

X-002 - ESTUDO DA CONTRIBUIÇÃO REGIONAL NOS NÍVEIS DO MP10 EM PIRACICABA USANDO ANÁLISE MULTIVARIADA

Danilo Covaes Nogarotto⁽¹⁾

Estatístico pela Universidade Estadual de Campinas (UNICAMP). Mestre em Estatística pela UNICAMP. Doutorando em Tecnologia pela Faculdade de Tecnologia na UNICAMP.

Simone Andrea Pozza⁽²⁾

Professora Doutora na Faculdade de Tecnologia na UNICAMP.

Endereço⁽²⁾: Rua Paschoal Marmo, 1888 - Jd. Nova Itália - Limeira - SP - CEP: 13484-332 - Brasil - Tel: (19) 2113-3407 - e-mail: simone.pozza@ft.unicamp.br

RESUMO

O Material Particulado inalável (MP10) é um dos principais poluentes atmosféricos estudados na literatura. A concentração diária do MP10 em Piracicaba, interior do Estado de São Paulo, nos anos de 2014 e 2015 foi estudada usando a Análise de Regressão de Trajetória (ART) e a Análise de Cluster (AC). Na ART, dois diferentes modelos foram abordados sendo eles: modelo Normal e Auto-Regressivo de Médias Móveis (ARMA). Medidas comparativas foram usadas para comparar cada um deles, além de uma análise dos resíduos. Por fim, os modelos com processo AR de ordem 2, AR(2), se mostrou melhor desempenho para o MP10. Regiões locais contribuíram com aproximadamente 75% na reconstituição da concentração do MP10. Com uso da AC, foi possível encontrar 4 grupos de dias, identificando um perfil temporal de acordo com dados obtidos da ART, e dados meteorológicos do período do estudo. Regiões próximas ficaram agrupadas dentro de um mesmo grupo. E a baixa umidade relativa no inverno, foi um dos fatores relacionados com altos níveis de MP10 em Piracicaba.

PALAVRAS-CHAVE: Análise de Regressão de Trajetória, Análise de Cluster, MP10, HYSPLIT.

INTRODUÇÃO

O Material Particulado (MP) é todo material sólido e líquido que fica suspenso no ar devido ao seu pequeno tamanho, podendo ocasionar sérios problemas, principalmente respiratórios e cardiovasculares (Vallero, 2008; Rodopoulou et al., 2015; Breitner et al., 2014). Trajetórias de massas de ar são utilizadas para estudar o transporte do MP pela atmosfera. Estudos mostraram que o MP, proveniente de longas distâncias (>500 km), exerce influência significativa nos níveis de concentração do MP em determinado local (Dimitriou e Kassomenos, 2014; Kavouras et al., 2013). Estudos apontam que a quantidade de tempo que o ar gasta sobre uma região é linearmente relacionada com a proporção de contribuição desta região na concentração do MP (Xu et al., 2006; Chalbot et al., 2013; Dimitriou e Kassomenos, 2014). A Análise de Regressão de Trajetória (ART) é aplicada para descrever a dependência da concentração de poluentes atmosféricos num local em relação ao tempo de residência da massa de ar gasto sobre cada região específica. Um estudo feito em Paris identificou a contribuição das massas de ar de diversas regiões na concentração diária do MP10 e MP2,5. Obteve-se que cerca de 50% da contribuição diária é de fontes locais (Dimitriou e Kassomenos, 2014). Esses modelos foram usados para identificar as fontes de emissão nas cidades de Amsterdã, Atenas, Birmingham e Helsinki (Kavouras et al., 2013). Rodopoulou et al. (2014) reforçaram o uso da ART, aplicando seus resultados para estimar o aumento das internações cardiovasculares e respiratórias, em Doña Ana County (EUA), devido ao MP e O₃.

Análise de Cluster (AC) pode ser usada para agrupar períodos de tempo, como dias ou horas, por exemplo, a fim de identificar padrões temporais da qualidade do ar. Targino e Krecl (2016) usaram AC para agrupar dias com perfis meteorológicos similares. Em Boston, durante 2004 e 2009, Austin et al. (2012) classificaram dias em grupos baseados no perfil da concentração de poluentes (gases e particulados).

