**Universidade de São Paulo
Faculdade de Saúde Pública**

**Análise de agravos à saúde e possíveis
associações aos produtos da queima de
cana-de-açúcar**

Fábio Silva Lopes

Tese apresentada ao programa de Pós-Graduação em Saúde Pública para obtenção do título de Doutor em Saúde Pública.

Área de concentração: Saúde Ambiental

Orientadora: Profa. Dra. Helena Ribeiro

São Paulo
2010

Análise de agravos à saúde e possíveis associações aos produtos da queima de cana-de-açúcar

Fábio Silva Lopes

Tese apresentada ao programa de Pós-Graduação em Saúde Pública para obtenção do título de Doutor em Saúde Pública.

Área de concentração: Saúde Ambiental

Orientador: Profa. Dra. Helena Ribeiro

São Paulo
2010

É expressamente proibida a comercialização deste documento, tanto na sua forma impressa como eletrônica. Sua reprodução total ou parcial é permitida **exclusivamente** para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano da tese

AGRADECIMENTOS:

Muitas pessoas colaboram diretamente ou indiretamente na construção de uma tese. Eu diria que é quase impossível caminhar sozinho. No entanto, não há como dividir a responsabilidade do equilíbrio entre momentos de stress e delírio. Mas, estes momentos foram minimizados com o apoio de pessoas que despenderam tempo e dividiram conhecimento, contribuindo para que este trabalho chegasse a um bom termo em sua finalização. A todos, a minha eterna gratidão.

À Sandra Perdoná, minha esposa que desempenhou um impagável trabalho de contra-regras, com apoio incondicional, sem ela não seria possível.

À Dona Carminha, entre orações e conselhos, sempre acreditou nesta jornada.

À Dona Neusa Perdoná e S. Ilso Perdoná, muito me enriqueceram com histórias e problemas do cotidiano que eles vivem dia-a-dia, no meio do mar de cana.

À professora Helena Ribeiro, com sua disponibilidade irrestrita e precisão nos apontamentos, foi, antes de tudo, em modelo de orientação que hoje procuro me espelhar.

Aos professores João Vicente Assunção, Adelaide Nardocci, Edelci Nunes e Raul Borges Guimarães que colaboraram com das muitas discussões que complementam o cotidiano do Departamento de Saúde Ambiental.

Aos professores Maria Regina Alves Cardoso, Grácia Anacleto, Péricles Turnes Jr. e Arnaldo Vallim, que me ajudaram nos momentos difíceis de imersão na estatística.

À professora Paula Laxer pelo suporte ao inglês.

Ao professor Vivaldo José Breternitz, doutorando contemporâneo, que compartilhou dúvidas e ansiedades nesta jornada, além de contribuir com muitas dicas sobre metodologia.

Ao Fernando Henrique Inocêncio B. Ferreira, meu aluno de graduação no Mackenzie, pelo apoio na elaboração de programas para manipulação das bases de dados.

Aos colegas Thomas, Marco, Samuel, Gustavo e Roberto do Laboratório de Geoprocessamento da FSP, sempre próximos e colaborativos quando o assunto é SIG.

À Faculdade de Computação e Informática da Universidade Presbiteriana Mackenzie, em especial ao professor Luciano Silva, pelo apoio ao trabalho e projetos paralelos desenvolvidos neste período de tese.

Ao VIGIAR- MS pelo apoio ao tema e suporte a projetos que caminharam em paralelo ao desenvolvimento deste trabalho.

RESUMO:

Introdução. O uso de tecnologias computacionais em pesquisas na área de saúde pública é cada vez mais acentuado devido às facilidades providas ao pesquisador. Em particular, o SIG (Sistema de Informação Geográfica) auxilia na fusão de dados dispostos em tabelas e mapas contendo aspectos geográficos do local de estudo, por meio da sobreposição de camadas. Tal fusão possibilita uma nova leitura dos dados associados à dimensão do espaço. Neste aspecto, os programas SIG são ferramentas ideais para estudos de geografia da saúde. No entanto, observa-se a escassez de estudos sobre os possíveis efeitos deletérios à saúde humana, causados pelos produtos das queimadas em culturas canavieiras. Neste contexto, os SIGs, associados a técnicas de análise espacial podem auxiliar na busca de uma melhor compreensão sobre as interações do homem e seu meio, no âmbito da saúde pública. **Objetivo.** Este estudo avaliou correlações entre os agravos à saúde respiratória e queimadas de cana-de-açúcar, utilizando análise espacial e modelos estatísticos. **Metodologia.** O desenvolvimento deste trabalho consistiu em coletar e organizar dados de queimadas e internações por afecções respiratórias, gerando um repositório de dados para duas regiões brasileiras produtoras de cana-de-açúcar: a região denominada Centro-Sul, com os estados de Goiás, Mato Grosso, Mato Grosso do Sul, Paraná e São Paulo, e a região Nordeste com os estados de Alagoas, Bahia, Sergipe, Pernambuco e Paraíba, em um período de 2000 a 2008. As análises espaciais foram elaboradas com a ferramenta Hot Spot do software Arcview 9.3 e as análises estatísticas foram feitas por meio de regressão linear e análise de Pearson, no software Stata versão 10. **Resultados.** As análises mostraram que a heterogeneidade entre as regiões estudadas constitui um limitador para comparações. No entanto, os resultados validaram a hipótese fomentada. Os mapas coropléticos apresentaram agrupamentos de focos de calor próximos ou sobrepondo agrupamentos de internações por asma ou enfisema. As regressões apresentaram resultados com significância estatística. Em todos os modelos analisados, a variável foco de calor apresentou correlação positiva e as variáveis IDH-M e área urbanizada apresentaram correlações negativas. Concluiu-se que as taxas de internação estão relacionadas com queimadas, aspectos sócio-econômicos e climáticos, porém, os modelos explicam parcialmente a variabilidade. É possível que outras variáveis não analisadas nesta tese também exerçam influência nos fenômenos estudados, assim como os efeitos podem ser sub-clínicos, não chegando a impactar significativamente as internações em curto prazo.

ABSTRACT:

Introduction: The use of computing technology in public health researches is increasing because of the facility it provides. Particularly, GIS helps the fusion between tabular data and maps which contain geographical aspects of a locality, through overlapping layers. This fusion may offer a new interpretation of data associated to spatial dimension. In this aspect, GIS programs are good tools for studies of health geography. However, it is observed that there is a lack of studies about the harmful effects caused by burned straw in sugarcane plantation. In this context, GIS and spatial analysis can help to obtain better knowledge about the man and his environment, in the scope of public health. **Aims:** This study evaluated correlations between respiratory health problems and burned sugarcane, using spatial analysis and statistical models. **Methodology:** The development of this work consisted of collecting and organizing data of hospital admissions due to respiratory problems and the burning of sugarcane, generating a data repository for two Brazilian regions which produce sugarcane. The region called Central-South with the states Goiás, Mato Grosso, Mato Grosso do Sul, Paraná and São Paulo, and the Northeast region with the states Alagoas, Bahia, Sergipe, Pernambuco and Paraíba, in a period between 2000 and 2008. The spatial analysis was made by the Hot Spot tool in Arcview 9.3 and the statistical analysis was made by linear regression and Pearson analysis in Stata version 10. **Results:** The analysis showed that the heterogeneity among the regions studied limits comparisons. However, the results validated the hypothesis evaluated. The choropleth maps showed hotspot clusters close or overlapping clusters of asthma or emphysema hospital admissions. The regressions showed statistically significant results. The hotspot variable was positively correlated and the IDH-M and urbanized area showed negative correlations in all models analyzed. It was concluded that the hospital admissions are related to burned sugarcane, socio-economic and climate aspects, however, the models partially explain the variability. It is possible that other variables not analyzed in this thesis also adversely influence the phenomena studied, as so as the effects can be sub-clinical, not getting to significantly impact the hospital admissions in a short term.

ÍNDICE:

1	INTRODUÇÃO	16
1.1	Aspectos Gerais	16
1.2	Justificativa do Estudo	25
1.3	Problema e Hipótese da Pesquisa	28
1.4	Objetivos	29
1.5	Estrutura do Trabalho	30
2	EMBASAMENTO TEÓRICO	31
2.1	A Geografia no Contexto da Saúde Ambiental	31
2.2	Abordagens da Estatística Espacial	36
2.3	O Uso da Análise Espacial de Dados na Saúde	48
2.4	SIG e Estudos da Qualidade do Ar e a Saúde	50
2.5	Estudos Envolvendo Queima de Biomassa e Saúde	57
3	CASUÍSTICA E MÉTODOS	66
3.1	Aspectos Gerais	66
3.2	Método de Pesquisa	66
3.3	Delineamento da Pesquisa	70
3.4	Definição dos Sujeitos da Pesquisa	79
3.5	Descrição das Bases de Dados e Adequações	80

3.6 Ferramentas Utilizadas	86
3.7 Limitações da Pesquisa	88
3.8 Aspectos Éticos	90
4 TRATAMENTO DOS DADOS, RESULTADOS E DISCUSSÃO	91
4.1 Aspectos Gerais	91
4.2 Análises Descritivas.....	91
4.2.1 Área do Estudo.....	91
4.2.2 Focos de Calor	93
4.2.3 Dados Climáticos.....	99
4.2.4 Relevo	101
4.2.5 IDH-M.....	102
4.2.6 Áreas de Cana e Áreas Urbanizadas	103
4.2.7 População.....	105
4.2.8 Internações.....	106
4.3 Análises Espaciais.....	115
4.3.1 Análise da Região Centro-Sul	115
4.3.2 Análise da Região Nordeste	125
4.3.3 Análise longitudinal para o estado de São Paulo	126
4.4 Análises Estatísticas	129
4.4.1 Distribuição Normal	129
4.4.2 Análises por Correlação	131

4.4.3 Análises por Regressão Linear.....	134
5 CONCLUSÕES	142
6 REFERÊNCIAS	148
7 APÊNDICES.....	158
Apêndice A – Dicionário de dados de estrutura dos arquivos “DR” (Reduzidas) – AIH – DATASUS.....	158
Apêndice B - Estruturas das tabelas criadas no repositório de dados integrado ao SIG.....	167

LISTA DE FIGURAS:

1- Serviços prestados ao homem pelos ecossistemas	16
2- Relação entre análise espacial e teorias disciplinares.....	44
3- Os quatro estágios do método científico e seus relacionamentos	45
4- Representação de objetos com pontos, linhas e polígonos.....	47
5- Modelo geográfico risco-exposição-efeito	54
6- Resumo dos delineamentos encontrados nas pesquisas sobre queima de biomassa e efeitos adversos à saúde humana.....	64
7- Tipos de pesquisa e grupos de delineamentos	70
8- Diagrama esquemático do delineamento da pesquisa	71
9- Diagrama esquemático das fontes de dados utilizadas no estudo e seus relacionamentos no repositório de dados	84
10- Histograma das variáveis incasma04 e qtdbf10, antes e depois da transformação com função Ln().....	130

LISTA DE GRÁFICOS:

1- Estimativa de produção e destinação de cana-de-açúcar no Brasil Safra 2007/2008	19
2- Evolução dos focos de calor nas regiões do estudo por média móvel e curva polinomial, no período de 2000 a 2008.....	94
3- Comparação entre número de comunicações de queimadas e focos de calor para o estado de São Paulo, em setembro de 2008.....	97
4- Distribuição de médias do IDHM por estados das regiões de estudo em 2001..	103
5- Curva de evolução da somatória de internações por asma, bronquite e enfisema na rede pública, Regiões Centro-Sul e Nordeste, período de 2000 a 2008	109
6- Evolução do número de internações com diagnóstico de asma, por faixa etária, na região Centro-Sul, no período de 2000 a 2008.....	111
7- Evolução do número de internações com diagnóstico de asma, por faixa etária, na região Nordeste, no período de 2000 a 2008.	111
8- Evolução do número de internações com diagnóstico de bronquite, por faixa etária, na região Centro-Sul, no período de 2000 a 2008	112
9- Evolução do número de internações com diagnóstico de bronquite, por faixa etária, na região Nordeste, no período de 2000 a 2008.....	112
10- Evolução do número de internações com diagnóstico de enfisema, por faixa etária, na região Centro-Sul, no período de 2000 a 2008	113
11- Evolução do número de internações com diagnóstico de enfisema, por faixa etária, na região Nordeste, no período de 2000 a 2008.....	113
12- Matriz de correlação entre variáveis dependentes e independentes para a região CS	133
13- Matriz de correlação entre variáveis dependentes e independentes para a região NE	133

LISTA DE MAPAS:

1- Concentrações de usinas e plantações de cana-de-açúcar, segundo dados oficiais do IBGE (Instituto Brasileiro de Geografia e Estatística), NIPE - UNICAMP (Núcleo interdisciplinar de Planejamento Energético da Universidade Estadual de Campinas – SP) e CTC (Centro de Tecnologia Canavieira)	18
2- Área de estudo e localização das áreas de plantio de cana-de-açúcar	93
3- Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 1º trimestre de 2008.....	117
4- Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 2º trimestre de 2008.....	118
5- Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 3º trimestre de 2008.....	119
6- Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 4º trimestre de 2008.....	120
7- Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 1º trimestre de 2008.....	121
8- Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 2º trimestre de 2008.....	122
9- Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 3º trimestre de 2008.....	123
10- Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 4º trimestre de 2008.....	124
11-Interseção dos <i>hotspots</i> de focos de calor e internações (asma e enfisema) para o estado de São Paulo, no período de 2000 a 2008.....	127

LISTA DE QUADROS:

1- Classificação de tipologia de sistemas	37
2- Principais poluentes provenientes da queima de biomassa	53
3- Variáveis utilizadas nos desenhos metodológicos em estudos de poluição do ar, na modalidade <i>outdoor</i>	63
4- Número de variáveis por dimensões utilizadas no repositório de dados	85
5- Síntese das ferramentas utilizadas na pesquisa.....	88
6- Altitudes médias, mínimas e máximas dos estados da área de estudo	102
7- Sumarização dos dados sobre internações para asma, bronquite e enfisema, baseado na tabulação diária por faixa etária, na área de estudo (CS + NE), no período de 2000 a 2008	114
8- Teste Shapiro-Wilk aplicado para as variáveis: focos de calor (qtdbf10 e qtdbf20) e incidências de internações (incasma04, incasma65 e incenfi65), nas regiões CS e NE.....	130
9- Coeficientes de Correlação de Pearson calculados entre as variáveis focos de calor (qtdbf10 – <i>buffer</i> 10km e qtdbf20 – <i>buffer</i> 20km) e incidências de internações (incasma04, incasma65 e incenfi65), na região CS e NE	132
10- Variáveis independentes utilizadas na análise de regressão	134
11- Valores de R ² para as diversas combinações dos modelos testados.	135
12- Modelo de regressão para a amostra CS/Município/Trimestre/ Incidência de internações por asma infantil.....	136
13- Modelo de regressão para a amostra NE/Município/Trimestre/ Incidência de internações por asma infantil.....	136
14- Modelo de regressão para a amostra CS/Município/Trimestre/ Incidência de internações por asma em idosos.....	138
15- Modelo de regressão para a amostra NE/Município/Trimestre/ Incidência de internações por asma em idosos.....	138
16- Modelo de regressão para a amostra CS/Município/Mês/ Incidência de internações por enfisema em idosos	139
17- Modelo de regressão para a amostra NE/Município/Mês/ Incidência de internações por enfisema em idosos	139

LISTA DE TABELAS:

1- Dimensão e população da área de estudo	92
2- Variações médias de temperatura e precipitações mínimas e máximas por estado, nas regiões incluídas no estudo.....	100
3- Área total, urbanizada e de plantio de cana por estado.....	105
4- Contagem de internações e taxa de incidência para asma, bronquite e enfisema, categorizados por idade, registrados por AIH no SIH/SUS por UF, no período de 2000 a 2008	108

1 INTRODUÇÃO

1.1 Aspectos Gerais

O termo “sistema” talvez seja aquele utilizado de forma mais heterogênea, na tentativa de representar uma combinação de meios coordenados, projetados ou criados para produzir resultados nos diversos domínios do conhecimento humano.

De forma muito similar e próxima, a literatura traz o termo “equilíbrio”, explicitando uma forte ligação, como se fosse necessário indicar o importante estado de equivalência de forças que ocorrem entre os elementos da composição sistêmica.

Em 1997, o economista Robert Constanza calculou em 55 trilhões de reais/ano o valor dos serviços prestados ao mundo pelo meio ambiente. Tal valor foi mensurado com base nas atividades vitais exercidas pelos ecossistemas, que vão desde a regulação de composições atmosféricas até a produção de alimentos, matéria-prima e recursos energéticos conforme apresenta a Figura 1 (LE MONDE DIPLOMATIQUE BRASIL, 2009).

Figura 1 – Serviços prestados ao homem pelos ecossistemas. Fonte: LE MONDE DIPLOMATIQUE, 2009).

Desta forma, os produtos gerados pelos diversos ecossistemas são geradores de insumos provedores de qualidade de vida para o ser humano.

Os diversos sistemas, frágeis ou robustos, simples ou complexos, seguem na tentativa de manter equilíbrio e coexistir com outros sistemas, interagindo ou não, formando uma complexa estrutura, que o ser humano tenta compreender desde que surgiram os primeiros traços de inteligência no processo evolutivo do planeta Terra.

No presente momento, o mundo discute o aquecimento global e as ações antrópicas causadoras, em um cenário de irreversibilidade do problema legado às gerações futuras. Entre as variáveis que interagem neste sistema, cita-se a combustão generalizada de combustíveis fósseis, em contraposição a alternativas denominadas “verdes”, como é o caso do etanol, proveniente da cana-de-açúcar.

O Brasil produziu, na safra 2007/2008, 20,88 bilhões de litros de álcool e 29,65 milhões de toneladas de açúcar (CONAB, 2007), provenientes de 623 milhões de toneladas colhidas de 7,9 milhões de hectares de área plantada de cana-de-açúcar (IEA, 2009). No que representa a produção de álcool anidro (8,19 bilhões de litros) (CONAB, 2007), este montante equivale a 3% da produção de petróleo da Arábia Saudita ou ainda um quarto da produção do Iraque antes da guerra (GOLDEMBERG, 2005).

O Mapa 1 apresenta a distribuição das áreas de produção de cana-de-açúcar ao longo do território brasileiro. Observam-se as duas maiores áreas de concentração de plantio, o Nordeste e o Centro-Sul. No entanto, pontos indicativos da atual expansão podem ser observados, a partir do centróide Sudeste, em São Paulo, em direção à região Centro-Oeste do Brasil, através dos estados de Mato-Grosso, Minas Gerais e Goiás.

Mapa 1 - Concentrações de usinas e plantações de cana-de-açúcar, segundo dados oficiais do IBGE (Instituto Brasileiro de Geografia e Estatística), NIPE - UNICAMP (Núcleo interdisciplinar de Planejamento Energético da Universidade Estadual de Campinas – SP) e CTC (Centro de Tecnologia Canavieira). Fonte: ÚNICA, 2009.

O Gráfico 1 apresenta a estimativa de produção e destinação de cana-de-açúcar no Brasil para a safra 2007/2008. Nesta visão, é possível observar o distanciamento existente entre as produções das regiões Norte/Nordeste comparada à região denominada Centro-Sul. No caso, a região Centro-Sul detém 87,03% da produção nacional e vem se expandindo em novas áreas nos estados de Goiás, Mato-Grosso e Mato Grosso do Sul, na direção do conhecido “arco do desmatamento” nos limites da Floresta Amazônica.

Gráfico 1 - Estimativa de produção e destinação de cana-de-açúcar no Brasil Safra 2007/2008. Fonte: CONAB, 2007.

Estes resultados representam um aumento de 10,62% em relação à safra anterior. Considerando as ações de governos estaduais como a Lei Estadual nº 11.879/2005 de incentivo às micro destilarias para o estado de São Paulo (SÃO PAULO, 2005), o Decreto 10.936/2008 de incentivos fiscais para produção de etanol na Bahia (BAHIA, 2008) e o Decreto Federal 6.539/2008, que reduz o imposto de renda pessoa jurídica no caso de ampliação ou modernização parcial de empreendimentos prioritários para o desenvolvimento regional (BRASIL, 2008), este setor deve prosseguir em expansão nas próximas safras.

Este cenário de crescimento está motivado pela produção mundial de automóveis *flexfuel* (combustível flexível) iniciada em 2003, além de legislações em outros países que permitem adicionar álcool à gasolina, ou ainda aumentar a proporção de álcool à mistura com gasolina (GOLDEMBERG, 2005).

Até 2010, estima-se que 70% dos veículos automotores produzidos no Brasil serão *flexfuel* e utilizarão álcool como opção de abastecimento. Esta previsão exigirá um incremento de 7 bilhões de litros/ano somente para a demanda brasileira.

O movimento denominado “retomada do programa de álcool” apresenta alguns diferenciais em relação ao Proálcool na década de 1970. É fomentado pela iniciativa privada e ocorre em um momento em que não há crise de abastecimento de gasolina. O usuário pode ter um carro híbrido e optar por gasolina ou álcool com preço mais acessível (INÁCIO, 2005). Um estudo recém publicado pelo BNDES (Banco Nacional de Desenvolvimento Econômico e Social) estima ser necessária a construção de 100 novas usinas até 2010 para suportar a demanda interna brasileira (LAGE E ZAFALON, 2007).

Aumentar a produção de álcool pode ser bom para a economia, porém contribui para o aumento das concentrações de CO₂ na atmosfera em função das queimadas que são adotadas como parte do processo de colheita da cana-de-açúcar. O aumento de concentrações impacta o clima afetando a dinâmica da circulação atmosférica e o ciclo da água, minimizando a precipitação de chuvas na região (ANDREAE et al., 2004).

Outra abordagem deste assunto diz respeito à questão energética. O bagaço de cana já é utilizado em sistemas de co-geração para suprir as demandas eletromecânicas e térmicas em usinas e destilarias (LEME, 2005). Considerando a estimativa de que a energia potencial da palha seja equivalente à do bagaço de cana, há um montante de energia desperdiçada com a queima da palha devido aos procedimentos de colheita, ainda largamente utilizados pelo setor canavieiro.

Neste ponto, há que se considerar o impacto das emissões provenientes da geração termoelétrica. Mesmo que as queimadas nas áreas de plantio sejam finalizadas, a emissão de poluentes como subproduto da geração de energia, ainda que produzida em processos controlados,

demanda estudos de impacto sobre os efeitos deletérios à saúde e ao meio ambiente, no entorno das usinas.

Considerando o ciclo de vida dos produtos álcool e açúcar, os resíduos do processo, como a palha, o bagaço e o restilo, ainda carecem de estudos sobre o impacto ambiental e alternativas de destinação menos agressivas ao meio ambiente.

Do ponto de vista da saúde pública, a maior parte da literatura que trata de emissões de poluentes atmosféricos produzidos pela queima de biomassa aponta que estas causam efeitos deletérios à saúde humana.

Os estudos dos efeitos que o material particulado e gases produzidos pelas queimadas podem causar à saúde humana são escassos. Em 2006, uma nota do escritório regional da OMS em Manila, nas Filipinas, recomendou a redução de material particulado nas cidades, estimando que este é responsável pela morte de 2 milhões de pessoas por ano (O ESTADO DE SÃO PAULO, 2006).

DOCKERY e POPE (1994) afirmam que, desde a década de 1970, estudos epidemiológicos seguem no sentido de prover maior quantificação de efeitos à saúde associados a níveis de poluentes. No entanto, estes resultados dependiam do avanço da bioestatística que, mais recentemente, proporcionou maior rigor analítico desses estudos.

Por outro lado, os fenômenos humanos não ocorrem em função de uma ordem semelhante ao que se observa no mundo físico, acarretando em imprevisibilidade e dificuldade de quantificação, o que torna difícil a comunicação de resultados das investigações (GIL, 1994).

Diversos estudos têm apontado associações positivas entre a exposição de contaminantes do ar e efeitos adversos à saúde. Poluentes provenientes da queima de combustíveis fósseis, tais como SO_2 , NO_x , CO , O_3 , VOC (compostos orgânicos voláteis) e HPA (hidrocarbonetos policíclicos aromáticos) estão presentes nas pesquisas, porém, seus efeitos são pouco documentados (OPAS, 2005).

No entanto, efeitos sobre a saúde infantil associados à exposição por longos períodos de fumaça proveniente de queima de biomassa *indoor*, foram bem documentados em países em desenvolvimento. Entre os efeitos estão relatadas infecções respiratórias agudas, doença pulmonar obstrutiva crônica (DPOC), pneumoconiose, catarata e cegueira, tuberculose pulmonar e efeitos adversos na gestação (ARBEX et al, 2004).

A OPAS (Organização Pan-Americana da Saúde) analisou 85 estudos realizados entre 1994 e 2004 na América Latina, sobre poluentes atmosféricos e saúde. Os resultados sugerem associações positivas relativas ao impacto sobre a morbi-mortalidade de crianças e idosos. No entanto, a maioria dos estudos está associada à poluição urbana, com características diferenciadas à produzida pela queima de biomassa (OPAS, 2005).

Ainda sobre a saúde das pessoas, os efeitos deletérios à saúde humana das populações que habitam áreas de entorno onde ocorre queima de biomassa são pouco conhecidos ou pesquisados. Os trabalhos de ARBEX et al (2000), CANÇADO (2006), RIBEIRO e ASSUNÇÃO (2002), e ZANCUL (1998) são pioneiros nesta temática.

A correlação entre problemas ambientais e saúde pública neste âmbito torna-se objeto de pesquisa com relevada importância nos tempos atuais, ainda que provoquem alertas contrários à ordem econômica do agropênisio. Soma-se a isso, a variedade de aspectos que envolvem o problema, dificultando a identificação de causas isoladas para um determinado efeito.

Considerando o reduzido número de estudos que abordam os agravos à saúde humana em função dos produtos da queima de biomassa em áreas de cultura canavieira, pouco se sabe sobre as condições de risco oferecidas às populações que vivem no entorno de tais áreas, ou ainda, que se submetem a condições de trabalho sub-humanas nesta lavoura.

O momento atual sugere atenção sob o ponto de vista da saúde pública, pois, a produção de etanol acena para o mundo como alternativa

renovável e ecologicamente correta aos combustíveis fósseis. No entanto, seu processo produtivo ainda persiste com questões não respondidas sobre a qualidade de vida dos entes envolvidos.

Portanto, existem lacunas a serem preenchidas sobre as condições de risco à saúde das populações que sistematicamente são expostas aos produtos da queima da palha de cana-de-açúcar.

A considerar a extensão do assunto, esta pesquisa focou o uso de técnicas de análise espacial de dados e sua eficiência na obtenção de resultados para o problema dissertado. Entende-se por análise espacial, os métodos utilizados para “*mensurar propriedades e relacionamentos, levando em conta a localização espacial do fenômeno em estudo de forma explícita*” (DIAS et al, 2002).

A computação disponibiliza várias ferramentas que auxiliam o manuseio de bancos de dados, proporcionando a identificação de padrões espaciais, de difícil visualização sem o devido tratamento. No entanto, vários caminhos podem ser seguidos na busca da comprovação estatística de um fenômeno.

Fatores como a democratização da informação, baixo custo de software e hardware e o desenvolvimento de interfaces mais amigáveis, estimulam a produção de estudos que incorporam os SIG como parte integrante do instrumental de trabalho.

Porém, estes fatores também podem facilitar a geração de resultados espúrios, pois os programas não são capazes de apontar erros no método aplicado, somente validam o formato dos parâmetros de entrada e geram uma saída que, obviamente, deverá ser objeto de análise do pesquisador. Isto posto, o usuário destes recursos deve ter bom conhecimento teórico sobre recursos e métodos, bem como sobre os eventuais problemas que podem ocorrer neste processo.

Sendo assim, esta pesquisa estruturou um ambiente com o propósito experimental de avaliar resultados gerados por métodos e ferramentas distintas, na perspectiva de aprofundar e validar a eficiência deste

instrumental tecnológico, em estudos que envolvem questões de saúde e ambiente.

Para alcançar os objetivos desta tese, o referencial teórico teve como pano de fundo a geografia da saúde e a análise espacial em estudos de saúde ambiental, na expectativa de reunir elementos norteadores para formalizar um modelo adequado de tratamento e análise das variáveis envolvidas.

A título de minorar a influência de determinadas variáveis de confusão, este estudo foi conduzido no sentido de produzir a análise de dados nas duas regiões brasileiras com maior produtividade sucroalcooleira: o estado de São Paulo, na região sudeste e os estados de região nordeste.

Sabe-se que os produtores de ambas as regiões utilizam a queima da palha da cana-de-açúcar na pré-colheita e, o período de safra ocorre em períodos distintos do ano. Este fato permite verificar uma possível ocorrência dos mesmos fenômenos em regiões com elementos geográficos e climáticos distintos.

A análise deste conjunto de dados seria impraticável sem o auxílio de ferramentas computacionais, pois, considerando o volume de dados, cálculos e variáveis envolvidas, o manuseio do conjunto demandaria muito tempo e recursos humanos e financeiros, ampliando a margem de erro dos resultados, desconsiderando fatores como aquisição de dados e validação, entre outros.

Programas como o Spring, Terraview, Geoda, Gslib, SatScan, Mapinfo e ArqGis, entre outros, são largamente utilizadas em estudos na área de saúde por oferecerem recursos estatísticos ou de geoprocessamento, necessários para o desenho de tais estudos. No entanto, foram desenvolvidos com interesses distintos, por pessoas e recursos diferenciados. Estes fatores proporcionaram uma grande variação de produtos, que demanda tempo de análise por parte do eventual usuário, para decidir que ferramenta utilizar.

Observa-se, ainda, que o uso das atuais ferramentas disponíveis no mercado demanda muita dedicação por parte do pesquisador, no sentido de seu aprendizado, bem como certa familiaridade com informática. Estes fatores ampliam as chances de insucesso no uso destas ferramentas. Este cenário tem melhorado na medida em que os novos pesquisadores são “nativos tecnológicos”, ou seja, nasceram na era da informação e já são iniciados no uso de tecnologias desde criança.

Por outro lado, avanços da computação, como a Web Semântica e a Inteligência Artificial, vem contribuindo para melhorar a usabilidade e a precisão das ferramentas computacionais. Explora-se cada vez mais a modelagem em que os sistemas observados no mundo real são transportados para o mundo digital, viabilizando simulações e predições sobre o objeto de estudo.

O crescimento do uso de dispositivos de posicionamento global (GPS – *Global Positioning System*) também aponta para uma maior disseminação das geotecnologias, evidenciando a preocupação do ser humano em conhecer o lugar onde vive e o que há em seu entorno.

Por fim, geotecnologias podem ser muito úteis no diagnóstico de problemas no âmbito da saúde ambiental. A simulação computacional de ocorrências ambientais pode contribuir para uma melhor compreensão de eventos no meio ambiente e seus impactos na saúde das pessoas.

1.2 Justificativa do Estudo

As pesquisas que relacionam poluentes atmosféricos, em ambientes externos, aos efeitos adversos à saúde humana, além dos benefícios diretos à sociedade, contribuem para refinamento do instrumental da estatística e da análise espacial, melhorando a formulação e validação de hipóteses em pesquisas que envolvem saúde e ambiente.

Tal argumento se vale para o fomento de indicadores que subsidiam a tomada de decisão no âmbito das políticas públicas, embasando de forma mais adequada, a legislação nesta agenda e as ações de vigilância ambiental.

