

Gustavo Garcia Manzato

**REGIÕES URBANAS HOMOGÊNEAS E
OFERTA DE TRANSPORTES**

Orientador:

Prof. Assoc. Antônio Néelson Rodrigues da Silva

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Civil: Planejamento e Operação de Sistemas de Transportes.

São Carlos
2007

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

M296r Manzato, Gustavo Garcia
 Regiões urbanas homogêneas e oferta de transportes /
Gustavo Garcia Manzato ; orientador Antônio Néilson
Rodrigues da Silva. -- São Carlos, 2007.

 Dissertação (Mestrado) - Programa de Pós-Graduação em
Engenharia Civil e Área de Concentração em Planejamento
e Operação de Sistemas de Transportes -- Escola de
Engenharia de São Carlos da Universidade de São Paulo.

 1. Regiões urbanas homogêneas. 2. Estatística
espacial. 3. Modelagem espacial. 4. ESDA. 5. Autômatos
celulares. 6. Uso e ocupação do solo. I. Título.

*Aos meus pais, Lucia e Valdomiro,
pelo amor incondicional.*

Agradecimentos

A Deus, por proporcionar mais esta valiosa oportunidade de engrandecimento intelectual e moral em minha vida.

Aos meus pais, com muito amor, pelo inestimável apoio em mais esta grande conquista de minha vida. Sou eternamente grato pela constante dedicação na construção dos meus valores intelectuais e morais.

Aos meus queridos irmãos, Marcelo e Daniel, pela nossa eterna amizade e companheirismo.

À minha tia, madrinha e segunda mãe, Josefa, a quem eu tenho como exemplo de vida. Sou grato por todo o apoio que sempre me deu.

Aos meus avós, Maria e Henrique, Gema e José (em memória), por sempre acreditarem e torcerem para que tudo de melhor acontecesse em minha vida.

À Eliana, por toda a paciência, apoio, amor e carinho que sempre teve comigo. Sou grato pelos momentos de amizade e aprendizado mútuo.

Em especial, ao meu orientador, Prof. Antônio Néson Rodrigues da Silva, por sua amizade, paciência e dedicação. Por acreditar no meu potencial proporcionando inúmeras oportunidades e desafios que contribuíram para o meu progresso acadêmico e profissional.

A todos os professores do Departamento de Transportes, que compartilharam seus conhecimentos e contribuíram para o meu desenvolvimento intelectual. Em especial, ao Prof. João Alexandre Widmer, pela recomendação no processo seletivo do doutorado.

Em especial também, à Prof^a. Eliane Viviani, por “alavancar” meu desenvolvimento acadêmico durante a graduação. Pela presença nas várias etapas da pós-graduação (Exame de Qualificação, recomendação para o mestrado e agora para o doutorado, dentre outras) e pela amizade construída ao longo desse tempo.

À Prof^a. Rita David, pela recomendação para o processo seletivo do mestrado, ao Prof. Carlos Felipe Grangeiro Loureiro, pelas contribuições no Exame de Qualificação do mestrado e ao Prof. Rui António Rodrigues Ramos, pela contribuição para o doutorado.

Aos “irmãos acadêmicos da família Silva”, pela amizade e ótima convivência na nossa sala: Bia (companheira de todas as horas, inclusive do Bandejão), Andréa Júlia (nossa “*mamma*”), Vanessa Santos (madrinha em *Cellular Automata*), Alexandra Akamine (madrinha em Algoritmos Genéticos), Michael Saunders, Fabíola Aguiar, Renata Magagnin, Vanessa Yuassa, Marcela Costa, Simone Becker, Leonardo Hotta e Fabiana Arruda.

A todos os amigos do departamento com os quais eu tive uma agradável convivência durante esse período do mestrado: Márcia Lika (pelas sugestões, incentivos e auxílios, principalmente no início do mestrado), Cira Pitombo (por compartilhar a expectativa no período “pré-PLURIS 2006”, pela companhia durante o PLURIS 2006 e pelos incentivos mútuos no “pós-PLURIS 2006”), Daniel Anijar, Bruno Bertoncini, Diogo Colella, Thaís Guerreiro, Cíntia Egami, Mateus Araújo, Sérgio Rodrigues (Serginho), Leandro Piassi, Adriana Ferreira, Adriana (gauchinha), Vivian Bardini, Karênina, Camila Modenese, Camilla Carrara, Jesner, Juliana, André, Fred, Jorge, Marcos e Adalberto. Se eu estiver me esquecendo de alguém, por favor, me desculpem!

A todos os funcionários do departamento de Transportes, especialmente Heloísa, Beth, Magaly, Alexandre e Suely.

À amizade do pessoal do CETEPE: Márcia, Nídia, Martinho e Umberto.

Às minhas grandes amigas Suzana Teixeira e Landa Costa, sempre presentes! E as minhas irmãs do coração: Cida, Maira e Maísa, pelo carinho e por sempre vibrarem com a minha presença.

À dona Gina, Sr. Néelson, Eloíza e Eliza, por sempre me incentivarem e desejarem que acontecesse tudo de bom na minha vida.

Ao Departamento de Estradas de Rodagem do Estado de São Paulo, pela cessão dos dados referentes à malha rodoviária do estado, necessários para esta pesquisa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro sem o qual este trabalho não teria sido realizado.

E a todos aqueles que, direta ou indiretamente, colaboraram para este trabalho.

*"Tudo tem seu tempo e até certas manifestações mais
vigorosas e originais entram em voga ou saem de moda.
Mas a sabedoria tem uma vantagem: é eterna."
(Baltasar Gracián)*

Resumo

MANZATO, G. G. *Regiões Urbanas Homogêneas e Oferta de Transportes* Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2007.

O objetivo deste trabalho é identificar regiões urbanas homogêneas por meio da aplicação de duas vertentes da análise espacial: a estatística espacial e uma estratégia de modelagem espacial baseada na comparação de informações oriundas de diferentes entidades espaciais, em níveis diversos de informação. Um método baseado em fluxos de viagens seria a melhor alternativa para o problema em questão, mas não há dados disponíveis para sua aplicação no Brasil. Em virtude disso, o método aqui apresentado identifica regiões que podem ser consideradas como uniformes em relação a uma variável a partir de técnicas de análise exploratória de dados espaciais, como por exemplo, o gráfico e o mapa de Moran. Em um estudo de caso para o estado de São Paulo, analisando-se as distribuições espaciais dos valores da densidade populacional por meio de sua representação em mapas temáticos classificados segundo os quadrantes do gráfico de Moran (ou *Box map*), esse indicador permite caracterizar razoavelmente bem as regiões urbanas homogêneas existentes (inclusive as oficiais). Entretanto, ao tentar representar o seu comportamento em uma análise temporal por meio de modelos, o indicador populacional não foi capaz de descrever esse comportamento e, conseqüentemente, não serviu para elaborar estratégias de previsão para o futuro. Por outro lado, ao combinar essas informações com um indicador que representa a oferta de transportes, os resultados obtidos permitiram observar o alto desempenho dos modelos, dada a forte influência recíproca entre uso e ocupação do solo e oferta de transportes. Ao permitir a identificação de padrões e a projeção de tendências, este tipo de análise pode ser útil para o planejamento urbano e regional, tanto no contexto estudado como em uma visão mais abrangente.

Palavras-chave: regiões urbanas homogêneas, estatística espacial, modelagem espacial, ESDA, Autômatos Celulares, uso e ocupação do solo.

Abstract

MANZATO, G. G. *Homogeneous Urban Regions and Transportation Supply* Dissertation (Master) – São Carlos School of Engineering, University of São Paulo, São Carlos, 2007.

The objective of this work is to identify homogeneous urban regions through the application of two branches of spatial analysis: spatial statistics and a modeling strategy based on the comparison of information from different spatial entities and at distinct levels. A commuting-based approach would be the best alternative in that case, but there is no data available for its application in Brazil. Thus, the method presented here identifies uniform regions regarding a particular variable through exploratory spatial data analysis techniques, such as the Moran scatter plot and Box maps. In a case study carried out in the state of São Paulo, in which the spatial distribution of the values of population density was analyzed through the representation in Box maps, a reasonable identification of the existing homogeneous urban regions (including the official ones) was performed. However, the models based only on the population density distribution did not perform well for analyses through time and therefore they were not adequate for forecasting strategies. In contrast, when combining population density information with an indicator of transportation supply the performance of the models was noticeably improved, what was likely caused by the strong reciprocal influence between land use and transportation supply. As a conclusion, the method developed in this work shall be useful for urban and regional planning at different scales, given its potential for patterns recognition and trends forecasting.

