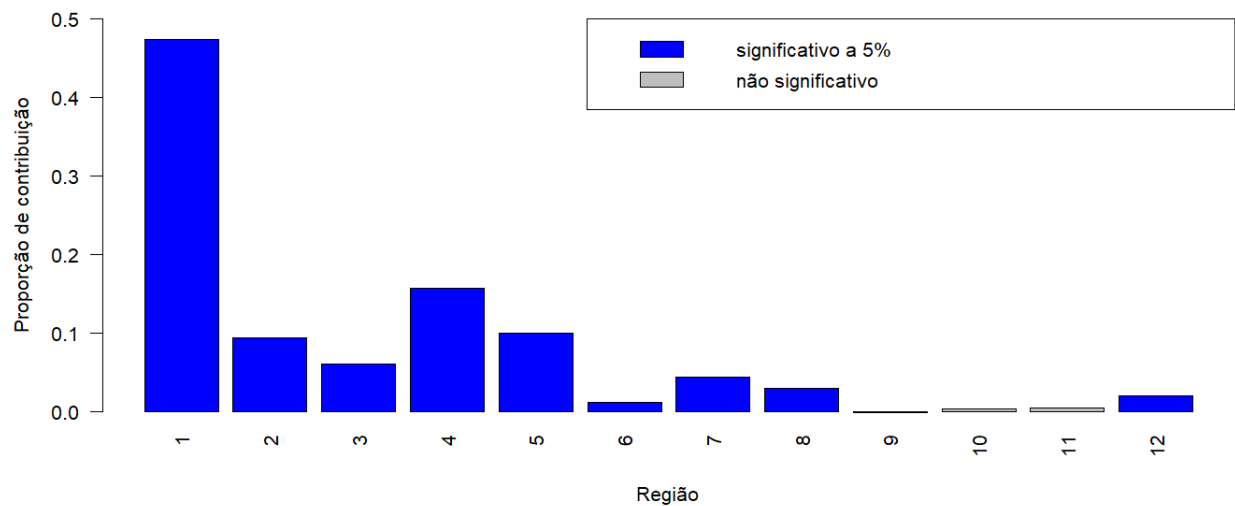
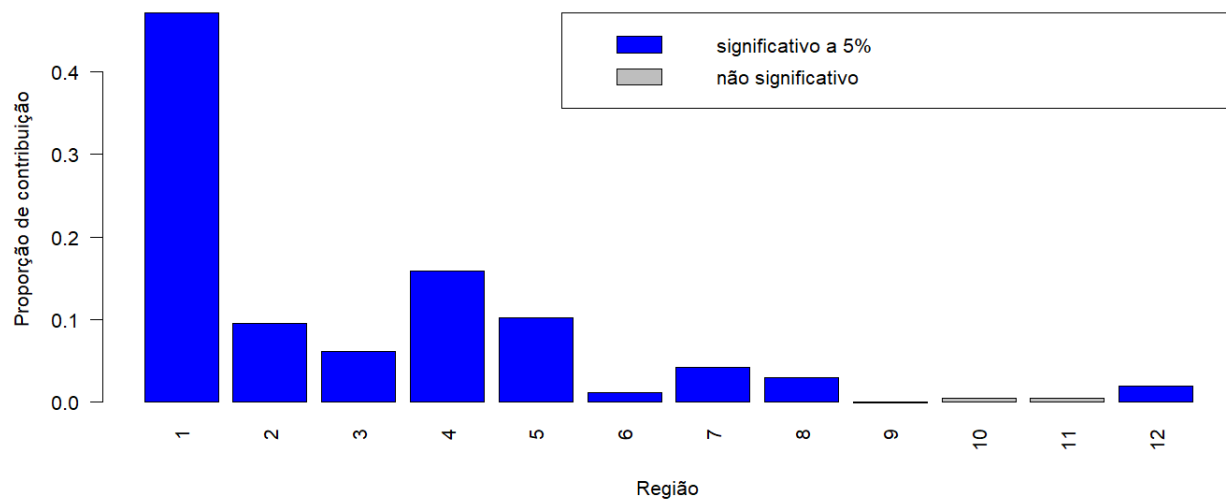
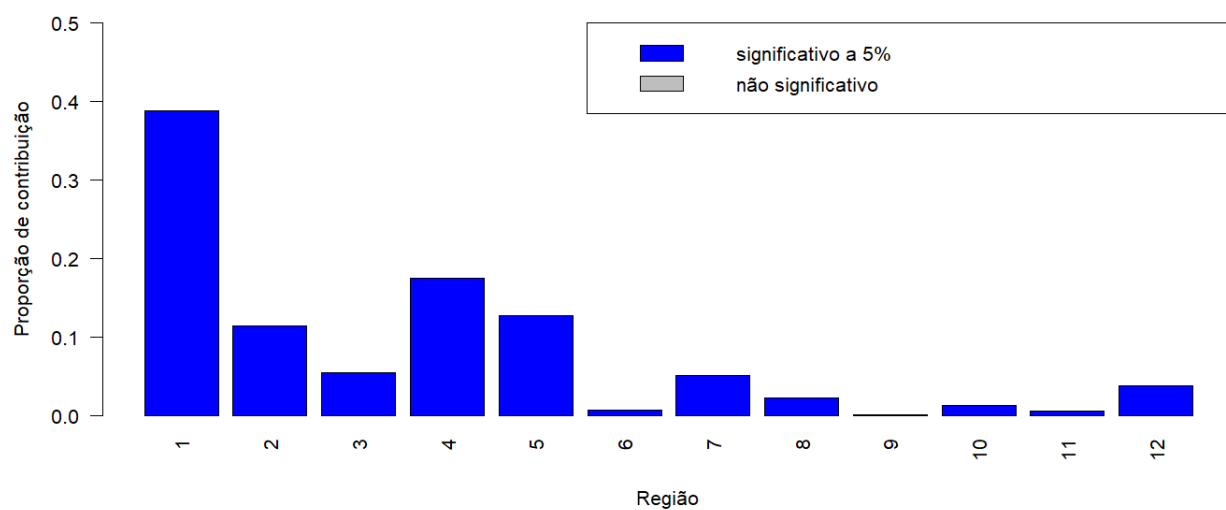