A ideia deste presente trabalho é abordar de outras maneiras a ART, incluindo uma estrutura de séries temporais por meio de um modelo Auto-Regressivo de Médias Móveis (ARMA). Assim, os principais

objetivos deste trabalho são identificar os locais de possíveis fontes de emissão do MP₁₀, e quantificar a contribuição de cada local para a cidade de Piracicaba-SP, usando ART. Por fim, a AC será utilizada para identificar um perfil temporal dos dados obtidos da ART e das variáveis meteorológicas no período.

MATERIAIS E MÉTODOS

Dados horários da concentração diária do MP₁₀ ($\mu\text{g}/\text{m}^3$), Temperatura ($^{\circ}\text{C}$), Umidade Relativa (%) e Velocidade do Vento (m/s) da estação automática da cidade de Piracicaba, nos anos de 2014 e 2015 foram usados neste trabalho. Tais dados foram coletados da base do QUALAR (Qualidade do Ar) disponível no site da CETESB (Companhia Ambiental do Estado de São Paulo) (Qualar, 2016). O modelo HYSPLIT (*Hybrid Single-Particle Lagrangian Integrated Trajectory*), do laboratório NOAA – ARL (*National Oceanic and Atmospheric Administration – Air Resources Laboratory*) foi usado para obter as trajetórias de massas de ar que chegam a Piracicaba. Considerou-se um total de 12 regiões de acordo com características similares entre elas (Figura 1 e Tabela 1).

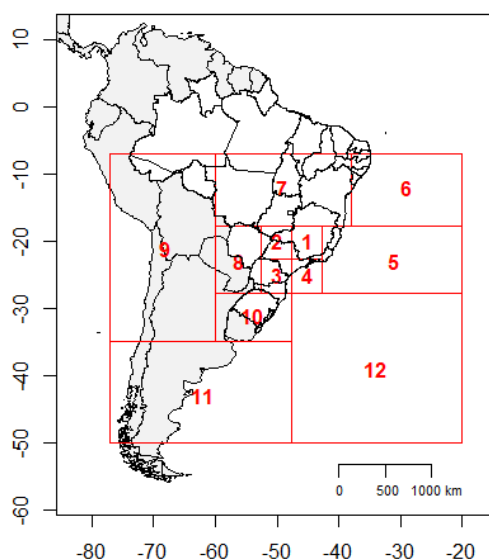


Figura 1: Regiões de residência de massa de ar em Piracicaba-SP

Neste trabalho considerou-se 2 modelos diferentes para abordagem da ART, descritos a seguir:

- i) Modelo Normal: é o modelo usual de regressão (Weisberg, 2005), o mesmo utilizado nos trabalhos de Xu et al. (2006), Huang et al. (2010), Kavouras et al. (2013) e Dimitriou e Kassomenos (2014), em que se considera a independência dos dados ao longo do tempo, ou seja, a concentração medida num dia, é independente da concentração medida em qualquer outro dia; e,
- ii) Modelo ARMA (Auto-Regressivo de Médias Móveis): Neste modelo se leva em consideração que a variável resposta é uma série temporal (Morettin e Toloi, 2006; Goyal et al., 2006). A ideia é usar a abordagem de Box-Jenkins para modelar a estrutura temporal dos dados, através de um processo ARMA, que não foi levada em consideração no modelo i).

A Correlação de Pearson (CP) e a raiz quadrada do erro quadrático médio (REQM) entre o valor observado e o valor ajustado foram calculadas para cada um dos 2 modelos a fim de comparar o desempenho deles. Também foi considerado o valor do Critério de Informação de Akaike (AIC) e a log-verossimilhança (LOG). Assim, valores menores do REQM e AIC, enquanto que valores maiores de CP e LOG são esperados para os modelos com melhor desempenho. A fim de comparação entre os modelos também foi considerada uma análise dos resíduos (Goyal et al., 2006; Weisberg, 2005).