Keywords: homogeneous urban regions, spatial statistics, spatial modeling, ESDA, Cellular Automata, land use.

Lista de figuras

Figura 3.1: Inter-relação entre os quatro grupos de ferramentas de análise espacial (Fonte: Teixeira, 2003).....	19
Figura 3.2: Matriz de proximidade espacial normalizada pela linha (Fonte: Câmara et al., 2002a).....	25
Figura 3.3: Gráfico e mapa de espalhamento de Moran.....	30
Figura 4.1: Exemplo de um CA simples (Fonte: adaptado de Santos, 2005).....	35
Figura 4.2: Estrutura básica de uma RNA (Fonte: Silva et al., 2004).....	39
Figura 5.1: Esquema ilustrativo do processo de atualização dos dados para municípios que sofreram desmembramento ao longo do tempo.....	51
Figura 5.2: Bases geográficas referentes às malhas rodoviárias do estado de São Paulo no período de 1960 a 2000.....	53
Figura 5.3: Esquema ilustrativo dos tipos de modelos desenvolvidos.....	57
Figura 6.1: Box map representando a densidade populacional em 1960.....	67
Figura 6.2: Box map representando a densidade populacional em 1970.....	67
Figura 6.3: Box map representando a densidade populacional em 1980.....	68
Figura 6.4: Box map representando a densidade populacional em 1990.....	68
Figura 6.5: Box map representando a densidade populacional em 2000.....	69
Figura 6.6: Box map representando o indicador de cobertura espacial de estrutura de circulação viária em 1960.....	71
Figura 6.7: Box map representando o indicador de cobertura espacial de estrutura de circulação viária em 1970.....	71
Figura 6.8: Box map representando o indicador de cobertura espacial de estrutura de circulação viária em 1980.....	72
Figura 6.9: Box map representando o indicador de cobertura espacial de estrutura de circulação viária em 1990.....	72
Figura 6.10: Box map representando o indicador de cobertura espacial de estrutura de circulação viária em 2000.....	73
Figura 6.11: Box map representando a densidade populacional estimada pelo Modelo I em 2000....	76
Figura 6.12: Box map representando a densidade populacional estimada pelo Modelo II em 2000..	76
Figura 6.13: Box map representando a densidade populacional estimada pelo Modelo III em 2000	77
Figura 6.14: Box map representando a densidade populacional estimada pelo Modelo IV em 2000.	77
Figura 6.15: Box map representando a densidade populacional estimada pelo Modelo V em 2000..	78
Figura 6.16: Box map representando a densidade populacional estimada pelo Modelo VI em 2000.	78
Figura 6.17: Box map representando a densidade populacional estimada pelo Modelo VII em 2000	79
Figura 6.18: Box map representando a densidade populacional estimada pelo Modelo I em 2010....	81
Figura 6.19: Box map representando a densidade populacional estimada pelo Modelo II em 2010..	82
Figura 6.20: Box map representando a densidade populacional estimada pelo Modelo III em 2010	82
Figura 6.21: Box map representando a densidade populacional estimada pelo Modelo IV em 2010.	83
Figura 6.22: Box map representando a densidade populacional estimada pelo Modelo V em 2010..	83
Figura 6.23: Box map representando a densidade populacional estimada pelo Modelo VI em 2010.	84
Figura 6.24: Box map representando a densidade populacional estimada pelo Modelo VII em 2010	84
Figura A. 1: Ativação dos marginais da matriz.....	106
Figura A. 2: Exportação da coluna marginal para uma tabela.....	107
Figura A. 3: Importação da coluna marginal para a matriz.....	107
Figura A. 4: Matriz com os valores importados da coluna marginal.....	108
Figura A. 5: Janela para o cálculo da matriz de adjacência normalizada.....	108
Figura A. 6: Matriz de adjacência normalizada.....	109
Figura A. 7: Importação dos valores de Z do dataview na matriz.....	110
Figura A. 8: Matriz com os valores de Z.....	111
Figura A. 9: Janela para o cálculo de W_z	112
Figura A. 10: Matriz com os valores de W_z	112
Figura A. 11: Exportação da coluna marginal com o valor de W_z	113
Figura A. 12: Gráfico de espalhamento de Moran.....	115
Figura A. 13: Mapa de espalhamento de Moran.....	115

Lista de tabelas

<i>Tabela 5.1: Método proposto para o estudo.....</i>	<i>48</i>
<i>Tabela 5.2: Modelos espaciais desenvolvidos.....</i>	<i>60</i>
<i>Tabela 5.3: Variáveis utilizadas em cada modelo e em cada fase da modelagem por RNA.....</i>	<i>61</i>
<i>Tabela 6.1: Proporções de acertos parciais (por quadrante) e totais de cada modelo.....</i>	<i>79</i>

Lista de equações

<i>Equação 3.1: Expressão para a determinação da média móvel espacial.....</i>	<i>26</i>
<i>Equação 3.2: Expressão para a determinação da média móvel espacial a partir da matriz de proximidade espacial normalizada.....</i>	<i>26</i>
<i>Equação 3.3: Índice I de Moran (global)</i>	<i>28</i>
<i>Equação 3.4: Índice de Moran (global) na forma matricial.....</i>	<i>28</i>
<i>Equação 5.1: Índice de cobertura de estrutura de circulação viária.....</i>	<i>55</i>

Sumário

1. INTRODUÇÃO	1
1.1. CARACTERIZAÇÃO DO PROBLEMA.....	1
1.2. OBJETIVOS	3
1.3. ESTRUTURA DO TRABALHO	4
2. REGIÕES URBANAS HOMOGÊNEAS	5
2.1. INTRODUÇÃO.....	5
2.2. CRITÉRIO POLÍTICO-ADMINISTRATIVO	8
2.3. CRITÉRIO BASEADO EM FLUXOS DE VIAGENS	12
2.4. CRITÉRIO BASEADO NA DENSIDADE POPULACIONAL.....	13
3. ANÁLISE ESPACIAL	15
3.1. INTRODUÇÃO.....	15
3.2. ESTATÍSTICA ESPACIAL.....	20
3.3. ANÁLISE EXPLORATÓRIA DE DADOS ESPACIAIS APLICADA AO ESTUDO DE DADOS DE ÁREAS.....	23
3.3.1. <i>Matriz de proximidade espacial (W)</i>	24
3.3.2. <i>Média móvel espacial</i>	26
3.3.3. <i>Indicadores de autocorrelação espacial</i>	27
3.3.4. <i>Diagrama e mapa de espalhamento de Moran</i>	29
4. MODELAGEM ESPACIAL	31
4.1. INTRODUÇÃO.....	31
4.2. CELLULAR AUTOMATA (CA)	32
4.3. REDES NEURAIS ARTIFICIAIS (RNA)	38
5. MÉTODO.....	45
5.1. TÉCNICAS EMPREGADAS.....	45
5.2. ESTUDO DE CASO	48
5.2.1. <i>Tratamento e formatação dos dados</i>	50
5.2.2. <i>Análise exploratória dos dados espaciais</i>	53
5.2.3. <i>Modelos espaciais</i>	56
5.3. SÍNTESE DAS ATIVIDADES DESENVOLVIDAS	62
6. ANÁLISE DOS RESULTADOS	65
6.1. ANÁLISE EXPLORATÓRIA DOS DADOS DE DENSIDADE POPULACIONAL	65
6.2. ANÁLISE EXPLORATÓRIA DOS DADOS DE INFRA-ESTRUTURA	69
6.3. MODELOS ESPACIAIS OBTIDOS PARA O ANO DE 2000.....	73
6.4. MODELOS ESPACIAIS PARA A PREVISÃO NO ANO 2010	80
7. CONCLUSÕES E RECOMENDAÇÕES.....	85
REFERÊNCIAS BIBLIOGRÁFICAS.....	91
ANEXO: ESDA POR MEIO DO SOFTWARETRANSCAD	105

1. Introdução

Neste capítulo são apresentados, de maneira sintética, o problema estudado, a justificativa, os objetivos e a estrutura do trabalho.

