

Aplicação de regressão baseada no uso do solo para prever a concentração de material particulado inalável no município de São Paulo, Brasil

Application of land use regression to predict the concentration of inhalable particular matter in São Paulo city, Brazil

Mateus Habermann

Departamento de Medicina Preventiva da Faculdade de Medicina da Universidade de São Paulo (USP) – São Paulo (SP), Brasil.

Nelson Gouveia

Departamento de Medicina Preventiva da Faculdade de Medicina da USP – São Paulo (SP), Brasil.

Resumo

O estudo teve por objetivo construir um modelo de regressão baseada no uso do solo para prever a concentração material particulado inalável (PM_{10}) no município de São Paulo, Brasil. O estudo se baseou na média de PM_{10} de 2007 de 9 estações de monitoramento. Obtiveram-se dados demográficos, viários e de uso do solo em círculos concêntricos de 250 a 1.000 m para compor o modelo. Calculou-se regressão linear simples para selecionar as variáveis mais robustas e sem colinearidade. Quatro variáveis entraram no modelo de regressão múltipla. Somente tráfego leve em círculos concêntricos ≤ 250 m permaneceu no modelo final, que explicou 63,8% da variância de PM_{10} . Verificou-se que o método de regressão baseada no uso do solo é rápido, de fácil execução. Entretanto, este modelo se baseou em medições de PM_{10} de poucos locais.

Palavras-chave: poluição do ar; exposição ambiental; material particulado; impacto ambiental; saúde ambiental.

Abstract

The study intended to develop a land use regression model to predict inhalable particulate matter (PM_{10}) concentrations in São Paulo, Brazil. The model was based on the 2007 average of PM_{10} available from 9 monitoring stations and demographic, roads and land use data in buffers of 250 to 1,000 m. Simple linear regression were used to select the more robust independent variables and those without collinearity. Four variables were selected for the multiple regression analysis. Only light traffic in buffers ≤ 250 m remained in the final model which explained 63.8% of the PM_{10} variance. The land use regression approach is a quick and easy method to predict air pollution levels. However, our model was based on measurements of only a few sites.

Keywords: air pollution; environmental exposure; particulate matter; environmental impact; environmental health.

Introdução

Exposição é um termo amplo, usado para denotar quaisquer dos atributos de um sujeito ou agente (contaminantes) que pode ser causa ou preditor de um agravo. Ela consiste no contato entre o indivíduo (via inalação, ingestão, contato dérmico) e o contaminante, em diferentes cenários ambientais (ar, água e solo) e envolve o quanto este foi ingerido ou absorvido pelo organismo (ZOU *et al.*, 2009).

A abundância de um poluente, definido como massa ou fração do material por volume ou área no ambiente, aumenta a chance de

contato com o contaminante com a população alvo em um dado tempo/espaço. O ambiente pode agir, portanto, como um facilitador da exposição aos poluentes e, com isso, desencadear diversos desfechos (NUCKOLS; WARD; JARUP, 2004).

As avaliações de exposição abordam o grau de interação entre as fontes contaminantes em vários níveis de detalhe, podendo incluir três abordagens de monitoramento: biológico (biomarcadores), pessoal (medições pessoais) e ambiental (modelagem ambiental) (ZOU *et al.*, 2009).

As diferentes abordagens para avaliar exposição foram hierarquizadas pela *US National Research Council* e as abordagens baseadas

em informações individuais foram consideradas as melhores formas de estimar exposição real. Em contrapartida, aquelas baseadas em dados residenciais ou ocupacionais em áreas geográficas definidas foram consideradas menos acuradas (NIEUWENHUIJSEN; PAUSTENBACH; DUARTE-DAVIDSON, 2006). No entanto, devido aos altos custos das avaliações individuais, tem-se buscado aperfeiçoar os métodos relacionados a medições e modelagens ambientais, assumindo que essas medições sejam uma boa aproximação da exposição individual nos estudos epidemiológicos.

Alguns estudos avaliam a poluição atmosférica utilizando as concentrações médias dos poluentes medidos pelas estações de monitoramento da qualidade do ar. No entanto, por motivos de caráter econômico e administrativo há limitação no número e na distribuição dessas estações, o implica na mensuração apenas da concentração regional dos poluentes sem detectar a heterogeneidade espacial dos mesmos, que possuem alta variação espacial a poucas distâncias da fonte geradora (SAHSUVAROGLU *et al.*, 2006).

Há locais em que a concentração da poluição atmosférica é mais elevada. As causas para isso podem ser a dispersão limitada de poluentes, como corredores de prédios (*street canyons*), ou emissões elevadas originárias do tráfego intenso, indústrias, estações ferroviárias, aeroportos e portos (WHO, 2006). Por isso, modelos baseados apenas em medições de estações de monitoramento não refletem com exatidão a exposição à poluição atmosférica de pessoas residentes próximas a estas áreas.