Entre os dois modelos, foi escolhido o de melhor desempenho. Com os resultados obtidos desse modelo, foi realizada a AC. Foi considerada a AC aglomerativa hierárquica, usando o método de Ward, enquanto que a distância euclidiana foi usada para medir similaridade (Han e Kamber, 2006; Austin et al., 2012). Nesta etapa, todos os 730 dias (2 anos de medição) foram agrupados de acordo com comportamentos similares em relação as regiões de maior contribuição no MP10, e também em relação as variáveis meteorológicas (Temperatura, Umidade Relativa e Velocidade do Vento).

Tabela 1: Descrição e percentual de contribuição das regiões 1-12

Região	Fonte	Descrição	Percentual de contribuição
1	Local	Interior e sul de Minas Gerais e parte do Rio de Janeiro	43,1%*
2	Local	Interior de São Paulo e Triângulo mineiro	12,7%*
3	Local	Sul de São Paulo, Paraná e Santa Catarina	4,7%*
4	Local	Região metropolitana e Litoral Paulista	14,5%*
5	Litorânea	Espírito Santo e Litoral do Sudeste	12,3%*
6	Litorânea	Litoral do Nordeste	1,1%*
7	Continental	Norte/Nordeste/Centro-Oeste	6,0%*
8	Continental	Mato Grosso do Sul e Paraguai	1,0%*
9	Continental	Estados do Acre e Rondônia, e Bolívia, Peru, Chile e Argentina	<0,1%
10	Continental	Rio Grande do Sul e Uruguai	0,5%
11	Continental	Argentina e litoral argentino	<0,1%
12	Litorânea	Atlântico Sul	4,0%*

*: Região significativa estatisticamente a 5%.

RESULTADOS E DISCUSSÃO

Os dados da concentração do MP10, em relação ao tempo de residência das trajetórias de massa de ar, foram ajustados nos modelos Normal e AR(2). A Tabela 2 apresenta os valores de CP, REQM, LOG e AIC para os 2 modelos ajustados. Nota-se que o modelo AR(2) mostrou-se melhor em todas as medidas. Isso demonstra que esse modelo tem um desempenho de predição eficiente e satisfatório. A Figura 2 mostra os valores ajustados em relação aos valores observados da série temporal da concentração de MP10 para o ano de 2014. Observa-se que o ajuste do modelo Normal, 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. O fato do modelo AR(2) considerar a estrutura temporal dos dados, enquanto o modelo Normal não, explica a diferença no ajuste dos dados da Figura 2. Além disso, os resíduos do modelo AR(2) não apresentaram nenhum tipo de padrão em seu comportamento, mais uma vez indicando que este modelo apresenta um bom ajuste.

Tabela 2: Medidas comparativas dos modelos

Modelo	CP	REQM	LOG	AIC
Normal	0,36	18,68	-3173	6372
AR(2)	0,78	12,56	-2884	5797

A Tabela 1 e a Figura 3 mostram a proporção de contribuição de cada região pré-determinada para o modelo AR(2). Se a região é significativa, tem-se que essa região influencia na composição da concentração diária do MP10, caso contrário, não há contribuição significativa da região. Assim, nota-se que as regiões 9, 10 e 11 não foram significativas.

As maiores contribuições ocorreram das regiões locais (Regiões 1 a 4), com 75%. Sendo que a Região 1 (Interior e Sul de Minas Gerais e parte do Rio de Janeiro) foi a de maior contribuição (43,1%). Dentre as fontes locais, a região 3 foi a de menor contribuição, 4,7%. As regiões 1 a 4 representam grande parcela da região sudeste e sul do Brasil, conhecidamente as regiões mais desenvolvidas do país, com um grande número de indústrias e alto tráfego de veículos automotores, que contribuem significativamente na concentração do MP10.

Influências litorâneas na concentração do MP10 em Piracicaba foram de 17,4%, referentes às regiões 5, 6 e 12. Outras regiões (7 e 8) tiveram uma contribuição média de 7%. Não houveram contribuições significativas de outros países da América do Sul (regiões 9 a 11).