Figura B.10: Contribuição média do $MP_{2,5}$ por área - Modelo Normal Inversa - Piracicaba

Figura B.11: Contribuição média do $MP_{2,5}$ por área - Modelo Gama - PiracicabaFigura B.12: Contribuição média do $MP_{2,5}$ por área - Modelo Log-Normal - Piracicaba

B.4 $MP_{2,5}$ - Santos

As Figuras B.13, B.14, B.15 e B.16 mostram as contribuições médias de cada região nos modelos Normal, Normal inversa, Gama e Log-normal, respectivamente, para os dados de $MP_{2,5}$ em Santos.

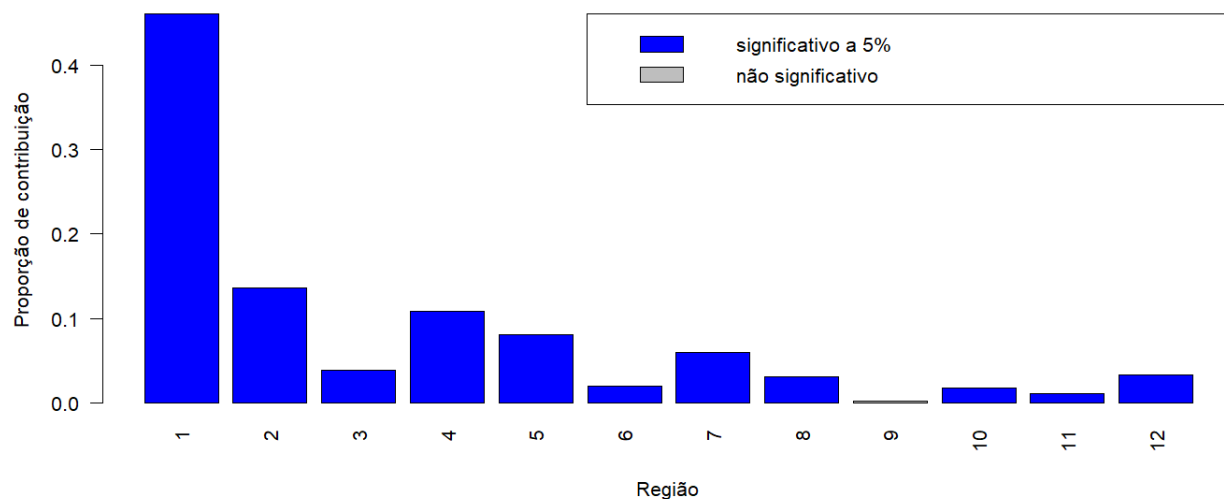


Figura B.13: Contribuição média do $MP_{2,5}$ por área - Modelo Normal - Santos

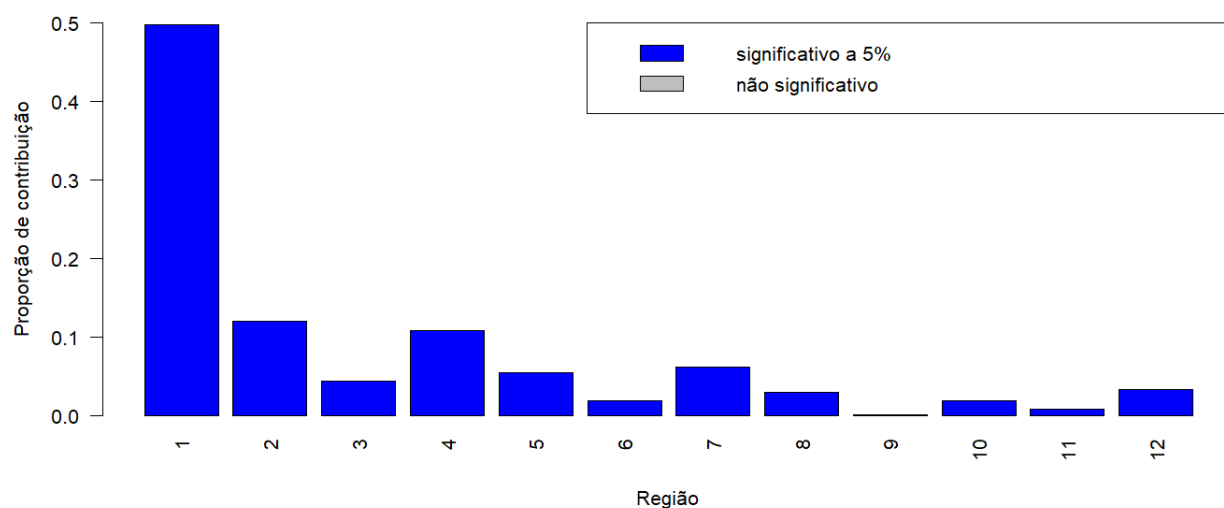
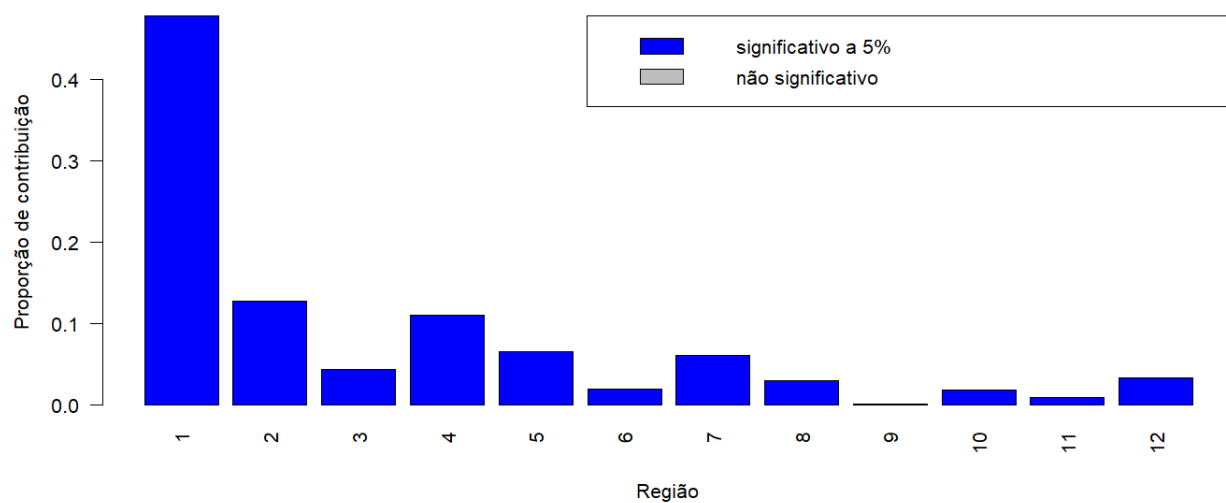
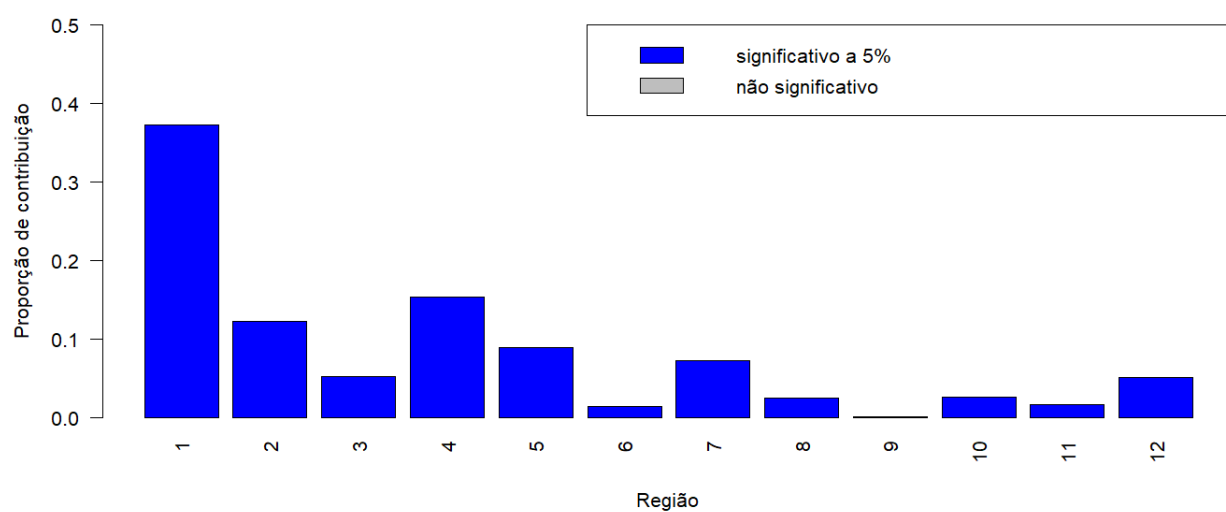