1.1. Caracterização do problema

Citadas por Gouvêa (2005), as estatísticas da ONU (Organização das Nações Unidas) de 2000 revelam que dentre as nações mais populosas do mundo, o Brasil ocupa a quinta posição, com uma taxa de crescimento demográfico bastante elevada. Entretanto, o que torna o país diferente nesse contexto é o fato de que o crescimento brasileiro se verifica sobretudo nas áreas urbanas, com mais de 81 % da população, segundo o IBGE (Instituto Brasileiro de Geografia e Estatística) e esse crescimento reflete um intenso processo de ocupação das áreas urbanas.

A crescente urbanização não somente do Brasil, mas de outras regiões do mundo, constitui uma das mais evidentes causas no aumento dos estudos e estatísticas urbanas observado nos últimos anos. Os planejadores enfrentam o desafio

de elaborar modelos capazes de representar as inter-relações presentes nessas estruturas multidimensionais e complexas.

É importante observar que essa questão não se restringe a apenas um município. O processo de ocupação do território muitas vezes acaba desenvolvendo regiões adjacentes e contínuas, levando a formação de grandes aglomerações urbanas, aqui tratadas de maneira geral, como regiões urbanas homogêneas (RUH). Nesse contexto reside um dos maiores desafios para os planejadores e administradores urbanos: a caracterização dessas RUH.

Dentre os diversos critérios que podem ser adotados, poderia se considerar o aspecto político-administrativo. Entretanto, para os processos de planejamento, esse critério muitas vezes não reflete a realidade necessária que envolve as questões resultantes da junção dos municípios.

Um outro critério, adotado nos Estados Unidos, se refere aos fluxos de viagens observados entre as regiões que compõem a área urbana. Entretanto, dados dessa natureza são difíceis de serem obtidos na maioria das cidades, particularmente no Brasil.

Uma alternativa, então, seria a utilização da distribuição populacional como forma de se caracterizar as RUH. A caracterização de RUH por meio da densidade populacional é um bom indicador, conforme pode ser verificado em Ramos e Silva (2003a; 2003b; 2007); Ramos, Silva e Miranda (2004).

No entanto, essa abordagem pode não refletir de maneira adequada as inter-relações presentes nessas áreas urbanas. Sabe-se que existe uma forte inter-relação entre uso e ocupação do solo e a oferta de infra-estrutura, principalmente a de transportes, como pode ser observado em Lemos, Santos e Portugal (2004; 2005) e em Silva e Vala (2001).

É justamente nesse ponto que o trabalho aqui apresentado se baseia, tentando explicar de forma quantitativa a inter-relação entre oferta de transportes e desenvolvimento urbano. Tendo em vista que um dos problemas que afetam as cidades brasileiras é a falta de ferramentas para o gerenciamento do atual crescimento acelerado de áreas urbanas, a contribuição do presente trabalho está em apresentar instrumentos que sejam eficazes nas estratégias de planejamento e, ainda, definidos em uma dimensão regional.

Nesse sentido, o estudo da inter-relação entre uso e ocupação do solo e oferta de transportes pode compreender um instrumento útil na tomada de decisões. Isto se soma ao fato do trabalho ter como objeto de estudo a Estatística Espacial e a Modelagem Espacial, temas ainda relativamente pouco estudados e difundidos no país, principalmente na área do planejamento urbano e de transportes.

1.2. Objetivos

O objetivo principal deste trabalho é identificar regiões urbanas homogêneas utilizando técnicas de análise espacial. Além disso, por meio de uma estratégia de modelagem espacial, pretende-se desenvolver um modelo da evolução dessas regiões

urbanas homogêneas, considerando-se a distribuição populacional e a oferta de transportes.

1.3. Estrutura do trabalho

O presente trabalho apresenta, primeiramente, os resultados de uma revisão bibliográfica referente a regiões urbanas homogêneas no Capítulo 2. Em seguida, nos Capítulos 3 e 4, são discutidos alguns conceitos referentes à Análise Espacial em suas duas vertentes estudadas: a Estatística Espacial e a Modelagem Espacial, respectivamente.

No Capítulo 5 é apresentado o método adotado nesta pesquisa, em que são descritas as etapas de tratamento e formatação dos dados, análise exploratória dos dados espaciais e a etapa de modelagem, com detalhes dos modelos espaciais aqui desenvolvidos.

Os resultados e análises obtidos por meio da aplicação do método proposto são apresentados no Capítulo 6 e no Capítulo 7 estão as conclusões e recomendações. Este trabalho finaliza com a apresentação das referências bibliográficas e de um anexo que trata da análise exploratória de dados espaciais utilizando-se o *software* TransCAD.

2. Regiões urbanas homogêneas

Este capítulo apresenta uma breve revisão bibliográfica sobre regiões urbanas homogêneas, tratando de alguns critérios para a sua caracterização.

2.1. Introdução

Regiões urbanas homogêneas (RUH) podem ser consideradas como áreas urbanizadas que concentram um grande poder de atração de pessoas e de bens, de empregos e de serviços, formadas pelo agrupamento de dois ou mais municípios que possuem interesses econômicos, sociais e político-administrativos em comum. Além disso, como em muitas regiões do mundo, o processo de formação dessas RUH no Brasil se deve à ocupação do território no passado.

Pode-se observar que durante muito tempo esse processo de ocupação apresentou uma mesma característica: grandes concentrações populacionais em regiões específicas, notadamente aquelas próximas ao litoral, resultado do processo de colonização. Como conseqüência, surgiram regiões díspares do ponto de vista da ocupação do território, com o

norte e o centro-oeste pouco povoados e com municípios de grandes extensões territoriais, enquanto que as demais regiões (sobretudo sudeste e sul) contam com expressivas concentrações populacionais em municípios de pequenas extensões (MANZATO; SOARES; SILVA, 2006).

Esse fato acabou por criar e fortalecer alguns desequilíbrios: nas regiões de maiores concentrações populacionais estão as melhores redes de infra-estrutura, seja pela quantidade ou pela qualidade, enquanto que nas regiões menos povoadas a infra-estrutura é freqüentemente precária, não favorecendo o seu desenvolvimento. Como redes de infra-estrutura podem ser citadas, neste caso, as de energia elétrica, saneamento básico, transporte e telecomunicações, todas essenciais para o crescimento de um município, região ou estado. Os sistemas de transportes, em particular, têm como característica fundamental promover a integração nacional e a redução das desigualdades regionais, sendo os órgãos públicos importantes agentes neste processo.

A opção de planejamento de transportes no século XX no país foi atribuir as funções anteriormente citadas aos modos rodoviários. Entretanto, embora se perceba um contínuo esforço governamental para que se efetive a ocupação territorial, ainda é possível encontrar, além de regiões com significativa oferta de infra-estrutura de transportes, outras com poucas vias de transportes, baixa ocupação e sistemas produtivos ineficientes. Uma consequência dessas disparidades é a existência de regiões com elevado poder de atração (de empregos e de habitantes), que levaram à formação de regiões urbanas homogêneas e regiões metropolitanas (MANZATO; SOARES; SILVA, 2006).

Atualmente, o fato de que os limites dessas RUH geralmente são diferentes dos limites administrativos oficiais dos municípios que as compõem, constitui um grande

desafio para os planejadores e administradores urbanos. Essas regiões na maioria das vezes englobam municípios vizinhos ultrapassando os limites de uma simples cidade, entretanto são menores que a divisão administrativa superior (os estados, por exemplo). Nesse contexto, é fundamental a necessidade de se caracterizar novas fronteiras para a resolução de problemas relativos à expansão urbana. No entanto, a definição das fronteiras dessas RUH não é geralmente simples, principalmente devido a incompatibilidades entre interesses políticos e econômicos das partes envolvidas (RAMOS; SILVA, 2003a; 2003b; RAMOS; SILVA; MIRANDA, 2004).

Este não é um problema recente e pode-se encontrar na literatura várias propostas de métodos para a delimitação dessas RUH. Como por exemplo, *Office of Management and Budget* (1998; 1999; 2000); *Metropolitan Area Standards Review Committee* (2000); NUREC (1994); Lacour e Puissant (1999); Ferreira e Rosado (1999); Ferrão, Rodrigues e Vala (2002); Ferrão e Vala (2001).