Neste cenário, observa-se uma grande abertura para o uso de tecnologia computacional, como agente facilitador e agilizador de tarefas que tomam precioso tempo no processo de pesquisa. A redução do custo dos itens de informática, associada à disponibilidade de informação cada vez mais ampla, vem aproximando o pesquisador aos recursos computacionais, fato este que promove a interdisciplinaridade e proporciona resultados interessantes para a ciência.

Esta aproximação não ocorre somente entre pesquisadores e equipamentos, mas entre pesquisadores de áreas distintas, que passam a cooperar em pesquisas de caráter interdisciplinar. Observa-se neste ponto, a possibilidade de novas visões, construídas a partir da vivência compartilhada de profissionais de áreas distintas. Este enfoque favorece uma caminhada mais rápida na busca do conhecimento.

Por outro lado, entende-se que a questão ambiental e da saúde pública, no contexto da cultura canavieira, ainda carece de estudos mais conclusivos sobre os procedimentos aplicados e os possíveis efeitos deletérios causados à saúde das coletividades que vivem no entorno das áreas de plantio. Considerando que a atual conjuntura aponta para o crescimento da produção de álcool combustível, os efeitos também crescerão em proporção desconhecida.

O estado de São Paulo regulamentou a eliminação da queima da palha de cana-de-açúcar através da Lei nº 11.241 de 2002. Sua regulamentação foi feita pelo decreto nº 47.700 de 2003 que apresenta um cronograma de eliminação gradativa no prazo de 30 anos (finalizando em 2031).

O governo de São Paulo vem promovendo iniciativas que mostram a preocupação com a questão das queimadas. O projeto de lei nº 14, de 8 de

fevereiro de 2007, propôs a alteração do cronograma de eliminação gradativa reduzindo o prazo para 2013, mas não foi aprovado. No entanto, o governo de São Paulo e a entidade representante dos produtores assinaram em junho de 2007 um protocolo de cooperação reduzindo o prazo de 2021 para 2014 (áreas mecanizadas) e de 2031 para 2017 (áreas não mecanizadas).

Outra iniciativa foi a implantação de um sistema de cadastramento eletrônico de comunicações de queima, no âmbito do sistema SIGAM da Secretaria de Estado de Meio Ambiente (SMA), com o propósito controlar requerimentos e autorizações de queima, em consonância com o decreto estadual nº 47.700 de 2003.

Recentemente, em agosto de 2008, o promotor da comarca de Jaú, condenou o estado de São Paulo a pagar uma multa 2,5 milhões de reais para quatro cidades da sua comarca (Jaú, Mineiros do Tietê, Bocaina e Itapuí), através de uma ação movida em 2007 alegando que as queimadas contribuem para a poluição do meio ambiente podem ocasionar problemas de saúde pública em função da exposição à fumaça emanada dos canaviais (SIQUEIRA, 2009).

O evento supracitado expõe um fato novo neste cenário, no qual o estado torna-se réu ao cumprir atos legitimados por decreto. A situação mostra o evidente descontentamento que há na mediação de interesses entre os envolvidos: usineiros e população. Tal mediação, muitas vezes favorável para usineiros, esteve pautada na argumentação de que os estudos não são suficientes para provar os efeitos deletérios causados pelas queimadas à saúde humana e ao meio ambiente.

Embora exista a preocupação das autoridades governantes com a questão das queimadas no estado de São Paulo, observa-se a expansão da cultura canavieira para outros estados como Goiás, Mato Grosso e Minas Gerais e Bahia, que carecem de iniciativas similares. Esta situação indica que o problema pode migrar para outras regiões ao invés de ser mitigado.

Por fim, está no âmbito da saúde pública, a identificação dos fatores de risco, a pesquisa, a discussão e o fomento de políticas públicas, propondo estratégias mais apropriadas para o controle, no sentido de minorar o risco em populações envolvidas.

1.3 Problema e Hipótese da Pesquisa

Do ponto de vista metodológico, o problema e a hipótese têm a função de orientar e determinar a natureza dos dados que serão coletados. Os dados buscam explicar os fenômenos associados com o intuito de validar ou refutar a hipótese do estudo (ALMEIDA FILHO e ROUQUAYROL, 2002).

Considerando os aspectos contextuais abordados até este ponto, é possível evidenciar o conjunto de elementos que compõe o problema da pesquisa. As questões colocadas representam um escopo multidisciplinar que relaciona intrinsecamente a saúde e o meio-ambiente, demandando maior aprofundamento sobre a compreensão dos fenômenos envolvidos.

No âmbito deste estudo, o problema da pesquisa foi identificado por meio de fatores conhecidos pelo senso comum, que estão associados ao processo produtivo da cana-de-açúcar. Por exemplo, a percepção da população que vive no entorno de áreas de queimadas, sobre o aumento de admissões hospitalares por problemas respiratórios nas regiões onde ocorre esta prática.

O desenho traçado para o encaminhamento deste trabalho partiu do estudo exploratório, onde se organizou um repositório de dados sobre queimadas e internações hospitalares por problemas respiratórios no período de 2000 a 2004, no estado de São Paulo (LOPES, RIBEIRO, 2006).

A análise dos dados produzida no estudo supracitado apontou coincidências na distribuição espacial dos dados, quando estes foram apresentados em mapas coropléticos. A incidência de internações

hospitalares apresentou-se superior à mediana do estado de São Paulo nos municípios que possuem cultura canavieira e queimadas.

Sendo assim, os resultados desse trabalho reforçaram a hipótese fomentada a partir do senso comum, sobre a possível correlação existente entre as variáveis estudadas (queimadas e problemas respiratórios), que ficou formulada da seguinte forma:

“A saúde respiratória da população que vive próxima de áreas de plantio de cana-de-açúcar é negativamente impactada pelos produtos das queimadas que lá ocorrem”.

Esta afirmativa busca sintetizar as conjecturas evidenciadas. Sua formulação marca uma etapa importante no método científico e permite a continuidade do processo que segue com a definição dos objetivos gerais e específicos.

1.4 Objetivos

Objetivo Geral

Este estudo descritivo objetiva avaliar a correlação entre doenças respiratórias e queimadas de cana-de-açúcar, utilizando técnicas de geoprocessamento e análise estatística.

Objetivos Específicos

-Estudar e comparar os dados sobre queimadas e internações hospitalares em duas áreas do país onde há cultivo extensivo de cana em condições climáticas e períodos de safra diferentes.

-Contribuir para uma melhor usabilidade do instrumental estatístico e de geoprocessamento nesta área de pesquisa.

1.5 Estrutura do Trabalho

Objetivando uma melhor exposição dos elementos analisados, esta tese foi estruturada da seguinte forma:

Este capítulo apresenta os elementos do escopo do trabalho, passando pelo contexto, problema da pesquisa, hipótese fomentada e objetivos do trabalho.

Na sequência, o capítulo 2 reuniu os fundamentos teóricos inerentes à pesquisa, elencando temas como a geografia no contexto da saúde ambiental, aspectos da estatística espacial, e pontos relacionados com a problemática da poluição atmosférica e queima da biomassa. Por fim, finalizou-se com um levantamento sobre o estado da arte em estudos com poluição envolvendo a queima de biomassa na modalidade *outdoor*.

A tese segue com o capítulo “Casuística e Métodos”, que objetiva esclarecer, a linha metodológica da pesquisa, abordando o delineamento do estudo, bem como variáveis, sujeitos e técnicas utilizadas para concretizar os objetivos propostos.

O capítulo “Tratamento dos Dados, Resultados e Discussão” descreve o seqüenciamento do estudo e os produtos finalizados, juntamente com as respectivas análises, que formam o conjunto de subsídios que sustentam as conclusões sintetizadas no capítulo “Conclusões”.

A título de material complementar, foram inseridos ao final do texto, apêndices desenvolvidos ou reunidos durante a pesquisa para um melhor embasamento dos resultados obtidos.

O conjunto de elementos aqui descritos foi reunido na expectativa de obter uma criteriosa análise sobre a associação causal que pode haver entre queimadas em canaviais e problemas respiratórios, no entanto existem muitas limitações e dificuldades que foram descritas ao longo desta tese.

2 EMBASAMENTO TEÓRICO

2.1. A Geografia no Contexto da Saúde Ambiental

No que tange as questões de saúde e ambiente, a variável geográfica dos fenômenos estudados assume papel relevante nos vários desenhos de estudos epidemiológicos, com a finalidade de identificar os fatores de risco potencial da relação causal em foco. OPAS (2007) coloca que “*ambiente e saúde são interdependentes e inseparáveis*” enfatizando que o lugar onde o ser humano habita e desenvolve suas atividades, deve ser favorável à saúde.

Os estudos de saúde que agregam recursos geográficos estão no contexto de uma área denominada Geografia Médica ou Geografia da Saúde. Segundo LACAZ (1972), esta área engloba o conjunto de disciplinas que estudam as doenças à luz dos conhecimentos geográficos.

Desde Hipócrates já se ponderava a importância do ambiente associada ao processo saúde-doença. No entanto, o espaço como variável começou a ser trabalhado de forma mais sistemática a partir da obra de Ludwig Finke, escrita em 1792 (BOUSQUAT e COHN, 2004).

No século XIX, surgiram os trabalhos sistemáticos que aproximaram a geografia da epidemiologia, voltados à descrição minuciosa da distribuição regional das doenças, contando com amplo emprego de recursos cartográficos (FERREIRA, 1991). De lá para os tempos atuais, o espaço tem mantido relevância em estudos no âmbito da saúde pública.

Na concepção de SANTOS (2002), a geografia poderia ser construída considerando o espaço como um conjunto de fixo e fluxos, onde os elementos fixos permitem ações que modificam o lugar e os fluxos são resultados das ações que atravessam ou se instalam nos fixos. O mesmo autor coloca, em um contexto mais atual, que:

“O espaço é formado por um conjunto indissociável, solidário e contraditório, de sistemas de objetos e sistemas de ações, não considerados isoladamente, mas como quadro único no qual a história se dá.” (SANTOS, 2002, p. 63).

Os objetos, antes somente naturais, agora também são de natureza artificial, ampliando a complexidade deste sistema, que segue interagindo com os sistemas de ações, de forma indissociável, transformando o espaço.

Para a geografia clássica, “o espaço foi entendido como o substrato de fenômenos naturais, tais como o clima, a topografia ou a vegetação”. Já no âmbito da epidemiologia, o espaço foi compreendido como resultado de uma interação entre organismo e natureza bruta, independente da ação ou percepção humana (CZERESNIA e RIBEIRO, 2000).

As pressões que acompanharam o desenvolvimento da lógica de conhecimento do objeto da geografia, por força das distinções dicotômicas, culminaram na cisão das ciências naturais e sociais. A epidemiologia acompanhou o desenvolvimento teórico-conceitual da geografia, de forma especial na geografia médica, buscando incluir na compreensão do processo da doença, dimensões sociais, culturais e simbólicas (CZERESNIA e RIBEIRO, 2000).

Sob outra ótica, FERREIRA (1991) coloca o espaço como um elemento formado pelo produto de decisões que orientam sua organização, em uma dada formação econômico-social. Assim sendo, este objeto de estudo fomenta questões interdisciplinares, promovendo o intercâmbio e a evolução do conhecimento entre áreas distintas, fato muito peculiar em questões de saúde pública.

No entanto, a virtualização do espaço em mapas, sejam eles impressos ou em formato digital, não é tarefa simples, pois, mapas envolvem uma interface multimídia e multimodal de informação geoespacial, integradas em um *framework* de representação. Desta forma, é essencial

conhecer as vantagens deste instrumental, que abrange o entendimento de regras e uso de ferramentas de geovisualização exploratória, conceitos de visualização da informação, aprimoramentos dos mapas em atividades colaborativas, e outros, a fim de proporcionar facilidades para a ciência e os processos de tomada de decisão (MacEACHREN, 2004).

Nos últimos anos, a elevada sofisticação das ferramentas computacionais, aliada ao crescente poder de processamento dos computadores, impulsionou o desenvolvimento de poderosas e versáteis ferramentas estatísticas, as quais implementam modelos e métodos cuja aplicabilidade manual é trabalhosa. Assim sendo, a tecnologia da informação tornou-se uma grande provedora de recursos que visam facilitar o trabalho do pesquisador.

Do conjunto de recursos computacionais disponíveis, as ferramentas que permitem a apresentação de dados em mapas, são parte integrante do instrumental denominado Geoprocessamento, que engloba vários assuntos como: cartografia digital, sensoriamento remoto, estatística espacial e banco de dados geográficos.

BARCELLOS e BASTOS (1996) afirmam que as tecnologias disponíveis de Geoprocessamento beneficiam trabalhos que buscam fixar um modelo de agregação de dados por unidade espacial.

Mais amplo que Geoprocessamento, o termo SIG (Sistema de Informação geográfica, ou GIS – *Geographic Information System*), é adotado para classificar sistemas que efetuam um tratamento computacional aos dados geográficos, agregando atributos descritivos a geometrias dos diversos tipos de dados geográficos (CÂMARA et al, 2001).

Em particular os SIGs, aliados à geografia da saúde, facilitam o trato de grandes volumes de dados, minimizando o trabalho do pesquisador que, desta forma, ganha flexibilidade e tempo para desenvolver tarefas mais importantes da pesquisa, como a avaliação e análise dos dados coletados. É um caso de junção de dados e técnicas com origem em áreas distintas, com o intuito de visualizar informação por meio de mapas.

Neste contexto, BAILEY (2001) usa o termo “epidemiologia geográfica” para referir-se a esta área do conhecimento e a fragmenta em quatro grandes áreas de interesse estatístico:

- mapeamento de doenças: foco na produção de mapas de distribuição da incidência de doenças;
- estudos ecológicos: associação da incidência observada e os fatores de risco em potencial para coletividades;
- estudos de agrupamentos de doenças: focado na identificação de áreas com risco significativo da doença;
- análise e monitoramento ambiental: preocupação com a verificação da distribuição espacial de fatores ambientais relevantes para a saúde humana, a fim de estabelecer medidas de controle preventivo.

BAILEY (2001) complementa esta classificação indicando a possibilidade de subáreas em função do contexto epidemiológico: se os dados disponíveis são casos individuais da doença ou se estão agregados em uma área geográfica, ou ainda se há uma dimensão temporal ou espacial para análise.

Quando os dados são agregados em áreas, classifica-se o estudo como ecológico, conforme define ALMEIDA FILHO E ROUQUAYROL (2002):

“Os estudos ecológicos abordam áreas geográficas ou blocos de populações bem delimitados, analisando comparativamente variáveis globais, quase sempre por meio de correlação entre indicadores de condição de vida e indicadores de situação de saúde” (ALMEIDA FILHO e ROUQUAYROL, 2002, p.174).

Na mesma linha, CASTRO et al (2003) definem que nos estudos ecológicos a unidade de análise não é o indivíduo, mas sim um grupo de indivíduos, ou seja, os dados disponíveis no estudo dizem respeito a um

agregado de indivíduos. Com efeito, a análise baseia-se na comparação dos indicadores de exposição à poluição do ar e doenças entre as unidades.

ALMEIDA FILHO e ROUQUAYROL (2002) citam como vantagens dos estudos ecológicos a facilidade de execução, o baixo custo relativo, a simplicidade analítica e a capacidade de geração de hipóteses. Na contrapartida, os problemas inerentes são o baixo poder analítico, as técnicas de análise de dados pouco desenvolvidas e a vulnerabilidade à falácia ecológica.

No que tange à tipologia, ALMEIDA FILHO e ROUQUAYROL (2002) afirmam que os estudos ecológicos constituem desenhos agregados-observacionais-transversais e seguem as abordagens analíticas da comparação de indicadores, análise de correlação linear ou da análise gráfica. As agregações podem ser institucionais ou territoriais. Esta última, diretamente associada ao Geoprocessamento.

Os produtos do Geoprocessamento são dados convertidos em informação, através de convenções e representações indicadas por pontos, linhas ou polígonos. A informação, quando vinculada a um juízo de valor, torna-se conhecimento e pode subsidiar decisões que convertem-se em ações. Tais ações, no âmbito das políticas públicas, são essenciais para a melhoria da qualidade de vida da população.

OPAS (2007) afirma que uma política pública saudável pode ser percebida quando o conjunto de sistemas que abrangem determinada população, favorece o desenvolvimento humano. Entre eles, as condições de saúde são de grande importância, pois, a ausência desta implica no enfrentamento de situações negativas que comprometem as potencialidades necessárias para o desenvolvimento humano.

Considerando o uso das ferramentas de Geoprocessamento para tomada de decisão em saúde pública, MÜLLER et al (2007), colocam que estes recursos constituem um instrumento poderoso e eficiente para o gestor público, formulador de políticas e tomador de decisão. No entanto,

ainda são pouco utilizados neste contexto há carência de estudos mais aprofundados sobre a validação de sua utilização.

LEMOS e LIMA (2002) sustentam que a compreensão do processo de organização do espaço nas coletividades, em diferentes lugares e períodos, contribuem para um melhor entendimento do processo saúde-doença. A elaboração de sistemas que consolidam dados ambientais e de saúde pública possibilita a implantação de programas de vigilância ambiental em saúde.

Estabelecido pelo decreto nº 3.450 de 10 de maio de 2000, o sistema nacional de vigilância ambiental brasileiro, sistematiza informações sobre fatores biológicos, qualidade da água para consumo humano, contaminantes ambientais químicos e físicos que possam comprometer a qualidade da água, ar e solo e os riscos decorrentes de desastres naturais e de acidentes com produtos perigosos (FUNASA, 2000).

Portanto, a produção de informação sobre saúde e ambiente, além de proporcionar um instrumental valioso para o gestor público para tomada de decisão em questões de saúde pública, está prevista na legislação brasileira, de modo a permitir um monitoramento mais criterioso, sobre a qualidade do espaço, bem como as condicionantes ambientais que estão integradas à vida da população.

2.2 Abordagens da Estatística Espacial

É provável que, desde a concepção do vocábulo sistema, o ser humano tente estruturar modelos que os simulem. Assim, desde Leibniz, no século XVII, discute-se a definição do termo sistema e seus aspectos de unidade, multiplicidade, totalidade, diversidade, organização e complexidade (MORIN, 1977).

Nesta linha, uma das principais dificuldades no entendimento de determinado sistema está na identificação das variáveis e suas inter-relações. Definir estes pontos compreende a delimitação do que se quer observar, considerando as implicações da exclusão de elementos do contexto.

Com relação à tipologia, diversos autores desenvolveram esta temática. Assim, sistemas podem ser classificados conforme diversos critérios conforme apresenta o Quadro 1.

Critério	Classificação	Sub-Classes
Funcionalidade	Isolados	
	Não isolados	Abertos
		Fechados
Complexidade	Morfológicos	
	Em sequência ou encadeantes	Caixa Branca
		Caixa Cinza
		Caixa Preta
	Processo-resposta	
	Controlados	Simples
		Complexos mas desorganizados
		Complexos e organizados

Quadro 1 – Classificação de tipologia de sistemas. Fonte: Adaptado de CHRISTOFOLETTI, 1999.

Considerando o escopo deste trabalho, a definição de sistemas isolados e não-isolados é relevante, pois agrega características importantes no contexto da modelagem de sistemas.

Já a expressão “modelo”, ou “modelagem”, é um pouco mais recente. Na década de 1940, engenheiros, físicos, estatísticos e economistas utilizavam o termo “modelo matemático” para caracterizar a representação de uma situação dada (NOGUEIRA et al, 2008) .

CHRISTOFOLETTI (1999) coloca que a abordagem sistêmica é fundamental na orientação dos procedimentos de modelagem em sistemas ambientais.

NOGUEIRA et al (2008) colocam ainda que os fenômenos das ciências e processos tecnológicos, entre outros, são situações compostas de etapas bem definidas que podem ser “matematizadas”. Contudo, a modelagem consiste em um processo de maior grandeza, que significa “*desvelar possíveis estruturas matemáticas contidas na situação*”. Assim sendo, estruturas seriam aportes para obter novas informações, previsões, projeções e estratégias de atuação sobre o fenômeno modelado.

BERTALANFFY (1975) já mencionava que os sistemas que excedem de muito a matemática convencional podem ser submetidos à computação. Desta forma, experiências de laboratório ou campo podem ser substituídas pela simulação em computadores, com posterior submissão a prova pelos dados experimentais.

No entanto, deve-se ressaltar a dificuldade em gerar tais modelos computacionais para simulação. CHRISTOFOLETTI (1999) acrescenta que:

“Embora amplamente usados e mencionados, os modelos estão longe de construir uma panacéia universal e devem ser usados considerando as suas limitações, procurando-se discernir as maneiras mais adequadas de superá-las” (CHRISTOFOLETTI, 1999, p. 32).

Desta forma, há que se considerar eventuais problemas na apuração dos resultados a partir de modelos matemáticos, mesmo implementados em ambientes computacionais.

A modelagem está diretamente relacionada à estatística, ciência esta que objetiva colecionar, analisar e interpretar dados numéricos provindos de fenômenos coletivos, permitindo inferir sobre as leis que regem tais fenômenos (MICHAELIS, 2008). Estas análises são necessárias para a composição do *status quo* sobre os fatos considerados, sejam eles de um determinado período ou de uma determinada região.

Segundo CORREA (2003), a análise dos resultados pode seguir por três distintas formas de apresentação: a Estatística Descritiva, a Estatística Inferencial e a Probabilidade.

Na Estatística Descritiva, os dados são coletados, organizados e analisados de forma a descrever uma população ou amostra. A Estatística Inferencial, ou Indutiva, procura obter informações sobre uma população, a partir dos dados observados. Por fim, a Probabilidade visa explicar os fenômenos estudados através de modelos matemáticos em um processo de generalização, associado a uma margem de incerteza (CORREA, 2003).

No entanto, são diversos os fatores que envolvem um problema de estudo. Esta diversidade era limitante nos estudos e deixou de ser a partir dos progressos tecnológicos da computação que propiciaram maior habilidade de se manipular dados. Este cenário crescente de domínio e análise dos dados vem favorecendo a estatística como um todo, mas em particular a Inferencial e a Análise Multivariada (HAIR Jr. et al, 2005).

HAIR Jr. et al (2005) argumentam que é difícil definir o termo Análise Multivariada, mas coloca que este “refere-se a todos os métodos estatísticos que simultaneamente analisam múltiplas medidas sobre cada indivíduo ou objeto em investigação”. Ou seja, *lato sensu*, uma análise que envolve duas ou mais variáveis, pode ser considerada Multivariada, estendendo-a, assim, à Análise Univariada (apenas uma variável) e à Bivariada (classificação cruzada, por correlação, análise de variância e regressão simples para usar duas variáveis) (HAIR Jr. et al, 2005).

As análises aqui descritas figuram entre os diversos modelos de processos de resposta que a estatística provê, nos quais o objetivo é capturar as principais características de processos empíricos sob observação.

Por outro lado, a coleta de dados deve observar dois fatores aglutinadores: o tempo e o espaço. O fator temporal ou cronológico refere-se ao período em que o fenômeno é observado. O fator espacial ou geográfico está associado ao local onde os fenômenos acontecem. As séries históricas

ou temporais e as séries geográficas ou territoriais são de grande interesse para a epidemiologia (BERQUÓ et al, 2002).

O fator temporal insere o caráter longitudinal do estudo quando se analisa a evolução do conjunto de dados ao longo de um período estabelecido. Esta abordagem é utilizada em estudos de tendência, séries temporais, estudos prospectivos (Coorte), estudos retrospectivos (caso-controle) e ensaios clínicos (MS, 2007).

Já o espaço é uma instância da sociedade, composta por elementos como homem, instituições, e outros atores, bem como as instâncias econômica e cultural-ideológica. Tais elementos estão submetidos a variações qualitativas e quantitativas. Estas variações e relações entre os elementos são objetos de estudo na geografia (COSTA e TEIXEIRA, 1999).

É comum utilizar o termo “espaço geográfico” para referenciar o local que se deseja estudar, pois este termo define uma coleção de entidades geográficas referenciadas em uma parte da Terra que participa de um processo geográfico, ou ainda, uma coleção de locações terrestres onde um fenômeno geográfico pode ocorrer (WORBOYS, 1995).

Nesta mesma linha, GOLLEDGE e STIMSON (1997 citado por CROMLEY e McLAFFERTY, 2002), usam o termo “*activity space*” para identificar a área onde as pessoas gastam tempo e são consideradas para distribuição geográfica em estudos de saúde pública. Esta área pode ser uma residência, local de trabalho, restaurante, escola, entre outros.

O ramo da estatística que permite analisar a localização espacial dos eventos é denominado Estatística Espacial. Explora-se, nesta área, a identificação, localização e visualização da ocorrência de fenômenos distribuídos no espaço, com auxílio de modelos que permitem a identificação de padrões a partir da distribuição espacial (MS, 2007).

O termo “espacial” caracteriza que o dado possui referência geográfica, ou seja, o local onde o fato gerador do dado ocorreu é conhecido (HAINING, 2004). Esta indexação dos dados no espaço é relevante para análise, quando a geografia é variável dependente do estudo.

Sendo assim, a Estatística Espacial provê ferramentas para o estudo de fenômenos, em caráter quantitativo, distribuídos em uma área de interesse. O estudo é direcionado para uma melhor compreensão dos processos geradores dos fenômenos observados.

Outro termo correlato é Geoestatística, que DIGGLE et al (2003) definem como a parte da Estatística Espacial preocupada com a contínua variação espacial, onde o foco científico é o estudo de um fenômeno espacial que existe em uma área contínua e pode ser tratado como um processo estocástico.

A este raciocínio, DODGE (2008) complementa afirmando que geoestatística é a estatística aplicada especialmente em geologia e hidrologia, utilizada para avaliar e prover a distribuição de valores de uma variável no espaço e no tempo.

Nesta mesma abordagem, DIGGLE e RIBEIRO Jr. (2000) complementam afirmando que os objetivos da análise geoestatística estão alinhados a duas vertentes: predição e estimativa. A predição está relacionada a inferências sobre sinais não observados e a estimativa refere-se a inferências sobre parâmetros de um modelo estocástico.

A estatística espacial sobrepõe-se à estatística comum quando o fenômeno a ser investigado tem a geografia como variável dependente do estudo. Esta característica é denominada dependência espacial.

Nesta linha, a dependência espacial é a chave para a compreensão de fenômenos espaciais. A conhecida lei cunhada por Waldo Tobler diz que *“todas as coisas são parecidas, mas coisas mais próximas se parecem mais que coisas mais distantes”*, e refere-se ao conceito de dependência espacial. Tal conceito indica que a maioria das ocorrências apresenta uma relação entre si que depende da distância entre elas (DRUCK et al, 2004).

HAINING (2004) observa que existem diferenças importantes entre espaço e tempo relacionadas à dependência. O tempo evolui em fluxo unidirecional fazendo com que o passado exerça influência sobre o presente e este, sobre o futuro, mas não o inverso. Já o espaço estende-se por duas

dimensões, ocasionando dependências isotrópicas (em ambos os eixos) ou anisotrópicas (somente em um dos eixos). Por fim, o recorte da periodicidade é encontrado somente nas séries temporais.

Por outro lado, HAINING (2004) acrescenta que, assim como ocorre com a dependência espacial, uma variável tende a manter similaridade quando observada suas variações em um curto espaço de tempo, tendendo a perder esta característica quando observada em períodos mais extensos.

Portanto, tempo e espaço apresentam estruturas de dependência e se complementam na busca do entendimento sobre os fenômenos do lugar, ao longo do tempo.

A dependência espacial é quantificada por meio de funções de autocorrelação espacial. Estas funções geram medidas que variam entre -1 e 1, onde os valores próximos a zero indicam inexistência de autocorrelação espacial significativa entre os objetos vizinhos. Valores próximos a 1 e -1 indicam existência de autocorrelação, positiva ou negativa, respectivamente (MS, 2007).

A autocorrelação espacial pode ser expressa através de Correlograma, que é a representação gráfica de um conjunto de dados em uma determinada distância onde a distância é indicada no eixo X e os valores da autocorrelação (entre 1 e -1), no eixo Y.

Outro aspecto inerente aos estudos espaciais é a estacionariedade. Um processo pode ser classificado como estacionário se for constante em toda a região estudada, ou seja, o fenômeno não se altera na área de estudo, não registrando tendências. Porém, se as medidas de similaridade mudam em cada parte da área analisada, o processo é classificado como não estacionário e a estrutura de dependências é heterogênea (MS, 2007; HAINING, 2004).

Vale ressaltar a importância de observar e classificar os processos, pois a estacionariedade pode influenciar os estimadores estatísticos (MS, 2007). Em particular para a Saúde Pública, a caracterização de não

estacionária pode representar focos de interesse para estudos mais detalhados.

A análise espacial é desenvolvida por meio de um conjunto de procedimentos, com objetivo de escolher um modelo inferencial para explicar um determinado fenômeno. Inicia-se com métodos genéricos de análise exploratória, como a construção de mapas, a fim de visualizar *outliers* (observações atípicas) acerca da distribuição de uma ocorrência em estudo ou sobre agrupamentos vizinhos.

Estes procedimentos, geralmente, são desenvolvidos em estudos ecológicos, objetivando o fomento de hipóteses que permitirão desenvolver estudos mais aprofundados com a utilização de modelos inferenciais espaciais (DRUCK et al, 2004).

HAINING (2004) observa a importância do processo de entendimento dos resultados da análise espacial. Deste, faz parte o relacionamento entre a realidade do meio estudado e como os dados foram capturados no banco de dados. A usabilidade das operações de análise espacial aplicadas sobre o conjunto de dados depende de quão bem representada está a realidade capturada na matriz de dados.

O mesmo autor cita também o crescimento do interesse sobre o tema análise espacial no âmbito da comunidade acadêmica, em função de hipóteses e questões bem elaboradas, bem como a qualidade dos dados e metodologias estatísticas apropriadas para uma área multidisciplinar denominada *Geographic Information Science* (GISc – Ciência da Informação Geográfica) (HAINING, 2004).

Reforçando este ponto, as ferramentas SIG são largamente empregadas em várias áreas do conhecimento humano, sempre que a variável geográfica é determinante do problema em estudo. Esta amplitude é alcançada devido aos requisitos funcionais pré-estabelecidos para a construção deste tipo de software.

CROMLEY e McLAFFERTY (2002) ressaltam que os dados armazenados no repositório de dados não estão organizados como planilhas

eletrônicas, mas de forma que é possível localizar variáveis na superfície da terra e analisar o mesmo espaço sob diferentes aspectos, além de permitir estabelecer relacionamentos de vizinhança entre os mesmos.

Ferramentas SIG despertam grande interesse por permitirem implementar diferentes metodologias e possuírem várias habilidades inseridas no software como (COWEN,1988; GOODCHILD,1995b citados em CROMLEY e McLAFFERTY,2002):

- Armazenar, processar e mostrar relacionamentos espaciais entre objetos;
- Armazenar diferentes atributos de um objeto;
- Analisar atributos espaciais e tabulares com simplicidade de gerenciamento e recuperação de dados;
- Integrar dados espaciais e tabulares de diferentes fontes.

A teoria sobre um fenômeno em uma dada área do conhecimento humano pode ser fomentada por hipóteses com suporte visual e analítico produzido por ferramenta SIG em estudos exploratórios. A Figura 2 resume o processo de utilização da análise espacial em pesquisas.

Figura 2 - Relação entre análise espacial e teorias disciplinares.
(Fonte: DRUCK et al, 2004).

Seguindo os quadros da Figura 2, as “Teorias” são desenvolvidas nas respectivas áreas de domínio do conhecimento. Neste momento, os SIGs dão suporte através de técnicas de visualização disponíveis em software específico. Aqui, os estudos exploratórios desempenham papel importante na formulação de hipóteses e problemas de pesquisa.