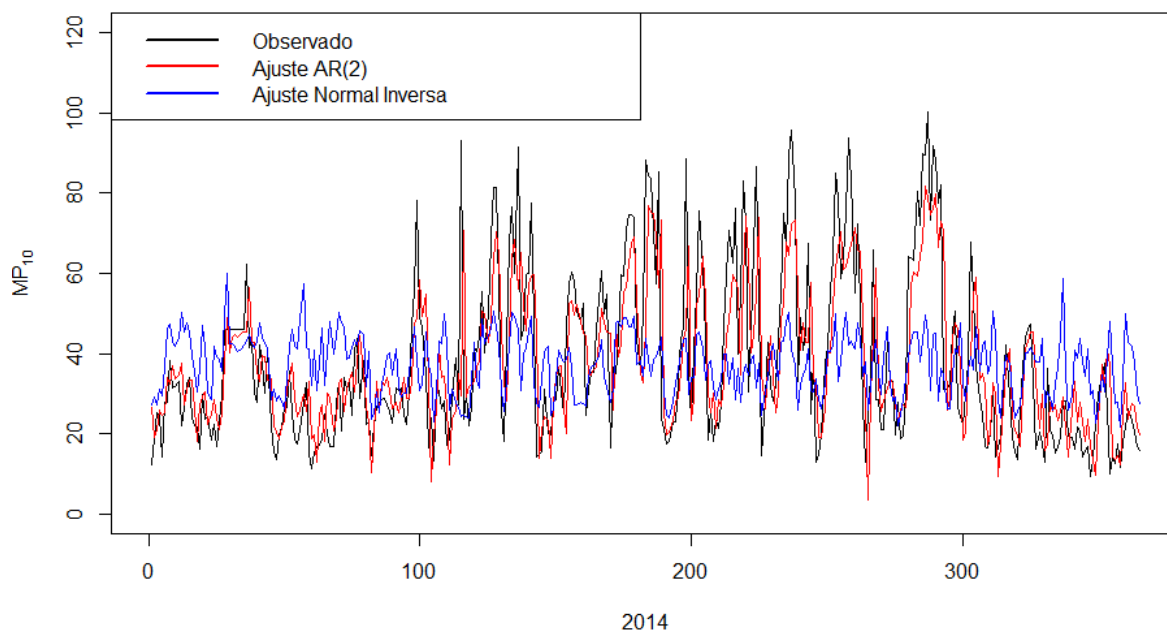


Figura 2: Ajuste comparativo dos dois modelos para dados do ano 2014

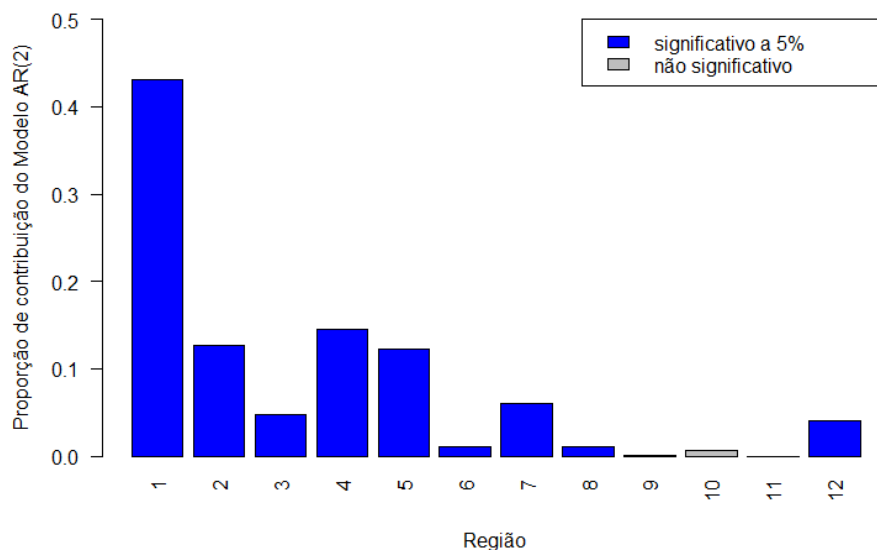


Figura 3: Percentual de contribuição das regiões 1-12

A AC foi realizada considerando as variáveis meteorológicas (Temperatura, Umidade Relativa e Velocidade do Vento) e a proporção de contribuição diária do MP10 por região (somente as regiões significativas 1-8 e 12). Os 730 dias de medição foram divididos em 4 grupos. 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 3 mostra os percentuais de dias dentro de cada estação do ano em cada um dos