Nesse sentido, a identificação dessas RUH, em geral contíguas e supostamente homogêneas, pode ser feita de diversas maneiras. Com base na literatura, apresentam-se aqui três critérios usuais, a saber: uma primeira abordagem em que se consideram relações político-administrativas entre os municípios; segundo, um critério considerado ideal, que é baseado em fluxos de viagens; e, na ausência destas informações, um terceiro critério (razoável) baseado na densidade populacional. A seguir, são discutidos alguns detalhes a respeito de cada uma dessas abordagens.

2.2. Critério político-administrativo

Gouvêa (2005) se refere aos municípios brasileiros mostrando que desde a época do Brasil Colônia até a promulgação da Constituição Federal de 1988, eles oscilaram de elementos meramente administrativos a entes federativos, vivenciando períodos muito diferentes em termos de autonomia política, econômica e administrativa.

A questão metropolitana (aqui tratada como RUH) só foi mencionada pela primeira vez nos textos constitucionais de 1967, autorizando a União a criar regiões metropolitanas, as quais receberam maior destaque na Emenda Constitucional de 1969. Em 1973, em um intenso ritmo de crescimento urbano em torno das principais capitais estaduais, foram criadas as oito regiões metropolitanas brasileiras (Belém, Belo Horizonte, Curitiba, Fortaleza, Porto Alegre, Recife, Salvador e São Paulo). No ano seguinte (1974) com a fusão dos estados do Rio de Janeiro e da Guanabara, surgiu a região metropolitana do Rio de Janeiro.

Entretanto, com a transferência da responsabilidade de criar regiões metropolitanas da União para os estados, por meio da Constituição Federal de 1988, observa-se que em 2000 o Brasil possuía 22 regiões metropolitanas. Além disso, em 1989, com a promulgação das novas Constituições estaduais, muitas regiões metropolitanas acabaram se ampliando, incorporando um grande número de novos municípios sem atender a quaisquer critérios urbanísticos ou econômicos que justificassem tal operação.

Nesse sentido, o termo “região metropolitana” tem sido utilizado de forma indevida nos processos de instituição de agrupamentos municipais. Ou seja, estão sendo classificados

como regiões metropolitanas grupos de municípios que não registram nenhuma situação real de conurbação e adensamento urbano, nem expressividade política, econômica e populacional.

Como a Constituição de 1988 não caracteriza nenhuma das três possibilidades de agrupamentos municipais, a saber: região metropolitana, aglomeração urbana e microrregião, cabe à Constituição de cada estado estabelecer seus critérios para definir cada uma dessas possibilidades.

Por exemplo, a Constituição de São Paulo conceitua de maneira precisa cada uma das possibilidades mencionadas:

- **Região metropolitana:** “agrupamento de municípios limítrofes que assuma destacada expressão nacional, em razão de elevada densidade demográfica, significativa conurbação e de funções urbanas e regionais com alto grau de diversidade, especialização e integração sócio-econômica (sic), exigindo planejamento integrado e ação conjunta permanente dos entes públicos nela atuantes”;
- **Aglomeração urbana:** “agrupamento de municípios limítrofes que apresente relação de integração funcional de natureza econômico-social e urbanização contínua entre dois ou mais municípios ou manifesta tendência nesse sentido, que exija planejamento integrado e recomende ação coordenada dos entes públicos nela atuantes”;
- **Microrregião:** “agrupamento de municípios limítrofes que apresentem, entre si, relações de interação funcional de natureza físico-territorial, econômico-social e

administrativa, exigindo planejamento integrado com vistas a criar condições adequadas para o desenvolvimento e integração regional”.

Por outro lado, está em discussão na Câmara dos Deputados desde 2004 um Projeto de Lei que cria o Estatuto da Metrópole, de autoria de Walter Feldman. O Projeto de Lei estabelece diretrizes para a execução da Política Nacional de Planejamento Regional Urbano e institui o Sistema Nacional de Planejamento e Informações Regionais Urbanas. Considera unidades regionais urbanas as regiões metropolitanas, aglomerações urbanas, microrregiões e regiões integradas de desenvolvimento (Rides). Pela proposição, a criação de unidades regionais urbanas é de competência dos estados e da União. As unidades regionais urbanas passariam a vigorar com as seguintes definições:

- **Região metropolitana** – agrupamento de municípios limítrofes que apresente, cumulativamente, as seguintes características: um núcleo central com, no mínimo, 5% da população do País ou dois núcleos centrais que apresentem, conjuntamente, no mínimo, 4% da população nacional; taxa de urbanização acima de 60%, para cada um dos municípios integrantes da região; população economicamente ativa residente nos setores secundário e terciário de, no mínimo, 65%, considerado cada um dos municípios integrantes da região; e urbanização contínua em, no mínimo, 50% dos municípios componentes da região;
- **Aglomeração urbana** – agrupamento de municípios limítrofes que apresente, cumulativamente, as seguintes características: um núcleo central com, no mínimo, 2,5% da população do País ou dois núcleos centrais que apresentem, conjuntamente, no mínimo, 2% da população nacional; taxa de urbanização acima de 60%, para cada um dos municípios integrantes da região; população economicamente ativa residente nos setores secundário e terciário de, no mínimo, 65%, considerado cada

um dos municípios integrantes da região; urbanização contínua em, no mínimo, 20% dos municípios componentes da região;

- **Microrregião** – agrupamento de municípios limítrofes, que apresente, cumulativamente, as seguintes características: um núcleo central com, no mínimo, 0,5% da população do País; taxa de urbanização acima de 40% no município central; população economicamente ativa residente nos setores secundário e terciário do município central de, no mínimo, 45%.
- **Região integrada de desenvolvimento econômico (Ride)** – agrupamento de municípios limítrofes, localizado em mais de um estado ou entre estados e o Distrito Federal que, de acordo com as características estabelecidas, será denominada, respectivamente, região metropolitana (Ride-RM), aglomeração urbana (Ride-AU) ou microrregião (Ride-MR).

À parte de qualquer legislação elaborada com o propósito de caracterizar RUH e devido à sua heterogeneidade, dentre as suas conclusões Gouvêa (2005) alerta para a necessidade de se avaliar, com base em critérios que comprovem a efetiva vinculação de um determinado município às municipalidades vizinhas, se está ou não inserido em alguma RUH. O autor coloca esses critérios como vínculos existentes entre os municípios metropolitanos, por meio dos quais possam ser identificados determinados fluxos de pessoas e mercadorias, presença de núcleos dormitórios, fenômenos de complementação e de suplementação de atividades etc.

Além disso, observa-se que os critérios baseados em relações político-administrativas nem sempre refletem a realidade necessária para os processos de planejamento e gerenciamento urbano da área total resultante da junção dos municípios. Nesse contexto,

podem ser inseridos os dois outros critérios aqui apresentados para a caracterização de RUH, os quais são detalhados a seguir.

2.3. Critério baseado em fluxos de viagens

O segundo critério para caracterização de RUH se baseia em fluxos de viagens observados entre zonas, que aqui podem ser entendidas como municípios. É apresentado pelo *Office of Management and Budget* (1998; 1999; 2000), um órgão governamental dos Estados Unidos, como um dos critérios mais indicados para a identificação de áreas urbanas homogêneas.

A hipótese deste critério é baseada no fato de que os municípios interagem entre si compartilhando suas funcionalidades, como por exemplo, habitação, emprego, serviços, lazer etc. Os padrões de viagens inter-zonais observados refletem essa interação e, portanto, a caracterização dos fluxos dessas viagens pode ser utilizada para a definição de RUH.

No entanto, o próprio órgão reconhece que dados confiáveis atualizados de viagens são difíceis de serem obtidos na maioria dos municípios. Neste caso, um critério considerado razoável é a utilização de densidade populacional como uma alternativa para caracterização de RUH, conforme é detalhado a seguir.

2.4. Critério baseado na densidade populacional

Conforme apresentado pelo *Office of Management and Budget* (1998), “a densidade populacional pode servir como uma alternativa em relação a outras medidas de atividade, na falta de dados nacionalmente consistentes e confiáveis que descrevam todos os movimentos diários e semanais de indivíduos” para a definição de RUH.