Nesta linha, STOCKBURGER (1996) corrobora a importância do processo, afirmando que o método científico é um processo de construção e verificação de modelos que passa quatro estágios: simplificação, representação, transformação e verificação. As implicações derivadas de estágios anteriores são comparadas através de experimentos ou observações do mundo real. A Figura 3 mostra uma síntese deste raciocínio.

Figura 3 – Os quatro estágios do método científico e seus relacionamentos. Fonte: Adaptado de STOCKBURGER (1996).

Contudo, a discussão que permeia o modelo, não é a sua perfeita concordância com o mundo real e sim, se o mesmo é adequado para a proposição do estudo (STOCKBURGER, 1996).

O pesquisador deverá, então, formular “Modelos Inferenciais” quantitativos que serão testados e validados por algum procedimento de “Análise Espacial”. Os resultados poderão auxiliar na tomada de decisão sobre rejeitar ou não os conceitos formulados na teoria, retroalimentando o sistema. BAILEY (1994 citado por CROMLEY e McLAFFERTY, 2002) define análise espacial como a habilidade de manipular dados espaciais em diferentes formatos para extrair significado adicional como resultado.

A caracterização dos problemas abordados em análise espacial leva em consideração diferentes tipos de dados. Estes podem ser representados como pontos, superfícies contínuas ou áreas com contagem e taxas agregadas (DRUCK et al, 2004).

Estes tipos de dados encapsulam informações que podem representar diversos fatos do mundo real, que são selecionados como objeto de interesse para eventuais estudos, com ampla cobertura nas diversas áreas do conhecimento humano.

Os pontos são utilizados para identificar a localização de um evento ou objeto no mapa. Tais eventos podem representar uma ocorrência criminal, um caso de dengue, ou ainda, determinado cliente de uma organização. Este tipo de representação é utilizado quando é necessário conhecer as coordenadas nas quais o objeto está localizado no mapa.

Os objetos podem ser representados no modelo na forma de ponto, linha ou polígono. Estes objetos, inseridos em um plano de modelagem, podem ser submetidos a diversos tipos de análise, que vão desde o distanciamento entre eles, até a influência que eles podem exercer ou pressionar outros objetos do modelo. Estas inferências são potencializadas com o uso de camadas que permitem novas visualizações através das sobreposições.

A Figura 4 ilustra tipos de representação dos objetos, considerando formas discretas e contínuas. No exemplo da representação contínua, apresenta-se a construção de mapas temáticos, a partir de amostras de uma localidade. Os valores coletados são base para inferir sobre os demais pontos, não coletados, gerando a modelagem da superfície, na figura “Uso do solo” e dos contornos na figura “Densidade Populacional”.

Neste caso, a análise espacial procura descrever a variação espacial dos valores atribuídos aos objetos (atributos) ao longo da área de estudo. Testa-se aqui, hipóteses sobre padrões espaciais dos objetos em função de outras variáveis. (DRUCK et al, 2004; HAINING, 2004). Por exemplo,

verificar se existe algum fator ambiental associado aos casos de doença pontuados em uma determinada área de estudo.

Figura 4 – Representação de objetos com pontos, linhas e polígonos.

Fonte: Adaptado de HAINING, 2004.

Assim, a análise espacial pode ser trabalhada em função da forma que os objetos são representados no modelo computacional. Taxonomicamente, os estudos podem ser classificados como análise de eventos, análise de superfícies, análise espacial de áreas, análise temporal e análise de redes. Ainda há a possibilidade de trabalhar estes formatos de maneira integrada.

A análise de eventos é caracterizada pelo uso de pontos nas camadas do SIG. Cada evento é representado por uma localização (coordenada X,Y),

Concluindo, a utilização da análise espacial permite uma melhor representação dos fenômenos, considerando que a visualização dos dados em mapas amplia a sensação de proximidade do problema analisado e traz maior domínio sobre as variáveis de determinado fenômeno.

2.3. O Uso da Análise Espacial de Dados na Saúde

MORAN (1994) afirmou que o ser humano pode ser avaliado segundo sua adaptabilidade, através de parâmetros demográficos, energéticos e nutricionais. Em particular os parâmetros demográficos mais utilizados são o equilíbrio entre natalidade e mortalidade, morbidade ou incidência de doenças e a taxa de reprodução da população. Nesta linha, observa-se grande quantidade de trabalhos na área de saúde que utilizam estes mesmos parâmetros incorporando a variável espaço à análise que se deseja realizar.

Neste aspecto, as ferramentas de SIG despontam como referência na associação de dados chamados “tabulares”, ou seja, dispostos em tabelas a mapas digitais, atribuindo assim, os dados dos elementos do estudo aos objetos geográficos. Esta fusão permite construir visualizações temáticas que tornam mais perceptível a observação de fenômenos em um determinado contexto estudado.

De forma complementar, RYTKÖNEN (2004) coloca que, embora apresentem uma evolução importante relacionada à computação, os mapas de doença vêm desempenhando um importante papel na geografia médica e na epidemiologia desde o século XVIII, com os seguintes objetivos:

- descrever a variação espacial da incidência de doenças;
- formular hipóteses etiológicas;
- identificar áreas de risco para ações preventivas;
- prover mapas de risco de doenças mais confiáveis;
- prover melhor alocação de recursos e análise de riscos.

DRUCK et al (2004) afirmam que *“além da percepção visual da distribuição espacial do problema, é muito útil traduzir os padrões existentes com considerações objetivas e mensuráveis”*. Tal colocação sintetiza a

importância de buscar a semântica através de métricas, e estimula a produção de trabalhos que apresentem resultados embasados não apenas na visualização de mapas, como também na comprovação estatística.

No esforço de atender às demandas neste campo de conhecimento, recentes avanços na área da estatística, trazem novas perspectivas com os modelos aleatórios e a investigação de efeitos denominados de vizinhança. Tais perspectivas apontam para resultados que possibilitam intervenções inovadoras nas comunidades estudadas (CARVALHO e SOUZA-SANTOS, 2005).

BAILEY (1995 citado por CARVALHO, 1997) esclarece que as técnicas de análise espacial podem ser classificadas em função do objeto e do tipo de dado disponível. As classificações são:

- distribuição de pontos: quando o objeto da análise é a localização dos eventos;
- geoestatística: técnicas que pressupõem a continuidade espacial do objeto;
- dados de áreas: quando a ocorrência do fenômeno é mensurada a partir de dados agregados em áreas;
- deslocamento: quando o objeto do estudo é o fluxo e o acesso entre regiões.

Entre as classificações, aquela denominada “dados de áreas”, desperta particular interesse no âmbito da geografia médica na medida em que o estudo das coletividades é estimulado, pois o ser humano vive em grupos e a análise individual não capta as interações entre pessoas, como transmissão de infecções, comportamentos e valores (CARVALHO, 1997).

Neste contexto, CARVALHO e SOUZA-SANTOS (2005) observam que a maioria dos estudos epidemiológicos está baseada no individualismo metodológico, ou seja, na idéia de que a distribuição dos fenômenos de saúde pode ser explicada através das características dos indivíduos estudados. À luz deste debate, os autores completam o raciocínio explicitando a necessidade de uma abordagem que considere o tempo e o espaço para clareza de um fenômeno.

“Padrões de saúde como mortalidade ou morbidade, propagação de epidemias, transmissão sexual de doenças ou a transferência de comportamentos e valores não podem ser explicados sem uma abordagem que, além de considerar os grupos estude o espaço e o tempo” (CARVALHO e SOUZA-SANTOS, 2005, p. 362).

Portanto, o tempo é um fator a considerar na produção de trabalhos que envolvem o espaço. Muitos estudos utilizam representações estáticas dos fenômenos espaciais. Tal constatação se deve ao fato de o principal produto do SIG ser um mapa.

Contudo, fenômenos como ocupação do solo, poluição e epidemias são dinâmicos e as representações estáticas não são suficientes para exibir determinados aspectos de forma adequada. Nestas situações, uma representação espaço-temporal poderia clarificar variações do fenômeno, tanto no espaço como no tempo.

2.4. SIG e Estudos da Qualidade do Ar e a Saúde

O crescente interesse pela área de saúde ambiental tem impulsionado o uso de aplicações SIG em saúde pública, dado que o gerenciamento ambiental está muito próximo das ferramentas de geoprocessamento, utilizando-o para representar o ambiente e suas condicionantes. Exemplos do uso de SIG envolvendo saúde e ambiente são diversos, abordando questões como exposição de substâncias tóxicas, ressurgimento de doenças infecto-contagiosas, efeitos associados à radiação, acesso a equipamentos públicos de saúde coletiva, questões relativas à saúde do trabalhador, entre outros.

Os estudos que envolvem poluentes atmosféricos e efeitos à saúde são de grande importância no âmbito da saúde pública, em especial, onde as concentrações de material particulado e outros poluentes excedem os padrões estabelecidos para qualidade do ar (OPAS, 2005).

OPAS (2005) ainda relata que os grupos mais suscetíveis a efeitos adversos da contaminação do ar são as crianças e idosos. Os sistemas mais afetados são o respiratório e o cardiovascular, apresentando efeitos como a diminuição da função pulmonar, variações do ritmo cardíaco e resposta a inflamações.

Vale ressaltar que outros agentes etiológicos estão diretamente relacionados aos efeitos adversos da poluição do ar, ampliando a complexidade de se chegar a uma associação causal.

Estudos comprovam a importância do fumo em doenças obstrutivas crônicas (asma, bronquite e enfisema), e doenças cardiovasculares como hipertensão arterial, *angina pectoris* e infarto do miocárdio. Outros agentes etiológicos são freqüentemente citados na literatura como, por exemplo, poluição domiciliar – *indoor pollution*, combustão de biomassa, produtos gerados pela combustão de gás de cozinha, pesticidas, solventes e asbestos utilizados para isolamento de incêndios, entre outros (DUCHIADE, 1992).

Segundo CETESB (2009):

“Poluente atmosférico é toda e qualquer forma de matéria ou energia com intensidade e em quantidade, concentração, tempo ou características em desacordo com os níveis estabelecidos em legislação, e que tornem ou possam tornar o ar impróprio, nocivo ou ofensivo à saúde, inconveniente ao bem estar público, danoso aos materiais, à fauna, e à flora ou prejudicial à segurança, ao uso e gozo da propriedade e às atividades normais da comunidade” (CETESB, 2009).

No que tange as questões de poluição do ar, entende-se por poluente atmosférico, toda e qualquer forma de matéria sólida, líquida ou gasosa e de energia que, presente na atmosfera, possa alterar suas condições normais levando à ocorrência de efeitos deletérios potenciais (CETESB, 2009).

Os estudos de saúde podem ser classificados segundo a fonte de emissão. Existem diversas fontes de emissão, sejam elas naturais (vulcões, vegetação, ressuspensão pelo vento) ou antrópicas (transporte, processos industriais, queimadas, etc.).

GATRELL (2002) distingue três classes de agrupamento para fontes de emissão, da mesma forma que os objetos geográficos utilizados em geoprocessamento, áreas, linhas e pontos. Os poluentes considerados de origem difusa seriam classificados como “áreas”, por exemplo, estudos que abordam níveis de ozônio em uma dada região. Estudos que envolvem emissões veiculares seriam classificados como “lineares”, devido ao comportamento dos processos que ocorrem ao longo das vias públicas. Por último, as emissões observadas a partir de uma fonte fixa seriam as “pontuais”, como chaminés de indústrias.

No entanto, esta forma de classificação pode não ser suficiente, pois as fontes de emissão muitas vezes somam-se, criando áreas de interseção e ampliando a complexidade e os fatores de risco na área de estudo.

Soma-se a isso a complexidade relativa a estudos que envolvem eventos de queimadas, pois, estes diferem em forma e conteúdo de fontes como chaminés de fábricas ou veículos. Embora uma área em processo de queimada seja fixa, a combustão avança em função de ventos e biomassa disponível. Dependendo da extensão, este processo ocorrerá de forma móvel até que seja cessado.

Outro diferencial está na regularidade das emissões durante o processo, pois a área queimada pode ser um talhão de floresta com certo grau de biodiversidade, ou uma área de monocultura, como acontece no caso da cana-de-açúcar. Por fim, a duração do evento é variável podendo durar horas (plantação de cana) ou até meses (uma floresta).

Vale salientar que a queima de biomassa ocorre de forma natural, acidentada ou deliberada, podendo ser controlada ou não. Um evento deste tipo pode atingir grandes extensões de florestas e savanas ou vegetações menos densas (ARBEX et al, 2004).

Assim como a fonte de emissão, é necessário observar os poluentes e os efeitos que se deseja estudar. GLOAG (1981) descreve três categorias de poluentes atmosféricos: as partículas sólidas, os gases e a radioatividade. Entre as partículas sólidas encontram-se a fumaça (*smoke*), o sulfato de amônia e as cinzas. Os gases incluem os óxidos sulfúricos (SO), de nitrogênio (NOx), hidrocarbonetos e HPA (hidrocarbonetos policíclicos aromáticos).

Destes, o material particulado gerado na combustão de biomassa é o poluente de maior toxicidade para o ser humano, conforme descrito na revisão de ARBEX et al (2004). No entanto, outros compostos produzidos no processo de queima de biomassa, também são estudados à luz dos efeitos deletérios à saúde. O Quadro 2 apresenta uma síntese acerca destes compostos.

Compostos	Exemplos	Fonte	Notas
Partículas	Partículas inaláveis (PM ₁₀)	Condensação após combustão de gases; combustão incompleta de material inorgânico; fragmentos de vegetação e cinzas	Partículas finas e grossas. Partículas grossas não são transportadas e contêm principalmente cinzas e material do solo
	Partículas respiráveis	Condensação após combustão de gases; combustão incompleta de material orgânico.	No caso de fumaça proveniente da queima de biomassa comporta-se como partículas finas
	Partículas finas (PM _{2,5})	Condensação por combustão de gases; combustão incompleta de material orgânico	Transportadas através de longas distâncias. Produção primária e secundária
aldeídos	acroleína	Combustão incompleta de material orgânico	
	formaldeído	Combustão incompleta de material orgânico	
Ácidos inorgânicos	Monóxido de carbono (CO)	Combustão incompleta de material orgânico	Transportado através de longas distâncias
	ozônio	Produto secundário do óxidos de nitrogênio e hidrocarbonetos	Presente somente adiante do fogo, transportado através de longas distâncias
	Dióxido de nitrogênio(NO ₂)	Oxidação em altas temperaturas do nitrogênio do ar	Espécies reativas; a concentração diminui com a distância do fogo
Hidrocarbonetos	benzeno	Combustão incompleta de material orgânico	Transporte local; também reage com outras formas de aerossol orgânico
Hidrocarbonetos aromáticos policíclicos(PAHs)	Benzopireno (BaP)	Condensação após combustão de gases; combustão incompleta de material orgânico	Compostos específicos que variam de acordo com a composição da biomassa

Quadro 2 – Principais poluentes provenientes da queima de biomassa.

Fonte: ARBEX, et al, 2004.

As variáveis: fonte de emissão, agente ou poluente e efeitos adversos à saúde humana, observados sob uma área de estudo e período de exposição, compõem o conjunto mínimo de entrada para um modelo de análise de risco em saúde e ambiente. A Figura 5 apresenta um modelo geográfico para processos de risco-exposição-efeito.

Algumas questões merecem discussão sobre o modelo apresentado. Primeiramente, é raro encontrar dados sobre poluentes, efeitos, e outras variáveis ambientais no mesmo banco de dados. Muitas vezes os dados apresentam problemas de intermitência por vários fatores, como quebra de equipamentos e perda de dados, entre outros, comprometendo a qualidade da informação.

Figura 5 – Modelo geográfico risco-exposição-efeito. Fonte: Adaptado de CROMLEY e MCLAFFERTY, 2002).

Embora o modelo apresente consistência no processo, ele não prevê a possibilidade de observar grupos ao invés de indivíduos. PAWITAN e STEEL (2006) colocam a necessidade de determinar onde a dependência espacial é esperada para definir o modelo. Isto evita resultados inapropriados para as análises.

Ocorre que, em certos casos, os dados sobre o indivíduo não estão disponíveis devido a questões de preservação de confidência, muito comum na área da saúde.

Ferramentas SIG permitem trabalhar com dados em diferentes níveis de agregação. Esta característica viabiliza a análise de padrões ou diferenças em escala local ou regional. A identificação de grupos de risco e áreas críticas é facilitada pela possibilidade de processar e analisar múltiplas variáveis de forma simultânea (OPAS, 2002).

No entanto, a possibilidade de gerar agregações distintas pode acarretar resultados diferentes. Este fenômeno é conhecido como problema da unidade de área modificável (MAUP – *Modifiable Areal Unit Problem*) (DIAS et al, 2002; PAWITAN e STEEL, 2006).

Outro problema conhecido é denominado Falácia Ecológica. Este caso, também conhecido como *bias* ou viés de agregação, ocorre quando as conclusões para dados agregados não refletem a realidade dos indivíduos utilizados na agregação. Duas abordagens foram desenvolvidas para minimizar o viés. A primeira foi a formulação de modelos estatísticos e procedimentos de estimativa para redução do viés, a segunda, foi o desenvolvimento de controles mais efetivos através de parametrizações para gerar polígonos (DUQUE et al, 2006).

Sobre o mesmo tema, ALMEIDA FILHO e ROUQUAYROL (2002) colocam que a Falácia Ecológica ocorre quando existe a suposição de que os coeficientes de uma determinada área correspondem à população total da mesma, mas, no entanto, este valor corresponde à média da variação dos subgrupos ali existentes. Tal problema pode ser minimizado com a utilização de agregados menores.

Os mesmos autores acrescentam que uma área ecológica pode sintetizar um conjunto considerável de variáveis e processos. Assim, no mesmo tempo em que esta situação amplia a complexidade do estudo, aproxima o escopo para uma realidade social mais concreta. Este raciocínio leva à característica que dá identidade própria aos estudos de agregados na epidemiologia. Desta forma, para o âmbito da epidemiologia, os autores sugerem a troca do termo para “efeito agregado”, minimizando a conotação de erro a ser evitado ou controlado para este fenômeno (ALMEIDA FILHO e ROUQUAYROL, 2002).

Além da Falácia Ecológica, o baixo poder analítico e a falta de desenvolvimento de técnicas de análise de dados, são apontados na literatura problemas associados aos estudos ecológicos, cuja denominação corresponde aos desenhos agregados-observacionais-transversais (ALMEIDA FILHO e ROUQUAYROL, 2002).

Neste contexto, CARVALHO e SOUZA-SANTOS (2005) colocam que a compreensão do que afeta a saúde de coletividades carece de estudos com base em variáveis agregadas, por meio de seleção, distribuição, interação, adaptação, entre outras possibilidades. Medidas em nível individual podem não ser suficientes para entender processos que envolvem grupos populacionais.

Complementando, GATRELL (2002) afirma que estudos individuais podem contribuir para a identificação de variáveis de confusão e oferecer melhores chances de detecção de impacto causado por exposição a um poluente de interesse. O mesmo autor acrescenta que estes estudos devem ser confrontados com investigações sobre morbi-mortalidade, assim como internações hospitalares por doenças associadas ao objeto do estudo.

2.5. Estudos Envolvendo Queima de Biomassa e Saúde

O interesse por estudar associações causais entre os produtos da queima de biomassa e os efeitos à saúde humana não é novo. Literatura científica e história mostram isso de forma longa, na eminência de obter resposta para diversas questões inerentes ao tema.

Estas questões são ou foram estudadas em frentes de pesquisa, com distintas abordagens atreladas aos mesmos descritores. Isto ocorre em função da riqueza de possibilidades que esta área interdisciplinar proporciona.

Cabe aqui ressaltar que os problemas respiratórios não estão restritos a divisões geopolíticas. Existem relatos nos 5 continentes, alguns estudando efeitos em escala local, outros em escala global, ou seja, o fator gerador ocorreu muito distante de onde se observou o efeito.

Embora haja interesse e relevância envolvidos nesta área do conhecimento, ainda são poucos os estudos relatados na literatura científica e muitas questões sobre o tema ainda não foram resolvidas (RIBEIRO, 2008).

Em primeira instância, os estudos desta linha podem ser categorizados quanto à exposição, *indoor* ou *outdoor*¹. Os estudos *indoor*, *lato sensu*, estão atrelados ao problema do uso de biomassa como combustível para aquecimento de residências, cozimento de alimentos ou como fonte de energia.

Nesta linha, ARBEX et al (2004) fazem uma extensa revisão de trabalhos que relatam doenças como DPOC (doença pulmonar obstrutiva crônica), pneumoconiose, catarata, cegueira, tuberculose e efeitos adversos na gestação, associadas à exposição em ambientes fechados à fumaça de fogões que funcionam com queima de biomassa. Estudos realizados na Bolívia, Índia, México, Nepal, Nova Guiné e Paquistão, abordam este tema.

¹ Estes termos são utilizados para definir ambientes internos ou abertos, respectivamente.

Assim, são consideradas algumas facilidades no controle de variáveis de confusão, pois existe regularidade na exposição e o desenho do estudo é focado no indivíduo, geralmente em escala local, permitindo identificar fatores de confusão como tabagismo e doenças crônicas, entre outros. No entanto, os estudos realizados nos países em desenvolvimento apresentam limitações metodológicas como melhor determinação da exposição ao poluente, o caráter observacional dos estudos, e os elementos de confusão não avaliados adequadamente (ARBEX et al, 2004).

Já a abordagem *outdoor* tem gerado estudos de difícil condução, e, conforme coloca BRAUER (1999, p. 225), “... mesmo os melhores estudos irão prover resultados equivocados devido a variáveis de confusão e classificações errôneas de exposição.”

Um fator que contribui para a ocorrência de resultados equivocados foi descrito por SAMET e SPEIZER (1993) quando relatam que as investigações que associam poluição do ar e morbidade coletam rotineiramente dados de estatística vital da população de uma localidade alvo. Já a exposição ao poluente é estimada a partir dos dados de monitoramento do ar, e assume-se que são representativos para aquela população.

Complementarmente, é fato de que muitos estudos *outdoor* estão associados a episódios acentuados de poluição atmosférica, que ocorrem em áreas não isentas de outras fontes poluidoras, ou fatores geradores de um determinado efeito adverso.

BRAUER (1999) analisou diversos estudos relacionados aos episódios de queimadas no sudeste da Ásia e salienta ser provável uma associação causal entre os referidos episódios e o aumento da taxa de mortalidade diária. No entanto, ele pondera que estes estudos encontram-se relativamente em uma única categoria de exposição, na qual os efeitos agudos são altamente prováveis, mas os níveis de exposição são muito baixos, especialmente em termos de duração.

O mesmo autor finaliza seu raciocínio colocando ser razoável considerar nos estudos os impactos agudos ao invés dos crônicos, pois são mais fáceis de medir e representam maior grau de confiança.

Ainda sobre a classificação das doenças, GATRELL (2002) explica que as doenças crônicas são aquelas como as do coração, diabetes ou reumatismo, que têm duração prolongada, muitas vezes, por toda a vida, enquanto que as agudas são aquelas que acontecem sob condições abruptas e são resolvidas em alguns dias, como enfarte do miocárdio, apendicite ou acidentes vasculares. No entanto, exemplos como asma podem ser classificados em ambos os casos, pois determinados casos são tratados ao longo da vida do paciente, mas, em outros casos, pacientes são diagnosticados com um súbito ataque agudo.

Determinados desenhos de pesquisa necessitam identificar a população exposta e, para tanto, utilizam algum critério para classificação de exposição da população amostrada. Tal classificação pode ser dicotômica, por gradiente, ou níveis prováveis de exposição (RIBEIRO e WÜNSCH FILHO, 2004). É comum a ocorrência de viés gerado por erro de classificação da população em exposição, assim como, na identificação se o evento específico resultou na variabilidade da exposição.

Assim como é evidente a dificuldade de determinar a população exposta, existe ainda a complexidade em se obter o registro das medições dos poluentes aos quais a população esteve exposta.

Nesta linha, pesquisadores vêm desenvolvendo estudos que envolvem medições de poluentes em diversas escalas geográficas, objetivando associar episódios, com medição acima dos padrões considerados seguros, e impactos na saúde da população exposta.

As medições são planejadas e realizadas no local do estudo, ou adquiridas por estações de medição automática. Em ambos os casos, existem limitações. As medições em campo são realizadas em períodos de campanha planejados para o desenho do estudo, assim, não há continuidade de aquisição dos dados por períodos muito extensos devido

ao custo das análises, disponibilidade de equipamento, ou pessoal de operação. No caso da medição automática, existe a continuidade de leitura que permite amostras para análises de series temporais, porém, são poucas as unidades em funcionamento, limitando os estudos em muitas localidades desprovidas de dados desta natureza.

Os estudos sobre queimadas e efeitos adversos à saúde elaborados no Brasil, em sua maioria, estão associados às queimadas de cana-de-açúcar na região sudeste ou, mais recentemente, às queimadas na Amazônia.

Avanços significativos na área de sensoriamento remoto em conjunção com modelos computacionais estão permitindo a construção de modelos de simulação, como é o caso do CATT-BRAMS (*Coupled Aerosol and Tracer Transport to the Brazilian developments on the Regional Atmospheric Modelling System*), que foi desenvolvido inicialmente para estudar fluxo de superfícies e transporte atmosférico de emissões provenientes de queima de biomassa na América do Sul (FREITAS et al, 2006). Este modelo prognostica concentrações de CO e PM_{2,5} através de dados obtidos por satélite, que são inseridos no modelo acoplado de forma simultânea e consistente com as condições atmosféricas geradas pelo modelo BRAMS.

Nesta linha, o trabalho de IGNOTTI et al (2007) apresenta um estudo sobre indicadores de morbidade para definir ranking de municípios prioritários para ações de vigilância em saúde na região do arco do desmatamento no Brasil (ao norte de Mato Grosso) e sugere a aquisição dos dados de PM e O₃ por estimativas do CATT-BRAMS, ou por meio de medições locais.

As questões relativas ao arco do desmatamento vêm motivando novas pesquisas no campo da saúde ambiental. O trabalho de MASCARENHAS et al (2008), avaliou, por meio de um estudo ecológico de série temporal, as concentrações de PM_{2,5} em Rio Branco no estado do Acre em setembro de 2005 e comparou com atendimentos diários de emergência,

concluindo que houve correlação positiva entre o aumento de concentrações e o número de atendimentos de asma em crianças com menos de 10 anos.

No estado de São Paulo, as cidades de Araraquara e Piracicaba foram alvo de diversos estudos envolvendo medições de poluentes atmosféricos, queimadas e efeitos adversos à saúde humana (ZANCUL, 1998; ARBEX, 2000; CETESB, 2002; CANÇADO, 2006).

Araraquara representa uma situação típica, considerando a existência de indústrias e ausência de cidades de grande porte no seu entorno. A área rural é utilizada, em sua maioria, para o cultivo da cana-de-açúcar. A cidade de Piracicaba está inserida na região chamada “mar de cana”, repleta de plantações de cana-de-açúcar no seu entorno, é uma cidade de porte maior, que apresenta fontes poluidoras diferenciadas de Araraquara.

A área objeto de estudos encontrados na bibliografia pertinente ao tema tem sido desde pequenos povoados até extensas áreas, como um estado ou país. Alguns estudos foram focados em polígonos municipais, como em Araraquara ou Piracicaba. Outros trabalhos adotaram macro regiões ou o estado inteiro, como o de LOPES e RIBEIRO (2006), que analisou correlações entre áreas de plantio de cana-de-açúcar, queimadas e internações hospitalares para o estado de São Paulo e para a macro-região administrativa de Bauru, no mesmo estado.

Entre os diversos desenhos de estudos epidemiológicos, BRAUER (1999) elenca os mais comuns, no contexto das queimadas e efeitos adversos à saúde humana, conforme descritos a seguir:

- estudos de Coorte onde os dados estão disponíveis em nível individual, desenvolvidos em locais onde os dados de monitoramento do ar estão disponíveis ou onde as concentrações podem ser estimadas;
- estudos envolvendo grandes populações, onde a variação de exposição aos produtos da queima de biomassa pode ser particularmente utilizada. Se disponíveis, dados de saúde podem ser associados aos de poluição do ar;

- comparações transversais de incidência de doenças respiratórias, cardiovasculares e de mortalidade em áreas com diferentes exposições aos poluentes;
- estudos semi-individuais com membros de populações demograficamente homogêneas, onde eventos de doenças e variáveis de confusão podem ser avaliados e medidos. As exposições seriam analisadas segundo o perfil individual para depois serem comparadas às medições dos poluentes (exemplos com militares ou estudantes de universidades);
- estudos de caso-controle foram utilizados para estimar risco de doenças crônicas. No entanto a dificuldade de conduzir tais estudos inicia-se na seleção dos casos e dos controles. No caso de impactos da poluição do ar, a definição dos grupos poderia requerer situações de maior similaridade, como grupo social, tabagismo, idade, entre outros.

Estudos epidemiológicos transversais e caso controle demonstraram que a exposição à fumaça de biomassa em ambientes internos é o principal fator de risco de Doenças Pulmonares Crônicas entre mulheres, conduzindo ao *cor pulmonale*. Isto ocorre em países como China, Índia onde mais de 75% das pessoas cozinham utilizando combustível sólido. Em partes da América do Sul e África este percentual cai para 50-70%. Grosso modo, são estimadas 1,6 milhões de mortes prematuras no mundo devido a esta poluição interna (WHO, 2002).

Os estudos na modalidade *indoor* estão mais difundidos, com considerável volume de publicações. Vários trabalhos têm alertado para os efeitos negativos à saúde de queima de biomassa em ambientes internos, sobretudo nos Países em Desenvolvimento, nos quais ainda se usa muito fogão à lenha ou carvão sem exaustão adequada para a fumaça.

Nesta modalidade, autores descrevem desenhos metodológicos característicos de estudos epidemiológicos como Coorte, Caso-controle, estudos transversais e longitudinais, cujos dados, coletados por questionários ou fichas hospitalares, são submetidos ao instrumental da estatística descritiva.

Já a modalidade *outdoor* ainda é deficitária de estudos. O estudo de LOPES e RIBEIRO (2009) avaliou 46 trabalhos citados nas revisões de

desenvolvidas por RIBEIRO e ASSUNÇÃO (2002), ARBEX et al (2004) e RIBEIRO (2008) e selecionou 27 estudos que envolviam pesquisas com poluentes atmosféricos na modalidade *outdoor* e utilizaram ferramentas de geoprocessamento.

A análise dos delineamentos permitiu identificar e agrupar os tipos de variáveis que foram utilizadas nos trabalhos. Os desenhos metodológicos consideraram duas ou mais variáveis dos grupos descritos no Quadro 3.

Variáveis geográficas:	<ul style="list-style-type: none"> -polígono municipal; -polígono da macro-área administrativa; -polígono estadual; -hospitais próximos das áreas de impacto.
Variáveis de saúde:	<ul style="list-style-type: none"> -internações hospitalares selecionadas por diagnóstico Capítulo X da CID 10; -internações de causa cardiorespiratória; -internações causadas por determinadas doenças como diagnóstico principal; -procedimentos hospitalares, como inalações; -entrevistas / questionários (morbidade referida); -testes com função pulmonar e função espirométrica.
Outras	<ul style="list-style-type: none"> -medições de poluentes atmosféricos (índices diários / partículas depositadas)

Quadro 3 - Variáveis utilizadas nos desenhos metodológicos em estudos de poluição do ar, na modalidade *outdoor*.