Muitos modelos de exposição ambiental à poluição atmosférica se baseiam na proximidade da fonte poluidora, por exemplo, a proximidade ao tráfego (JERRET; FILKENSTEIN, 2005). Embora esta abordagem seja mais simples, ela é limitada por considerar que os sujeitos estão igualmente expostos e desconsidera outros parâmetros que influenciam a dispersão dos poluentes, por exemplo, topografia, uso do solo, temperatura e estagnação atmosférica (JERRET; FILKENSTEIN, 2005; RYAN; LEMASTERS, 2007). Portanto, esses modelos vêm se apresentando cada vez mais refinados ao longo do tempo, incluindo variáveis que influenciam na dispersão dos poluentes (ZOU *et al.*, 2009) capazes de identificar variações, em pequena escala, dos níveis dos poluentes, como tipo de rodovia/via, fluxo de tráfego e uso do solo, densidade populacional etc.

Um método que se tornou frequente na literatura sobre poluição atmosférica é a regressão baseada no uso do solo (*land use regression*). Ele estima os níveis de poluição em uma dada área, baseado em características que se atribuem estar relacionadas com a concentração e dispersão dos poluentes (SU *et al.*, 2009) e utiliza medições de poluição em um determinado local como variável dependente e o tipo de uso do solo, tráfego e características geográficas dentro de círculos concêntricos (*buffers*) ao redor deste local como preditores das concentrações medidas (ARAIN *et al.*, 2007).

Este estudo teve por objetivo construir um modelo de regressão baseado no uso do solo para prever as concentrações ambientais de MP_{10} no município de São Paulo.

Metodologia

O estudo foi realizado no município de São Paulo (MSP), que se constitui como centro de uma grande metrópole, com uma população de 11.324.102 habitantes, alta densidade demográfica (≈ 7.500 habitantes por km^2), numa área de 1.509 km^2 e um grau de urbanização de 98,9% (SÃO PAULO, 2011). A cidade caracteriza-se pela elevada emissão de poluentes, principalmente atribuída à frota veicular superior a 7 milhões de veículos (SÃO PAULO, 2011) circulando em ≈ 16.300 km de vias.

Dados da média anual de MP_{10} para o ano de 2007 foram obtidos de medições das estações automáticas de monitoramento da qualidade do ar localizadas no município de São Paulo, e disponibilizadas no sítio eletrônico da Companhia de Tecnologia de Saneamento Ambiental (CETESB), cujas medições eram representativas para o ano em questão. Nesse caso, a adoção de critérios de representatividade dos dados para uma determinada estação ou período significa que as medições não apresentaram falhas que comprometam a interpretação do resultado obtido. No caso da média diária, é representativa a informação que engloba dois terços das medições horárias válidas no dia (≥ 16 horas). A média anual é considerada representativa se atingirem pelo menos metade das medições diárias válidas para os quadrimestres janeiro-abril, maio-agosto e setembro-dezembro (SÃO PAULO, 2008).

Foram utilizadas informações do Censo 2000 do Instituto Brasileiro de Geografia e Estatística (IBGE), como total de residentes em domicílios particulares e total de domicílios particulares, por setor censitário. As informações de altitude do município, em intervalos de 5 m, foram obtidas da base de dados Geolog da Prefeitura de São Paulo.

A base cartográfica de uso do solo, por quadra fiscal, do MSP, foi elaborada pela Secretaria Municipal de Planejamento (SEMPPLA) e conta com 15 classes de predominância de uso, por tipologia de ocupação (horizontal e vertical), padrão das edificações (baixo, médio e alto), tipo de uso (residencial, industrial, comércio e serviços, garagens, escolas, equipamentos de uso público, outros usos), terrenos vagos e sem predominância.

A Companhia de Engenharia de Tráfego (CET) realiza contagem de veículos em algumas vias do município, nos horários de pico da manhã e utilizam o programa canadense EMME-2 para simular o tráfego nas demais vias do município. Baseando-se nas contagens e simulações, a CET forneceu informações do tráfego médio de 2007. Também foram incluídos dados das vias, por tipo (vias locais, coletoras, arteriais e de trânsito rápido), número de faixas de rolamento e velocidade média do fluxo.

Todas as variáveis independentes de uso do solo, demográficas, tráfego e altitude foram construídas utilizando o programa de sistema de informação geográfica (SIG) Mapinfo (Professional version 8.5; Mapinfo Corporation, New York, NY, USA) e se baseou em círculos concêntricos de 250, 500, 750 e 1.000 m de raio no entorno das estações de monitoramento da CETESB (Figura 1). Foi possível obter

um total de 113 variáveis independentes exploradas na construção do modelo final e são apresentadas de forma resumida no Quadro 1.

Foi criado um mapa de 6.098 pontos regulares de 500x500 m do MSP, no qual se aplicou a fórmula obtida pelo modelo de regressão final, de forma a prever os níveis de MP₁₀ para cada ponto. O mapa da concentração de MP₁₀ no município em 2007 (em µg.m³) foi elaborado no SIG ArcGIS (versão 9.3.1 ESRI, Redlands, CA, USA) por meio da ferramenta de interpolação, que se baseou no método denominado Kriging, atribuindo valores da variável de interesse em pontos desconhecidos utilizando valores de pontos conhecidos.