Esse terceiro critério é encontrado com frequência na literatura, como por exemplo, Ramos e Silva (2003a; 2007); Ramos, Silva e Miranda (2004) e, ainda, com algumas variações, como Ramos e Silva (2003b) e em Ferrão e Vala (2001) em que se testou uma alternativa proposta pela NUREC (1994), por meio da consideração da continuidade e densidade de edificações.

Por meio dessas informações, verifica-se nesse contexto uma alternativa adicional possível, apresentada por Manzato, Soares e Silva (2006). A adoção do terceiro critério (densidade populacional) como substituto daquele que seria o mais indicado (fluxos de viagens) não considera, no entanto, a possibilidade de se adotar uma variável *proxi* para descrever os fluxos de viagens entre municípios. Como as viagens dependem necessariamente da existência de infra-estrutura viária, a caracterização do nível de oferta dessa infra-estrutura deveria refletir com razoável fidelidade os fluxos observados. Assim, na ausência de dados de fluxos de viagens, uma forma de identificar RUH seria a partir de dados de oferta de infra-estrutura de transportes ou, mais especificamente, através da utilização de indicadores que representem essa oferta.

Por meio dos resultados obtidos, os autores sugerem então uma combinação de indicadores de oferta de infra-estrutura e populacionais para a identificação de regiões urbanas homogêneas. Manzato e Silva (2006), em uma seqüência daquele estudo, apresentaram uma metodologia que contemplava essa questão, mais especificamente, a definição de critérios para delimitação de RUH. Essa metodologia se baseava na combinação de indicadores de oferta de infra-estrutura viária e populacional, fazendo uso de recursos de duas técnicas de análise espacial: a estatística espacial e a modelagem espacial.

Com a apresentação desse contexto e conforme foi observado no item 2.1, o desenvolvimento de RUH, ou de maneira geral, o desenvolvimento regional está intimamente relacionado ao nível de oferta de infra-estrutura, principalmente a de transportes. De acordo com Araújo (2006), o setor de transportes exerce forte influência sobre o crescimento econômico, e conseqüentemente sobre o regional, dado o seu efeito multiplicador. Estudos sobre a dependência entre o sistema de transportes e o uso do solo também podem ser encontrados em Lemos, Santos e Portugal (2004; 2005) e em Silva e Vala (2001).

A partir dos resultados obtidos, Manzato e Silva (2006) sugerem a continuidade das investigações sobre a inter-relação entre uso e ocupação do solo e oferta de transportes. Nesse sentido, o trabalho aqui proposto apresenta uma contribuição para essa questão, trazendo novas alternativas para a combinação dessas variáveis, conforme também sugerido pelos autores.

Como essa proposta envolve duas técnicas de análise espacial, a saber: a estatística espacial e a modelagem espacial, alguns conceitos sobre esses temas são apresentados, respectivamente, no Capítulo 3 e no Capítulo 4.

3. Análise espacial

Este capítulo tem como objetivo apresentar os principais conceitos relacionados à Análise Espacial, principalmente no que se refere a um dos temas estudados no presente trabalho: a Estatística Espacial.

3.1. Introdução

Ramos (2002), citando alguns autores (CÂMARA; MONTEIRO, 2001; BAILEY; GATTREL, 1995; LONGLEY, 1996) define análise espacial como a aplicação de técnicas matemáticas e computacionais em conjuntos de dados geo-referenciados. Estas técnicas envolvem operações simples, como a sobreposição de camadas de informação (*overlays*), e outras mais complexas, como por exemplo, álgebras definidas para operar sobre representações digitais de mapas. De maneira geral, qualquer processo de apresentação, manipulação, análise, inferência e estimação de dados espaciais pode ser considerado como análise espacial. Seu objetivo principal é mensurar propriedades e relacionamentos desses dados espaciais, os quais englobam todas as informações que possam ser caracterizadas no espaço, em função de um

sistema de coordenadas. Isto significa que a posição que um atributo estudado ocupa no espaço é característica fundamental para sua análise, ou seja, a idéia principal é incorporar o espaço à análise que se deseja efetuar (LOPES, 2005; QUEIROZ, 2003; TEIXEIRA, 2003).

Basicamente os dados espaciais utilizados na análise espacial são representações de entidades encontradas no mundo real. A representação da realidade em ambientes computacionais é geralmente efetuada por meio de três tipos de análises de dados, conforme sugerido por Cressie (1991): pontuais, em áreas ou superfícies. Entretanto, há um quarto tipo especialmente importante na área de transportes: a análise de redes.

Anselin (1992), citando a sugestão de Cressie (1991), e Câmara *et al.* (2002b), a exemplo de outros autores, detalham os três primeiros tipos de dados em análise espacial, conforme apresentado a seguir:

- **Análise de padrões pontuais:** são fenômenos expressos por meio de ocorrências identificadas como pontos localizados no espaço, como por exemplo, a localização de crimes e a ocorrência de acidentes de trânsito;
- **Análise de dados em áreas:** são dados agregados em unidades delimitadas por polígonos fechados (setores censitários, municípios etc.), geralmente para garantir a confidencialidade dos dados que se referem a indivíduos localizados em pontos específicos do espaço. Exemplos desses dados são: levantamentos populacionais, níveis de renda, estatísticas de saúde, pesquisas O-D.
- **Análise de superfícies:** são eventos estudados de forma contínua, estimadas a partir de um conjunto de amostras de campo. Geralmente este

tipo de dados é resultante de levantamentos de recursos naturais, como mapas geológicos, topográficos, ecológicos, pedológicos etc.

O quarto tipo de análise de dados é apresentado, com base na literatura, por Teixeira (2003) e Carvalho (1997). A **análise de redes** se refere aos objetos que são representados por entidades lineares, como por exemplo, o acesso e fluxo entre regiões, otimizando trajetórias e estudando a localização de equipamentos urbanos (discutidos em todas as áreas de transportes).

Há alguns conceitos básicos na análise espacial, fundamentais para o entendimento de sua natureza. Câmara *et al.* (2002b), Queiroz (2003) e Teixeira (2003), a exemplo de outros autores, apresentam um detalhamento desses conceitos:

- **Dependência espacial:** se refere ao entendimento dos fenômenos espaciais, em que a maior parte destes (naturais ou sociais) apresenta entre si uma relação de dependência. Essa dependência está intimamente ligada à distância, ou seja, “a dependência espacial está presente em todas as direções e fica mais fraca à medida que aumenta a dispersão na localização dos dados” (Cressie, 1991);
- **Autocorrelação espacial:** se refere à expressão computacional do conceito de dependência espacial. Derivada do conceito estatístico “correlação”, que mede a relação entre duas variáveis, a autocorrelação mede a relação de uma mesma variável, mas em locais diferentes. É mensurada por meio de diferentes indicadores, os quais são baseados na idéia de se verificar como varia a dependência espacial a partir da comparação dos valores de um atributo e de seus vizinhos.

- **Estacionariedade e isotropia:** se referem aos conceitos estatísticos que definem a estrutura espacial dos dados. Há estacionariedade se um conjunto de dados possui uma média constante e seus valores flutuam sobre essa média com uma variância constante. Há isotropia se, além de estacionário, a covariância depender somente da distância entre os pontos e não da direção deles. Caso contrário, se o processo também depender da direção, ele é considerado anisotrópico.

Além desses conceitos básicos e dos tipos de dados espaciais, segundo Anselin (1992) a análise espacial apresenta uma subdivisão, em quatro grupos, quanto às ferramentas utilizadas:

- **Seleção:** processo mais simples de consultas a banco de dados, amostragem e agrupamentos de dados analisados e apresentação dos dados em mapas temáticos, além de métodos gráficos e numéricos para sumarizar dados;
- **Manipulação:** permite a criação de novos dados espaciais usando propriedades e relacionamento entre entidades espaciais;
- **Análise exploratória:** tem como objetivo visualizar e descrever distribuições espaciais, identificar padrões de associação espacial (aglomerados espaciais ou *clusters*), identificar observações atípicas (valores extremos ou *outliers*) ou a existência de instabilidades espaciais (não-estacionariedade). Os métodos são descritivos e não confirmatórios, servindo para detectar padrões, elaborar hipóteses e estimar modelos espaciais; as técnicas são visuais e poucas transformações são efetuadas sobre os dados originais;

- **Análise confirmatória:** agrupa os processos quantitativos de modelagem, estimação e validação necessários à análise de componentes espaciais. Destacam-se, neste grupo, o ferramental disponibilizado pela estatística espacial.