Sobre as análises espaciais, observou-se carência de variação. Embora as ferramentas de geoprocessamento disponibilizem diversos recursos para análises desta natureza, seu uso não se reflete nos trabalhos analisados. Entre os trabalhos analisados, somente foram citados o uso de Análise *Cluster*, Índice I de Moran e Mapas Coropléticos. A Figura 6 apresenta a síntese dos delineamentos encontrados.

Figura 6 - Resumo dos delineamentos encontrados nas pesquisas sobre queima de biomassa e efeitos adversos à saúde humana.

Os trabalhos na modalidade *indoor* tendem a utilizar a abordagem da análise estatística com desenhos epidemiológicos conhecidos como estudos de Coorte, Caso-Controle e Comparações Transversais com os dados individualizados, fato este que difere dos estudos outdoor, onde os dados agregados são mais utilizados. Isso requer a generalização de algumas variáveis, que pode incorrer em resultados equivocados.

Este fato é advertido por autores em trabalhos nas duas modalidades, quando esclarecem que a intensidade e gravidade dos efeitos à saúde dependem de: características dos poluentes, características da população exposta, exposição individual, susceptibilidade do indivíduo exposto e fatores de confusão

Embora exista certa variabilidade nos estudos analisados, o uso de ferramentas de análise espacial ainda está em fase embrionária neste campo de conhecimento. Os poucos trabalhos existentes ainda não exploram todo o potencial que estas ferramentas disponibilizam ao pesquisador.

Ressalta-se ainda a necessidade de futuros estudos de orientação empírica, a luz de um aprofundamento maior sobre estas questões, na eminência de reduzir vieses e ampliar a confiabilidade dos desenhos de pesquisa que trabalham com estimativas e dados agregados.

3 CASUÍSTICA E MÉTODOS

3.1 Aspectos Gerais

Entre muitas definições, o termo método representa o caminho para chegar a um fim e, especialmente para chegar a um conhecimento científico (MICHAELIS, 2008). Do mesmo modo que esta afirmação é verídica, também é verdadeira a necessidade de definir, detalhar e seguir tais caminhos no sentido de preencher novas lacunas do conhecimento humano.

Descartes, em seu conhecido Discurso do Método, tenta estabelecer um método universal baseado nas regras de evidência, análise e síntese. Segundo ele, a razão é formalmente igual em todos, o que muda é a sua aplicação, pois depende de costumes, religião e conhecimentos adquiridos (DESCARTES, 2004).

No entanto, o conhecimento é mais amplo, e o ser humano precisa de diferentes caminhos que vão formando uma intrínseca rede de conexões, muitas vezes de caráter interdisciplinar, para efetivar as sinapses da ciência.

Este capítulo descreve os “caminhos” percorridos no presente estudo, bem como o detalhamento dos insumos utilizados para sua realização.

3.2 Método de Pesquisa

O método, visto como processo organizado, lógico e sistemático, sempre esteve presente no âmbito da experiência humana para aquisição de conhecimento.

A palavra método é originária do termo grego *méthodos* que, por sua vez, é a composição das palavras *metá* (através de) e *hodos* (caminho). Já o

termo “método científico” configura um caso particular dentre os diversos tipos de métodos, constituído de etapas como: identificação do fenômeno (observação), a explicação provisória (hipótese), procedimentos de testes da explicação (experimentação), e por fim, a análise e conclusão, visando aplicar aquela verdade em outros contextos (generalização) (APPOLINÁRIO, 2006).

Já o termo “metodologia” posiciona-se como o estudo e a avaliação dos métodos a fim de determinar aplicabilidade no processo da pesquisa científica (DIEHL e TATIM 2004).

Portanto, existem vários métodos que podem ser utilizados na pesquisa, cabe ao pesquisador exercer a prática da metodologia para definir o melhor “caminho” a seguir na busca do conhecimento.

Segundo CERVO, BERVIAN e Da SILVA (2007), o método científico depende da evidência como arrimo, pois, toda investigação nasce de algum problema observado ou sentido, que se condensa em uma hipótese ou pressuposição norteadora e, ao mesmo tempo, delimitadora do assunto a ser investigado.

A natureza humana é um fator motivador que ativa a curiosidade e o interesse em diversos aspectos e dimensões. No âmbito da ciência, foi necessário desenvolver diversas formas de abordar objetos de estudo, de modo a possibilitar o aprofundamento adequado para as diversidades.

Nesta linha, objetos devem ser delimitados com relação à finalidade do estudo com circunscrição de tempo e espaço. Esta delimitação implica em certo conhecimento do objeto de estudo, de modo que não inviabilize a sua observação (PEREIRA, 1999).

Assim, os estudos exploratórios marcam um passo inicial no processo de pesquisa através de descrições precisas sobre o objeto alvo. Este modelo metodológico permite a formação de um repositório de dados, ideal quando há pouco conhecimento sobre o problema abordado (CERVO, BERVIAN e Da SILVA, 2007).

DIEHL e TATIM (2004) acrescentam que os estudos exploratórios diferem dos descritivos, pois este outro preconiza a descrição das características do objeto do estudo, ou ainda, o estabelecimento de relações entre as variáveis da pesquisa.

As variáveis são propriedades do processo, fato ou fenômeno observado, cuja finalidade é determinar a forma pela qual os elementos do conjunto são diferenciados (ALMEIDA FILHO e ROUQUAYROL, 2002).

Nesta tese, optou-se por utilizar os termos “dependente” e “independente” para caracterizar as variáveis selecionadas. Assim, no conjunto observado, as internações hospitalares constituem a variável dependente deste estudo, enquanto que as queimadas, dados sobre área plantada, área do município e população, foram inseridas no modelo como variáveis independentes.

Retomando a linha metodológica, tanto os estudos exploratórios amparados pela metodologia científica, como o senso comum, enquanto saber informal, não são suficientes para explicar fenômenos e generalizá-los, mas agregam contribuições na formulação de hipóteses mais significativas.

A hipótese nula (H_0) é a inexistência da correlação entre as variáveis citadas. Desta forma, partiu-se do estabelecimento da hipótese alternativa (H_1) que afirma a existência de correlação positiva entre as variáveis: queimadas (VI – variável independente) e internações por problemas respiratórios (VD – variável dependente).

A hipótese apresentada no tópico 1.3 desta tese é a idéia central a ser demonstrada no trabalho, conforme afirma SEVERINO (2002). Tal idéia remete a uma questão não resolvida, que pode estar associada à sustentação de uma afirmação, ou à necessidade de por em prova uma suposição.

Segundo GRAYBILL, IYER e BURDICK (1998), a inferência é o cerne da questão em estatística e consiste no ato de tomar decisões racionais baseadas em evidências.

Nesta linha, é razoável ponderar que as decisões são tomadas com base em informações originadas de uma amostra da população. Tais informações são processadas no âmbito da estatística analítica, onde está o instrumental adequado para a produção de inferências sobre uma população com base em amostras.

No âmbito desta pesquisa, optou-se por utilizar a análise multivariada como ferramenta analítica. Considerou-se que o fenômeno em estudo pode ser sensível a diversos fatores, os quais são passíveis de serem expressos em variáveis numéricas para compor modelos de regressão.

Os modelos de regressão são úteis para a predição da VD em função de alterações das VI. Para tanto, são baseados em um tipo de distribuição estatística. No caso de variáveis baseadas em taxas, o mais comum é o uso da distribuição normal, pois permite adaptações das variáveis às suas características, denominadas transformações.

Com relação à prova, é sensato considerar métodos de análise das variáveis sob a luz da estatística, fato este que remete à pesquisa descritiva. Conforme CERVO, BERVIAN e Da SILVA (2007) definem, que neste tipo de estudo o pesquisador “observa, registra, analisa e correlaciona fatos ou fenômenos sem manipulá-los”, sendo desta forma, inversamente ao que ocorre na pesquisa experimental.

Vale ponderar que em certos casos, a pesquisa descritiva vai além da identificação das relações entre as variáveis. Na pretensão de determinar a natureza das relações, chega ao limiar da pesquisa explicativa. Da mesma forma, em outros casos, o desfecho dos objetivos proporciona uma nova visão do problema, aproximando mais da pesquisa exploratória (GIL, 1994).

Assim, observadas as sutilezas e imprecisões que envolvem uma classificação, optou-se por classificar esta pesquisa como descritiva, posto que esta, encontra-se relacionada com questões sociais que objetivam ser melhor explicadas através de classificações, correlações e comparações.

Por fim, a hipótese conduziu à coleta de dados sobre causa (queimadas) e efeito (internações hospitalares por problemas respiratórios),

assim como, subsidiou o delineamento da pesquisa, melhor descrito no próximo tópico desta tese.

3.3 Delineamento da Pesquisa

Segundo APOLINÁRIO (2006), o delineamento da pesquisa constitui o planejamento daquilo que se pretende realizar. Sendo assim, trata-se de formalizar o plano de ações a ser desenvolvido no trabalho.

Alguns atores classificam o delineamento segundo o tipo de pesquisa a ser realizada. O autor supracitado identifica quatro grupos de delineamento, dois associados a pesquisas descritivas e outros dois ligados a pesquisas experimentais, conforme descreve o esquema da Figura 7.

Figura 7 – Tipos de pesquisa e grupos de delineamentos.
Adaptado de APOLINÁRIO, 2006.

Assim, seguindo o constructo das pesquisas descritivas, este estudo foi delineado em etapas desde a aquisição e tratamento de dados até a construção do repositório SIG e testes da hipótese.

A abordagem metodológica conduziu no sentido de gerar um banco de dados geográficos, agregando dados ambientais, demográficos e de saúde, para as duas principais regiões brasileiras onde se desenvolve o plantio de cana-de-açúcar: Na região Nordeste, os estados de Alagoas,

Bahia, Paraíba, Pernambuco e Sergipe. E, na região Centro-Oeste/Sudeste (denominada Centro-Sul), os estados de Goiás, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraná, e São Paulo.

Optou-se por analisar estas regiões, pois apresentam características diferenciadas em relação ao clima e ao período de safra. Assim, buscou-se minimizar o viés que se refere ao período de seca e inverno, que coincide com a safra na região Centro-Sul.

Os procedimentos desenvolvidos neste trabalho seguiram por três fases distintas, que permitiram uma melhor organização das atividades executadas e resultados obtidos. São elas: Estruturação, Análise e Resultados. O diagrama da Figura 8 apresenta as fases, conexões e seqüenciamento.

Figura 8 - Diagrama esquemático do delineamento da pesquisa.

Os tópicos seguintes explicitam as atividades executadas no projeto, bem como os respectivos desdobramentos.

I - Revisão Bibliográfica

Esta fase consistiu no levantamento das contribuições científicas sobre os temas no contexto desta pesquisa: análise espacial de dados, poluição atmosférica, queimadas em cultura canavieira e problemas respiratórios. Observando a interseção destes assuntos, esta revisão reuniu material suficiente para construir o atual estágio do conhecimento humano sobre o uso do instrumental de geoprocessamento nos estudos de poluição *outdoor* e de saúde pública, descrito no Capítulo 1 desta tese.

Para efeitos de levantamento bibliográfico foram considerados os documentos publicados em periódicos indexados nas bases de dados Biblioteca Virtual em Saúde (BVS, 2009), Web of Knowledge (WEB OF KNOWLEDGE, 2009), Scielo (SCIELO, 2009), Medline (MEDLINE, 2009) e bancos de teses, com as palavras-chaves: *sugarcane*, *health*, *burnt*, *air pollution*, bem como, combinações entre elas e variações em português.

Foram utilizados também os termos “GIS/SIG” e “geoprocessamento” associados com o intuito de refinar a identificação trabalhos que utilizaram programas SIG. Os termos também foram testados na busca textual do Google (GOOGLE, 2009), capaz de localizar artigos indexados nas bases anteriormente citadas de forma mais ágil, porém com limitações no refinamento dos links selecionados. Este procedimento permitiu localizar artigos que não foram inicialmente localizados nas bases.

II - Definição das Variáveis

As variáveis coletadas nesta pesquisa são públicas e disponíveis para *download* na internet. Com relação à abrangência espacial, optou-se por coletar dados das regiões Nordeste e Centro-Sul, conforme mencionado no tópico anterior, acrescentando-se aqui a unidade geográfica municipal, como

concentrador de dados, devido à escala da pesquisa. Portanto, na região Centro-Sul, utilizaram-se 2363 polígonos municipais somados a mais 1002 da região Nordeste.

Com relação à variável dependente, optou-se por eleger os registros de internações hospitalares por doenças respiratórias descritas no Capítulo 10 da CID (Classificação Internacional de Doenças).

Tal escolha embasou-se na literatura, como por exemplo, os trabalhos relacionados e analisados em ARBEX (2001) e ARBEX et al (2004) que associam as internações por doenças respiratórias à questão da poluição do ar.

Vale aqui ressaltar que as doenças do Capítulo 10 não têm como único fator gerador a poluição do ar. Este fato amplifica a interferência de vieses ao modelo proposto. Outra questão inerente aos dados de saúde é a natureza secundária da fonte, com problemas de imprecisão do diagnóstico, conforme relatam VERAS e MARTINS (1994).

Desta forma, optou-se por selecionar três classes de doenças frequentemente citados nos estudos observados, de forma a minimizar vieses. São eles: asma, bronquite e enfisema.

Os registros de internações hospitalares da rede pública são coletados pelo sistema SIH-SUS (Sistema de Informações Hospitalares do SUS), criado para viabilizar o pagamento dos serviços hospitalares prestados pelo SUS, por meio da captação dos formulários AIH (Autorização de Internação Hospitalar).

Dos arquivos da AIH, foram utilizadas variáveis como: diagnóstico primário (CID-10), idade, escolaridade, código do município de residência, data da internação, óbito (Sim/Não). As agregações foram feitas considerando o município de residência.

Os dados da rede privada, coletados pela CIH (Comunicação de Internação Hospitalar) não foram utilizados nesta tese porque não estão disponíveis para a consulta pública, e a aquisição destes dados deve ser por

intermédio da Secretaria de estado da Saúde. Este fato gerou dificuldades na aquisição de dados para toda a área de estudo e, por este motivo, optou-se pela não utilização desta base de dados.

As variáveis independentes estão associadas ao processo de queimada de cana-de-açúcar. Portanto, optou-se por utilizar os dados sobre queimadas, disponibilizados pelo INPE (Instituto Nacional de Pesquisas Espaciais), no formato de pontos (coordenadas de localização de focos de calor, detectadas por sensoriamento remoto).

De forma complementar para a região Centro-Sul, também foram utilizados os dados sobre áreas de plantio de cana, disponibilizado pelo Projeto CANASAT do INPE (CANASAT, 2009).

Deve-se considerar que no período de inverno a incidência de doenças respiratórias aumenta no estado de São Paulo, devido às condições climáticas que a estação proporciona, como clima seco ou ocorrência de inversões térmicas bruscas. Este quadro interfere diretamente no aumento da procura por serviços de saúde, em função de problemas relacionados com afecções das vias respiratórias.

Em estudo realizado anteriormente, tal fenômeno foi identificado como variável de confusão, uma vez que as queimadas na região Sudeste ocorrem justamente no mesmo período (LOPES, 2005). Assim, com vistas a eliminar este viés, optou-se por ampliar as áreas de estudo, incluindo a região Nordeste, cujas condições climáticas e período de safra são diferentes.

III - Aquisição de Dados

Os dados utilizados para a construção do SIG são secundários, disponibilizados gratuitamente na Internet, por instituições que os geraram:

Dados sobre queimadas: INPE (Fonte: INPE, 2009)

- Período: 2000 a 2008 (diário);

Uso do solo: CANASAT / INPE (Fonte: CANASAT, 2009)

- Período: 2004 a 2008 (somente para a região Centro-Sul);

Dados da AIH – Autorização de Internações Hospitalares: DATASUS

- Período: 2000 a 2008 (diário por município);

Dados censitários: IBGE (Fonte: IBGE, 2009)

- Ano: 2000 e estimativas por município – 2001 a 2008;

Malha municipal digital do Brasil - 2007: IBGE (Fonte: IBGE, 2007)

- Escala de 1: 2.500.000 / Ano de referência: 2007;

Dados climáticos: IBAMA (Fonte: IBAMA, 2008)

- Médias anuais para 2008;

Áreas urbanizadas: EMBRAPA (Fonte: EMBRAPA, 2009b)

- Dados censitários de 2000 e imagens de satélite de 2000 e 2001;

Índice de Desenvolvimento Humano – IDH (Fonte: PNUD, 2009)

- Ano: 2000.

IV - Preparação dos Dados

Esta atividade consistiu na modelagem e construção do SIG, ou seja, o banco de dados geográficos que foi projetado e construído para agregar as informações adquiridas para este estudo.

Dada a diversidade de fontes de dados disponíveis em formatos variados, assim como o considerável volume do banco de dados o trabalho de preparação consolidou os dados tabulares em um repositório único, no sistema gerenciador de banco de dados MS-SQL Server 2005 (MICROSOFT, 2009).

Já os dados geográficos foram inseridos no geodatabase da Ferramenta ArcView 9.3, que consiste em um repositório de armazenamento de dados tabulares e mapas, para uso em geoprocessamento (ARCGIS, 2009). Posteriormente, os dados armazenados no MS-SQL Server 2005 foram conectados ao ArcView 9.3, via ODBC (*Object DataBase Connection*) para integrar os dados tabulares aos geográficos.

O produto final desta atividade foi o geodatabase, concluído para as análises estatísticas e espaciais.

V - Análise Espacial:

Esta fase foi desenvolvida integralmente no ArcView 9.3, onde estruturou-se um geodatabase com as variáveis envolvidas, distribuídas espacialmente nos polígonos municipais e polígonos de micro-região.

O modelo de variação espacial discreta parte da hipótese nula (H_0) na qual cada área apresenta uma distribuição de probabilidade distinta das demais, assim, a hipótese alternativa (H_1) infere a continuidade espacial.

Uma forma de analisar a variação espacial e suas possíveis dependências é através da identificação de agrupamentos, baseados em uma variável que pondera o peso desta variação através da distância entre os eventos.

Nesta linha, optou-se em utilizar a análise *Hot spot*, disponível no Arcview 9.3. Este recurso parte da variável estudada e calcula o Gi estatístico para organizar os agrupamentos. O índice Gi é similar ao Z score da estatística e representa a significância estatística do agrupamento para uma distância especificada, ou seja, é um teste de significância estatística, com base na distribuição normal, que considera uma matriz de pesos baseada nos valores atribuídos aos vizinhos. A soma de um polígono e seus vizinhos é comparada proporcionalmente à soma da área toda. Se um polígono tem G alto e seus vizinhos também, então, esta área é um *Hot Spot*. O resultado demarca áreas no mapa com valores altos ou baixos que tendem a formar agrupamentos (ARCGIS, 2009).

O índice Gi é categorizado por níveis de desvio padrão, baseado no valor de $p = 0,05$. Assim, os índices são estatisticamente significativos quando apresentam valores com desvio padrão acima ou abaixo de 1,96 (hipótese bicaudal), os demais não apresentam significância estatística para a hipótese da continuidade espacial (H_1).

O cálculo da distância foi baseado no centróide do polígono e optou-se pela distância euclidiana para a análise. O método conceitual utilizado para os relacionamentos espaciais o denominado “inverso da distância”

(*Inverse Distance Method*), onde o impacto de uma feição sobre outra aumenta com a proximidade.

Por fim, as análises foram feitas distintamente para cada estado integrante do estudo. Esta abordagem foi escolhida, pois o cálculo do Gi leva em consideração os índices de todos os elementos selecionados para determinar os agrupamentos, fato este que gera viés no resultado final.

Isto ocorre porque existem variações elevadas entre contagens de queimadas nos estados. Estados com contagens altas apresentam os agrupamentos altos, eliminando os agrupamentos que seriam altos em outros estados.

Desta forma, optou-se por agrupar os municípios por estado para efetuar as análises e, posteriormente agrupá-los em um mosaico de resultados.

VI - Mapas Temáticos:

O índice “Gi” foi apresentado nos mapas temáticos categorizados por desvio-padrão. Tal categorização atribui valores acima e abaixo da média, a partir da hipótese da distribuição normal da variável utilizada (DRUCK et al, 2004).

Os mapas foram desenvolvidos na ferramenta ArcView 9.3 e os dados analisados foram agregados nos polígonos municipais ou micro-região, a partir da malha municipal digital do IBGE. Considerou-se que as primitivas gráficas em uma representação cartográfica, pode ser feita tornando linhas, pontos e áreas mais ou menos perceptíveis conforme identificou o criador da semiologia gráfica, Jaques Bertin (LOCH, 2006).

No ambiente SIG estruturado, os mapas coropléticos com variações de cores para a categorização em feições poligonais se mostrou mais adequado para a apresentação dos resultados. Optou-se também por apresentar recortes da região Centro-Sul e Nordeste isoladamente.

VII - Análise Estatística:

Esta atividade foi prioritariamente desenvolvida na ferramenta Stata, versão 10. Os *datasets* foram previamente preparados no MS-SQL Server 2005 e acessados pelo Stata 10, via ODBC.

Desenvolveram-se então, análises em duas linhas. A primeira, focada na verificação da qualidade da base de dados, consistindo em sumarizações e testes de qualidade para o conjunto de dados. Após esta etapa, os modelos estatísticos foram executados sobre o *dataset* validado.

Quando VD e VI são contínuas é possível aplicar o teste de correlação de Pearson para verificar força e direção de associação entre ambas. Trata-se de um teste simples cujo resultado gera um valor entre -1 e 1 que pode ser interpretado da seguinte forma:

- 0,7 ou mais – forte correlação;
- 0,3 a 0,69 – correlação moderada;
- 0 a 0,29 – correlação fraca;
- Valores positivos: correlações positivas;

Valores negativos: correlações negativas.

Portanto, aplicou-se o teste de correlação de Pearson para as duas regiões. Em seguida, utilizou-se a Regressão Linear para o estudo transversal. Foram criados vários modelos a partir das combinações de agregações das dimensões espacial (município, micro-região, meso-região) e temporal (dia, mês, trimestre, ano) para os 9 anos do estudo.

O teste de Shapiro-Wilk foi aplicado para verificar a normalidade das variáveis. Este teste baseia-se na H_0 de população normalmente distribuída e gera a variável Prob>z com valores entre 0 e 1. Valores inferiores a α (significância estatística – 95%) indicam que a H_0 deve ser rejeitada.

VIII - Análise dos Resultados / Conclusão

Finalização das etapas anteriores.

3.4 Definição dos Sujeitos da Pesquisa

O termo “sujeito de pesquisa” pode ser definido como sendo o ente objeto da investigação. Esta unidade observacional pode ser: um ser humano, uma empresa, uma área de plantio, ou agrupamentos de cada exemplo anterior (APOLINARIO, 2006).

No âmbito da epidemiologia, trata-se das relações de ocorrência de saúde-doença em massa, envolvendo agregações de seres humanos em coletividades classificadas de diversas formas. Estas relações são analisadas à luz do conceito do risco (ALMEIDA FILHO, ROUQUAYROL, 2002).

Assim, optou-se por considerar sujeitos da pesquisa, as coletividades que vivem em municípios localizados próximos, ou onde ocorrem queimadas na pré-colheita de cana-de-açúcar, dentro do território nacional.

Para a variável população, utilizou-se a estimativa populacional do IBGE (IBGE, 2009), no período de 2000 a 2008.

Devido às características dos problemas de saúde relacionados à poluição atmosférica, optou-se, também, por analisar isoladamente as internações ocorridas com crianças (≤ 4 anos) e idosos (≥ 65 anos). Estas faixas etárias são relatadas em vários estudos como as mais susceptíveis às doenças respiratórias. Da mesma forma, a Proporção de internações hospitalares por grupos de causas do IDB (Indicadores de Dados Básicos), disponível no Tabnet (DATASUS, 2009) mostra as internações por doenças respiratórias com números mais altos para a faixa etária ≤ 4 anos e, em

segundo (menor apenas que as internações por doenças do aparelho circulatório) para idosos (> 60 anos), no período de 2000 a 2007.

3.5 Descrição das Bases de Dados e Adequações

Por tratar-se de um estudo epidemiológico, que pretende explicar as causas da incidência de agravos em grupos populacionais e comparar diferentes populações, utilizaram-se dados secundários de internação hospitalar agregados por município. No entanto, as agregações foram geradas a partir de dados individuais, disponibilizados pelo DATASUS.

Esses dados estão disponíveis no site do DATASUS (DATASUS, 2009), na área MS-BBS, Janela: Informações de Saúde - Epidemiológicas e Morbidade. Aqui, selecionou-se a opção Morbidade Hospitalar, Geral, por local de residência – de 1995 a 2007. Em seguida, selecionou-se a UF do estado para ter acesso aos arquivos de AIH.

Optou-se por adquirir os arquivos denominados “Reduzidas” ou RD. Cada arquivo RD contém dados sobre internações hospitalares de um estado da federação, por mês de competência. A descrição da estrutura deste arquivo está descrita no Apêndice A deste trabalho.

Esses dados são armazenados em arquivos no formato DBF (dBase *file*) e compactados no formato DBC (formato de compressão desenvolvido pelo DATASUS para arquivos DBF). Os arquivos contêm internações por mês de referência de pagamento.

O Apêndice A apresenta a estrutura dos arquivos RD disponíveis para *download* e a lista dos campos selecionados, com as respectivas descrições.

Cada arquivo RD está disponível em um formato comprimido (extensão DBC). Antes de manipulá-los é necessário descompactá-los,

utilizando a ferramenta TABWIN, disponibilizada gratuitamente pelo DATASUS.

Optou-se por selecionar as internações cujo diagnóstico principal foi identificado com um dos códigos, segundo classificação CID-10:

J20 – Bronquite Aguda (incluídas as sub-categorias)

J40 – Bronquite não especificada como aguda ou crônica

J41 – Bronquite crônica simples e a mucopurulenta (incluídas as sub-categorias)

J42 – Bronquite crônica não especificada

J43 – Enfisema (incluídas as sub-categorias)

J45 – Asma (incluídas as sub-categorias)

As bronquites foram somadas em conjunto. Assim, foram calculadas as internações diárias para 3 classes de doenças, sub-divididas nas faixas etárias:

- 0 a 4 anos;

- 65 anos ou mais.

Os dados foram convertidos para o formato MS-Access e posteriormente para o MS-SQLServer 2005. Esta sequência foi necessária, pois existem limitações na conversão direta do formato DBF para o SQLServer 2005.

Nesta etapa, utilizou-se um programa desenvolvido para facilitar o processo, pois foram processados dados de 9 anos, 12 meses e 11 estados, totalizando 1188 arquivos.

O banco de dados final soma 4 Gb de espaço físico, no formato SQLServer 2005. As estruturas das tabelas geradas estão descritas no Apêndice B desta tese.

Os focos de calor adquiridos de forma gratuita por *download* no site do INPE já estão no formato *shapefile* e puderam ser inseridos no SIG diretamente, bastando apenas fazer a alteração do sistema de coordenadas

geográficas para SIRGAS 2000, para manter compatibilidade com a malha municipal, que já estava neste sistema.

Foram geradas nove feições de ponto com focos de calor segmentados por ano. A estas tabelas acrescentou-se o código do município para possibilitar relacionamentos com as demais tabelas do repositório. Aqui vale ressaltar que a tabela original possui o nome do município. No entanto, fazer uma junção com a tabela de municípios somente pelo nome é uma tarefa que necessita de certos ajustes, pois os nomes dos municípios nem sempre são iguais nas duas tabelas. Diferenças de acentuação na grafia dos nomes foram encontradas nas bases como, por exemplo, o município de Chavantes que antes era escrito com “X”.

Desta forma, optou-se por gerar os códigos de município através de junções espaciais, ou seja, atribuiu-se o código do município cujo foco de calor estava localizado dentro do respectivo polígono. Com os códigos de municípios adicionados, foi possível contar focos por município e, posteriormente, contar por micro-região administrativa. Estes procedimentos foram executados no ArcView 9.3.

Assim, optou-se por contabilizar focos de calor ocorridos no entorno da sede municipal, inferindo que aqueles inseridos em um dado raio, seriam impactantes para a população do referido município.

Para tanto, foram elaboradas duas feições geradas a partir de *buffers* de raio igual a 10km e 20km, medidos a partir da sede do município. Optou-se por utilizar as duas distâncias para avaliar diferenças entre ambas. Desta contagem, calculou-se a intensidade de focos de calor, utilizando como denominador a extensão da área não urbanizada do município conforme descrito na fórmula:

$$\text{Intensidade de focos} = \frac{\text{Contagem de focos no } \textit{buffer}}{\text{Área urbanizada do município}}$$

Por fim, na composição dos dados por micro-região, utilizou-se a soma das contagens dos municípios e calculou-se a intensidade de focos utilizando como denominador, a somatória de áreas não urbanizadas. Os valores gerados deste procedimento foram utilizados posteriormente na análise multivariada.

Os dados coletados foram organizados em um repositório SIG cujo esquema de tabelas está representado na Figura 9. Sua composição foi organizada segundo critérios estabelecidos para a modelagem de dados tipo *snowflake*, que prevê a possibilidade de hierarquizar as dimensões (ELMASRI e NAVATHE, 2005), que torna a modelagem mais adequada para este estudo.

No modelo final, duas tabelas denominadas “fato” foram criadas e populadas com internações hospitalares e focos de calor. Estas tabelas estão associadas a três grupos de dados. Dois deles representam as dimensões “espacial” e “temporal”, respectivamente.

Como as tabelas “fato” possuem registros diários, a dimensão temporal foi modelada com agregações para mês, trimestre e ano, bem como, segmentações por períodos de safra/não safra, nas regiões de estudo (CS e NE).

Figura 9 - Diagrama esquemático das fontes de dados utilizadas no estudo e seus relacionamentos no repositório de dados.

Para não haver lacunas entre os dias, que poderiam ocorrer em virtude da ausência dos eventos estudados, optou-se por criar uma tabela contendo os dias do período considerado no estudo. A esta, foram acrescentados campos para identificar dia da semana (numerados de 1 a 7) e estação do ano (1-Verão, 2-Outono, 3-Inverno e 4-Primavera).

Complementarmente, foram inseridos nesta tabela, dois campos dicotômicos para identificação do período de safra (sim/não), um para cada região (CS e NE).

Para a dimensão espacial, considerou-se o polígono municipal como a menor unidade de informação geográfica. Em seguida, com base nos identificadores de região administrativa, foi possível criar as agregações poligonais para micro e meso-regiões, bem como agregações por estado da federação e área de estudo (CS e NE).

Os dados tabulares foram vinculados ao polígono municipal e, para as agregações, foram feitas sumarizações (somatórias ou médias, dependendo do caso). O Quadro 4 sintetiza os montantes de objetos considerados para as duas dimensões.

Dimensão Espacial				
Regiões	Municípios	Micro-Regiões	Meso-Regiões	Estados
CS	2363	220	52	7
NE	1002	100	22	5
Total	3365	320	74	12
Dimensão Temporal				
Período	Dias	Meses	Trimestres	
2000 – 2008	3288	108	36	

Quadro 4 - Número de variáveis por dimensões utilizadas no repositório de dados.

No terceiro grupo de tabelas, foram colecionadas as co-variáveis utilizadas na análise multivariada, assim como os dados de estimativa populacional para os anos estudados e as bases não tabulares que serviram de base para a construção dos mapas e análise espaciais. No caso, foram

utilizados os polígonos da malha municipal brasileira – IBGE de 2007 e os pontos das sedes municipais, disponível no DATASUS (DATASUS, 2009).