quatro grupos formados na AC. Já a Figura 4 mostra os boxplots da concentração do MP10, e das variáveis meteorológicas divididos nos 4 grupos. A Tabela 4 apresenta os percentuais de contribuição do grupo dentro de cada região.

Tabela 3: Percentual de dias dentro de cada estação do ano por grupo

Grupo	Primavera	Verão	Outono	Inverno
1	28,8%	23,4%	21,6%	26,1%
2	25,4%	17,1%	26,3%	31,1%
3	20,5%	27,4%	21,9%	30,1%
4	28,5%	32,0%	29,1%	10,5%
Total	25,2%	24,7%	24,9%	25,2%

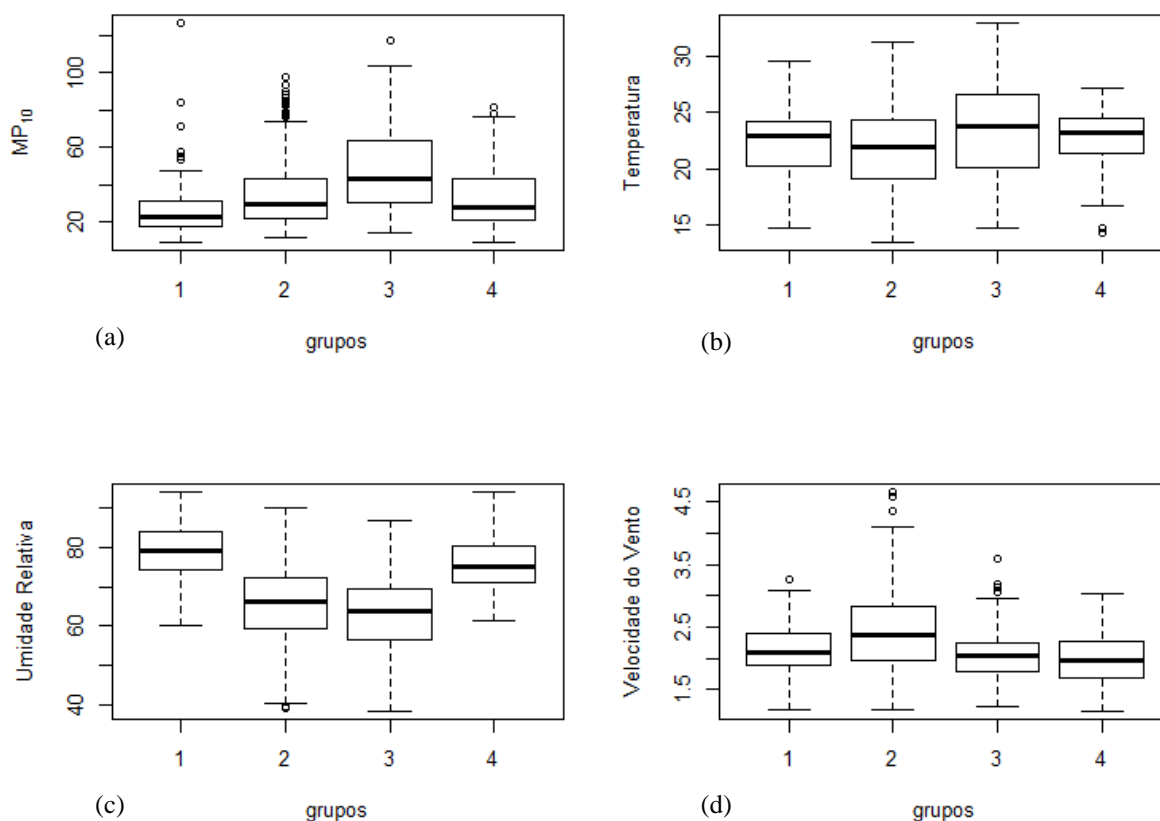


Figura 4: Box-plots da concentração do MP10 (a) e das variáveis meteorológicas (b-d) por grupo