É importante observar que estes quatro grupos interagem entre si no processo da análise espacial e essa inter-relação pode ser compreendida por meio da Figura 3.1. Além disso, os dois últimos grupos de ferramentas apresentados integram uma vertente da análise espacial, objeto de estudo do presente trabalho: a estatística espacial. No próximo item serão apresentados maiores detalhes a respeito desse tema.

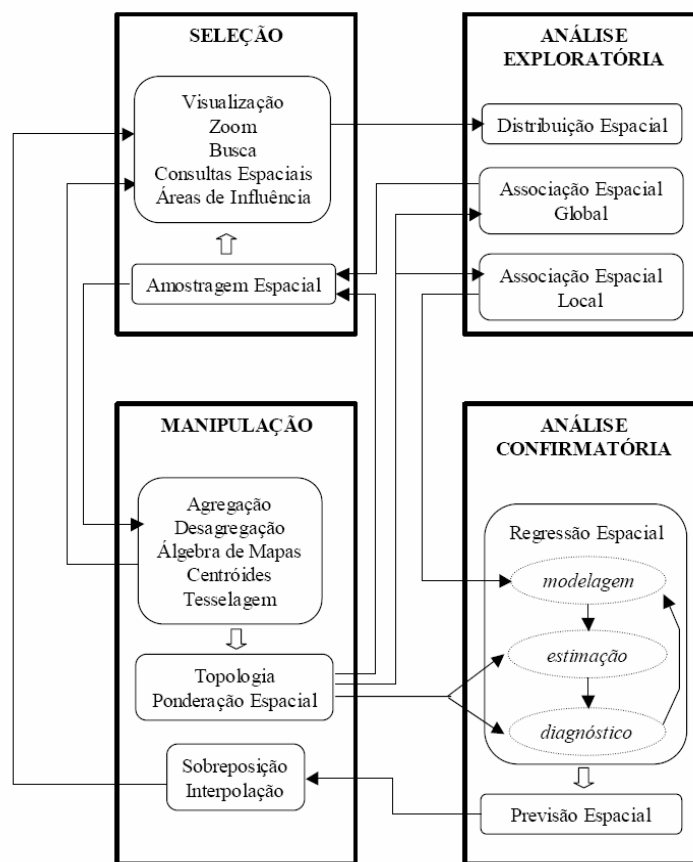


Figura 3.1: Inter-relação entre os quatro grupos de ferramentas de análise espacial (Fonte: Teixeira, 2003)

3.2. Estatística espacial

Conforme visto no item 3.1, a estatística espacial faz parte dos dois últimos grupos de ferramentas da análise espacial: a análise exploratória e a análise confirmatória. Trata-se de um desdobramento da análise espacial, cujo objetivo principal é a quantificação da dependência espacial presente em um conjunto de dados. Por meio de um amplo conjunto de técnicas matemáticas, a estatística espacial permite ao planejador a capacidade de estabelecer critérios quantitativos de análise exploratória e confirmatória.

Segundo Krempi (2004), a estatística espacial difere da estatística clássica pelo fato de levar em consideração a localização espacial do atributo considerado. Nesse sentido, um dos pressupostos da estatística clássica, que é o da independência entre as variáveis, não se aplica à estatística espacial, uma vez que a localização espacial, a qual pode ser tratada como uma variável e ainda dependente, constitui uma informação fundamental.

Essa área de estudos em crescente desenvolvimento, que mescla conceitos da análise espacial e da estatística convencional, vem se mostrando uma ferramenta de elevado potencial para diversas análises, com muitas aplicações encontradas na literatura.

Alguns exemplos da aplicação de estatística espacial em diversas áreas do conhecimento podem ser encontrados em: Kampel, Câmara e Quintanilha (2000); Kabos e Csillag (2002); Ramos e Silva (2003a; 2003b; 2007); Queiroz (2003);

Teixeira (2003); Ramos, Silva e Miranda (2004); Krempi (2004); Silva (2004); Ping *et al.* (2004); Orford (2004); Premo (2004); Lopes (2005); Manzato, Soares e Silva (2006); Manzato e Silva (2006); Grigolon, Manzato e Silva (2007).

Bailey (1994) *apud* Queiroz (2003) e Teixeira (2003) destaca o grande potencial de contribuição do SIG como plataforma de aplicação da estatística espacial, agrupando algumas das técnicas da seguinte forma:

- **Análise descritiva simples, transformação e caracterização de dados:** se referem à aplicação de métodos estatísticos, numéricos e gráficos, para manipular e caracterizar conjuntos de dados (histogramas, diagramas de dispersão etc.);
- **Autocorrelação espacial e estrutura de covariância:** se referem à descrição de como e quanto são semelhantes os atributos dos pontos posicionados geograficamente próximos;
- **Modelagem econométrica espacial:** se refere à explicação da variável espacial por meio de um conjunto de variáveis independentes, como na regressão linear tradicional, mas considerando uma ponderação da autocorrelação espacial destas variáveis;
- **Modelagem espacial linear:** se refere a uma extensão das técnicas de regressão espacial aplicadas a variáveis categóricas;
- **Técnicas multivariadas:** se referem à incorporação do caráter espacial na modelagem de múltiplas variáveis dependentes.

Nesse sentido, incorporada aos conceitos de estatística espacial e essencial ao desenvolvimento das etapas da modelagem estatística, Anselin (1995) trata o conjunto dessas técnicas como análise exploratória de dados espaciais, ESDA, do inglês: *Exploratory Spatial Data Analysis*.

Segundo Câmara *et al.* (2002a), geralmente as técnicas de ESDA empregadas são adaptações das ferramentas usuais. Como por exemplo, se na investigação de valores extremos de uma determinada variável se utiliza histogramas, na análise espacial é importante investigar esses valores não só no conjunto dos dados, mas também o seu comportamento em relação aos vizinhos. Para isso, seguindo a classificação da ESDA, as técnicas envolvidas neste processo são apresentadas a seguir:

- **Visualização de dados:** é a forma mais simples e intuitiva de análise dos dados, por meio de mapas;
- **Gráficos de médias e medianas:** permitem explorar simultaneamente a presença de tendências (não-estacionariedade), em que a variância e a covariância entre vizinhos não se mantêm constantes;
- **Análise de autocorrelação espacial:** tem o objetivo de identificar a estrutura de correlação espacial que melhor descreva os dados, por meio de indicadores, como os índices (global e local) de Moran, o índice de Geary e o variograma;
- **Matriz de proximidade espacial:** utilizada para estimar a variabilidade espacial de dados;

- **Média móvel espacial:** utilizada para explorar a variação da tendência espacial dos dados, por meio da avaliação da média dos valores dos vizinhos;
- **Diagrama e mapa de espalhamento de Moran:** constituem maneiras adicionais de se visualizar a dependência espacial.

Após a apresentação das técnicas de ESDA de maneira geral, é importante observar que o trabalho aqui desenvolvido utiliza algumas dessas técnicas para a caracterização de regiões urbanas homogêneas. Além disso, como se trata de um estudo em que todas as variáveis envolvidas são sempre representadas em dados espaciais de áreas, apresenta-se a seguir, um detalhamento de algumas dessas técnicas de ESDA aplicadas neste caso.

3.3. Análise exploratória de dados espaciais aplicada ao estudo de dados de áreas

Dentre as técnicas de ESDA discutidas anteriormente, cujo objetivo é mensurar a dependência espacial, algumas delas são apresentadas a seguir com um detalhamento quando aplicadas a dados espaciais de áreas. São elas: a matriz de proximidade espacial, a média móvel espacial, os indicadores de autocorrelação espacial e o diagrama e mapa de espalhamento de Moran.

3.3.1. Matriz de proximidade espacial (W)

A matriz de proximidade espacial, também chamada de matriz de adjacência ou matriz de peso/ponderação espacial é o elemento principal no estudo da dependência espacial de dados de áreas. Esta matriz estima a variabilidade espacial dos dados, descrevendo o arranjo espacial dos objetos.