A associação entre características de uso do solo, demográficas, tráfego e altitude obtidas pelas 113 variáveis independentes com os níveis de MP₁₀ da respectiva estação de monitoramento foi avaliada por regressão linear simples ou univariada (erro tipo I ou α=5%). Aquelas com valor preditivo (valor p) ≤0,2 nesta análise foram selecionadas para cálculo do modelo final. A colinearidade das variáveis foi examinada pelo teste de correlação de Pearson. No caso de haver colinearidade, foi selecionada aquela mais robusta (menor valor p). Esses cálculos foram realizados no programa SPSS para Windows, versão 13.0.

O modelo final foi obtido por regressão linear múltipla, pelo método *bootstrap* (MOORE *et al.*, 2007), no qual as análises foram replicadas 150 vezes. As análises múltiplas eram refeitas, com a exclusão das variáveis sem significância estatística. Esta prática foi repetida até se obter a(s) variável(s) com p≤0,05. A regressão linear múltipla foi realizada no programa *Statistics Data Analysis 9.0* (STATA).

Por fim, a fórmula obtida na regressão múltipla foi aplicada no mapa de pontos 500x500 m do MSP para o mapeamento por interpolação.

Resultados

Das 12 estações de monitoramento da CETESB que mediram a concentração de MP₁₀, em 2007, somente 9 foram representativas e utilizadas para construção do modelo. A concentração média de MP₁₀ obtida dessas medições das estações de monitoramento foi de 40,1 µg.m⁻³ (desvio padrão [DP]=4,6 µg.m⁻³). As estações de monitoramento Cambuci e Congonhas foram as que apresentaram maior média de MP₁₀, seguida pela estação Centro (Tabela 1).

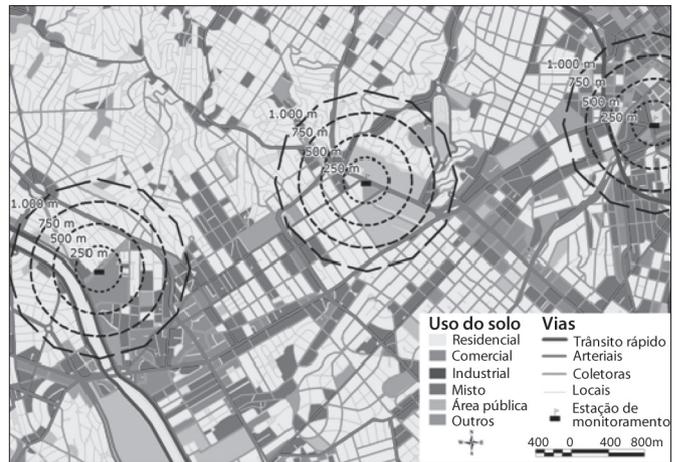


Figura 1 – Representação das estações de monitoramento da CETESB e dos círculos concêntricos construídos para obtenção das variáveis do modelo de regressão baseada no uso do solo.

Quadro 1 – Variáveis utilizadas para construção do modelo de regressão baseada no uso do solo no município de São Paulo.

USO DO SOLO	TRÁFEGO	VIAS	DEMOGRAFIA	ALTITUDE
Círculos concêntricos com raios de 250, 500, 750 e 1.000 m				ALTITUDE
Área (km ²) dos tipos de uso	Média de tráfego das vias	Comprimento das vias (m)	População e domicílios	
Residencial horizontal baixo padrão.	Locais — tráfego de todos os tipos de veículos.	Vias locais	Total de pessoas residentes em domicílios particulares permanentes.	
Residencial horizontal médio e alto padrões.		Coletoras		
Residencial horizontal.	Arteriais, coletoras e trânsito rápido — tráfego de veículos leves.			
Residencial vertical baixo padrão.		Arteriais		
Residencial vertical médio e alto padrões.				
Residencial vertical				
Residencial baixo padrão.	Arteriais, coletoras e trânsito rápido — tráfego de veículos pesados.	Altitude (m) das estações de monitoramento.		
Residencial médio e alto padrões.				
Residencial, comercial e serviços.	Arteriais, coletoras e trânsito rápido — tráfego de veículos pesados.	Trânsito rápido	Total de domicílios particulares permanentes.	
Industrial e armazéns.				
Residencial e comercial/serviços.	Arteriais, coletoras e trânsito rápido — tráfego de todos os tipos de veículos.			
Residencial e industrial/armazéns.				
Comercial/serviços e industrial/armazéns.	Arteriais, coletoras e trânsito rápido — tráfego de todos os tipos de veículos.			
Garagens.				
Equipamentos de público.	Arteriais, coletoras e trânsito rápido — tráfego de todos os tipos de veículos.			
Escolas.				
Terrenos vagos.	Arteriais, coletoras e trânsito rápido — tráfego de todos os tipos de veículos.			
Outros usos.				