Para as análises, optou-se por estudar os dados das duas dimensões espaciais: por município e por micro-região. Em cada uma, os dados foram agregados temporalmente por trimestre e por mês. Assim, todas as análises foram feitas em quatro recortes para a região CS e mais quatro para a região NE.

Observou-se, no entanto, que os recortes com maior nível de agregação apresentavam melhores resultados. Isso era esperado, pois, a redução de escala tende a homogeneizar os dados, bem como, reduzir a flutuação aleatória reforçando as correlações (DRUCK et al, 2004). Desta forma, os resultados apresentados capítulo 4 correspondem aos agregados por micro-região e trimestre, mas as análises foram realizadas de forma mais ampla, considerando os demais agregados, que geraram resultados inferiores aos aqui citados.

3.6 Ferramentas Utilizadas

Considerando o conjunto de variáveis envolvidas, faz-se necessária a utilização de ferramentas adequadas para o tratamento e análise do banco de dados.

Segundo CROMLEY e McLAFFERTY (2002), ferramentas de SIG vêm sendo utilizadas em estudos de Saúde Pública, entre outras aplicações, para modelar onde as pessoas vivem e suas experiências no ambiente ao longo de suas vidas. No entanto, esta é uma tarefa complexa, pois a saúde não se trata somente da doença, mas do estado físico, social e bem-estar emocional.

Assim, ferramentas SIG são úteis em estudos desta natureza, pois são capazes de agregar um conjunto considerável de variáveis e associá-las

ao espaço geográfico, proporcionando a exploração de novas conexões e dinâmicas entre as variáveis coletadas.

No entanto, durante o processo de aquisição de dados, foram efetuados procedimentos como: a adequação das bases de dados para um mesmo padrão, descompactação de arquivos, limpeza de dados inválidos, padronização de valores como datas no formato dd/mm/aaaa, ajustes de tipo de dado (ex. mudança de *float* para *integer* ou redução do tamanho de campos tipo texto, criação de campos adicionais para facilitar consultas, etc.).

Esses procedimentos demandaram o uso de outras ferramentas de apoio, que foram muito úteis nesta fase da preparação do repositório de dados.

Outros procedimentos específicos de SIG como: mudança de sistema de coordenadas, construção de novas feições com operações espaciais, junção de tabelas de dados com feições, sumarizações, construção de *buffers*, layout de mapas, etc. foram executadas no ArcView 9.3.

Para os gráficos de evolução temporal, utilizou-se a visualização da curva baseada no cálculo de médias móveis. Esta opção foi utilizada para filtrar as oscilações e suavizar a representação da tendência. O parâmetro de ajuste da janela do tempo foi explicitado como observação nos gráficos.

Assim, um conjunto de ferramentas computacionais foi utilizado neste trabalho, de forma integrada ou complementar, conforme as necessidades descritas no delineamento do projeto.

O Quadro 5 sintetiza o uso dos recursos computacionais e suas contribuições.

Ferramenta	Fase	Aplicação
TabWin	Aquisição de dados	Descompressão de arquivos RD (Internações Hospitalares – DATASUS). Originalmente estes arquivos são disponibilizados para <i>download</i> no formato DBC (<i>Database compressed</i>).
MS-Access 2007	Aquisição de dados	Transformação dos arquivos RD, do formato .DBF (dbase File) para MDB (Access na versão 2000).
MS-SQL Server 2005	Aquisição de dados	Transformação dos arquivos RD, do formato MDB para o SQL Server 2005 e unificação das tabelas mensais em apenas uma, contendo os nove anos da abrangência do estudo. Foram executados ajustes nas colunas para permitir a união das tabelas.
MS-SQL Server 2005	Aquisição de dados	Limpeza dos dados: eliminação dos registros duplicados ou inválidos. Eliminou-se também os registros pertencentes a pessoas fora das faixas etárias, assim como as internações cuja CID não pertencia ao capítulo J, conforme tabela de códigos considerados (vide Apêndice B).
ArcView 9.3	Aquisição de dados	Recepção dos dados sobre queimadas do INPE, em formato SHP (<i>Shapefile</i>). Transformação dos dados para geodatabase (formato de banco de dados geográfico).
Stata 10	Análise estatística	Validação dos conjuntos de dados, sumarizações, análise descritiva e análise multivariada.
ArcView 9.3	Análise espacial	Análise <i>hotspots</i> , construção de <i>buffers</i> , mapas coropléticos, finalização de mapas.
MS-Excel	Caracterização dos sujeitos	Elaboração de gráficos, médias móveis e tabelas, resultantes das análises.

Quadro 5 – Síntese das ferramentas utilizadas na pesquisa.

3.7 Limitações da Pesquisa

Este trabalho possui uma abordagem interdisciplinar, visto que o campo de Saúde Pública, no qual se encaixa o objeto desta pesquisa, não é passível de ser englobado ou reduzido a uma única disciplina (BRUYNE, 1982 citado por ALVARENGA, 1994). Para tanto, dada a complexidade do problema, este estudo foi encaminhado agregando conhecimentos nas áreas:

- Geografia da Saúde e Epidemiologia;
- Poluição Atmosférica;
- Análise Espacial e Geoprocessamento.

Considerando as características do estudo, bem como as variáveis envolvidas, faz-se necessário ressaltar os aspectos limítrofes da pesquisa.

Inicialmente, é importante esclarecer que a pesquisa apresenta recorte transversal e longitudinal, ou seja, os resultados comparativos refletem a evolução dos eventos no período de nove anos. Dada a dinâmica dos fenômenos, os resultados podem variar na continuidade, em decorrência de fatores econômicos ou climáticos.

Complementarmente, fatores como condições climáticas e padrão sócio-econômico do local de estudo, devem ser considerados em pesquisas que envolvem comparações entre localidades.

No entanto, sob o ponto de vista da climatologia, o estudo considerou apenas a sazonalidade dos eventos, posto que o período da safra é diferente nas duas regiões de estudo, assim como as condições climáticas, que são características no Centro-Sul e no Nordeste, anteriormente região de mata atlântica.

Na escala do estudo, a base de dados climatológicos por município é inexistente em muitas localidades. Este fato implicou em algumas generalizações que pode ter comprometido as análises estatísticas.

Os aspectos de relevo também não foram considerados, pois demandariam outras análises que estão além do escopo deste trabalho. A pesquisa considerou a proximidade das populações às áreas de queimada. Questões como relevo, direção dos ventos e dispersão da pluma de poluentes, devem ser tratados em estudos empíricos, de escala local.

Outra questão diz respeito aos eventos de saúde considerados na pesquisa. Trata-se do efeito agudo que as queimadas podem causar à saúde das pessoas. Neste aspecto, a opção por utilizar os dados de internações é plausível, pois, a variabilidade da contagem de internações pode ser sensível aos eventos de queimadas em uma dada localidade.

Por outro lado, eventos agudos de internações por problemas respiratórios são menos frequentes e podem estar associados a problemas

crônicos. Sabe-se que o complexo bronquite-enfisema pode estar associado a desencadeantes como a exposição ocupacional, insuficiência de auto-imunidade, infecções ou pré-disposição genética, além dos fatores climáticos e poluentes atmosféricos (RATTO, SANTOS e BOGOSSIAN, 1981).

Por fim, o delineamento conduzido, a partir de variáveis agregadas em polígonos municipais, minimiza a possibilidade de generalização dos resultados da pesquisa para outras localidades. No entanto, a reprodução do método proposto é recomendável para efeitos comparativos.

3.8 Aspectos Éticos

Está assegurado aos sujeitos, o uso dos dados e produtos obtidos exclusivamente no projeto, vedando-se qualquer outro fim, que não os objetivos propostos, que visam o benefício da comunidade.

Somente foram usados dados agregados, não permitindo a identificação de sujeitos e resguardando a confidencialidade dos dados individuais dos pacientes.

4 TRATAMENTO DOS DADOS, RESULTADOS E DISCUSSÃO

4.1 Aspectos Gerais

Objetivando a consonância com a metodologia proposta, texto a seguir busca apresentar, de forma sintética, os resultados obtidos, abordando três enfoques:

- Análises Descritivas: exploram a caracterização dos dados adquiridos;
- Análises Espaciais: apresentam mapas e aspectos espaciais sobre o fenômeno estudado e colaboram para a fomentação de hipóteses, dada a sua natureza exploratória;
- Análises Estatísticas: buscam explicações através de correlações entre as variáveis selecionadas.

4.2 Análises Descritivas

4.2.1 Área do Estudo

Para uma melhor compreensão sobre os dados coletados, apresentam-se informações sobre a área de estudo. A Tabela 1 descreve alguns aspectos sobre os estados da federação inseridos nesta pesquisa. Tais estados foram selecionados por apresentarem áreas de plantio de cana-de-açúcar.

O Distrito Federal, embora esteja no estado de Goiás, não foi considerado na análise da região Centro-Sul, devido ao fato de não apresentar área de plantio de cana-de-açúcar.

Tabela 1 - Dimensão e população da área de estudo. Fonte: IBGE, 2009.

Região	Unidade da Federação	Total de Municípios	Área (Km ²)	População Estimada (2008)	Densidade Demográfica
Centro-Sul	Goiás	246	340.086,69	5.844.996	17,18
	Mato-Grosso	141	903.357,90	2.957.732	3,27
	M. Grosso do Sul	78	357.124,96	2.336.058	6,54
	Minas Gerais	853	586.528,29	19.850.072	33,84
	Paraná	399	199.314,85	10.590.169	53,13
	São Paulo	645	248.209,42	41.011.635	165,23
	Distrito Federal	1	5.801,93	2.557.158	440,74
<i>Sub-Total</i>		2.363	2.640.424,04	82.147.820	32,25
Nordeste	Alagoas	102	27.767,66	3.127.557	112,63
	Bahia	417	564.692,66	14.502.575	25,68
	Paraíba	223	56.439,83	3.742.606	66,31
	Pernambuco	185	98.311,61	8.734.194	88,84
	Sergipe	75	21.910,34	1.999.374	91,25
<i>Sub-Total</i>		1.002	769.122,10	32.106.306	41,74
<i>Total</i>		3.365	3.409.546,14	117.254.126	34,39

Os dados da Tabela 1 mostram que a região Centro-Sul é 3,4 vezes maior que a Nordeste em extensão e 2,6 vezes maior em população. No entanto, a média da densidade demográfica é maior na região Nordeste. Sabe-se ainda que esta distribuição populacional é irregular, com maior concentração na região litorânea e menores proporções no interior. É possível observar este fenômeno por meio das taxas calculadas para Mato-Grosso e Mato Grosso do Sul, que não estão na costa brasileira e apresentam as menor valor para este indicador.

O Mapa 2 apresenta a localização geográfica das regiões do estudo. Observa-se contigüidade entre ambas, dado que o estado de Minas-Gerais é vizinho ao estado da Bahia. Porém, do ponto de vista das áreas de plantio, esta contigüidade não ocorre, pois as áreas de plantio da região Nordeste encontram-se próximas ao litoral, enquanto que, no caso de Minas Gerais, existem alguns pontos distribuídos no estado, mas a concentração maior é

próxima ao estado de São Paulo e ao estado de Goiás, onde se dá a expansão da cana.

Mapa 2 - Área de estudo e localização das áreas de plantio de cana-de-açúcar.

4.2.2 Focos de Calor

O Gráfico 2 apresenta a evolução quantitativa dos focos de calor, nas regiões de estudo. Estas curvas de tendência mostram dois pontos importantes. Primeiramente, a série mostra que o número de focos aumenta ao longo dos 9 anos, como é possível observar pela linha de tendência polinomial. O segundo ponto está associado à sazonalidade dos picos de eventos que se mostram similares nas duas regiões.

A leitura da linha polinomial informa que as queimadas não diminuíram neste período. Embora seja perceptível a queda de valores a partir de 2005, existe uma retomada em 2007. Em parte, estas variações podem ser explicadas por mudanças de metodologia e substituição de sensores orbitais que ocorreram neste período, esclarecidas em INPE (2009).

Gráfico 2 - Evolução dos focos de calor nas regiões do estudo por média móvel e curva polinomial, no período de 2000 a 2008.

* As médias móveis foram calculadas com variável de ajuste $N = 50$.

Com vistas à sazonalidade por região, esperava-se observar picos em épocas distintas, compatíveis com os períodos de safra em ambas as regiões. Observa-se que as curvas possuem padrão similar e os picos para a região NE ocorrem com certa defasagem em relação aos da região CS, provavelmente em função do período de safra no NE.

Desta forma, é plausível inferir que a safra de cana-de-açúcar na região NE contribui com os eventos acentuados de queimadas, sendo que a safra de cana-de-açúcar tradicionalmente ocorre entre dezembro e março naquela região. Porém, observa-se que o período de safra pode estar iniciando antes e alcançando os picos nos meses de novembro e dezembro.

Ressalta-se ainda que a análise dos mapas de queimadas na região NE permitiu observar que há ocorrência de focos de calor em áreas diferentes daquelas onde há cultura canavieira, o que permite concluir que as queimadas não são exclusivamente geradas pela queima de cana-de-açúcar.

Na região Centro-Sul, a curva começa a subir no mês de abril e perde intensidade no final do ano, coincidindo com a safra de cana desta região. Os picos ocorrem na época mais intensa da safra, ou seja, entre julho e novembro.

Neste ponto, vale acrescentar que outras questões estão relacionadas à coleta e processamento dos dados de queimadas.

Um foco indica a existência de fogo em um *pixel* que pode variar de 1km por 1km até 5km por 4km, fazendo que possa haver várias queimadas indicadas em apenas um foco. Os trabalhos de validação apresentam erro de localização na ordem de 400m, com desvio padrão próximo de 3km (INPE, 2009).

Portanto, existem limitações no uso desta base de dados. O princípio de detecção do foco é baseado na emissão de energia. Os sensores do satélite captam energia emitida na faixa termal-média de 3,7 μm a 4,1 μm do espectro óptico. O procedimento de detecção seleciona as imagens

coletadas por sensores nesta faixa e selecionam os *pixels* que apresentam saturação do sensor, ou seja, onde se detectou mais calor (INPE, 2009).

Ainda, INPE (2009) lista os seguintes itens como fatores que prejudicam a detecção de queimadas:

- Frentes de fogo com menos de 30 metros;
- Fogo apenas no chão de uma floresta densa, sem afetar a copa das árvores;
- Nuvens cobrindo a região (exceto nuvens de fumaça);
- Queimada de pequena duração, ocorrendo entre as imagens disponíveis;
- Fogo em uma encosta de montanha, enquanto o satélite só observou o outro lado;
- Imprecisão na localização do foco de queima, que, no melhor caso, é de cerca de 1 km, mas podendo chegar a 6 km;

Acrescenta-se ainda que a queimada é uma prática tropical utilizada em várias culturas agrícolas e para renovação de área de pastagem. Em outras áreas, como o cerrado, ela ocorre naturalmente e possibilita a quebra de dormência de determinadas espécies vegetais.

A não existência de um mecanismo de validação dos dados gerados, observações *in loco* e relatos de terceiros sobre queimadas ocorridas que não constavam na base de dados, fomentaram dúvidas relacionadas à viabilidade de usar os dados no contexto deste estudo.

Um estudo paralelo comparou focos de queimada disponibilizados pelo INPE e as comunicações de queimadas no estado de São Paulo na perspectiva de encontrar similaridades entre os dados. No entanto, os resultados para a amostra de um mês mostraram significativas diferenças entre estes bancos de dados. A mediana das amostras aponta que os focos de queimada correspondem a 10,9% das comunicações sobre queimadas no mesmo período e região, assim como o coeficiente de correlação calculado para a amostra apontou correlação negativa moderada com $r = -0,62$ (LOPES, 2009).

O Gráfico 3 apresenta o número de eventos de focos de calor e comunicações de queimadas no período de um mês, no qual observa-se a não similaridade das curvas.

Existem questões a serem ponderadas. Como já fora dito, um foco de queimada pode significar um conjunto de talhões. O proprietário é obrigado a informar talhão por talhão que será queimado. Isso faz com que a base de comunicações seja mais precisa que a do INPE.

Outra questão está associada ao horário das queimadas. O decreto 47.700 de 2003 prevê horários para queimada, que deve ocorrer no período noturno, horário que proporciona melhores condições para dispersão dos poluentes, além de coincidir com os horários de passagem dos sensores orbitais no estado de São Paulo. No entanto, várias queimadas são testemunhadas durante o dia, minimizando a captação dos dados pelos satélites.

Gráfico 3: Comparação entre número de comunicações de queimadas e focos de calor para o estado de São Paulo, em setembro de 2008. Fonte: Lopes, 2009.

Ainda relacionado às comunicações, estas, muitas vezes, são canceladas pelos solicitantes devido a fatores climáticos como muita chuva ou ventos fortes. A SMA/SP também cancela a solicitação quando a umidade relativa do ar está abaixo de 20% no município onde ocorreria a queima.

Embora existam fatores limitantes, é importante salientar os esforços do INPE no sentido de melhorar a capacidade de detecção e a precisão da localização destes focos, diretamente associado à qualidade das imagens coletas pelos sensores orbitais e modelos computacionais para tratamento das mesmas.

Aferidos estes resultados tentou-se utilizar as comunicações de queimadas ao invés dos focos de calor do INPE. Após sucessivas tentativas de obtenção dos dados e sua não disponibilização pela Secretaria do Meio Ambiente e por força do prazo para conclusão deste trabalho, a opção foi seguir este estudo usando os focos de calor.

Neste ponto, é razoável ressaltar que não foram localizados estudos que dimensionam a área de impacto de um evento de queimada de cana-de-açúcar. Tão pouco é possível saber o tamanho da área impactada por meio dos dados do INPE. Um foco de calor corresponde às coordenadas do centróide onde o evento ocorreu.

Uma das formas utilizadas para estimar área de impacto de uma queimada poderia ser modelagem da pluma resultante da queimada, levando em consideração variáveis como vento, umidade, pressão atmosférica, entre outros, assim como o tamanho do talhão queimado. Mas estes dados não estão disponíveis ou não são coletados em todas as localidades.

Contudo, é plausível inferir que um foco de calor pode causar impacto em dois ou mais municípios, a considerar a distância entre local da queimada e as distâncias das áreas urbanizadas do entorno.

A opção por utilizar a quantidade de focos de calor nos *buffers*, embora seja uma abordagem mais realista, não apresentou diferenças significativas quando comparada à quantidade no polígono municipal. O mesmo efeito foi possível observar entre os *buffers* de 10km e 20 km.

4.2.3 Dados Climáticos

Os dados climáticos foram obtidos por meio dos mapas de temperaturas (mínimas, máximas e médias) e precipitações (mínimas, máximas e médias), no formato *shapefile*, disponibilizado pelo IBAMA (IBAMA, 2009) e convertidos para dados tabulares por meio do cálculo de média para os polígonos municipais e micro-região. Este procedimento foi executado no ArcView 9.3.

Infelizmente, a disponibilidade de dados climáticos é limitada. Além dos dados do IBAMA, existe um banco de dados disponibilizado pela Embrapa Satélites (EMBRAPA, 2009a), que organizou médias mensais de temperatura e precipitações para 500 municípios brasileiros. Dada a escala deste estudo, a base da Embrapa não é suficiente para cobrir a área pesquisada com homogeneidade. Desta forma, optou-se pela utilização da base de dados do IBAMA, descartando-se a base da Embrapa. As médias de temperatura e precipitações para o ano de 2008, que utilizadas no estudo, estão sumarizadas na Tabela 2.

Observam-se aqui algumas diferenças entre as regiões selecionadas. A região Nordeste possui temperaturas mais elevadas que a maior parte dos estados da região Centro-Sul, com médias anuais cerca de 3 graus mais elevadas. No entanto, as precipitações pluviométricas são mais intensas na região Centro-Sul.

Tabela 2 - Variações médias de temperatura e precipitações mínimas e máximas por estado, nas regiões incluídas no estudo. Fonte: IBAMA, 2008.

Região	UF	Temperatura (C°)			Prec. Mínimas (mm)			Prec. Máximas (mm)		
		Mínima	Máxima	Média	Mínima	Máxima	Média	Mínima	Máxima	Média
CS	DF	20,00	20,00	20,00	1500,00	1500,00	1500,00	1750,00	1750,00	1750,00
	SP	17,00	23,00	20,97	1000,00	2833,33	1155,88	1250,00	3083,33	1405,49
	MT	22,00	25,00	24,08	1125,00	2750,00	1864,36	1375,00	3000,00	2114,36
	PR	14,00	23,00	19,30	1000,00	1750,00	1341,79	1250,00	2000,00	1591,79
	MG	16,00	25,00	20,25	750,00	2000,00	1273,89	1000,00	2250,00	1523,89
	MS	21,00	25,00	22,77	1000,00	1500,00	1262,82	1250,00	1750,00	1512,82
	GO	20,00	25,00	22,29	1250,00	1750,00	1626,52	1500,00	2000,00	1876,52
Mínima		14,00	20,00	19,30	750,00	1500,00	1155,88	1000,00	1750,00	1405,49
Máxima		22,00	25,00	24,08	1500,00	2833,33	1864,36	1750,00	3083,33	2114,36
Média		18,57	23,71	21,38	1089,29	2011,90	1432,18	1339,29	2261,90	1682,13
D. Padrão		3,06	2,01	1,80	262,07	558,59	264,17	262,07	558,59	264,27
NE	BA	20,00	27,00	23,21	300,00	1750,00	973,40	500,00	2000,00	1221,42
	PE	20,00	26,00	22,97	300,00	1750,00	755,01	500,00	2000,00	991,28
	PB	20,00	27,00	23,84	150,00	1750,00	787,18	400,00	2000,00	1032,77
	SE	23,00	26,00	24,59	300,00	1125,00	774,67	500,00	1375,00	1019,44
	AL	22,00	26,00	24,30	525,00	1750,00	951,47	750,00	2000,00	1200,78
	Mínima	20,00	26,00	22,97	150,00	1125,00	755,01	400,00	1375,00	991,28
	Máxima	23,00	27,00	24,59	525,00	1750,00	973,40	750,00	2000,00	1221,42
Média		21,00	26,40	23,78	315,00	1625,00	848,35	530,00	1875,00	1093,14
D. Padrão		1,41	0,55	0,69	134,16	279,51	105,07	130,38	279,51	108,96

Os dados da Tabela 2 mostram que a região Nordeste possui temperaturas mais elevadas e pluviosidade menor, em relação à região Centro-Sul.

No âmbito dos domínios climáticos do Brasil, as áreas de cultivo de cana-de-açúcar da região Centro-Sul são classificadas com clima tropical úmido-seco e apresentam subtipos relativos à seca, que variam desde a ausência de seca (litoral e áreas próximas ao Trópico de Capricórnio a subtipos com até 5 meses de seca (Mato-Grosso, Goiás e parte de Minas Gerais). Já a região Nordeste, o domínio é o Tropical Litorâneo do Nordeste Oriental, com variações de período de seca entre 1 e 5 meses (1 mês no

litoral da Bahia, e até 5 meses no Rio Grande do Norte) (MENDONÇA e DANNI-OLIVEIRA, 2007).

A pluviosidade varia em função das estações do ano. No caso do Brasil, a época de chuvas intensas ocorre durante o verão na região Sudeste e durante o inverno na região Nordeste (MENDONÇA e DANNI-OLIVEIRA, 2007). Esta variação favorece os períodos de safra de cana-de-açúcar, para ambas as regiões.

É importante salientar que estas classificações estão imbuídas de elevado grau de generalização, que dificulta a comparação mais apurada, dado que a extensão das regiões é ampla. Porém, são evidentes as diferenças climáticas existentes entre o litoral e a região central do Brasil.

Precipitações comprometem o ciclo da safra de cana-de-açúcar porque dificultam a colheita e interferem no processo de amadurecimento da planta. Por outro lado, diminuem a poluição atmosférica por meio da sedimentação do material particulado e diluição de gases.

No que tange à temperatura, as oscilações térmicas interferem na incidência de doenças respiratórias, porém, morbidade e mortalidade também dependem de fatores como idade, fatores sócio-econômicos e ambientais (GOUVEIA; HAJAT; ARMSTRONG, 2003).

A variabilidade dos dados na região Centro-Sul sugere maior heterogeneidade, do ponto de vista climático, nesta região. De forma inversa, a região NE apresenta dados mais homogêneos.

4.2.4 Relevô

A altitude é um fator que influencia a temperatura e a velocidade do vento, interferindo na dispersão dos poluentes atmosféricos.

A heterogeneidade também é identificada no relevo em toda área estudada. Uma das variáveis utilizadas foi altitude. Este dado provindo do IBGE traz a altitude medida na sede do município e mostra que os estados

do estudo possuem variação de altitude diferenciada nas duas regiões. Tal diferenciação pode ser observada no Quadro 6.

Região	UF	Média	Mínima	Máxima
CS	DF	1171	1171	1171
CS	SP	553	0	1628
CS	MT	276	0	851
CS	PR	600	3	1202
CS	MG	684	0	1505
CS	MS	375	90	790
CS	GO	620	0	1189
NE	BA	373	0	1268
NE	PE	394	0	1004
NE	PB	339	0	815
NE	SE	126	0	351
NE	AL	161	0	633

Quadro 6 - Altitudes médias, mínimas e máximas dos estados da área de estudo.

No conjunto observado, nota-se variação de altitude mais acentuada nos estados da região Centro-Sul. O estado de São Paulo possui áreas de litoral e serra, que contribuem para uma variação de altitude entre 0 a 1628 metros. No entanto, estes extremos não possuem áreas de plantio de cana-de-açúcar e os problemas respiratórios nestas áreas não estão associados à questão da cana-de-açúcar, mas, provavelmente associados a esta variável.

Outrossim, a região Nordeste apresenta variabilidade de altitude mais estável, ponderando-se Bahia e Pernambuco, que possuem regiões serranas com altitudes mais elevadas. Aqui também a cultura canavieira está longe destes extremos concentrando-se nos municípios litorâneos e próximos.

4.2.5 IDH-M

Os dados do Índice de Desenvolvimento Humano Municipal – IDH-M foram coletados no site do PNUD Brasil (PNUD, 2009) com o objetivo de agregar dados sócio-econômicos e ampliar as variáveis independentes

utilizadas na análise multivariada. Conforme GOUVEIA, HAJAT e ARMSTRONG (2003), os aspectos sociais podem interferir na incidência de doenças respiratórias.

O IDH-M é um índice que varia de 0 a 1, calculado com base municipal a partir dos sub-índices de educação, renda e longevidade, cuja fonte é o Censo. Existem duas séries disponíveis: 1991 e 2001. O Gráfico 4 apresenta as médias dos índices por estado, calculadas a partir da série de 2001.

Gráfico 4 - Distribuição de médias do IDH-M por estados das regiões de estudo em 2001. Fonte: PNUD, 2009.

Observa-se que a região Centro-Sul apresenta índices melhores que os da região Nordeste. Em entre os estados de cada região, existe relativa homogeneidade.

4.2.6 Áreas de Cana e Áreas Urbanizadas

Os dados sobre área plantada de cana-de-açúcar estão disponíveis no site do Projeto CANASAT do INPE criado para fornecer informações sobre a distribuição espacial da área cultivada de cana-de-açúcar na região Centro-Sul (CANASAT, 2009).

Este dado foi selecionado com o intuito de analisar a localização das áreas canavieiras, bem como, de utilizar os dados quantitativos nas análises estatísticas, objetivando encontrar um bom denominador para calcular

intensidade de focos de calor para uma dada área impactada (município, ou micro-região, por exemplo). Porém, estes dados não estão disponíveis para a região Nordeste.

Isto posto, optou-se por utilizar o trabalho da Embrapa, Monitoramento de Satélites, que mapeou as áreas urbanizadas brasileiras por meio do cruzamento de dados censitários da população urbana do Brasil de 2000 com imagens do satélite LANDSAT ETM feitas em 2000 e 2001. Este estudo dimensionou as áreas urbanizadas em todo o território nacional e concluiu que mais de 80% da população brasileira vive nestas áreas, em geral, aglutinadas nas sedes municipais (EMBRAPA, 2009b).

Trata-se de um estudo interessante, pois expressa numericamente a heterogeneidade continental do Brasil, em particular, das regiões estudadas.

As variáveis: área total, área não urbanizada e área urbanizada do município foram analisadas na expectativa de encontrar possíveis correlações com as variáveis dependentes, obtidas das internações hospitalares.

Por fim, optou-se por utilizar a variável área urbanizada como denominador para o cálculo da taxa de intensidade de focos de calor, pois esta representa numericamente a área de maior concentração da população.

A Tabela 3 apresenta uma sumarização das áreas por estado e por região de estudo. Ocorre que as diferenças entre área não urbana e área de plantio de cana-de-açúcar são acentuadas na região Centro-Sul. Da mesma forma que as áreas urbanizadas representam uma parcela muito pequena em relação à área total.

Finalizados os testes, a utilização da área total ou da área urbanizada como denominador não surtiu diferenças significativas nos resultados da análise multivariada. Testes realizados somente na região Centro-Sul, com os dados de plantio, também não apresentam diferenças significativas.

Tabela 3 - Área total, urbanizada e plantio de cana por estado. Fontes: ⁽¹⁾, ⁽²⁾, ⁽³⁾ e ⁽⁴⁾

Região	UF	Área Total ⁽¹⁾	Urbana ⁽²⁾	Não Urbana ⁽³⁾	Plantio Cana ⁽⁴⁾
CS	DF	5.801,94	621,20	5.180,74	-
	GO	340.086,70	1.117,68	338.969,02	4.320,03
	MG	586.528,29	2.522,48	584.005,81	5.749,95
	MS	357.125,96	442,30	356.683,66	2.909,91
	MT	903.359,91	534,65	902.825,26	2.310,59
	PR	199.314,85	1.603,77	197.711,08	6.049,35
	SP	248.209,43	4.984,26	243.225,17	44.452,77
	Total	2.640.427,07	11.826,35	2.628.600,73	65.792,60
		100%	0,45%	99,55%	2,49%
NE	AL	27.767,66	203,11	27.564,56	n/d
	BA	564.692,67	764,85	563.927,82	n/d
	PB	56.439,84	320,68	56.119,16	n/d
	PE	98.311,62	511,47	97.800,15	n/d
	SE	21.910,35	118,66	21.791,69	n/d
	Total	769.122,13	1.918,76	767.203,37	n/d
		100%	0,25%	99,75%	

⁽¹⁾ Área total em Km²: (IBGE, 2009).

⁽²⁾ Áreas urbanizadas em km²: (EMBRAPA, 2009b).

⁽³⁾ Áreas não urbanizadas em km²: (Área total – Área urbanizada).

⁽⁴⁾ Área disponível para plantio de cana, safra 2008/2009 em km²: (CANASAT, 2009).

n/d = não definido.

A análise destes dados contribuiu no sentido de reforçar a idéia de quantificar focos de calor a partir de *buffers* no entorno da sede municipal, ao invés de considerar somente os focos dentro do polígono municipal. A área urbanizada é menor que a área rural e concentra a maior parcela da população.

4.2.7 População

Os dados de contagem populacional foram coletados diretamente na página do IBGE (IBGE, 2009). Foram utilizadas contagens de população total por município do censo 2000 e as estimativas para os demais anos do estudo (2001 a 2008).

Os dados foram transferidos para o SQL-Server 2005, onde se sumarizou totais por faixa etária de interesse (0 a 4 anos e 65 anos ou mais). Estes dados foram utilizados no cálculo das taxas de incidência.

Na avaliação destes dados, observou-se que a região CS possui 68,3% da população em áreas urbanizadas, diferindo da região NE com 51,0% da população nestas mesmas condições.