Tabela 4: Percentual de contribuição do grupo dentro de cada região de residência do ar

	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%

O grupo 1 é o menor grupo, formado por 111 dias. Neste grupo, os dias são predominantemente da estação da primavera (Tabela 3) e ocorre os maiores valores de Umidade Relativa, o que facilita a “limpeza” da atmosfera, pois o nível mediano de MP₁₀ é menor dentre todos os grupos (Figuras 4a e 4c). Note que o percentual de umidade relativa é próximo a 80% neste grupo. As regiões 2, 7 e 8 são as que tem as maiores contribuições nesse grupo, portanto com contribuições desde o Oeste ao Norte de Piracicaba, ou mais

detalhado, do Interior de São Paulo, Triângulo mineiro, Norte, Nordeste e Centro-oeste do Brasil, além do Mato Grosso do Sul e Paraguai (Figura 1 e Tabela 1).

O grupo 2 é o maior grupo, com 228 dias. O segundo maior valor mediano da concentração do MP10 ocorre neste grupo (Figura 4a). Este grupo tem predominância dos dias no inverno (Tabela 3), com os menores valores na temperatura e maiores velocidade do vento (Figuras 4b e 4d, respectivamente). As regiões 3, 4, 8 e 12 foram as de maiores contribuições nesse grupo. Essas regiões estão localizadas mais ao sul de Piracicaba, incluindo contribuição litorânea, do Atlântico Sul.

Nota-se que, em relação ao nível mediano de MP10, o grupo 3 foi o que apresentou maior valor (Figura 4a). Este grupo, formado por 219 dias, sendo a maioria no período de inverno (Tabela 3), o que aumenta a concentração do MP10, devido ao tempo seco, isto é, com baixa umidade relativa, próximas a 60% (Figura 4c). As regiões 1, 5 e 6 mais contribuem, formadas pelo interior de Minas Gerais e parte do Rio de Janeiro, 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 verão (Tabela 3), com umidade relativa alta (Figura 4c) e níveis do MP10 não tão altos (Figura 4a). 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 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 MP10. Além disso, as variáveis meteorológicas tiveram comportamento diferente nos grupos, também indicando um fator importante na concentração do MP10.

A ART permitiu explorar o perfil espacial dos dados de MP10, controlado pelo comportamento das trajetórias das massas de ar. Enquanto isso, a AC identificou um perfil temporal do MP10, de acordo com os fatores meteorológicos e a contribuição de cada região obtidas na ART. Esta abordagem conjunta possibilitou identificar o comportamento do MP10 em Piracicaba, de uma forma bem ampla, tanto identificando os locais das fontes emissoras deste poluente, assim como observando a tendência temporal relacionada com as variáveis meteorológicas.

CONCLUSÕES

A Análise de Regressão de Trajetória é uma ferramenta muito interessante para identificar possíveis locais de fontes emissoras do MP. Com a ART, é possível verificar a contribuição percentual na concentração do MP10 de cada região resultante da análise das trajetórias de massas de ar para a cidade de Piracicaba, no interior do Estado de São Paulo, durante o período de 2014 e 2015.

Neste estudo, 2 maneiras diferentes foram abordadas na utilização da ART. A primeira, foi o modelo de regressão linear usualmente aplicado em diversos estudos anteriores. E a segunda, foi um modelo de regressão combinado com um processo Auto-Regressivo de Médias Móveis (ARMA), introduzindo a dependência temporal dos dados na estrutura de regressão. O modelo AR(2) se mostrou mais satisfatório apresentando melhores resultados em termos das medidas comparativas, indicando um melhor poder de predição da concentração do MP10. A análise comparativa dos dois modelos mostrou que a abordagem com a metodologia de Box-Jenkins (modelos ARMA) parece ser a mais adequada.