Do total de 113 variáveis construídas por SIG, 24 delas apresentaram $p \leq 0,2$ na análise de regressão linear univariada. Estas se referiram ao tráfego de veículos, extensão de vias, total de domicílios, uso residencial, terrenos vagos e indústrias/armazéns dos círculos concêntricos. No entanto, 20 delas apresentaram colinearidade e foram excluídas da análise, restando apenas quatro para o cálculo de regressão multivariada. Três variáveis finais foram associadas ao aumento dos níveis de material particulado (coeficiente ou β positivo): tráfego de veículos leves; total de domicílios; uso comercial/serviços e industrial/armazéns. Inversamente, o uso residencial horizontal de médio e alto padrões foi negativamente associado com o nível de material particulado (Tabela 2).

Tabela 1 – Média anual de MP_{10} das estações de monitoramento do município de São Paulo, 2007.

Estação de monitoramento	média de MP_{10} (em $ug.m^{-3}$)
Cambuci	46
Congonhas	46
Centro	45
Parque Dom Pedro II	41
Cerqueira César	39
Ibirapuera	38
Nossa Senhora do Ó	36
Santo Amaro	36
Pinheiros	34
Mooça	s/r
Parelheiros	s/r
Santana	s/r

s/r: sem representatividade

Tabela 2 – Variáveis finais do modelo univariado ($p \leq 0,2$).

Variável	Buffer (m)	r	r ²	Coefficiente (β)	Valor p
Tráfego de veículos leves	250	0,799	0,638	$1,0 \times 10^{-3}$	0,01
Tráfego de veículos*	250	0,727	0,528	0	0,027
Extensão de vias arteriais*	250	0,717	0,514	$3,0 \times 10^{-3}$	0,03
Extensão de vias arteriais*	500	0,64	0,41	$1,0 \times 10^{-3}$	0,063
Uso residencial horizontal*	500	0,635	0,403	$-2,35 \times 10^{-5}$	0,066
Uso residencial horizontal — médio e alto padrões*	750	0,63	0,397	$-1,17 \times 10^{-5}$	0,069
Uso residencial horizontal*	750	0,63	0,397	$-1,10 \times 10^{-5}$	0,069
Uso residencial horizontal — médio e alto padrões*	500	0,628	0,395	$-2,45 \times 10^{-5}$	0,07
Extensão de vias coletoras*	1.000	0,579	0,335	0	0,102
Total de domicílios particulares permanentes	250	0,573	0,329	$4,0 \times 10^{-3}$	0,107
Extensão de vias arteriais*	1000	0,571	0,326	0	0,108
Extensão de vias arteriais*	750	0,555	0,308	$1,0 \times 10^{-3}$	0,121
Uso residencial*	250	0,632	0,399	$-9,62 \times 10^{-5}$	0,128
Total de habitantes*	250	0,537	0,288	$2,0 \times 10^{-3}$	0,136
Extensão de vias locais*	1.000	0,508	0,258	0	0,163
Uso residencial horizontal — médio e alto padrões	250	0,496	0,246	$-5,56 \times 10^{-5}$	0,174
Uso residencial horizontal*	250	0,497	0,247	$-5,55 \times 10^{-5}$	0,174
Uso comercial/serviços e industrial/armazéns	500	0,493	0,243	$8,16 \times 10^{-5}$	0,177
Total de domicílios particulares permanentes*	500	0,479	0,23	$1,0 \times 10^{-3}$	0,192
Uso Industrial e armazéns*	250	0,478	0,228	$1,0 \times 10^{-3}$	0,193
Terrenos vagos*	250	0,478	0,228	0	0,193
Extensão de vias locais*	750	0,473	0,224	0	0,198
Extensão de vias coletoras*	750	0,468	0,219	$1,0 \times 10^{-3}$	0,2
Tráfego de veículos leves*	500	0,471	0,221	$6,16 \times 10^{-5}$	0,2

*Variáveis excluídas por colinearidade.

Apenas a média diária do tráfego de veículos leves das vias coletoras, arteriais e de trânsito rápido até 250 m apresentou-se estatisticamente significativa ($p < 0,001$) e permaneceu no modelo de regressão multivariada, que reportou $r^2 = 0,638$, conforme se observa na Tabela 3.

No mapa de pontos regulares 500×500 m se aplicou a fórmula da regressão para predizer o valor de MP_{10} em 2007 ($PM_{10,2007} = 35,31484 + (0,0005114 * \text{traf_auto250})$). A média predita de MP_{10} pelo modelo de regressão foi de $36,35 \text{ ug.m}^{-3}$ ($DP = 3,24 \text{ ug.m}^{-3}$). A interpolação desses pontos gerou o mapa da concentração média diária deste poluente no município de São Paulo em 2007.

Conforme se observa na Figura 2, a concentração de MP_{10} é maior ao longo de vias com tráfego elevado e nos distritos centrais, onde há maior tráfego. Em contrapartida, as concentrações diminuem nas partes periféricas, principalmente no extremo sul do município de São Paulo.

Discussão

O modelo de regressão final do MP_{10} desenvolvido neste estudo incluiu apenas uma variável independente (média diária do tráfego

Tabela 3 – Resultados do modelo final da regressão linear multivariada.