Segundo a estimativa populacional de 2008, a população na faixa etária entre 0 a 4 anos, representa 6,2% na região CS e 7,6% na região NE. Muito similar à faixa etária de 65 anos ou mais, com 7% na região CS e 6,6% na região NE. Embora em números absolutos a população praticamente dobre na região CS, em números relativos observa-se similaridade entre ambas na distribuição por faixas etárias.

4.2.8 Internações

Conforme descrito anteriormente, os dados de saúde foram coletados do DATASUS por *download* dos arquivos denominados Reduzidas. Basicamente, cada linha de uma tabela reduzida pode corresponder a uma internação de caráter agudo ou a um período delimitado de uma internação de longa duração, na rede pública (PORTELA, et al, 1997).

Assim, os dados desta fonte não correspondem à totalidade de internações ocorridas, pois seria necessário somar aos dados da rede privada, que não estão disponíveis². Soma-se a isso, a existência de restrições relativas à qualidade dos dados. Contudo, esta base de dados é disponível e tem sido utilizada em muitos estudos. Cita-se como características positivas o fato de possuir abrangência nacional, com pequeno intervalo de atualização, e o amplo acesso ao seu conteúdo (COSTA, L.S.; PINHEIRO, R.S.; ALMEIDA, R.V.R., 2003).

² Os dados da AIH correspondem a informações da rede pública e são públicos. Os dados da rede privada conveniada com o SUS são coletados pela CIH – Comunicação de Internação Hospitalar, conforme mencionado anteriormente. Os dados da rede privada não conveniada, não são coletados.

A Tabela 4 apresenta uma sumarização dos dados coletados a partir das AIH, considerando as taxas calculadas para o acumulado de 9 anos, Observa-se que o estado do Paraná apresenta valores mais elevados para asma na Região Centro-Sul. No entanto, estas taxas não superam os valores apresentados por Bahia e Pernambuco na região Nordeste.

Quando se observam as taxas calculadas para enfisema, a situação se inverte, pois as taxas são mais elevadas em Goiás e Mato Grosso na região Centro-Sul.

Neste ponto vale ressaltar uma consideração relacionada aos dados de enfisema. Embora o principal fator etiológico seja o fumo, é muito habitual encontrar esta doença e a bronquite crônica presentes no mesmo paciente sugerindo que os fatores etiológicos sejam comuns. Seu principal sintoma é a falta de ar (RATTO, SANTOS e BOGOSSIAN, 1981). Dadas as características da doença, são esperados valores baixos para a faixa etária de 0 a 4 anos, conforme foi observado. Assim, para as demais análises eles foram descartados.

Do ponto de vista temporal, observou-se a evolução das grandezas no período do estudo, O Gráfico 5 mostra a curva baseada em quantidades de internações por ano, considerando todos as internações por asma, bronquite e enfisema, nas duas regiões estudadas. De forma geral, observou-se uma tendência de queda dos números de internações, que pode estar associada à melhora da qualidade de vida da população em geral ou a mudança de hábitos associados à procura de serviços de saúde, com a migração da procura do serviço público para o privado.

Tabela 4 - Contagem de internações e taxa de incidência para asma, bronquite e enfisema, categorizados por idade, registrados por AIH no SIH/SUS por UF, no período de 2000 a 2008.

Região	UF	População 2008		Asma 0 a 4 a.		Asma 65 ou + a.		Bronquite 0 a 4 a.		Bronquite 65 ou + a.		Enfisema 0 a 4 a.		Enfisema 65 ou + a.	
		0 a 4 anos	65 ou mais	Casos	Taxas*	Casos	Taxas*	Casos	Taxas*	Casos	Taxas*	Casos	Taxas*	Casos	Taxas*
CS	GO	394.874	339.300	26.145	66,21	11.606	34,21	2.619	6,63	1.193	3,52	131	0,33	10.898	32,12
	MG	1.268.462	1.495.296	95.594	75,36	16.145	10,80	3.527	2,78	645	0,43	129	0,10	8.361	5,59
	MS	164.409	146.931	8.745	53,19	4.394	29,91	516	3,14	137	0,93	48	0,29	2.894	19,70
	MT	218.600	140.869	16.006	73,22	4.697	33,34	152	0,70	13	0,09	25	0,11	3.397	24,11
	PR	613.216	748.865	59.908	97,69	27.982	37,37	1.549	2,53	666	0,89	326	0,53	16.337	21,82
	SP	2.435.990	2.980.753	121.006	49,67	26.100	8,76	10.138	4,16	2.766	0,93	343	0,14	8.450	2,83
Total		5.095.551	5.852.014	327.404	64,25	90.924	15,54	18.501	3,63	5.420	0,93	1.002	0,20	50.337	8,60
NE	AL	289.985	165.473	16.212	55,91	2.792	16,87	311	1,07	12	0,07	39	0,13	222	1,34
	BA	1.114.821	958.368	139.988	125,57	36.017	37,58	1.750	1,57	549	0,57	78	0,07	3.367	3,51
	PB	261.878	285.820	25.037	95,61	9.040	31,63	880	3,36	426	1,49	16	0,06	1.465	5,13
	PE	615.677	594.123	78.285	127,15	18.873	31,77	3.347	5,44	1.134	1,91	69	0,11	595	1,00
	SE	164.299	113.940	9.760	59,40	1.741	15,28	99	0,60	42	0,37	7	0,04	130	1,14
Total		2.446.660	2.117.724	269.282	110,06	68.463	32,33	6.387	2,61	2.163	1,02	209	0,09	5.779	2,73

* Casos / 1.000 habitantes sobre a população estimada de 2008, nas faixas 0 a 4 anos e 65 anos ou mais.

Gráfico 5 - Curva de evolução da somatória internações por asma, bronquite e enfisema na rede pública, regiões Centro-Sul e Nordeste, período de 2000 a 2008.

Embora o crescimento populacional nas duas regiões tenha mantido estável a proporção entre ambas (variando entre 37% e 38% no período), observa-se a tendência de queda mais acentuada na curva da região CS. A diferença de internações entre as regiões era de 21.358 eventos em 2000 e caiu para 7.027 em 2008, reduzindo a diferença para 1/3 da quantidade registrada no início do período estudado.

A diminuição do número de internações na região CS pode ser explicada, em parte, pelos melhores índices sócio-econômicos obtidos na região CS, no mesmo período. Índices sócio-econômicos positivos podem contribuir com a redução do número de internações em função da melhoria da qualidade de vida da população envolvida. Por outro, a melhoria sócio-econômica proporciona acesso ao sistema privado ou a planos de saúde, que pode também contribuir com a diminuição dos números, uma vez que aqui se contabilizou apenas as internações da rede pública.

Outra possível explicação pode estar relacionada à implementação da atenção básica do SUS – Sistema Único de Saúde nos municípios brasileiros. Este processo tem proporcionado maior resolução dos casos no pronto atendimento e diminuído a necessidade de internação. No que diz respeito aos agentes do PSF (Programa Saúde da Família), a atuação no

âmbito da promoção da saúde também contribui para a redução de casos que necessitam de internação hospitalar.

Contudo, no mesmo período, observou-se aumento dos focos de queimada, conforme discutido anteriormente no Gráfico 2. A análise deste contexto sugere inferir que a evolução das queimadas não prepondera na etiologia das doenças avaliadas. Porém, a análise sob as condições de generalização aqui colocadas podem carregar sobremaneira os vieses anteriormente discutidos, limitando a acurácia das inferências.

Na perspectiva de analisar as doenças isoladamente, foram gerados gráficos com estratificações para asma, bronquite e enfisema. Observou-se que os dados sobre bronquite apresentaram queda brusca a partir do ano de 2003, voltando a recuperar o padrão somente em 2008. Sendo assim, optou-se por descartar os dados de bronquite, uma vez que a análise ficaria comprometida com esta anomalia. Os Gráficos 6, 7, 8, 9, 10 e 11 apresentam estes dados.

As curvas para asma (Gráficos 6 e 7) e enfisema (Gráfico 10 e 11) apresentaram padrão esperado, exceto pela tendência de queda no período. Em particular a curva de asma para a faixa etária de 0 a 4 anos apresenta variação sazonal que eleva os valores no período de inverno.

Esta característica não se repete nas curvas para asma, na faixa etária de 65 anos ou mais e enfisema, na faixa etária de 65 anos ou mais, indicando que estas doenças são menos sensíveis aos fatores sazonais decorrentes do inverno nestas regiões. Ainda sobre este tema, a comparação dos gráficos por região não apresentaram variações significativas entre ambos, no que diz respeito à oscilação entre os meses. Porém os números de internações são bem mais elevados na região CS para os três grupos de doenças.

Gráfico 6 - Evolução do número de internações com diagnóstico de asma por faixa etária, na região Centro-Sul, no período de 2000 a 2008.

Gráfico 7 - Evolução do número de internações com diagnóstico de asma por faixa etária, na região Nordeste no período de 2000 a 2008.

* As médias móveis foram calculadas com variável de ajuste $N = 30$.

* As médias móveis foram calculadas com variável de ajuste $N = 30$.

* As médias móveis foram calculadas com variável de ajuste $N = 30$.

O Quadro 7 apresenta uma sumarização dos dados diários de internações, para as doenças do estudo.

Região CS	Asma04	Asma65	Bron04	Bron65	Enfi04	Enfi65
Total de registros	327.468	90.933	18.508	5.421	1.005	50.381
Valor Máximo	38	11	7	12	3	5
Valor Mínimo	0	0	0	0	0	0
Média	0	0	0	0	0	0
Desvio Padrão	1,45	0,45	0,25	0,12	0,05	0,36
Variância	2,128	0,211	0,062	0,016	0,002	0,130

Região NE	Asma04	Asma65	Bron04	Bron65	Enfi04	Enfi65
Total de registros	269.282	68.463	6.387	2.163	209	5.779
Valor Máximo	24	5	8	6	3	4
Valor Mínimo	0	0	0	0	0	0
Média	1	0	0	0	0	0
Desvio Padrão	1,08	0,47	0,17	0,09	0,02	0,15
Variância	1,178	0,225	0,032	0,009	0,0008	0,024

Quadro 7 - Sumarização dos dados sobre internações para asma, bronquite e enfisema, baseado na tabulação diária por faixa etária, na área de estudo (CS + NE), no período de 2000 a 2008.

Variáveis: Asma04 – internações por asma com faixa etária de 0 e 4 anos. Asma65 – internações por asma com faixa etária de 65 anos ou mais. Bron04 – internações por bronquite com faixa etária de 0 a 4 anos. Bron65 – internações por bronquite com faixa etária de 65 anos ou mais. Enfi04 – internações por enfisema com faixa etária de 0 a 4 anos. Enfi65 – internações por enfisema com faixa etária de 65 anos ou mais.

Neste quadro, observa-se variância elevada para asma de 0 a 4 anos em função das ocorrências sazonais anteriormente discutidas. Da mesma forma, observa-se que as médias calculadas apresentam valor igual a zero (excetuando-se Asma04 / NE). Tais valores resultam de contagens muito baixas para estas variáveis, esperado para enfisema na faixa de 0 a 4 anos, porém não esperados para bronquite.

Ressalta-se ainda que, embora os números sejam menores na região Nordeste, as taxas de asma são maiores no NE. Porém esta situação se inverte com relação às taxas calculadas para enfisema.

4.3 Análises Espaciais

Segundo DRUCK, et al (2004):

“As técnicas de análise exploratória aplicadas a dados espaciais são essenciais ao desenvolvimento de etapas da modelagem estatística espacial, em geral, sensível ao tipo de distribuição, à presença de valores extremos e à ausência de estacionariedade” (DRUCK et al, 2004, pág. 164).

É muito comum limitar-se a operações de visualização em análise espacial. No entanto, as ferramentas atuais de geoprocessamento permitem outras avaliações mais objetivas ao conjunto de dados inseridos em m SIG.

A sequência de mapas numerados de 3 a 10 apresenta os resultados da análise *Hotspot* para as variáveis do estudo: focos de calor, asma (faixas etárias de 0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais). Os mosaicos foram elaborados agrupando-se os dados por município e trimestres de 2008 nas regiões do estudo.

4.3.1 Análise da Região Centro-Sul

A região Centro Sul, apresenta um padrão esperado para a evolução temporal dos focos de calor que coincidente com o período de safra de cana-de-açúcar que ocorre entre março e novembro. O primeiro trimestre apresenta agrupamentos menores e as manchas vão se acentuando até o quarto trimestre. No entanto, este fenômeno não é observável em igual intensidade para as internações. Os agrupamentos estão mais dispersos e não se intensificam em função do período, o que se observa são alterações na distribuição, ou seja, na mudança dos trimestres, alguns agrupamentos desaparecem e outros surgem. Mesmo assim, alguns se mantêm ativos durante todo o período.

Na comparação entre os agrupamentos altos (com desvio padrão maior que 1,96) dos focos de calor e das intenações, observa-se que os agrupamentos das intenações não são exatamente coincidentes aos agrupamentos de focos de calor. No entanto, eles ocorrem, na maioria das vezes, com certa proximidade ou ainda com intersecções. Estas ocorrências são mais freqüentes nos estados de Goiás e São Paulo.

Em São Paulo, os agrupamentos de focos de calor são amplos e atingem o ápice no terceiro trimestre analogamente ao período mais intenso da safra de cana nesta região. O maior cobre desde a região de Botucatu, passando por Ribeirão Preto e o Triângulo Mineiro até o estado de Goiás, no terceiro trimestre. Dois outros agrupamentos estão localizados na região de Presidente Prudente e no Noroeste do Paraná, respectivamente.

Os demais agrupamentos estão fora das áreas de plantação de cana-de-açúcar, como os observados no norte de Minas Gerais, Mato Grosso e Mato Grosso do Sul.

Com relação às intenações, a região de Jales, no Noroeste de São Paulo, merece destaque. Juntamente com municípios de Minas Gerais, como Carneirinho e Iturama, na ponta do Triângulo, forma agrupamentos para os três tipos de intenação, que variam no período de forma similar aos focos de calor.

Outro ponto interessante é o Noroeste do Paraná. Embora exista um agrupamento de focos de calor nesta área de plantio de cana-de-açúcar, que se mantêm nos quatro trimestres, os agrupamentos de intenações não são visíveis na mesma área. No entanto, observam-se agrupamentos no entorno, sudeste do Paraná e municípios de Mato Grosso do Sul. Esta configuração permite fomentar hipóteses sobre um efeito regional, ao invés de processos locais, corroborando estudos de transporte de poluentes atmosféricos que indicam efeitos em escala regional ou global (ANDREAE, et al, 2004).

Região Centro-Sul

Pág. 117 / Mapa 3 – Mosaico da Análise Hot Spot para as variáveis: focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 1º trimestre de 2008.

Pág. 118 / Mapa 4 – Mosaico da Análise Hot Spot para as variáveis: focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 2º trimestre de 2008.

Pág. 119 / Mapa 5 – Mosaico da Análise Hot Spot para as variáveis: focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 3º trimestre de 2008.

Pág. 120 / Mapa 6 – Mosaico da Análise Hot Spot para as variáveis: focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 4º trimestre de 2008.

Região Nordeste

Pág. 121 / Mapa 7 – Mosaico da Análise Hot Spot para as variáveis: focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 1º trimestre de 2008.

Pág. 122 / Mapa 8 – Mosaico da Análise Hot Spot para as variáveis: focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 2º trimestre de 2008.

Pág. 123 / Mapa 9 – Mosaico da Análise Hot Spot para as variáveis: focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 3º trimestre de 2008.

Pág. 124 / Mapa 10 – Mosaico da Análise Hot Spot para as variáveis: focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 4º trimestre de 2008.

Por outro lado, os agrupamentos de internações estão dispersos em São Paulo, apesar da existência do agrupamento bem definido para focos de calor. No entanto, nas proximidades de Jaú, onde o cultivo de cana é muito intenso, observa-se a continuidade dos agrupamentos para asma infantil no período, à exceção do primeiro trimestre.

4.3.2 Análise da Região Nordeste

Como já fora visto anteriormente, os focos de calor na região Nordeste ocorrem com mais intensidade na mesma época que a região Centro-Sul. Este fato suscita a hipótese da menor intensidade de queimadas específicas na lavoura de cana, em relação àquelas geradas por outros fatores.

No estado da Paraíba, observam-se agrupamentos coincidentes para focos de calor, asma infantil e enfisema, no primeiro trimestre, em duas áreas: litoral e oeste do estado.

O norte de Sergipe e o litoral sul de Alagoas apresentam agrupamentos de focos de calor próximos de agrupamentos de internação, mas sem intersecção entre eles. Porém os agrupamentos de focos de calor se mantêm no ano, indicando que existem outras atividades que promovem queimadas nestas localidades.

A micro-região de Cotegipe apresenta um agrupamento de enfisema contínuo para os quatro trimestres, muito próximo do agrupamento de focos de calor. No terceiro trimestre, o agrupamento de focos regride e, de forma similar, o agrupamento de enfisema perde intensidade.

Com relação ao agrupamento de asma infantil, duas manchas se mantêm estáveis no estado da Bahia. A mais intensa está na faixa litorânea e não relacionada a queimadas, embora esteja muito próxima a um agrupamento de focos no primeiro trimestre. O segundo agrupamento, mais ao centro do estado, também não parece ter relação com os focos de calor.

Estas análises confirmam a heterogeneidade das regiões estudadas, e reforçam os argumentos da dificuldade de compará-las à luz destes fenômenos pesquisados.

4.3.3 Análise longitudinal para o estado de São Paulo

O estado de São Paulo apresenta um diferencial em relação aos demais estados do estudo, posto que as queimadas que ocorrem em sua extensão territorial são, principalmente, causadas em função da cultura de cana-de-açúcar.

As ilustrações do Mapa 11 representam a junção dos resultados da análise *Hotspot* para o estado de São Paulo, com agregados anuais de queimadas e grupos de doenças. Nesta análise, optou-se por sobrepor as camadas geradas afim de melhor identificar a interseção dos agrupamentos de ambos.

A sequência de mapas permite observar que havia quatro agrupamentos de queimadas, localizados nas meso-regiões de Assis, Ribeirão Preto, Bauru, Araraquara e Piracicaba. Este último, composto também por municípios da meso-região de Campinas.

No decorrer do período de estudo, estes agrupamentos se modificaram em função da dinâmica da cultura de cana-de-açúcar, baseada na expansão das áreas de cultivo e redução das queimadas em substituição à colheita mecânica sem queima da palha na área plantada.

Pode-se observar que os agrupamentos na meso-região de Assis e meso-região Piracicaba/Campinas não estão presentes no mapa de 2008. Este detalhe indica que a análise *Hotspot* não identificou um agrupamento de queimadas significativo nesta área em 2008. No entanto, isso não significa a ausência de queimadas na região.

Por outro lado, os agrupamentos de queimadas estão mais intensos e formam três grandes blocos ocupando as meso-regiões de Araçatuba,

Página reservada para o jogo de mapas da figura X

Mapa 11 – Interseção dos *hotspots* de queimadas e internações (asma e enfisema) para o estado de São Paulo, no período de 2000 a 2008.

(impresso em A3)

Bauru, Presidente Prudente, Ribeirão Preto e São José do Rio Preto. Os agrupamentos em Araçatuba e Presidente Prudente indicam a tendência de expansão no sentido oeste do estado de São Paulo, que já alcançou o noroeste do Paraná e o sudeste do Mato Grosso do Sul, conforme observado nos agrupamentos da região CS.

Vale ressaltar que na região de Araçatuba e Presidente Prudente, tradicionalmente pecuarista, os produtores vem substituindo áreas de pasto por plantações de cana nos últimos anos, contribuindo para o aumento de queimadas nos municípios desta região.

Portanto, os *hotspots* de queimadas apresentam compatibilidade com as áreas de plantio, bem como a tendência de expansão no estado de São Paulo e documentam sinais do processo de mecanização deste agronegócio.

De forma inversa, os agrupamentos calculados para as interações estão dispersos no estado. Observa-se que as manchas tendem a diminuir e se concentrar no noroeste do estado ao longo dos anos. Agrupamentos importantes nas regiões de Itapetininga e São João da Boa Vista perdem intensidade, praticamente desaparecendo até 2008.

Os agrupamentos de interações da meso-região de Presidente Prudente se mantêm no período do estudo, embora não sejam observáveis agrupamentos de queimadas com igual intensidade e estabilidade de outras áreas. Porém, em 2007 e 2008 eles intersectam com o agrupamento de queimadas.

O mesmo fenômeno ocorre, em maior intensidade nos encontro dos limites das meso-regiões de Araraquara, Bauru e Ribeirão Preto. Os agrupamentos de queimadas e interações apresentam interseções constantes para os nove anos estudados. Fato similar ocorre próximo a Ribeirão Preto, nos municípios de Ipuã, Morro Agudo e São Joaquim da Barra, a partir de 2005.

Por outro lado, um agrupamento de internações importante se formou no entorno do município de Jales, ao norte do estado. Nos limites do agrupamento de queimadas de São Paulo, este *hotspot* localiza-se nos limites do agrupamento de queimadas do Triângulo Mineiro, conforme observado anteriormente nos mosaicos.

As análises aqui descritas não são suficientes para explicar a variabilidade das taxas de internações na área de estudo, mas fomentam a hipótese das relações indissociáveis que ocorrem entre os objetos do espaço observado.

4.4 Análises Estatísticas

4.4.1 Distribuição Normal

Conforme descrito anteriormente, a análise estatística é o passo seguinte à análise espacial e é usada para validar ou refutar a hipótese fomentada nas etapas anteriores.

Taxas provenientes de contagens populacionais tendem a gerar uma curva de distribuição logarítmica, positiva ou negativa. No caso deste estudo, as taxas de incidência: asma04 – internações por asma na faixa etária de 0 a 4 anos, asma65 – internações por asma na faixa etária de 65 anos ou mais e enfi65 – internações por enfisema na faixa etária de 65 anos ou mais, assim como focos de calor: qtdbf10 – quantidade de focos de calor localizados no entorno da sede municipal, dentro do raio de 10 Km e qtdbf20 – quantidade de focos de calor localizados no entorno da sede municipal, dentro do raio de 20 Km, apresentaram curvas log negativas que foram transformadas para ajustar à curva normal.

A Figura 10 mostra o histograma das variáveis incasma04 (incidência de asma 0 a 4 anos) e qtdbf10 (quantidade de focos de calor com *buffer* de

10km do centro do município) antes e depois do processo de transformação executado com a função Ln() (logaritmo neperiano) no software Stata 10.

Figura 10 - Histograma das variáveis incasma04 e qt dbf10, antes e depois da transformação com função Ln().

Os testes Shapiro-Wilk, executados no Stata 10 para as regiões CS e NE separadamente, apresentaram valor zero, reafirmando a não normalidade da distribuição conforme visto anteriormente nos histogramas. O Quadro 8 apresenta os resultados do teste.

Região CS						
Variable	Obs	W	V	z	Prob>z	
qt dbf10	7920	0.32517	2743.977	21.059	0.00000	
qt dbf20	7920	0.27022	2967.383	21.267	0.00000	
INCASMA04	7920	0.69749	1230.060	18.925	0.00000	
INCASMA65	7920	0.66769	1351.224	19.175	0.00000	
INCENFI65	7920	0.38193	2513.180	20.825	0.00000	

Região NE		continuação...				
Variable	Obs	W	V	z	Prob>z	
qtdbf10	3600	0.28363	1446.449	18.894	0.00000	
qtdbf20	3600	0.20298	1609.275	19.170	0.00000	
INCASMA04	3600	0.56209	884.202	17.616	0.00000	
INCASMA65	3600	0.76650	471.467	15.983	0.00000	
INCENFI65	3600	0.54616	916.350	17.708	0.00000	

Quadro 8 - Teste Shapiro-Wilk aplicado para as variáveis: focos de calor (qtdbf10 e qtdbf20) e incidências de internações (incasma04, incasma65 e incenfi65), nas regiões CS e NE.

Após a etapa de transformação, os dados do conjunto apresentaram resultados mais apropriados para uso no modelo de regressão.

4.4.2 Análises por Correlação

Os testes executados no Stata 10, separadamente para as regiões CS e NE estão apresentados no Quadro 9.

As células com fundo cinza demarcam os valores de interesse, pois mostram as correlações entre focos de calor e as internações. Vale lembrar que foram testadas as variáveis de incidência de focos de calor, *buffer* de 10km e 20km, afim de verificar discrepâncias entre elas, considerando que o *buffer* de 20km possibilitou contabilizar mais focos por unidade espacial. No entanto, os resultados mostram poucas diferenças entre ambas, quando comparados com as internações.

Observa-se que os coeficientes de correlação para incidência de asma 0 a 4 anos são negativos indicando correlações inversas, ou seja, quando há aumento de focos de calor, há diminuição de internações por asma nesta faixa etária. Porém, estas correlações se mostram fracas entre as variáveis medidas em ambas as regiões.

Os resultados somente se mostram significativos para asma de 0 a 4 anos na região NE. Este fato é interessante, pois a intensidade de focos de

calor pode explicar parcialmente o aumento de internações por asma, apesar dos agrupamentos de queimadas coincidam parcialmente com os agrupamentos de internações nesta região.

Na região NE, as correlações são positivas e moderadas para enfisema 65 anos ou mais. No entanto, os valores de p são altos indicando que não há significância estatística.

Região CS	Incasma04	Incasma65	Incenfi65	Qtdfb10	Qtdbf20
Incasma04	1,0000 4170				
Incasma65	0,4963 0,0000 2390	1,0000 2975			
Incenfi65	0,4229 0,0000 1592	0,6206 0,0000 1268	1,0000 2050		
Qtdfb10	-0,0430 0,5050 243	-0,1215 0,1102 174	0,0527 0,5678 120	1,0000 526	
Qtdbf20	-0,0727 0,2345 269	-0,0733 0,3162 189	0,0691 0,4380 128	0,9802 0,0000 526	1,0000 587

Região NE	Incasma04	Incasma65	Incenfi65	Qtdfb10	Qtdbf20
Incasma04	1,0000 2639				
Incasma65	0,5366 0,0000 1683	1,0000 1940			
Incenfi65	0,3216 0,0000 262	0,5602 0,0000 207	1,0000 286		
Qtdfb10	-0,2236 0,0009*** 216	0,0229 0,7692 166	0,5192 0,1017 11	1,0000 321	
Qtdbf20	-0,2404 0,0003*** 218	-0,0026 0,9736 167	0,5337 0,0909 11	0,9769 0,0000 321	1,0000 329

Quadro 9 - Coeficientes de Correlação de Pearson entre as variáveis intensidade de focos de calor (qtdfb10 – *buffer* 10Km e qtdbf20 – *buffer* 20km) e incidências de internações (incasma04, incasma65 e incenfi65), nas regiões CS e NE.

Nota: os três valores apresentados nas células do quadro são respectivamente: r – coeficiente de correlação, valor de p – 95%, (considerou-se a hipótese de correlação igual a 0 e teste bicaudal) e número de observações.

As correlações fracas também podem ser visualizadas nos gráficos do tipo Matriz de Correlação. Os Gráficos 12 e 13 apresentam estes resultados para as regiões CS e NE gerados no Stata 10.

Gráfico 12 - Matriz de Correlação entre variáveis dependentes e independentes para a região CS.

Gráfico 13 - Matriz de Correlação entre variáveis dependentes e independentes para a região NE.

Os desenhos permitem observar que os pontos estão bem dispersos e não apresentam padrão característico de quando há correlação, positiva ou negativa.

4.4.3 Análises por Regressão Linear

O coeficiente de correlação de Pearson não é suficiente para explicar interações que ocorrem quando existem mais variáveis independentes.

Optou-se então por fazer a análise dos modelos de regressão, considerando que as transformações das variáveis apresentaram curvas similares à distribuição normal.

Desta forma, as covariáveis foram submetidas ao modelo de regressão multivariado, que se julgou ser adequado para analisar a variabilidade da doença na amostra. O Quadro 10 apresenta as variáveis utilizadas e as respectivas descrições.

Nome	Descrição
ALTITUDE	Altitude da sede do município
IDH	Índice de desenvolvimento humano
TEMPMED	Temperatura média anual
PRECMIN	Precipitação mínima anual
PRECMAX	Precipitação máxima anual
AREAU	Área urbanizada
INTFOCOS	Intensidade de focos de calor – <i>Polígono municipal</i>
INTFOCOBF10	Intensidade de focos de calor – <i>Buffer 10km</i>

Quadro 10 - Variáveis independentes utilizadas na análise de regressão.

As variáveis selecionadas foram testadas em diferentes formas de agrupamento, variando-se a combinação das dimensões espacial e temporal. Os testes foram executados separadamente para as regiões CS e NE. Utilizaram-se os dados coletados no período de nove anos.

As variáveis INCFOCOS e INCFOCOSBF10 foram testadas isoladamente a fim de observar possíveis diferenças das mesmas nos

resultados da análise. As variáveis dependentes A04, A65 e E65 também foram testadas separadamente.

Todos os modelos testados apresentaram valor de $p = 0,0000$ indicando a confiabilidade do mesmo. O Quadro 11 apresenta os resultados de R^2 para cada conjunto testado.

Dimensões*	Região Centro-Sul				Região Nordeste			
Polígono Municipal**	A04	A65	E65	Média	A04	A65	E65	Média
Micro/Trimestre	0,31	0,10	0,21	0,21	0,38	0,19	0,37	0,31
Município/Trimestre	0,42	0,41	0,43	0,42	0,46	0,34	0,44	0,41
Micro/Mês	0,11	0,11	0,25	0,16	0,05	0,14	0,61	0,26
Município/Mês	0,21	0,40	0,49	0,37	0,10	0,28	0,59	0,32
Buffer 10Km***	A04	A65	E65	MÉDIA	A04	A65	E65	MÉDIA
Micro/Trimestre	0,31	0,10	0,17	0,19	0,38	0,21	0,39	0,33
Município/Trimestre	0,44	0,45	0,47	0,45	0,46	0,36	0,40	0,40
Micro/Mês	0,07	0,08	0,18	0,11	0,05	0,11	0,60	0,25
Município/Mês	0,25	0,45	0,51	0,40	0,12	0,27	0,55	0,31

Quadro 11 – Valores de R^2 para as diversas combinações dos modelos testados.

Notas:

* Utilizou-se para a dimensão temporal, agregações por Mês e Trimestre. Para a dimensão Espacial, utilizou-se Município e Micro-região.

**No conjunto “Polígono Municipal” os testes foram feitos com a variável INCFOCOS, que representa a soma dos focos de calor dentro do polígono municipal, dividida pela área do mesmo.

***No grupo *Buffer 10Km*, utilizou-se a soma dos focos de calor dentro do *buffer* de 10km medido a partir da sede municipal, dividido pela área urbanizada do município.

Os modelos apresentam diferenças quanto ao poder de explicação da variabilidade da VD em ambas as regiões. Quando a comparação é feita entre as combinações que utilizaram INTFOCOS e INTFOCOSBF10, observa-se que os modelos gerados com este último possuem resultados mais significativos. Este fato corrobora a opção da escolha da variável INTFOCOSBF10 para as análises.

As tabelas a seguir apresentam resultados do modelo de regressão de mínimos quadrados, separados para as regiões CS e NE. Cada modelo

compara uma VD às demais VI. Optou-se por apresentar os modelos que obtiveram R^2 mais elevado. Desta forma, 5 dos modelos selecionados utilizaram a variável INTFOCOSBF10.

No caso do modelo para enfisema no Nordeste, optou-se por apresentar a análise gerada com a variável INTFOCOSBF10 com as dimensões Município/Mês para compatibilizar a comparação. Considerou-se que os valores de R^2 eram muito próximos e o modelo selecionado apresentou teste de significância estatística melhor para as variáveis independentes.