A matriz de proximidade espacial é uma matriz quadrada, não estocástica, cujos elementos w_{ij} representam a intensidade da interdependência existente entre as regiões i e j . O valor de cada elemento w_{ij} pode assumir algumas medidas de adjacência entre as zonas analisadas A_i e A_j , como por exemplo:

- $w_{ij} = 1$, se o centróide de A_i está a uma determinada distância de A_j , caso contrário $w_{ij} = 0$;
- $w_{ij} = 1$, se A_i compartilha um lado comum com A_j , caso contrário $w_{ij} = 0$;
- $w_{ij} = l_{ij}/l_i$, em que l_{ij} é o comprimento da fronteira entre A_i e A_j e l_i é o perímetro de A_i . caso contrário $w_{ij} = 0$;

Além disso, a matriz de proximidade espacial apresenta dois tipos básicos de efeitos de vizinhança (adjacência). Um deles se refere ao caso quando há apenas um ponto em comum entre A_i e A_j , em que se considera a vizinhança entre as zonas. O outro tipo está relacionado ao fato da necessidade de se ter realmente um lado comum entre as zonas para se considerar a vizinhança. Em geral, os pacotes computacionais utilizados para a obtenção da matriz de proximidade espacial apresentam as duas possibilidades. No entanto, é importante observar que quando se

trabalha com dados que não possuem uma geometria regular das zonas, esses tipos de efeitos de adjacência não apresentam expressivas diferenças.

Como a matriz de proximidade espacial é utilizada no cálculo de indicadores na fase de análise exploratória, é conveniente normalizar suas linhas, simplificando o cálculo de índices de autocorrelação espacial. A Figura 3.2 ilustra um exemplo de matriz de proximidade normalizada pelas linhas.

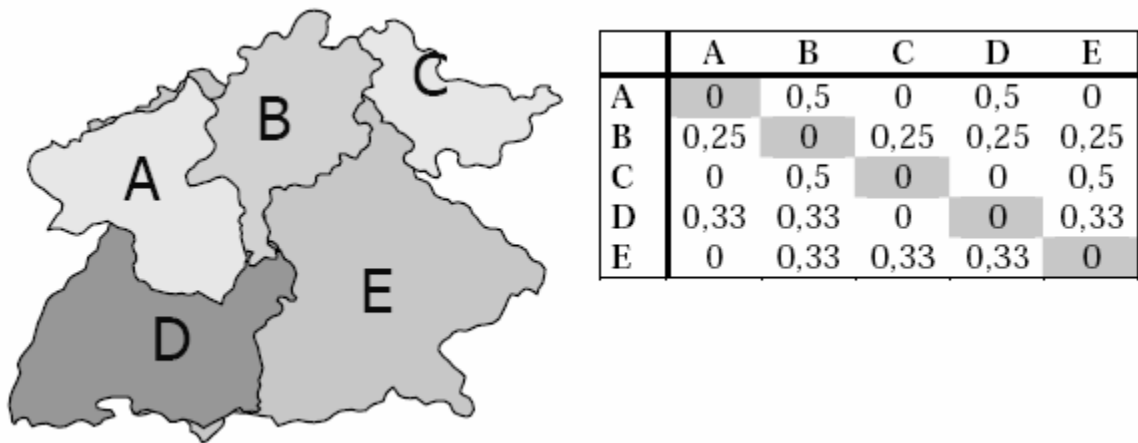


Figura 3.2: Matriz de proximidade espacial normalizada pela linha
(Fonte: Câmara *et al.*, 2002a)

É importante observar também que a idéia de proximidade espacial pode ser generalizada para vizinhos de maior ordem, ou seja, vizinhos dos vizinhos (Câmara *et al.*, 2002a). Isto significa diferentes graus de proximidade (primeira ordem, segunda ordem etc.) que podem ser obtidos no arranjo espacial estudado.

3.3.2. Média móvel espacial

A média móvel espacial, também chamada de vetor das médias ponderadas ou média dos valores dos vizinhos é utilizada para explorar a variação da tendência espacial dos dados, por meio da avaliação da média dos valores dos vizinhos.

Considerando-se um determinado atributo Y , a estimativa da média móvel espacial pode ser obtida por meio da Equação 3.1, a qual relaciona o atributo estudado com a matriz de proximidade espacial.

$$\hat{\mu}_i = \frac{\sum_{j=1}^n w_{ij} y_i}{\sum_{j=1}^n w_{ij}}$$

Equação 3.1: Expressão para a determinação da média móvel espacial

Se os elementos da matriz de proximidade espacial estiverem normalizados, pode-se determinar a média móvel espacial pela Equação 3.2, mostrada a seguir:

$$\hat{\mu}_i = \sum_{j=1}^n w_{ij} y_i$$

Equação 3.2: Expressão para a determinação da média móvel espacial a partir da matriz de proximidade espacial normalizada

3.3.3. Indicadores de autocorrelação espacial

A informação que se busca com a estatística espacial por meio da análise da autocorrelação espacial é a quantificação da dependência espacial, ou seja, quanto o valor de uma variável em uma área é parecido com o do vizinho mais próximo e quanto ele é diferente daquele mais distante. Para isso existem diversos indicadores, como por exemplo, os índices de Moran (global e local), o índice de Geary, o teste *Ipop*, dentre outros.

O índice de Moran é o indicador mais difundido e fornece uma medida geral da associação espacial existente em um conjunto de dados, testando se as áreas conectadas apresentam maior semelhança quanto ao indicador estudado do que o esperado em um padrão aleatório (ANSELIN, 1995; LOPES, 2005; NEVES *et al.*, 2000).

Seu valor varia de -1 a 1 e valores próximos de zero indicam a inexistência de autocorrelação espacial significativa entre os valores dos objetos e seus vizinhos. Valores positivos para o índice indicam autocorrelação espacial positiva, ou seja, o valor do atributo de um objeto tende a ser semelhante aos valores dos seus vizinhos. Valores negativos para o índice, por sua vez, indicam autocorrelação negativa. O grau de autocorrelação é positivo para correlação direta, negativo quando inversa, e nulo quando o indicador se distribui ao acaso entre as áreas, sem relação com a posição.

Considerando os elementos da matriz de proximidade espacial normalizada (w_{ij}), o valor do atributo estudado (y_i) e sua média (\bar{y}), o índice de Moran pode ser escrito conforme a Equação 3.3:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Equação 3.3: Índice I de Moran (global)

De forma alternativa, conforme proposto por Anselin (1996), o índice de Moran também pode ser escrito na sua forma matricial, por meio dos vetores Z e W_Z , segundo a Equação 3.4 (o expoente t define o vetor transposto):

$$I = \frac{Z^t W_Z}{Z^t Z}$$

Equação 3.4: Índice de Moran (global) na forma matricial

O vetor Z é o vetor de desvios, calculado como a diferença entre o valor do atributo observado e a média geral no conjunto de dados. O vetor W_Z é um vetor que representa a diferença entre as médias ponderadas (ou média móvel espacial) e a média geral do conjunto de dados. Alternativamente, este vetor pode ser calculado pelo produto da matriz de proximidade espacial com o vetor de desvios ($W \times Z$).

Embora existam outros indicadores de autocorrelação, conforme mencionado anteriormente, procurou-se apresentar com maiores detalhes apenas o índice de Moran, o qual foi utilizado no estudo aqui desenvolvido.

Para finalizar a apresentação das técnicas de ESDA aplicadas no presente trabalho, são discutidos a seguir o diagrama e o mapa de espalhamento de Moran.

3.3.4. Diagrama e mapa de espalhamento de Moran

Adicionalmente às técnicas de análise da dependência espacial, o diagrama (ou gráfico) e o mapa de espalhamento de Moran constituem maneiras de se visualizar o fenômeno estudado de forma gráfica, permitindo analisar o comportamento da variabilidade espacial.

Essas técnicas se baseiam nos vetores Z e W_z , cujos valores normalizados (divididos pelo desvio padrão do conjunto de dados) são apresentados em um gráfico ou um mapa. A idéia é comparar a distribuição espacial de uma variável com a média de sua vizinhança, construindo um gráfico bidimensional de Z por W_z . Este gráfico é subdividido em quatro quadrantes através da delimitação a partir dos valores nulos de cada eixo, podendo-se identificar quatro zonas distintas, normalmente definidas como Q1, Q2, Q3 e Q4.