Variável	Coefficiente (β)	Erro padrão (\hat{S}_{boot})	z	Valor p
Constante	35,31484	1,394	25,34	<0,001
Tráfego de veículos leves ≤ 250	$5,114 \times 10^{-4}$	$1,33 \times 10^{-4}$	3,85	<0,001

$r^2 = 0,6384$

de veículos leves das vias coletoras, arteriais e de trânsito rápido até 250 m). Em geral, a literatura sobre este tema tem reportado modelos finais que incluem diversas variáveis, mesmo assim o modelo foi capaz de explicar 63,8% da variabilidade na concentração média anual de MP_{10} no município de São Paulo.

Este percentual está de acordo com o Relatório Anual da Qualidade do Ar de 2007, da CETESB, no qual se estima que 40% da

concentração de MP_{10} na região metropolitana de São Paulo é originária da queima de combustível veicular e os outros 25% da ressuspensão de partículas (não detalhando se é ocasionado pelo tráfego), que somados totalizam 65% (SÃO PAULO, 2008).

Alguns autores relataram o uso de dados apenas da rede de monitoramento da qualidade do ar para gerar modelos de regressão baseada no uso do solo. Este foi o caso do modelo de Moore *et al.* (2007)

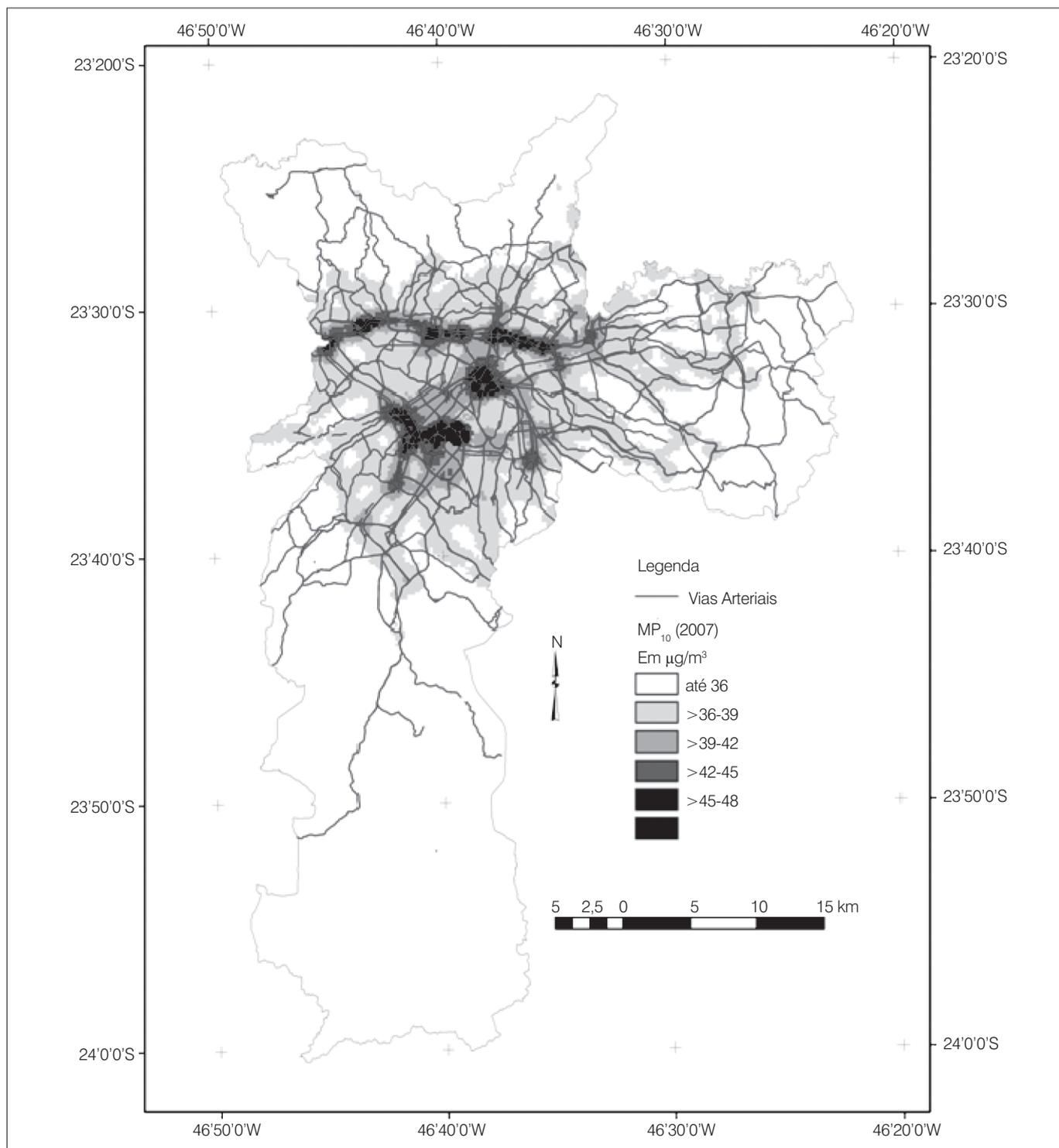


Figura 2 – Média anual da concentração de MP_{10} (em $\mu g/m^3$) no município de São Paulo, 2007.

para prever $PM_{2,5}$ na região metropolitana de Los Angeles (EUA), baseado em dados de 23 estações de monitoramento. O modelo final apresentou $r^2=0,63$ e se mantiveram as variáveis relativas a tráfego e áreas governamentais e industriais. O modelo foi aplicado em pontos aleatórios e os valores estimados foram mapeados por interpolação.