Análise: Internações por asma, faixa etária de 0 a 4 anos.

Source	SS	df	MS	Number of obs = 14822		
Model	12075.0305	7	1725.00435	F(7, 14815) = 1696.47		
Residual	15064.1759	14815	1.01681917	Prob > F = 0.0000		
Total	27139.2064	14822	1.83100839	R-squared = 0.4449		
				Adj R-squared = 0.4447		
				Root MSE = 1.0084		
incasma04	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
INTFOCOSBF10	.090968	.0058248	15.62	0.000	.0795507	.1023852
ALTITUDE	-.0003818	.0000505	-7.56	0.000	-.0004807	-.0002828
IDH	-4.922542	.163662	-30.08	0.000	-5.24334	-4.601744
TEMPMED	.0154	.0053302	2.89	0.004	.0049522	.0258479
PRECMIN	-.0178312	.0007283	-24.48	0.000	-.0192586	-.0164037
PRECMAX	.0181325	.0007263	24.97	0.000	.0167089	.0195561
AREAU	-.0059186	.0003731	-15.86	0.000	-.0066499	-.0051872

Quadro 12 - Modelo de regressão para a amostra CS/Município/Trimestre/ Incidência de internações por asma infantil.

Source	SS	df	MS	Number of obs = 7691		
Model	6650.8556	7	950.122228	F(7, 7684) = 939.80		
Residual	7768.39215	7684	1.01098284	Prob > F = 0.0000		
Total	14419.2477	7691	1.87482093	R-squared = 0.4612		
				Adj R-squared = 0.4608		
				Root MSE = 1.0055		
incasma04	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
INTFOCOSBF10	.0197273	.0079032	2.50	0.013	.0042349	.0352197
ALTITUDE	.0005246	.0000691	7.59	0.000	.0003891	.0006601
IDH	.2533325	.2175804	1.16	0.244	-.1731845	.6798495
TEMPMED	-.0516859	.008577	-6.03	0.000	-.0684991	-.0348727
PRECMIN	-.0063612	.0009902	-6.42	0.000	-.0083022	-.0044202
PRECMAX	.0066493	.0009658	6.89	0.000	.0047562	.0085425
AREAU	-.0110158	.002041	-5.40	0.000	-.0150166	-.007015

Quadro 13 - Modelo de regressão para a amostra NE/Município/Trimestre/Incidência de internações por asma infantil.

Os Quadros 12 e 13 apresentam os modelos para asma, faixa etária de 0 a 4 anos. Os testes do modelo indicam condições satisfatórias de análise. O R^2 indica que as variáveis independentes explicam 44% da variabilidade da VD no modelo CS e 46% para o modelo NE. O R^2 ajustado menor que R^2 , indica que as variáveis independentes não melhoram o modelo ao acaso. Os valores de p para as VI (na coluna $P > |t|$) indicam que os coeficientes calculados são estatisticamente significativos, a exceção do IDH para o modelo NE.

A variável INTFOCOSBF10 é positiva em ambos os modelos indicando correlação positiva, ou seja, para cada unidade acrescida de intensidade de focos de calor, a incidência de internações por asma infantil acresce 0,090968 no CS e 0,0197273 no NE, se as demais VI não se alteraram.

A ALTITUDE pouco influencia a VD, o coeficiente gerado tem valor próximo a zero em ambos os modelos, que também apresentam valor negativo para o modelo CS e positivo para o modelo NE

O coeficiente para variável IDH não apresenta significância estatística para o modelo NE, porém no modelo CS ela apresenta correlação negativa e indica que os municípios com melhor IDH-M tendem a apresentar menores taxas de internação por asma infantil.

A temperatura média apresenta correlação positiva no modelo CS e negativa no modelo NE indicando possíveis diferenças regionais relativas à influência das variáveis climáticas estudadas. Embora as correlações encontradas para precipitação mínima e máxima sejam as mesmas nos modelos, observa-se numericamente que a influencia destas, é maior na região CS.

A variável AREAU apresentou correlação negativa, indicando que os municípios com maior área urbanizada tendem a ter menores taxas de internação para asma infantil. Esta variável corrobora a análise do IDH-M.

Análise: Internações por asma, faixa etária de 65 anos ou mais.

Source	SS	df	MS	Number of obs = 11039		
Model	8968.64603	7	1281.23515	F(7, 11032)	=	1273.13
Residual	11102.1959	11032	1.00636293	Prob > F	=	0.0000
Total	20070.8419	11039	1.81817573	R-squared	=	0.4468
				Adj R-squared	=	0.4465
				Root MSE	=	1.0032

incasma65	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
INTFOCOSBF10	.1254449	.006776	18.51	0.000	.1121628	.138727
ALTITUDE	-.0003574	.0000607	-5.88	0.000	-.0004765	-.0002384
IDH	-8.391595	.2019435	-41.55	0.000	-8.78744	-7.995749
TEMPMED	.0582146	.0062798	9.27	0.000	.0459052	.0705241
PRECMIN	-.0213376	.0008798	-24.25	0.000	-.0230623	-.019613
PRECMAX	.0221477	.0008765	25.27	0.000	.0204297	.0238658
AREAU	-.0132797	.0003852	-34.48	0.000	-.0140348	-.0125247

Quadro 14 - Modelo de regressão para a amostra CS/Município/Trimestre/Incidência de internações por asma em idosos.

Source	SS	df	MS	Number of obs = 5737		
Model	2304.62362	7	329.231946	F(7, 5730)	=	452.98
Residual	4164.65061	5730	.726815116	Prob > F	=	0.0000
Total	6469.27423	5737	1.12764062	R-squared	=	0.3562
				Adj R-squared	=	0.3555
				Root MSE	=	.85253

incasma65	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
INTFOCOSBF10	.056195	.0078847	7.13	0.000	.040738	.071652
ALTITUDE	.0003689	.0000697	5.29	0.000	.0002322	.0005056
IDH	-2.963438	.240552	-12.32	0.000	-3.435011	-2.491865
TEMPMED	-.0149513	.0084425	-1.77	0.077	-.0315018	.0015992
PRECMIN	-.011926	.0009979	-11.95	0.000	-.0138823	-.0099697
PRECMAX	.0118972	.0009713	12.25	0.000	.0099932	.0138012
AREAU	-.0355542	.0018216	-19.52	0.000	-.0391252	-.0319831

Quadro 15 - Modelo de regressão para a amostra NE/Município/Trimestre/Incidência de internações por asma em idosos.

A análise dos Quadros 14 e 15 é semelhante ao que foi descrito para os Quadros 12 e 13, no que diz respeito aos testes p , R^2 e R^2 ajustado.

O modelo gerado para o NE tem menor poder de explicação da variabilidade da VD. A variável TEMPMED não apresentou significância estatística nesta análise.

As variáveis IDH e AREAU se mantiveram com correlação inversa corroborando as observações dos quadros anteriores, porém existem

variações numéricas que podem indicar pesos diferentes na influencia das variáveis nas duas regiões.

A variável ALTITUDE manteve comportamento semelhante ao observado para asma infantil, da mesma forma que as variáveis PRECMIN e PRECMAX.

Pode-se inferir, portanto, que as VI explicam a variabilidade de internação para asma em idosos de forma diferenciada nas regiões CS e NE.

Análise: Internações por enfisema, faixa etária de 65 anos ou mais.

Source	SS	df	MS	Number of obs = 5089		
Model	3983.54305	7	569.077579	F(7, 5082)	=	756.04
Residual	3825.28276	5082	.752712074	Prob > F	=	0.0000
Total	7808.82581	5089	1.53445192	R-squared	=	0.5101
				Adj R-squared	=	0.5095
				Root MSE	=	.86759

incenfi65	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
INTFOCOSBF10	.133667	.0094337	14.17	0.000	.1151729	.1521611
ALTITUDE	-.0001921	.0000804	-2.39	0.017	-.0003497	-.0000345
IDH	-10.99922	.2846487	-38.64	0.000	-11.55725	-10.44118
TEMPMED	.0659618	.0087045	7.58	0.000	.0488972	.0830263
PRECMIN	-.0256109	.0011541	-22.19	0.000	-.0278733	-.0233485
PRECMAX	.0265704	.0011499	23.11	0.000	.024316	.0288248
AREAU	-.0128651	.0003642	-35.32	0.000	-.0135791	-.0121512

Quadro 16 - Modelo de regressão para a amostra CS/Município/Mês/Incidência de internações por enfisema em idosos.

Source	SS	df	MS	Number of obs = 635		
Model	364.434237	7	52.0620339	F(7, 628)	=	108.17
Residual	302.261635	628	.481308337	Prob > F	=	0.0000
Total	666.695872	635	1.04991476	R-squared	=	0.5466
				Adj R-squared	=	0.5416
				Root MSE	=	.69376

incenfi65	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
INTFOCOSBF10	.109611	.0178193	6.15	0.000	.0746184	.1446037
ALTITUDE	.0002799	.0001679	1.67	0.096	-.0000498	.0006096
IDH	-9.584262	.6497386	-14.75	0.000	-10.86019	-8.308339
TEMPMED	-.013591	.0225275	-0.60	0.547	-.0578293	.0306474
PRECMIN	-.0276125	.0028882	-9.56	0.000	-.0332842	-.0219407
PRECMAX	.0273315	.0028442	9.61	0.000	.0217463	.0329167
AREAU	-.0182225	.0024261	-7.51	0.000	-.0229868	-.0134582

Quadro 17 - Modelo de regressão para a amostra NE/Município/Mês/Incidência de internações por enfisema em idosos.

A análise dos Quadros 16 e 17 reitera as observações anteriores sobre IDH e área urbanizada. De forma similar, a altitude mantém o padrão das análises anteriores, embora não tenha apresentado significância estatística no modelo para o NE.

Nos modelos anteriores, o número de observações consideradas para a região CS foi próximo do dobro, comparado ao NE. No caso dos modelos para enfisema este número foi dez vezes maior. Provavelmente porque o NE apresenta um reduzido número de internações por asma, se comparado ao CS, conforme apresentado anteriormente nos Gráficos 10 e 11.

De forma geral, todos os modelos apresentaram correlação positiva para a variável INTFOCOSBF10, indicando que as queimadas são importantes na explicação da variabilidade das taxas de internação hospitalar por asma e enfisema.

O mesmo ocorre com as variáveis IDH-M e área urbanizada, neste caso, de forma inversa.

No caso da variável ALTITUDE, observou-se correlação negativa em todos os modelos da região CS e positiva nos modelos da região NE. Porém, os coeficientes apresentados com números significativos na quinta casa decimal indicam pouca influência desta variável no conjunto observado.

De forma inversa, a variável TEMPMED apresentou correlação positiva nos modelos da região CS e negativa nos modelos da região NE. A correlação negativa era esperada para todos os modelos, visto que, as condições climáticas que acompanham temperaturas baixas propiciam a ocorrência de problemas respiratórios. Esta situação talvez possa ser explicada pelo uso das temperaturas médias anuais, que reduziram a variabilidade da curva e podem ter provocado viés nos resultados.

Por outro lado, utilizaram-se também as médias anuais para precipitações (mínima e máxima), e estas não se alteraram nos modelos, mantendo correlação negativa para precipitação mínima e positiva para precipitação máxima.

Pondera-se ainda que os resultados apresentados nos modelos de regressão podem conter imprecisões, uma vez que as variáveis de entrada representam contagens e geram um histograma de distribuição com curva log negativa, conforme apresentado anteriormente na Figura 10.

Embora as análises espaciais e estatísticas expliquem de forma parcial os fenômenos estudados, os modelos apresentam resultados que corroboram a hipótese desta tese.

5 CONCLUSÕES

Em se tratando de ciência, a construção do saber sempre conviveu com desafios. O contexto desta pesquisa não é diferente. Trata-se de uma área ampla, multidisciplinar, com limitações de dados disponíveis e carente de estudos conclusivos.

Retornando ao objetivo da pesquisa, a correlação entre queimadas e internações hospitalares por problemas respiratórios foi avaliada sob a perspectiva da análise espacial e da estatística. A análise espacial permitiu observar os agrupamentos mapeados para focos de calor e internações hospitalares e analisá-los à luz da dependência espacial. A análise estatística foi utilizada para explicar a variabilidade das internações hospitalares reunindo dados que poderiam traduzir resultados numéricos significativos.

São muitas as ferramentas que o Geoprocessamento provê para atender os mais diversos problemas, objetivando atribuir maior fidedignidade aos resultados que se deseja alcançar. No caso, a análise *Hotspot* apresentou resultados interessantes no apontamento das áreas mais impactadas pelas variáveis estudadas. Este fato foi claramente observado no estado de São Paulo, haja vista os agrupamentos de queimadas coincidentes às áreas onde a cultura canavieira é mais intensa.

Os agrupamentos observados indicaram as áreas de impacto nas meso-regiões de Araraquara, Bauru e Ribeirão Preto, bem como, as áreas de expansão nas meso-regiões de Araçatuba e Presidente Prudente. O estudo longitudinal no período de 2000 a 2008 permitiu observar o processo de expansão no sentido oeste paulista e a redução das queimadas em áreas onde a mecanização da lavoura está mais adiantada, como é o caso das áreas no entorno de Piracicaba.

Com relação aos agrupamentos de internações, outros fatores podem gerar os problemas respiratórios estudados, portanto, é natural que os agrupamentos não sejam totalmente similares aos de queimadas. Mesmo assim, os mapas temáticos mostraram que, quando não houve interseção, observou-se a proximidade dos limites poligonais dos agrupamentos.

O mini-mundo da pesquisa e as variáveis eleitas para representá-lo, são condicionantes da qualidade dos resultados. Assim, definir estes fatores depende de um bom conhecimento do próprio mini-mundo para uma escolha adequada. Esta escolha está associada a um nível de subjetividade que também está presente na análise dos resultados (MOITA NETO, 2004).

Pondera-se também, as dificuldades existentes no processo de transformação da informação em conhecimento. A ciência representa o mundo real através de modelos, os quais agregam uma fração daquilo que de fato ocorre. Portanto, os modelos se aproximam do mundo real, mas são limitados em sua representação.

De forma complementar, DRUCK et al (2004) ponderam que ao se desconsiderar a hipótese de dependência espacial, quando ela é verificada no fenômeno em estudo, as inferências estatísticas usuais perdem eficiência, refletindo em variâncias maiores e significância menor em testes de hipóteses, assim como, modelos estimados com ajustes piores comparados a dados da mesma dimensão que exibam dependência.

Nesta perspectiva as análises estatísticas corroboraram a explicação da variabilidade de internações por problemas respiratórios. Os modelos desenvolvidos demonstraram a influência das variáveis independentes sobre as dependentes, com destaque para as variáveis relacionadas com aspectos sócio-econômicos (IDH e áreas urbanizadas) e os focos de calor.

A partir da análise dos dados utilizados e outras informações obtidas durante a pesquisa, é razoável colocar algumas questões relacionadas ao tema.

Os dados de AIH apresentam limitações, ponderando-se os problemas encontrados para as internações por bronquite, bem como a tendência de queda destas internações. Este fenômeno pode ser explicado, em parte, pela migração dos pacientes para a rede privada, melhoria da qualidade de vida dos mesmos e a implantação do SUS nos municípios com ampliação dos atendimentos em UBS – Unidade Básica de Saúde e o PSF - Programa Saúde da Família.

As internações hospitalares por problemas respiratórios são eventos menos freqüentes comparados a outras causas. Isso contribuiu para muitas lacunas na base de dados, inviabilizando comparações por regressão. Mesmo assim, a ausência de dados sobre bronquite não estaria esclarecida.

Embora esta base de dados possa conter imprecisões no diagnóstico das doenças, a quantidade de eventos considerados no tempo e espaço suaviza este viés.

Fato análogo ocorre com os focos de calor. Além das dificuldades existentes e discutidas na aquisição deste dado, a comparação realizada com as comunicações de queimadas e focos de calor no estado de São Paulo mostrou que o uso desta base de dados deve ser ponderado com vistas à limitação do quanto ela representa a realidade do fenômeno.

A Secretaria de estado do Meio Ambiente de São Paulo disponibiliza a base de dados de comunicações de queimadas na internet por intermédio de consultas no sistema SIGAM, porém com limitações de acesso aos anos anteriores, impossibilitando análises comparativas. Soma-se a isso, o fato de que somente o estado de São Paulo possui dados de comunicações de queimadas em formato digital.

Mas ainda há a possibilidade do modelo construído ser incompleto para apresentar os resultados. Existe a carência de informações sobre a área de impacto de uma queimada. Os testes, considerando a situação

inversa, como foi feito com *buffers* de 10 e 20km, apresentaram resultados interessantes mas, carecem de estudos mais conclusivos.

Outros fatores gerados pelo processo de monocultura carecem de estudos mais aprofundados. Por exemplo, as áreas de plantio, após colheita e na fase de crescimento da planta, estão suscetíveis a ação do vento e geram imensas nuvens de poeira nesta época. Tais nuvens são relatadas na imprensa local, mas não se estima o impacto deste fenômeno na saúde da população.

Os dados climáticos também são limitados e não coletados em vários municípios. O mesmo é possível dizer dos dados de medição de poluentes atmosféricos. Embora existam iniciativas interessantes como o CATT-BRAMS do INPE, que faz estimativas de emissões, trata-se de modelagem computacional com *pixel* de 30km, ainda pouco preciso para estudos em escala local.

Cabe ainda ressaltar os esforços no sentido de pesquisar efeitos causados pelos compostos produzidos durante as queimadas, como o trabalho de PATERLINI (2007), entre outros. As associações entre queimadas e câncer ainda são muito primárias, mas estão na pauta de pesquisa, também convivendo com dificuldades na obtenção de dados, bem como, no isolamento das variáveis de confusão.

O Ministério da Saúde vem incentivando projetos nesta área, como é o caso do projeto VIGIAR, cujo objetivo é elaborar um sistema de alertas a eventos acentuados de queimadas, pois, sabe-se que quando eles ocorrem, a demanda por serviços de saúde aumenta nas localidades impactadas, tornando necessário ao gestor de saúde, ter informações em tempo real.

Para cumprimento do objetivo específico desta pesquisa, analisaram-se as regiões Centro-Sul e Nordeste separadamente, a fim de observar similaridades na relação queimadas e internações hospitalares em diferentes condições climáticas.

Neste aspecto, a análise transversal através de modelos de regressão permitiu, entre outros, observar as diferenças regionais. As variáveis temperatura e altitude apresentaram correlações invertidas para as regiões. Outros aspectos como população, densidade demográfica, clima são diferentes e interferem nas questões de saúde.

As variáveis temperatura média e precipitação (mínima e máxima) apresentaram correlações inversas na comparação entre as regiões estudadas. Este fato pode estar associado aos períodos de baixa umidade que são diferentes no Centro-Sul e no Nordeste.

A baixa umidade dificulta a dispersão dos poluentes atmosféricos. No Centro-Sul, o período seco corresponde ao inverno. No Nordeste, o inverno é chuvoso e o verão é o período de baixa umidade. Portanto, em ambas as regiões, safras e queimadas ocorrem em períodos de baixa umidade, contribuindo para o aumento do risco de problemas respiratórios.

O desenho longitudinal desenvolvido por intermédio da análise *Hotspot* indicou áreas críticas para queimadas e internações hospitalares e proveu a comparação visual das interseções e proximidades entre elas.

Por fim, desenho metodológico proposto foi executado, e as análises permitiram confirmação estatística e espacial sobre a correlação entre as variáveis, validando a hipótese deste trabalho.

Sendo assim, é razoável concluir que as queimadas contribuem com as internações hospitalares por problemas respiratórios como asma e enfisema. A variabilidade das taxas de internação está relacionada com aspectos sócio-econômicos e climáticos, porém, queimadas, aspectos sócio-econômicos e climáticos explicam parcialmente as taxas de internação. É certo que outras variáveis não analisadas nesta tese também exerçam influência nos fenômenos estudados.

Neste cenário, serão necessários novos estudos exploratórios com o intuito de avaliar novas variáveis, assim como a composição de bases de dados para análises e o aprimoramento de programas estatísticos mais adaptados para este campo de estudo.

Por fim, esta imersão possibilitou vivenciar as dificuldades de obter e manipular dados secundários para trabalhar com o instrumental da Estatística e do Geoprocessamento. Esta vivência ampliou a visão sobre este contexto e a necessidade de continuidade nas pesquisas, sobremaneira da construção de novas fontes primárias de dados para melhor subsidiar as análises aqui propostas, na busca por respostas ainda não encontradas.

6 REFERÊNCIAS

ALMEIDA FILHO, N.; ROUQUAYROL, M.Z. **Introdução à Epidemiologia**. 3 ed. Rio de Janeiro: MedSi, 2002.

ALVARENGA, A.T. A Saúde Pública como campo de investigação interdisciplinar e a questão metodológica. **Saúde e Sociedade** n. 3(2), p. 23-41, 1994. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-12901994000200003&lng=pt&nrm=iso>. Acesso em: 23 maio 2009. doi: 10.1590/S0104-12901994000200003.

ANDREAE, M.O. et al. Smoking Rain Clouds over de Amazon. **Science Magazine** n. 303, p. 1337-1342, 2004.

APPOLINÁRIO, F. **Metodologia da Ciência: Filosofia e Prática da Pesquisa**. São Paulo: Pioneira Thompson Learning, 2006.

ARBEX, M.B. et al. Assessment of the effects of sugarcane plantation burning on daily counts of inhalation therapy. **J Air Waste Manag Assoc.** n. 50(10), p. 1745-1749, 2000.

ARBEX, M.B. Avaliação dos efeitos do material particulado proveniente da queima da plantação de cana-de-açúcar sobre a morbidade respiratória na população de Araraquara – SP. 2001. 188p. Tese (Doutorado em Medicina) – Faculdade de Medicina da Universidade de São Paulo, São Paulo.

ARBEX, M.B. et al. Queima de biomassa e efeitos sobre a saúde. **Jornal Brasileiro de Pneumologia** n. 30(2), p. 158 – 175, Mar/Abr 2004.

ARCGIS 9.2 Desktop Help. Disponível em < <http://webhelp.esri.com/arcgisdesktop/9.2>>. Acesso em: 23 ago 2009.

BAHIA. Decreto nº 10.936 de 27 de fevereiro de 2008. Dispõe sobre o tratamento tributário nas operações com álcool etílico hidratado e anidro combustível e dá outras providências. **Diário oficial do Estado da Bahia**, Bahia, 28 fev 2008.

BAILEY, T.C. Spatial statistical methods in health. **Cadernos de Saúde Pública** n. 17(5), p. 1083 – 1098, 2001.

BARCELLOS, C.; BASTOS, F.I. Geoprocessamento Ambiente e Saúde: Uma união possível? **Cadernos de Saúde Pública** n. 12(3), p. 389 – 397, 1996.

BERQUÓ, E.S. et al. **Bioestatística**. 2. ed. São Paulo: E.P.U. 2002.

BERTALANFFY, L.V. **Teoria Geral dos Sistemas**. São Paulo: Ed. Vozes, 1975.

BOUSQUAT, A; COHN, A. dimensão espacial nos estudos sobre saúde: uma trajetória histórica. **História, Ciências, Saúde – Manguinhos** n.11(3), p. 549-568, 2004.

BRASIL. Decreto nº 6.539, de 18 de agosto de 2008. Estabelece critérios para o enquadramento de projeto de instalação, diversificação ou modernização de empreendimento, para efeito de redução de imposto sobre renda. **Diário Oficial da União**, Brasília, DF, 19 ago 2008; Seção 1, p.1.

BRAUER, M. Health Impacts of Biomass Air Pollution. **Health Guidelines for Vegetation Fire Events: Background Papers**. WHO – World Health Organization, 1999. Disponível em <http://www.who.int/docstore/peh/Vegetation_fires/vegetationfirbackgrtoc.htm> Acesso em: 27 junho 2009.

BVS – Biblioteca Virtual em Saúde. Disponível em <<http://regional.bvsalud.org/php/index.php>>. Acesso em: 15 dezembro 2009.

CAMARA, G. et al. **Introdução à Ciência da Geoinformação**. INPE, 2001. Disponível em <mtc-m12.sid.inpe.br/col/sid.inpe.br/sergio/2004/04.22.07.43/doc/publicacao.pdf>. Acesso em: 23 maio 2009.

CANASAT - Mapeamento da cana via imagens de satélite de observação da Terra. Disponível em <<http://www.dsr.inpe.br/canasat/>>. Acesso em: 12/maio 2009.

CANÇADO, J.E.D. et al. Repercussões clínicas da exposição à poluição atmosférica. **J Bras Pneumol**. n.32(Supl 1), p. S5 – S11, 2006.

CARVALHO, M.S. **Aplicação de métodos de análise espacial na caracterização de áreas de risco a saúde**. 1997. 138 p. Tese (Doutorado em Ciências) – COPPE da Universidade Federal do Rio de Janeiro, Rio de Janeiro.

CARVALHO, M.S; SOUZA-SANTOS, R. Análise de dados espaciais em saúde pública: métodos problemas e perspectivas. **Cadernos de Saúde Pública** n. 21(2), p. 361-378, 2005.

CASTRO, H.A. et al. Questões metodológicas para a investigação dos efeitos da poluição do ar na saúde. **Revista Brasileira de Epidemiologia** n. 6(2): p. 135-149, 2003.

CERVO, A.L.; BERVIAN, P.A.; DA SILVA, R. **Metodologia Científica**. 6 ed. São Paulo: Pearson Prentice Hall, 2007.

CETESB – Companhia de Tecnologia de Saneamento Ambiental. **Avaliação dos compostos orgânicos provenientes da queima de palha de cana-**

de-açúcar na região de Araraquara e comparação com medições efetuadas em São Paulo e Cubatão: Relatório Final- 2002,: relatório técnico. São Paulo: CETESB, 2002. 97 p.

CETESB – Companhia de Tecnologia de Saneamento Ambiental. Qualidade do Ar - Informações. Disponível em <http://www.cetesb.sp.gov.br/Ar/ar_saude.asp>. Acesso em: 31 maio 2009.

CHRISTOFOLETTI, A. **Modelagem de Sistemas Ambientais**. São Paulo: Editora Edgard Blücher LTDA., 1999.

CONAB – Companhia Nacional de Abastecimento. **Acompanhamento da Safra Brasileira Cana-de-Açúcar, Safra 2007/2008, Terceiro Levantamento**.: relatório técnico. Brasília: CONAB, 2007. 13 p. Disponível em <<http://www.conab.gov.br/conabweb/download/safra/3lev-cana.pdf>>. Acesso em: 06 abril 2009.

CORREA, S.M.B.B. **Probabilidade e Estatística**. 2.ed. Belo Horizonte: PUC Minas Virtual, 2003.

COSTA, L.S., PINHEIRO, R.S., ALMEIDA, R.V.R. Recuperação de informação de endereço na base de dados AIH: percentual de Aproveitamento. **Cadernos de Saúde Coletiva** n.11(2), p. 1311-141, 2003.

COSTA, M.C.N.; TEIXEIRA, M.G.L.C. A concepção de “espaço” na investigação epidemiológica. **Cadernos de Saúde Pública** n.15(2), p. 271-279, 1999.

CROMLEY, E.K.; McLAFFERTY, S.L. **GIS and Public Health**. New York: The Guilford Press, 2002.

CZERESNIA, D.; RIBEIRO, A.M. O conceito do espaço em epidemiologia: uma interpretação histórica e epistemológica. **Cadernos de Saúde Pública** n. 16(3), p.595-617, 2000.

DATASUS. Departamento de Informática do SUS. Disponível em <www.datasus.gov.br>. Acesso em: 25 ago 2009.

DESCARTES, R. **Discurso do Método**. Tradução P. Neves. Porto Alegre: L&PM Editores, 2004.

DIAS, T.L. et al. Problemas de escala e a relação área-indivíduo em análise espacial de dados censitários. **Informática Pública**, n. 4(1):89, p.104, 2002 Disponível em <http://www.dpi.inpe.br/gilberto/papers/ip_falacia.pdf>. Acesso em: 23 maio 2009.

DIEHL, A.A.; TATIM, D.C. **Pesquisa em ciências sociais aplicadas**. São Paulo: Pearson Pretence Hall, 2004.

DIGGLE, P.J.; RIBEIRO Jr., P.J. **Model-Based Geostatistics**. Caxambu: Associação Brasileira de Estatística, 2000.

DIGGLE, P.J. et al. An Introduction to Model-Based Geostatistics. In: MOLLER, J.(Coord.). **Spatial statistics and computational methods**. New York: Springer-Verlag, 2003.

DOCKERY, D.W.; POPE III, C.A. Acute respiratory effects of particulate air pollution. **Annual Reviews Public Health**, n. 15, p. 107-32, 1994.

DODGE, Y. **The Concise Encyclopedia of Statistics**. Swartzland: Springer, 2008.

DRUCK, S. et al (editores). **Análise espacial de dados geográficos**. Planaltina: Embrapa Cerrados, 2004.

DUCHIADE, M.P. Poluição do Ar e Doenças Respiratórias: Uma Revisão. **Cadernos de Saúde Pública**, n. 8(3), p. 311-330, julho/setembro, 1992.

DUQUE, J.C et al. The ecological fallacy in a time series context: evidence from Spanish regional unemployment rates. **J Geograph Syst**, n.8, p.391-410, 2006.

ELMASRI, R.E.; NAVATHE, S. **Sistemas de Banco de Dados**. 4 ed. São Paulo: Addison Wesley 2005.

EMBRAPAA – Empresa Brasileira de Pesquisa Agropecuária. Banco de Dados Climáticos do Brasil. Disponível em <<http://www.bdclima.cnpm.embrapa.br/>>. Acesso em 30 ago 2009.

EMBRAPAb – Empresa Brasileira de Pesquisa Agropecuária. Mapeamento e Estimativa da Área Urbanizada do Brasil. Disponível em <<http://www.urbanizacao.cnpm.embrapa.br/>>. Acesso em 20 set 2009.

FERREIRA, M.U. Epidemiologia e Geografia: O complexo patogênico de Max. Sorre. **Cadernos de Saúde Pública** n. 7(3), p. 301-309, 1991.

FREITAS, S.R. et al. The coupled aerosol and tracer transport model to the Brazilian developments on the regional atmospheric modeling system (CATT-BRAMS) – Part 1: Model description and evaluation. **Atmospheric Chemistry and Physics Discussions** n. 7, p. 8525 – 8569, 2007.

FUNASA - Fundação Nacional de Saúde. Vigilância ambiental em Saúde. Brasília 2000. Disponível em <www.funasa.gov.br/amb/amb00.htm>. Acesso em: 21 fev 2007.

GATRELL, A. **Geographies of Health: an introduction**. Oxford: Blackwell Publishers Inc., 2002.

GIL, A.C. **Métodos e Técnicas de Pesquisa Social**. 4 ed. São Paulo: Atlas, 1994.

GLOAG, D. Air Pollution: The “classical” pollutants. **British Medical Journal** n. 282, p. 723 – 725, 1981.

GOLDEMBERG J. Uma nova OPEP. **O Estado de São Paulo**. São Paulo, 18 jan 2005.

GOOGLE – Site de Buscas na Internet. Disponível em <<http://www.google.com.br/>>. Acesso em: 15 dezembro 2009.

GOUVEIA, N.; HAJAT, S.; ARMSTRONG, B. Socioeconomic differentials in the temperature-mortality relationship in São Paulo, Brasil. **Int J. Epidemiology**, n.32: p.390-397, 2003. Disponível em <<http://ije.oxfordjournals.org/cgi/content/full/32/3/390>>. Acesso em : 11 jan 2010.