Figura B.14: Contribuição média do $MP_{2,5}$ por área - Modelo Normal Inversa - Santos

Figura B.15: Contribuição média do $MP_{2,5}$ por área - Modelo Gama - SantosFigura B.16: Contribuição média do $MP_{2,5}$ por área - Modelo Log-Normal - Santos

Apêndice C

Dendograma

As Figuras C.1, C.2, C.3 e C.4 mostram o dendograma resultante da AC para os dados de MP_{10} e $MP_{2,5}$ em Piracicaba, e MP_{10} e $MP_{2,5}$ em Santos, respectivamente.

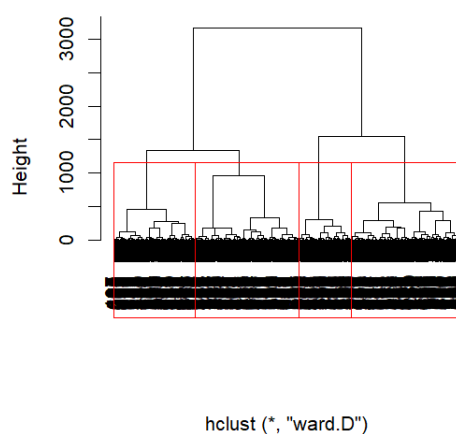


Figura C.1: Dendograma da Análise de Cluster do MP_{10} - Piracicaba

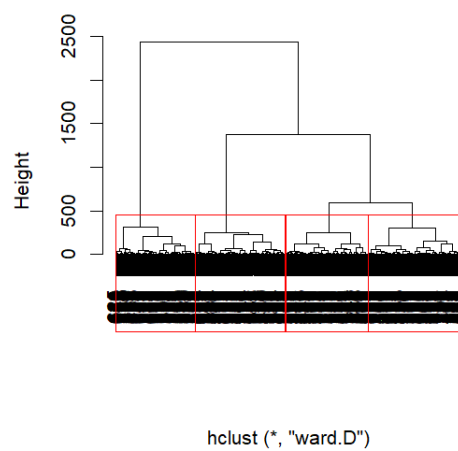


Figura C.2: Dendrograma da Análise de Cluster do MP₂₅ - Piracicaba

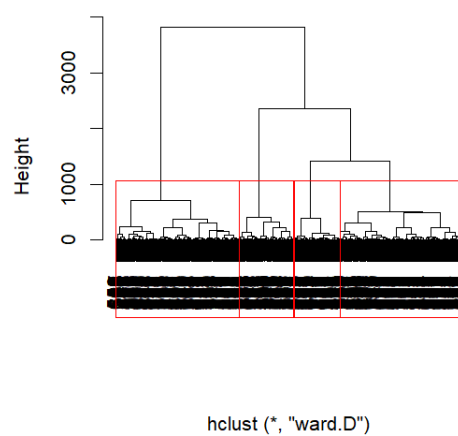


Figura C.3: Dendrograma da Análise de Cluster do MP₁₀ - Santos

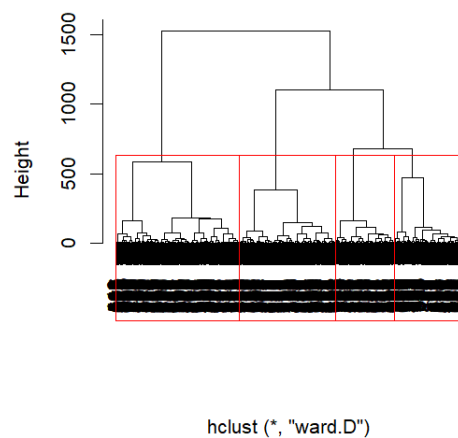


Figura C.4: Dendrograma da Análise de Cluster do MP_{2,5} - Santos