O outro foi o estudo desenvolvido em Shizuoka, Japão (KASHIMA *et al.*, 2009) e utilizou dados de medições da rede de monitoramento local. O modelo final de NO_2 apresentou poder preditivo de 52% e incluiu dados de densidade de vias, tráfego, distância do litoral e terreno cultivável. O modelo de partículas suspensas totais incluiu somente volume de tráfego de caminhões e o seu poder preditivo foi baixo ($r^2=0,11$).

Apesar da possibilidade de utilizar dados de estações de monitoramento, a literatura aborda o maciço uso de amostradores (*samplers*) para coleta de informações de poluição atmosférica. As medições são realizadas ao longo das áreas onde se pretende desenvolver modelos de regressão para estimar a concentração de poluentes.

Gilbert *et al.* (2009) mediram NO_2 em Montreal/Canadá por meio de 67 amostradores. O modelo final predisse 54% da variação deste poluente e incluiu o comprimento de vias movimentadas, fluxo veicular, uso de solo industrial e densidade populacional.

Ross *et al.* (2006) mediram NO_2 com 39 amostradores em San Diego (EUA). O modelo final ($r^2=0,79$) foi aplicado em 500 pontos aleatórios e foi feita interpolação para mapear a dispersão de NO_2 na área estudada.

O modelo desenvolvido em Roma por Rosenlund *et al.* (2009) se baseou em 70 locais de medição de NO_2 . O poder preditivo do modelo foi de 68% e incluiu a distância de vias movimentadas, tamanho do setor censitário, número de residentes, densidade populacional e altitude.

Brauer *et al.* (2003) desenvolveram um modelo de $PM_{2,5}$ na Holanda, Munique (Alemanha) e Estocolmo (Suécia), utilizando entre 40 e 42 amostradores em cada localidade. Os modelos finais explicaram 73, 56 e 50% da variabilidade média anual de $PM_{2,5}$ na Holanda, Munique e Estocolmo, respectivamente.

Su *et al.* (2009) modelaram concentrações de NO , NO_x e NO_2 em Los Angeles (EUA) com dados de medições de 181 amostradores. Os modelos finais explicaram 81, 86 e 88% da variação de NO , NO_2 e NO_x , respectivamente.

Brauer *et al.* (2008) estimaram fumaça preta, $MP_{2,5}$, NO e NO_2 em Vancouver (Canadá). O poder preditivo do modelo de NO ficou em 62,56% para NO_2 e fumaça preta e, por fim, um poder preditivo de 52% no modelo de $MP_{2,5}$.

Em Valência, Espanha, Iñiguez *et al.* (2009) desenvolveram um modelo de NO_2 com medições de 93 amostradores. O poder preditivo do modelo foi de 81% e houve tendência de decréscimo dos níveis de NO_2 da zona urbana para a rural e grande influência de uma rodovia que cruza a região.

Apenas alguns estudos utilizaram variáveis meteorológicas na construção dos modelos de regressão baseada no uso do solo (ARAIN *et al.*, 2007; BRIGGS *et al.*, 2000; MAUKO; TANG; GEORGE, 2008).

Assim como na regressão baseada no uso do solo desenvolvida no município de São Paulo, a grande maioria dos trabalhos verificados desconsidera a influência da direção/velocidade dos ventos na dispersão dos poluentes ou a sazonalidade dos níveis de poluição, o que pode diminuir a capacidade de predição dos modelos.

Briggs *et al.*, (2000) aplicaram um modelo de NO_2 em quatro cidades do Reino Unido, identificando características como temperatura, velocidade do vento e pressão atmosférica. Mauko, Tang e George (2008) construíram dois modelos: com dados de círculos concêntricos e outro com círculos distorcidos para a direção predominante dos ventos. No modelo de círculos concêntricos o r^2 foi 0,66 e no modelo baseado em círculos distorcidos o poder preditivo aumentou para 0,81.

No modelo desenvolvido em Toronto e Hamilton (Canadá) (ARAIN *et al.*, 2007), a inclusão de dados eólicos aumentou a predição do modelo de 65 para 69%. Em outro modelo de Hamilton (SAHSUVAROGLU *et al.*, 2006) com informações de direção do vento o r^2 foi de 0,76.

Embora a literatura não relate a confiabilidade e qualidade das informações de uso do solo, a forma de organização dos dados de uso do solo em São Paulo pode ter limitado o modelo, pois a informação é disponível em categorias prontas e agregadas por quadras, o que impossibilita o aumento do nível de detalhamento, por exemplo, por lotes. Além disso, a dinâmica da cidade faz com que o uso do solo se altere rapidamente e desordenadamente, principalmente nas periferias ou em antigas áreas industriais. Apesar disso, houve associação negativa entre o uso residencial horizontal de médio e alto padrões e MP_{10} no cálculo do modelo final. Este tipo de uso geralmente se caracteriza pela baixa densidade construtiva, menos domicílios e habitantes por quilometro quadrado, fatores que podem contribuir com menor circulação de veículos.