GRAYBILL, F.A.; IYER, H.K.; BURDICK, R.K. **Applied Statistics: A first course inference**. New Jersey: Prentice-Hall Inc., 1998.

HAINING, R. **Spatial Data Analysis: Theory and Practice**. Cambridge: Cambridge University Press, 2004.

HAIR Jr, J.F. et al. **Análise Multivariada de Dados**. 5 ed. São Paulo: Bookman, 2005.

IBAMA – Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis. **Temas Vetoriais Formato Shapefile – médias anuais de temperaturas e precipitações**. Disponível em <<http://siscom.ibama.gov.br/shapes/>>. Acesso em: 10 ago 2009.

IBGE – Instituto Brasileiro de Geografia e Estatística. **Malha Municipal Digital do Brasil - 2007**. Brasília, 2007. Disponível em <ftp://geoftp.ibge.gov.br/mapas/malhas_digitais/municipio_2007/Malha_Municipal_Digital_2007_2500/>. Acesso em 20 ago 2009.

IBGE - Instituto Brasileiro de Geografia e Estatística. Disponível em: <<http://www.ibge.gov.br/home/>>. Acesso em: 25 ago 2009.

IEA-Instituto de Economia Agrícola. Produção, área plantada e área colhida de cana-de-açúcar, por ano de safra, Brasil. Disponível em <<http://ciagri.iea.sp.gov.br/bioenergia/html/ibge20072008canabr.htm>>. Acesso em: 29 jun 2009.

IGNOTTI, E. et al. Efeitos das queimadas na Amazônia: método de seleção dos municípios segundo indicadores de saúde. **Revista Brasileira de Epidemiologia** n. 10(4), p. 453-464, 2007.

INÁCIO, J. A nova era da cana-de-açúcar. Agrinova Web. Disponível em <www.agrinovaweb.com.br/agw_imprimir_vxlpub?codnoticia=77521>. Acesso em: 27 jan 2005.

INPE – Instituto Nacional de Pesquisas Espaciais. Disponível em <<http://www.inpe.br/>> . Acesso em: 20 fev 2009.

LACAZ, C. **Introdução à Geografia Médica do Brasil**. São Paulo: Editora E. Blücher, 1972.

LAGE, J.; ZAFALON, M. País precisa de cem novas usinas de etanol até 2010. **Folha de São Paulo**, 07/02/2007. Disponível em <www.tools.folha.com.br>. Acesso em: 07 fev 2007.

LE MONDE DIPLOMATIQUE BRASIL. **Atlas do Meio Ambiente**. Instituto Polis, 2009.

LEME, R.M. **Estimativa das emissões de poluentes atmosféricos e uso de água na produção de eletricidade com biomassa de cana-de-açúcar**. 2005. Dissertação (Mestrado em Engenharia Mecânica) – Faculdade de Engenharia Mecânica da Universidade Estadual de Campinas, Campinas.

LEMOS, J.C.; LIMA, S.C. A geografia médica e as doenças infecto-parasitárias. **Caminhos da Geografia** – Revista on-line, n. 3(6), p. 74-86, jun. 2002. Disponível em: <http://www.ig.ufu.br/revista/volume06/artigo05_vol06.pdf>. Acesso em: 21 fevereiro 2007.

LOCH, R.E.N. **Cartografia: representação, comunicação e visualização de dados espaciais**. Florianópolis: Ed. da UFSC, 2006.

LOPES, F.S. **A utilização de sistemas de informação geográfica no estudo da exposição humana aos produtos da queima da palha de cana-de-açúcar (*Saccharum sp*) no estado de São Paulo**. 2005. 127p. Dissertação (Mestrado em Saúde Pública) – Faculdade de Saúde Pública da Universidade de São Paulo, São Paulo.

LOPES, F.S. **Uma análise sobre fontes primárias para aquisição de dados de queimadas no estado de São Paulo**. Anais do XIV Simpósio Brasileiro de Sensoriamento Remoto - SBSR. 2009.

LOPES, F.S.; RIBEIRO, H. Mapeamento de internações hospitalares por problemas respiratórios e possíveis associações à exposição humana aos produtos da queima da palha de cana-de-açúcar no estado de São Paulo. **Revista Brasileira de Epidemiologia**, n. 9(2), p. 215-225, junho 2006.

LOPES, F.S.; RIBEIRO, H. A evolução do uso de ferramentas de geoprocessamento em estudos sobre queima de biomassa em ambientes outdoor e efeitos adversos à saúde humana. In: Anais do II International

Congress of Geography Health e IV Simpósio Nacional de Geografia da Saúde. Uberlândia, 2009.

MacEACHREN, A.M. **How Maps Work – Representation, Visualization and Design**. New York: Guilford Press, 2004.

MASCARENHAS, M.D.M. et al. Poluição atmosférica devida à queima de biomassa florestal e atendimentos de emergência por doença respiratória em Rio Branco, Brasil – Setembro, 2005. **J Bras Pneumol**. n. 34(1), p. 42-46, 2008.

MEDLINE – Base de Dados de literatura internacional na área médica. Disponível em <<http://basesg4.bireme.br/cgi-bin/wxislind.exe/iah/online/?IscScript=iah/iah.xis&base=MEDLINE&lang=p>>. Acessado em: 15 dezembro 2009.

MENDONÇA, F.; DANNI-OLIVEIRA, I.M. **Climatologia: Noções básicas e climas do Brasil**. São Paulo: Oficina de Textos, 2007.

MICHAELIS. Moderno Dicionário da Língua Portuguesa. Disponível em <[HTTP://michaelis.uol.com.br](http://michaelis.uol.com.br)> Acesso em: 15 dez 2008.

MICROSOFT – SQL Server. Disponível em < <http://www.microsoft.com/brasil/servidores/sql/2005/default.msp>>. Acesso em 28 ago 2009.

MOITA NETO, J.M. Estatística Multivariada: Uma didática-metodológica. **Crítica: Revista de Filosofia**. 9 mai 2004. Disponível em <http://criticanarede.com/cien_estatistica.html> Acesso em: 24 out 2009.

MORAN, F. **Adaptabilidade Humana**. São Paulo: EDUSP. 1994.

MS – Ministério da Saúde. **Introdução à Estatística Espacial para a Saúde Pública. Série B. Textos Básicos de Saúde. Série Capacitação e Atualização em Geoprocessamento em Saúde; nº 3**. Brasília: Ministério da Saúde, 2007.

MORIN, E. **O Método 1. A natureza da natureza**. 2 ed. Portugal: Publicações Europa-América, LDA., 1977.

MÜLLER, E.P.L. et al. **Sistemas de Informação Geográfica em Políticas Públicas**. Disponível em <www.sbis.org.br/cbis/arquivos/1002.pdf>. Aceso em 21 fevereiro 2007.

NOGUEIRA, E.A. et al (organizadores). **Modelos Matemáticos nas Ciências não-Exatas**. São Paulo: Editora Blucher, 2008.

O ESTADO DE SÃO PAULO. OMS diz que poluição do ar mata 2 milhões por ano, **Estado de São Paulo**, São Paulo 5 out 2006. Ciência e Meio Ambiente. Disponível em <www.estadao.com.br>. Acesso em: 07 fev 2007.

OPAS - Organização Pan-Americana da Saúde. **Sistema de Informação Geográfica em Saúde: Conceitos Básicos**. Brasília: OPAS, 2002.

OPAS - Organização Pan-Americana da Saúde. **Evaluación de los efectos de La contaminación del aire en la salud de América latina y el Caribe**. Washington, DC: OPS, 2005.

OPAS - Organização Pan-Americana da Saúde. Saúde Coletiva. Promoção da Saúde / Ambientes Saudáveis. Disponível em <www.opas.org.br/coletiva/temas.cfm?_id=16&area=Conceito>. Acesso em: 23 janeiro 2007.

PATERLINI, W. C. **Fontes e composição das partículas atmosféricas na área urbana e rural da região central do estado de São Paulo**. 2007. 168 p. Tese (Doutorado em Química) – Instituto de Química da Universidade Estadual Paulista, Araraquara.

PAWITAN, G.; STEEL, D.G. Exploring a Relationship Between Aggregate and Individual Levels Spatial Data Through Semivariogram Models. **Geographical Analysis**, n.38, p. 310-325, maio 2006.

PEREIRA, J.C.R. **Análise de dados qualitativos: Estratégias Metodológicas para as Ciências da Saúde, Humanas e Sociais**. São Paulo: Edusp, 1999.

PNUD – Programa das Nações Unidas para o desenvolvimento. Desenvolvimento Humano e IDH. Disponível em <www.pnud.org.br/idh>. Acesso em: 15 set 2009.

PORTELA, M.C. et al. Algoritmo para a composição de dados por internação hospitalar a partir do sistema de informações hospitalares do sistema único de saúde (SIH/SUS) - Composição dos dados por internação a partir do SIH/SUS. **Cadernos de Saúde Pública**, n. 13(4), p. 771-774, out-dez 1997.

RATTO O.R.; SANTOS, M.L.; BOGOSSIAN, M. **Insuficiência Respiratória**. Rio de Janeiro: Livraria Atheneu, 1981.

RIBEIRO, F.S.N.; WÜNSCH FILHO, V. Avaliação retrospectiva da exposição ocupacional a cancerígenos: abordagem epidemiológica e aplicação em vigilância em saúde. **Cadernos de Saúde Pública**, n. 20(4), p. 881-890, ago 2004.

RIBEIRO, H. Sugar cane burning in Brazil: respiratory health effects. **Revista de Saúde Pública**, n. 42(2), p. 370-376, fev 2008.

RIBEIRO, H.; ASSUNÇÃO, J.V. Efeitos das queimadas na saúde humana. **Estudos Avançados**, n. 16(44), p. 125-148, 2002.

RYTKÖNEN, M.J.P. Not all maps are equal: GIS and spatial analysis in epidemiology. **International Journal of Circumpolar Health**, n. 63(1), p. 9-24, 2004.

SAMET, J.M.; SPEIZER, F.E. Assessment of Health Effects in Epidemiologic Studies of Air Pollution. **Environmental Health Perspectives Supplements**, n. 101, Supplement 4, Dez 1993.

SANTOS, M. **A Natureza do Espaço**. São Paulo: Edusp, 2002.

SÃO PAULO (Estado). Lei nº 11.879 de 19 de janeiro de 2005. Dispõe sobre a criação da Política Estadual de incentivo às microdestilarias de álcool e beneficiamento de produtos derivados da cana-de-açúcar e dá outras providências. **Diário Oficial do Estado de São Paulo**, São Paulo, 20 de janeiro de 2005. Seção 1, p. 1.

SCIELO – Scientific Electronic Library Online. Disponível em: <<http://www.scielo.org/php/index.php>>. Acesso em: 15 dezembro 2009.

SEVERINO, A.J. **Metodologia do Trabalho Científico**. São Paulo: Editora Cortez, 2002.

SIQUEIRA, F. Justiça condena SP a pagar multa por queimada. **O Estado de São Paulo**, São Paulo, 17 ago 2009. Disponível em <www.estadao.com.br/estadaodehoje/20090817/not_imp419809,0.php>, Acesso em 9 nov 2009.

STOCKBURGER, D.W. Introductory Statistics: **Concepts, Models, and Applications**. www version 1.0, 1996. Disponível em <www.psychstat.missouristate.edu/introbook/sbk00.htm>, Acesso em: 16 junho 2009.

ÚNICA. Portal da União da Indústria de Cana-de-Açúcar. Disponível em <www.unica.com.br>. Acesso em: 31 março 2009.

VERAS, C.M.T.; MARTINS, M.S. A confiabilidade dos dados nos formulários de Autorização de Internação Hospitalar (AIH), Rio de Janeiro, Brasil. **Cadernos de Saúde Pública**, n. 10(3), p. 339-355, set 1994.

WEB OF KNOWLEDGE – Plataforma de acesso a artigos em Ciências, Ciências Sociais, Artes e Humanidades. Disponível em: <<http://www.isiwebknowledge.com/>>. Acesso em: 15 dezembro 2009.

WHO – World Health Organization. The World Health Report. Reducing Risk Promoting Healthy Life. Relatório Técnico. WHO, 2002.

WORBOYS, M.F. **GIS: A Computing Perspective**. London: Taylor & Francis. 1995.

ZANCUL, A. **O efeito da queimada de cana-de-açúcar na qualidade do ar de região de Araraquara.** 1998. 99p. Dissertação (Mestrado em Saneamento) - Escola de Engenharia de São Carlos da Universidade de São Paulo, São Carlos.

APENDICE A:

DICIONÁRIO DE DADOS DA ESTRUTURA DOS ARQUIVOS “RD” (REDUZIDAS) – AIH – DATASUS.

Observação 1: O campo “Descrição” apresenta observações sobre mudanças ocorridas na estrutura do arquivo ao longo dos anos. Estes alterações foram identificadas neste trabalho, pois dificultam a unificação dos dados em uma única base.

Observação 2: os campos cujo nome está grafado com um asterisco no final, indica que o campo foi selecionado para compor a base de dados deste estudo.

Observação 3: determinados campos possuem códigos para preenchimento. Quando for o caso, o campo Descrição possui um informativo “Vide Tabela”. As tabelas com as opções estão descritas após este dicionário.

Campo	Tipo	Descrição
UF_ZI*	Texto (6)	Código da unidade da federação (35 = Estado de São Paulo). Inicialmente eram 2 dígitos, mudou o formato para 6 dígitos em 2006 e o dado passou a ser município gestor.
ANO_CMPT*	Texto (4)	Ano de competência da AIH, no formato aaaa.
MÊS_CMPT*	Texto (2)	Mês de competência da AIH, no formato mm.
ESPEC*	Texto (2)	Código da especialidade (vide tabela).
CGC_HOSP*	Texto (14)	CNPJ do Hospital.
N_AIH*	Texto (13)	Número da AIH.
IDENT*	Texto (1)	Identificação do tipo de AIH (vide tabela).
CEP*	Texto (8)	Código de CEP de residência do paciente.
MUNIC_RES*	Texto (6)	Código IBGE do município de residência do paciente.
NASC*	Data	Data de nascimento do paciente.
SEXO*	Texto (1)	Sexo do paciente (vide tabela).
UTI_MES_IN	Número	Quantidade de dias de UTI no mês inicial da internação.
UTI_MES_AN	Número	Quantidade de dias de UTI no mês anterior à alta.
UTI_MES_AL	Número	Quantidade de dias de UTI no mês da alta.
UTI_MES_TO	Número	Quantidade de atos – UTI.
MARCA_UTI	Texto (2)	Indica qual o tipo de UTI utilizado pelo paciente da AIH (vide tabela).
UTI_INT_IN	Número	Dias UTI, mês inicial (tipo ato=34 ou 56).

Campo	Tipo	Descrição (continuação...)
UTI_INT_AN	Número	Dias em UTI no mês anterior à alta (tipo ato = 34 ou 56).
UTI_INT_AL	Número	Dias em UTI no mês da alta (tipo ato = 34 ou 56).
UTI_INT_TO	Número	Total de dias em UTI (tipo ato = 34 ou 56)
DIAR_ACOM	Número	Quantidade de diárias de acompanhante.
PROC_SOLIC	Texto (10)	Código do procedimento solicitado. Obs. Acrescentado em 04/2006.
PROC_REA*	Texto (10)	Código do procedimento realizado. Mudou a tamanho de 8 para 10 em 04/2006.
VAL_SH	Número	Valor de serviços hospitalares.
VAL_SP	Número	Valor de serviços profissionais prestados por terceiros.
VAL_SADT	Número	Valor de SADT (serviços auxiliares de diagnose e terapia)
VAL_RN	Número	Valor de recém-nato (a partir de dezembro de 1994)
VAL_ACOMP	Número	Valor de acompanhante.
VAL_ORTP	Número	Valor de órtese e prótese
VAL_SANGUE	Número	Valor gasto com sangue; nos arquivos de 1992, este valor está somado em VAL_SADT
VAL_SADTSR	Número	Valor referente a tomografias e ressonância nuclear magnética
VAL_TRANSP	Número	Valor referente a transplantes (retirada de órgãos), incluindo: taxa de sala cirúrgica (SH), retirada de órgão (SP), exames no cadáver (SADT), avaliação auditiva (SADT).
VAL_OBSANG	Número	Valor de anestesia.
VAL_PED1AC	Número	Valor de pediatria.
VAL_TOT	Número	Valor total da AIH.
VAL_UTI	Número	Valor referente aos gastos em UTI.
US_TOT	Número	Valor total da AIH (VAL_TOT) em dólar.
DT_INTER*	Data	Data de internação no formato aaammdd
DT_SAIDA*	Data	Data de saída, no formato aaaammdd
DIAG_PRINC*	Texto (4)	Diagnóstico principal, segundo a CID-10 Tabela de Classificação Internacional de Doenças
DIAG_SECUN*	Texto (4)	Diagnóstico secundário, segundo a CID-10
COBRANCA	Texto (2)	Motivo da cobrança (vide tabela).
NATUREZA	Texto (2)	Natureza da relação do hospital com o SUS (Vide Tabela)
GESTAO	Texto (1)	Indica o tipo de gestão do hospital (vide tabela).

Campo	Tipo	Descrição (continuação...)
RUBRICA	Número	Número da rubrica. Obs. Acrescentado em 04/2006.
IND_VDRL	Texto (1)	Indica exame VDRL (SIM = 1). Obs. Acrescentado em 03/2006.
MUNIC_MOV	Texto (6)	Código IBGE do município onde se localiza o do hospital.
COD_IDADE*	Texto (1)	Unidade de medida da idade (campo complementar ao IDADE) (vide tabela).
IDADE*	Número	Idade, campo complementar ao "COD_IDADE"
DIAS_PERM*	Número	Dias de permanência no hospital
MORTE*	Número	Indica se o paciente teve saída com morte (vide tabela).
COD_ARQ	Texto (1)	Código identificador do registro reduzido de AIH. Foi retirado em 05/2006.
CONT	Número	Foi retirado em 05/2006.
NACIONAL	Texto (2)	Nacionalidade do paciente (vide tabela).
NUM_PROC	Texto (4)	Número do processamento.
CAR_INT	Texto (2)	Caráter da internação, (vide tabela)
TOT_PT_SP	Número	Número de pontos de Serviços Profissionais nesta AIH.
CPF_AUT	Texto (11)	CPF do auditor que autorizou o hormônio.
HOMONIMO	Texto (1)	Marca de hormônio (0=não / 1=sim).
NUM_FILHOS	Número	Número de filhos (Obs. – 0 inclui casos sem informação)
INSTRU*	Texto (1)	Grau de instrução (vide tabela)
CID_NOTIF	Texto (4)	CID de indicação para realização de laqueadura. Não utilizado em caso de vasectomia.
CONTRACEP1	Texto (2)	Tipo de contraceptivo utilizado (vide tabela)
CONTRACEP2	Texto (2)	Segundo tipo de contraceptivo utilizado
GESTRISCO	Texto (1)	Indicador se é gestante de risco (0=não / 1=sim).
INSC_PN	Texto (10)	Número da gestante no pré-natal. Obs. Acrescentado em 01/2001. Alteração de tipo de dado para Float em 07/2004.
SEQ_AIH5	Texto (3)	Sequencial da AIH tipo 5 – Longa permanencia. Obs. Acrescentado em 01/2001. Alteração de tipo de dado para Smallint em 07/2007.
CBOR	Texto (3)	Código da CBO – Classificação brasileira de ocupações. Obs. Acrescentado em 01/2002. Mantido até 12/2005.

Campo	Tipo	Descrição (continuação...)
CNAER	Texto (3)	Código de acidente de trabalho. Obs. Acrescentado em 01/2002. Mantido até 12/2005.
VINCPREV	Texto (1)	Vínculo com a previdência. Obs. Acrescentado em 01/2002. Mantido até 12/2005. Depois Reaparece somente para 04/2006.
GESTOR_COD	Texto (2)	Código da autorização de pagamento. Obs. Acrescentado em 08/2002
GESTOR_CPF	Texto (11)	Número do CPF do gestor. Obs. Acrescentado em 08/2002
GESTOR_DT	Data	Data da autorização dada pelo gestor. Obs. Acrescentado em 08/2002
CNES	Texto (7)	Código de identificação do Cadastro nacional de estabelecimentos de saúde. Obs. Acrescentado em 08/2003.
CGC_MANT	Texto (14)	CNPJ da entidade mantenedora. Obs. Acrescentado somente em 07/2004
COD_SEG	Texto (8)	Obs. Acrescentado em 07/2004. Mantido até 03/2006.
INFEHOSP	Texto (1)	Status de infecção hospitalar (0=não / 1=sim). Obs. Acrescentado em 04/2006
CID_ASSO	Texto (4)	CID causa. Obs. Acrescentado em 04/2006
CID_MORTE	Texto (4)	CID da morte. Obs. Acrescentado em 04/2006

TABELAS AUXILIARES:

ESPECIALIDADES DA AIH – Campo: ESPEC

- 1- Cirurgia geral
- 2- Obstetrícia
- 3- Clínica médica
- 4- Crônico e FPT (fora de possibilidade terapêutica)
- 5- Psiquiatria
- 6- Tisiologia
- 7- Pediatria
- 8- Reabilitação
- 9- Psiquiatria- hospital/dia

IDENTIFICAÇÃO DA AIH – Campo: IDENT

- 1- AIH normal
- 3- AIH de continuação
- 5- AIH de longa permanência e FPT (Fora de Possibilidade Terapêutica)

SEXO DO PACIENTE – Campo: SEXO

- 0- Ignorado
- 1- Masculino
- 3- Feminino

MARCA UTI – Campo: MARCA_UTI

- 0- Leito sem especialidade ou não utilizou UTI
- 1- UTI adulto nível II
- 2- UTI adulto nível III
- 3- UTI neonatal nível III
- 4- UTI neonatal nível II
- 5- UTI pediátrica nível II
- 6- UTI pediátrica nível III
- 7- Transplante pediátrico
- 8- Transplante adulto

Motivo da cobrança - Campo: COBRANCA

Alta:

- 11- Curado
- 12- Melhorado
- 13- Inalterado
- 14- A pedido
- 15- Internado para diagnóstico
- 16- Administrativa
- 17- Indisciplina
- 18- Alta por evasão
- 19- Complementação de tratamento em ambulatório

Permanência maior que 30 dias:

- 21- Características próprias da doença
- 22- Intercorrência
- 23- Motivo social
- 24- Doença crônica

25- Impossibilidade de vivência sócio-familiar

Transferência:

31- Tisiologia

32- Psiquiatria

33- Clínica médica

34- Cirurgia

35- Obstetrícia

36- Berçário

37- Pediatria

38- Isolamento

39- Outros (UTI, Medicina física, Radioterapia etc.)

Óbito com autopsia:

41- Ocorrido nas primeiras 48 horas após a internação em estado agônico ou pré-agônico

42- Ocorrido nas primeiras 48 horas após a internação em estado não agônico ou pré-agônico

43- Ocorrido a partir de 48 horas de internação

Óbito sem autopsia:

51- Ocorrido nas primeiras 48 horas após a internação em estado agônico ou pré-agônico

52- Ocorrido nas primeiras 48 horas após a internação em estado não agônico ou pré-agônico

53- Ocorrido a partir de 48 horas de internação

Alta por reoperação:

61- Em politraumatizado com menos de 24 horas da 1ª cirurgia

62- Em politraumatizado 24 a 48 horas após 1ª cirurgia

63- Em politraumatizado 48 a 72 horas após 1ª cirurgia

64- Em politraumatizado acima de 72 horas após 1ª cirurgia

65- Em cirurgia de emergência com menos de 24 horas da 1ª cirurgia

66- Em cirurgia de emergência 24 a 48 horas após 1ª cirurgia

67- Em cirurgia de emergência 48 a 72 horas após 1ª cirurgia

68- Em cirurgia de emergência acima de 72 horas após 1ª cirurgia

Alta da parturiente com permanência do recém nascido:

71- Alta da parturiente com permanência do recém nascido

Natureza da relação entre o hospital e o SUS – Campo: NATUREZA

10- Hospital próprio (extinto INAMPS)

20- Hospital contratado

22- Hospital contratado, optante pelo SIMPLES

30- Hospital federal

- 31- Hospital federal com verba própria (apresenta AIH, mas não recebe por elas)
- 40- Hospital estadual
- 41- Hospital estadual com verba própria (apresenta AIH, mas não recebe por elas)
- 50- Hospital municipal
- 60- Hospital filantrópico
- 61- Hospital filantrópico, isento de tributos e contribuições sociais
- 63- Hospital filantrópico, isento de Imposto de Renda e contribuição sobre o lucro líquido
- 70- Hospital universitário de ensino
- 80- Hospital de sindicato
- 90- Hospital universitário de pesquisa
- 91- Hospital universitário de pesquisa, isento de tributos e contribuições sociais
- 93- Hospital universitário de pesquisa isento de Imposto de Renda e contribuição sobre o lucro líquido

Tipo de gestão – Campo: GESTAO

- 0- Gestão estadual
- 1- Gestão plena municipal
- 2- Gestão plena estadual

Código da idade - Campo: COD_IDADE

- 0- ignorada
- 2- se for em dias
- 3- se for em meses
- 4- se for em anos

Saída do paciente com morte – Campo: MORTE

- 0- Não
- 1- Sim

Nacionalidade do paciente - Campo: NACIONAL

- 10- brasileiro
- 20- naturalizado brasileiro
- 21- argentino
- 22- boliviano
- 23- chileno
- 24- paraguaio
- 25- uruguaio
- 30- alemão
- 31- belga
- 32- britânico
- 34- canadense
- 35- espanhol
- 36- norte-americano (EUA)
- 37- francês
- 38- suíço
- 39- italiano
- 41- japonês
- 42- chinês
- 43- coreano
- 45- português
- 48- outros latino-americanos
- 49- outros asiáticos
- 50- outros

Caráter de Internação - Campo: CAR_INT

- 01- Eletiva
- 03- Urgência/Emergência, quando a AIH tiver sido emitida antes da internação
- 04- Internação em AIH de alta complexidade
- 05- Urgência/Emergência, quando a AIH tiver sido emitida após a internação.
- 06- quadro compatível com acidente no local de trabalho ou a serviço da empresa
- 07- quadro compatível com acidente de trajeto entre a residência e trabalho
- 08- quadro compatível com outros tipos de acidente de trânsito, não considerados acidentes de trajeto entre residência e trabalho
- 09- quadro compatível com outros tipos de lesões e envenenamentos, por agentes físicos ou químicos.
- 20- Urgência/Emergência em Unidade de Referência
- 26- Urgência/Emergência – quadro compatível com acidente no local de trabalho ou a serviço da empresa

27- Urgência/Emergência – quadro compatível com acidente de trajeto entre a residência e trabalho

28- Urgência/Emergência – quadro compatível com outros tipos de acidente de trânsito, não considerados acidentes de trajeto entre residência e trabalho

29- Urgência/Emergência – quadro compatível com outros tipos de lesões e envenenamentos, por agentes físicos ou químicos.

Grau De Instrução - Campo: INSTRU

0- SEM INFORMAÇÃO

1- Analfabeto

2- Primeiro grau

3- Segundo grau

4- Terceiro grau

Contraceptivo Utilizado - Campos: CONTRACEP1 e CONTRACEP2

00- Sem Informação

01- LAM

02- Ogino Knaus

03- Temperatura basal

04- Billings

05- Cinto térmico

06- DIU

07- Diafragma

08- Preservativo

09- Espermicida

10- Hormônio oral

11- Hormônio injetável

12- Coitus interruptus

APENDICE B:

ESTRUTURAS DAS TABELAS CRIADAS NO REPOSITÓRIO DE DADOS INTEGRADO AO SIG

1) Sumarização de Internações por dia, por município:

Esta tabela foi preenchida por programação, a partir dos dados dos arquivos RD. Considerou-se o código do município de residência, a data da internação e o campo diagnóstico principal para o seu preenchimento. Se um destes campos não estava preenchido, o registro foi desconsiderado.

Campo	Tipo	Descrição
CODMUN	Texto (6)	Código IBGE do Município
DATAINT	Data	Data da internação
ASMA04	Inteiro	Total de internações com diagnóstico principal = J45 e idade entre 0 e 4 anos.
ASMA65	Inteiro	Total de internações com diagnóstico principal = J45 e idade entre 65 anos ou mais.
BRON04	Inteiro	Total de internações com diagnóstico principal = J20, J40, J41, J42 e idade entre 0 e 4 anos.
BRON65	Inteiro	Total de internações com diagnóstico principal = J20, J40, J41, J42 e idade entre 65 anos ou mais.
ENFI04	Inteiro	Total de internações com diagnóstico principal = J43 e idade entre 0 e 4 anos.
ENF65	Inteiro	Total de internações com diagnóstico principal = J43 e idade entre 65 anos ou mais.

2) Lista de dias considerados no estudo:

Para trabalhar com dados diários, elaborou-se uma tabela contendo todos os dias do período pesquisado. A tabela resultante, com 3288 linhas, foi projetada para inserir outras variáveis necessárias para a modelagem estatística, conforme descrito a seguir.

Campo	Tipo	Descrição
DIA	Data	Dia/mês/ano do período considerado no estudo.
DIASEM	Inteiro	Identificador do dia da semana assumindo: 1-Domingo até 7-Sábado.
ESTACAO	Texto(1)	Identificador da estação do ano, assumindo: 1-Verão, 2-Outono, 3-Inverno e 4-Primavera
SAFRACS	Texto(1)	Identificador de período de safra na região Centro-Sul, assumindo: S-Período de safra e N-período fora da safra
SAFRANE	Texto(1)	Identificador de período de safra na região Nordeste, assumindo: S-Período de safra e N-período fora da safra

Nota: O campo ESTACAO foi preenchido conforme as datas de início das estações do ano, a saber: 21/março – Outono, 21/junho – Inverno, 23/setembro – Primavera e 21/dezembro – Verão.

3) Resultsets utilizados no Stata:

As tabelas de dados utilizadas para as análises geradas no Stata foram construídas no SQL Server 2005 através de uma composição de visões. Visões são estruturas de consultas ao banco de dados, que são previamente armazenadas com o propósito de minimizar a complexidade na construção de outras consultas ou tarefas que se deseja executar.

Assim, optou-se por criar consultas intermediárias que facilitaram a montagem dos *resultsets* finais utilizados no Stata.

A figura abaixo resume a concepção da estratégia implementada.

Onde:

FOCOS: são as contagens de focos de calor por município, por dia.

INTERDIA: contagens de internações por município, por dia.

POP2000 a POP2008: somatória de população, por faixa etária, por município, por ano.

INTERDIAINC: incidências de internações por município.

SET_MUNIC_DIA: visão utilizada no Stata para as análises nas dimensões município e dia.

Mapa 3 – Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 1º trimestre de 2008.

Mapa 4 – Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 2º trimestre de 2008.

Mapa 5 – Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 3º trimestre de 2008.

Mapa 6 – Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Centro-Sul, 4º trimestre de 2008.

Mapa 7 – Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 1º trimestre de 2008.

Mapa 8 – Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 2º trimestre de 2008.

Mapa 9 – Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 3º trimestre de 2008.

Mapa 10 – Agrupamentos espaciais dos focos de calor, internações por asma (0 a 4 anos e 65 anos ou mais) e enfisema (65 anos ou mais), região Nordeste, 4º trimestre de 2008.

Mapa 11 - Agrupamentos por HotSpot para Internações Hospitalares e Focos de Calor no estado de São Paulo no período de 2000 a 2008