As fontes locais correspondem a aproximadamente 75% da contribuição da concentração diária do MP10 em Piracicaba. As regiões litorâneas e regionais/continentais também tiveram contribuição significativa. Por outro lado, não houve contribuição significativa dos países sul-americanos, vizinhos ao Brasil.

Na Análise de Cluster foi possível identificar um perfil temporal dos dados de MP10, agrupando os dias em grupos com comportamento similares em relação as variáveis meteorológicas e o perfil espacial obtido na ART. A AC obteve 4 grupos de dias, sendo que o nível do MP10 teve maiores valores no inverno, onde os

menores valores de umidade relativa foram encontrados. Cada grupo, identificou perfis espaciais diferentes entre os grupos, e regiões próximas ficaram agrupadas dentro de um mesmo grupo. Este fato indica que as trajetórias de massa de ar (relacionadas com a ART, e consequentemente com o perfil espacial) são um fator relevante para os níveis do MP10 em Piracicaba.

Estudos futuros podem realizar esta abordagem em outras cidades paulistas e brasileiras. Dados da concentração de outros poluentes e outras variáveis meteorológicas poderiam enriquecer ainda mais a pesquisa.

REFERÊNCIAS BIBLIOGRÁFICAS

1. AUSTIN, E., COULL, B., THOMAS, D., KOUTRAKIS, P. A framework for identifying distinct multipollutant profiles in air pollution data. *Environment International*, v.45, 112–121, 2012.
2. BREITNER, S., WOLF, K., DEVLIN, R. B., DIAZ-SANCHEZ, D., PETERS, A., SCHNEIDER, A. 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 Total Environment*, v.485-486, p.49-61, 2014.
3. 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.
4. DIMITRIOU, K., KASSOMENOS, P. A study on the reconstitution of daily PM10 and PM2.5 levels in Paris with multivariate linear regression model. *Atmospheric Environment*, v.98, p.648-654, 2014.
5. GOYAL, P., CHAN, A. T., JAISWAL, N. Statistical models for the prediction of respirable suspended particulate matter in urban areas. *Atmospheric Environment*, v.40, p.2068-2077, 2006.
6. HAN, J., KAMBER, M. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2006.
7. HUANG, L., GONG, S. L., SHARMA, S., LAVOUÉ, D., JIA, C. Q. 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.
8. KAVOURAS, I. G., LIANOU, M., CHALBOT, M. -C., VEI, I. C., KOTRONAROU, A., HOEK, G., HAMERI, K., HARRISON, R. M. Quantitative determination of regional contributions to fine and coarse particle mass in urban receptor sites. *Environmental Pollution*, v.176, p.1-9, 2013.
9. MORETTIN, P. A., TOLOI, C. M. C. *Análise de séries temporais*, 2. Ed., Blucher, 2006.
10. QUALAR. Qualidade do Ar, CETESB, Dados meteorológicos. Acesso em 16 de Janeiro de 2016. Disponível em: <<http://ar.cetesb.sp.gov.br/qualar/>>, 2016.
11. RODOPOULOU, S., CHALBOT, M. -C., SAMOLI, E., DUBOIS, D. W., FILIPPO, B. D. S., KAVOURAS, I. G. 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.
12. RODOPOULOU, S., SAMOLI, E., CHALBOT, M. G., KAVOURAS, I. G. Air pollution and cardiovascular and respiratory emergency visits in Central Arkansas: A time-series analysis. *Science of Total Environment*, v.536, p.872-879, 2015.
13. TARGINO, A. C., 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.
14. VALLERO, D. *Fundamentals of Air Pollution*, 4 ed., Academic Press, 2008.
15. WEISBERG, S. *Applied linear regression*, 3 ed., John Wiley & Sons, 2005.
16. XU, J., DUBOIS, D., PITCHFORD, M., GREEN, M., ETYEMEZHIAN, V. 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.