Os modelos verificados na literatura, em sua maioria, estimaram a concentração de NO_2 . Ele é considerado um bom indicador da poluição gerada por tráfego veicular e é fácil de medir (ROSE *et al.*, 2009). No caso de São Paulo, havia apenas cinco estações de monitoramento cujos dados de NO_2 eram representativos em 2007, enquanto dados de MP_{10} eram representativos em nove estações.

A maior parte desses estudos apresentou poder preditivo acima de 50%. Mesmo obtendo um percentual preditivo satisfatório, os resultados do modelo aplicado em São Paulo devem ser observados com precaução, uma vez que este se baseou em medições de poucas estações de monitoramento da qualidade do ar. Os modelos de regressão baseada no uso do solo relatados se basearam em mais pontos, mesmo aqueles que utilizaram medições de estações de monitoramento, variando entre 20 (BRIGGS *et al.*, 2000) a 181 locais (SU *et al.*, 2009) e isso aumenta a precisão dos resultados obtidos.

A aplicação de regressão baseada no uso do solo permite a mensuração da contribuição das diversas fontes nos níveis de poluição do ar (fontes pontuais ou móveis) e isso pode auxiliar para que medidas de controle e combate à emissão de poluentes sejam tomadas.

A única variável que permaneceu estatisticamente significativa no modelo final foi o volume de tráfego de veículos leves até 250 m da estação de monitoramento, o que indica a influência do tráfego veicular nos níveis ambientais de material particulado em São Paulo. A associação com tráfego veicular está de acordo com outros estudos de regressão baseada no uso do solo aplicados com outros poluentes atmosféricos e que enfatizam a contribuição das fontes móveis nos níveis de poluição atmosférica urbana (ARAIN *et al.*, 2007; BRAUER *et al.*, 2003; GILBERT *et al.*, 2009; KASHIMA *et al.*, 2009; MAUKO; TANG; GEORGE, 2008; ROSENLUND *et al.*, 2009; SAHSUVAROGLU *et al.*, 2006; SU *et al.*, 2009).

Conforme se verificou no mapa de interpolação, as maiores concentrações de MP_{10} foram em locais de elevado fluxo veicular associado a áreas com maiores densidades populacionais. Isto evidencia o potencial impacto que a exposição ao MP_{10} pode acarretar à saúde, uma vez que as maiores concentrações atingem um número maior de pessoas.

Considerações Finais

O modelo de regressão baseada no uso do solo desenvolvido no município de São Paulo foi capaz de explicar mais de 60% da

variabilidade espacial da concentração média de MP_{10} no ano de 2007. Este poder preditivo está de acordo com estudos similares. A permanência da variável volume de tráfego de veículos leves até 250 m no modelo final destaca a importância da fonte veicular como contribuinte dos níveis de MP_{10} no município de São Paulo.

O pequeno número de pontos de medição de MP_{10} limita a exatidão deste modelo de regressão baseada no uso do solo. Por isso, recomenda-se a construção de um modelo baseado na medição de MP_{10} em diversos pontos da cidade por meio de amostradores, de forma a aumentar sua precisão. Também é importante que as informações de uso do solo estejam mais detalhadas.

Deve-se enfatizar que este estudo mostrou como é rápido e fácil desenvolver modelos de regressão de uso do solo para estimar exposição aos poluentes atmosféricos, quando há disponibilidade de dados que envolvem sua elaboração. Os pontos aleatórios criados para aplicação de interpolação poderiam ser substituídos por endereço de sujeitos, de forma a estimar a exposição dos mesmos em estudos epidemiológicos, como verificado em Brauer *et al.* (2008) e Iñiguez *et al.* (2009). Tendo em vista que a realização de medições pessoais em estudos epidemiológicos é cara e logisticamente difícil, destaca-se a importância de desenvolver métodos indiretos de avaliação da exposição, entre eles a regressão baseada no uso do solo.

Referências

- ARAIN, M.A.; BLAIR, R.; FINKELSTEIN, N.; BROOK, J.R.; SAHSUVAROGLU, T.; BECKERMAN, B.; ZHANG, L.; JERRET, M. (2007) The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. *Atmospheric Environment*, v. 41, p. 3453-3464.
- BRAUER, M.; HOEK, G.; VLIET, P.; MELIEFSTE, K.; FISCHER, P.; GEHRING, U.; HEINRICH, J.; CYRYS, J.; BELLANDER, T.; LEWNE, M.; BRUNEKREEF, B. (2003) Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology*, v. 14, n. 2, p. 228-239.
- BRAUER, M.; LENCAR, C.; TAMBURIC, L.; KOEHOORN, M., DEMERS, P.; KARR, C. (2008) A cohort study of traffic-related air pollution impacts on birth outcomes. *Environmental Health Perspectives*, v. 116, n. 5, p. 680-686.
- BRIGGS, D.J.; HOOGH, C.; GULLIVER, J.; WILLS, J.; ELLIOT, P.; KINGHAM, S.; SMALLBONE, K. (2000) A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of the Total Environment*, v. 253, p. 151-167.
- GILBERT, N.L.; GOLDBERG, M.S.; BECKERMAN, B.; BROOK, J.R.; JERRET, M. (2009) Assessing spatial variability of ambient nitrogen dioxide in Montréal, Canada, with a land-use regression model. *Air & Waste Management Association*, v. 55, p. 1059-1063.
- IÑIGUEZ, C.; BALLESTER, F.; ESTARLICH, M.; LLOP, S.; FERNANDEZ-PATIER, R.; AGUIRRE-ALFARO, A.; ESPLUGUES, A. (2009) Estimation of personal NO_2 exposure in a cohort of pregnant women. *Science of the Total Environment*, v. 407, p. 6093-6099.
- JERRET, M.; FINKELSTEIN, M. (2005) Geographies of risk in studies linking chronic air pollution exposure to health outcomes. *Journal of Toxicology and Environmental Health*, v. 68, p. 1207-1242.
- KASHIMA, S.; YORIFUJI, T.; TSUDA, T.; DOI, H. (2009) Application of land use regression to regulatory air quality data in Japan. *Science of the Total Environment*, v. 407, p. 3055-3062.
- MAUKO, M.E.; TANG, B.; GEORGE, L.A. (2008) A sub-neighborhood scale land use regression model for predicting NO_2 . *Science of the Total Environment*, v. 398, p. 68-75.
- MOORE, D.K.; JERRET, M.; MACK, W.J.; KÜNZLI, N. (2007) A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. *Journal of Environmental Monitoring*, v. 9, p. 246-252.
- NIEUWENHUIJSEN, M.; PAUSTENBACH, D.; DUARTE-DAVIDSON, R. (2006) New developments in exposure assessment: the impact on the practice of health risk assessment and epidemiological studies. *Environmental Health*, v. 32, p. 996-1009.

- NUCKOLS, J.R.; WARD, M.H.; JARUP, L. (2004) Using geographic information systems for exposure assessment in Environmental Epidemiology studies. *Environmental Health Perspectives*, v. 112, n. 9, p. 1007-1015.
- ROSE, N.; COWIE, C.; GILLET, R.; MARKS, G.B. (2009) Weighted road density: a simple way of assigning traffic-related air pollution exposure. *Atmospheric Environment*, v. 43, p. 5009-5014.
- ROSENLUND, M.; FORASTIERE, F.; PORTA, D.; SARIO, D.; BADALONI, C.; PERUCCI, C.A. (2009) Traffic-related air pollution in relation to respiratory symptoms, allergic sensitisation and lung function in schoolchildren. *Thorax*, v. 64, p. 573-580.
- ROSS, Z.; ENGLISH, P.B.; SCALF, R.; GUNIER, R.; SMOROKINSKY, S.; WALL, S.; JERRET, M. (2006) Nitrogen dioxide prediction in Southern California using land use regression modeling: potential for environmental health analyses. *Journal of Exposure Science and Environmental Epidemiology*, v. 16, p. 106-114.
- RYAN, P.H.; LEMASTERS, G.K.A (2009) review of land-use regression models for characterizing intraurban air pollution exposure. *Inhalation Toxicology*, v. 19, n. Suppl.1, p. 127-133.
- SAHSUVAROGLU, T.; ARAIN, A.; KANAROGLU, P.; FINKELSTEIN, N.; NEWBOLD, B. (2006) A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario. *Air & Waste Management Association*, v. 56, p. 1059-1069.
- SÃO PAULO (estado). Companhia de Tecnologia e Saneamento Ambiental (CETESB). (2008) *Relatório de qualidade do ar no estado de São Paulo 2007*. São Paulo: CETESB, 2008. Disponível em: <<http://www.cetesb.sp.gov.br/ar/qualidade-do-ar/31-publicacoes-e-relatorios>>. Acesso em: ago. de 2009.
- SÃO PAULO (estado). Departamento Estadual de Trânsito de São Paulo (DETRAN-SP). (2011) *Frota de Veículos – DETRAN-SP*. São Paulo: 2011. Disponível em: <<http://www.detran.sp.gov.br/>>. Acesso em: ago. de 2011.
- SÃO PAULO (estado). Secretaria de Planejamento e Desenvolvimento Regional. Fundação Sistema Estadual de Análise de Dados (SEADE). (2011) *Município de São Paulo – MSP* [internet]. São Paulo: 2011. Disponível em: <<http://www.seade.sp.gov.br/produtos/msp/index.php>>. Acesso em: ago. de 2011.
- SU, J.G.; JERRET, M.; BECKERMAN, B.; WILHELM, M.; GHOSH, J.K.; RITZ, B. (2009) Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. *Environmental Health*, v.109, p. 657-670.
- WORLD HEALTH ORGANIZATION (WHO). (2006) *WHO air quality guidelines global update 2005 – Report on a working group meeting*. Bonn, Germany.
- ZOU, B.; WILSON, J.G.; ZHAN, F.B.; ZENG, Y. (2009) Air pollution exposure assessment methods utilized in epidemiological studies. *Journal of Environmental Monitoring*, v. 11, p. 475-490.