Os pontos localizados em Q1 e Q2 indicam áreas em que o atributo possui valor semelhante ao da média das áreas vizinhas. Em Q1 ambos os valores são positivos, por serem superiores à média global, e em Q2 são ambos negativos, por serem inferiores à média global. Estas situações indicam uma autocorrelação espacial positiva. Os pontos localizados em Q3 e Q4 indicam áreas em que o atributo possui valor diferente ao da média das áreas vizinhas. Em Q3 a área possui um valor inferior à média global e as zonas vizinhas possuem um valor superior à média global. Em Q4 a área possui um valor superior à média global e as zonas vizinhas possuem um valor inferior à média global. Estas situações, ao contrário das duas primeiras, indicam uma autocorrelação espacial negativa, ou seja, são áreas que não seguem o padrão estabelecido pelos vizinhos (RAMOS; SILVA, 2003a; 2003b).

O gráfico de espalhamento de Moran também pode ser apresentado na forma de um mapa temático bidimensional (*Box map*), no qual cada polígono é apresentado indicando-se o seu quadrante no gráfico de espalhamento. A Figura 3.3 ilustra o exemplo de um gráfico e de um mapa de espalhamento de Moran.

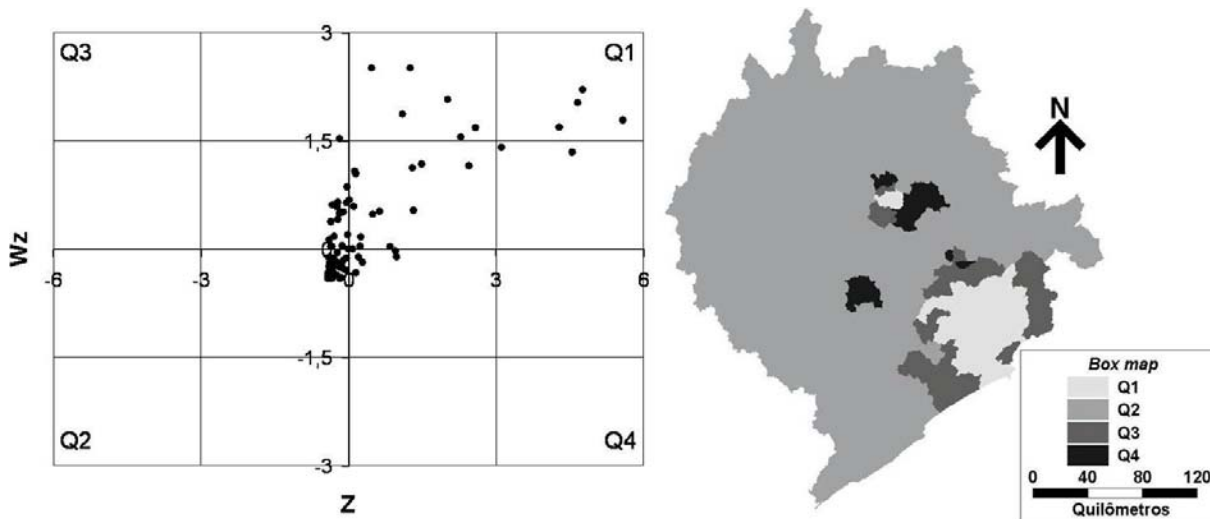


Figura 3.3: Gráfico e mapa de espalhamento de Moran

Após apresentados os principais pontos relacionados à análise espacial, no que se refere à aplicação de uma de suas vertentes, a estatística espacial, o próximo capítulo trata de uma outra área pertencente à análise espacial estudada no presente trabalho: a modelagem espacial.

4. Modelagem espacial

O objetivo deste capítulo é apresentar um referencial teórico sobre a modelagem espacial, que também constitui uma das áreas da análise espacial. O modelo tratado aqui se refere aos Autômatos Celulares utilizados juntamente com ferramentas de Redes Neurais Artificiais.

4.1. Introdução

Embora se observe uma ampla disseminação na utilização de recursos de análise espacial nas mais diversas áreas, pode-se perceber um ponto em comum na maioria delas: a representação dos fenômenos espaciais estudados de maneira estática.

Inicialmente a principal abstração proposta pela análise espacial, juntamente com os sistemas de informação geográfica, foi realmente a representação do mundo real por meio do auxílio matemático-computacional. No entanto, verifica-se a necessidade de uma forma adequada de representação para um significativo conjunto

de fenômenos espaciais que se comportam de maneira dinâmica. Este é o caso, por exemplo, do estudo sobre o escoamento de águas pluviais, da dispersão de sementes e, especialmente, de estudos sobre o planejamento urbano.

Devido à alta complexidade envolvida nos processos de planejamento regional e urbano, este tem recorrido crescentemente aos modelos dinâmicos para gerar cenários múltiplos de apoio à decisão. A explicação para o seu sucesso se deve à correspondência das simulações desenvolvidas com alguns dos fenômenos estudados e à possibilidade de se compreender dinâmicas entre elementos (SILVA, 2002).

Webster e Wu (2001) salientam que a modelagem dinâmica, ao contrário dos modelos convencionais, é capaz de representar os comportamentos complexos presentes no contexto urbano, permitindo a caracterização de fenômenos espaço-temporais de forma dinâmica. Nesse sentido, um grupo de modelos que está presente em algumas das mais recentes aplicações são os Autômatos Celulares (CA, do inglês *Cellular Automata*). Os principais conceitos a respeito dos CA's são apresentados no item a seguir.

4.2. *Cellular Automata* (CA)

Conforme mencionado anteriormente, o estudo de fenômenos complexos está diretamente relacionado à modelagem dinâmica, área intimamente ligada aos CA. O seu desenvolvimento está associado principalmente a três investigadores: von Neumann e Morgenstern's, com a publicação da teoria dos jogos e comportamento econômico em 1944, e Standislav Ulman, na década de 1950, que sugeriu a

possibilidade de se encontrar comportamentos de CA em regras locais gerando padrões globais (SILVA, 2002; YEH; LI, 2002).

Nesse sentido, conforme Silva (2002), o conceito principal da utilização de CA é a possibilidade de se extrair padrões regionais através do comportamento local de um reduzido número de elementos. Da observação desses comportamentos individuais extraem-se grupos que categorizam padrões, formas e intensidades no espaço e no tempo. De forma semelhante ao mundo real, em que o tempo e as dinâmicas entre elementos assumem características que variam de local para local, sendo imprevisíveis e complexas, no CA cada elemento comporta-se individualmente e os resultados são imprevisíveis, assumindo ao nível global, padrões diversos.

Taylor (1992) acrescenta ainda que os modelos de CA procuram explicar alguns fenômenos por meio de regras simples, fazendo com que os seus elementos interajam de forma dinâmica até um estado macroscópico emergir.

Dessa forma, conforme mencionado por Yeh e Li (2002), os CA tornam-se ferramentas poderosas de planejamento quando integradas com sistemas de informação geográfica, permitindo simular diversas variáveis, como por exemplo, dados sociais, econômicos, demográficos, físicos etc. Ao serem integrados ao sensoriamento remoto e SIG, os CA fornecem informações detalhadas a respeito do uso do solo e outras características sobre as cidades, permitindo simulações urbanas de forma realística.

Barredo *et al.* (2003) concordam nesse ponto, acrescentando que os CA se tornaram populares nos últimos anos como ferramentas de modelagem de processos

urbanos. Os autores mencionam a conveniência no uso dos CA nestes casos, em que as cidades podem ser entendidas como sistemas complexos, considerando suas características intrínsecas de emergência, auto-organização e comportamento não linear da dinâmica de uso do solo. Além disso, Liu e Phinn (2003) apresentam uma série de aplicações de CA em planejamento urbano, que vem ganhando rapidamente a preferência de muitos pesquisadores.

Santos (2005) cita ainda, em uma revisão detalhada da literatura, vários trabalhos de CA relacionados ao planejamento urbano: Almeida *et al.* (2002a, 2002b e 2003); Antoni (2002a e 2002b); Batty, Xie e Sun (1999); Cecchini e Rizzi (2001); Dewdney (1989); Diappi, Bolchi e Buscema (2004); Haque e Okamura (2003); Liu e Phinn (2003); Ohgai *et al.* (2001); Ramos e Silva (2003a); Roy e Snickars (1996); Sasaki, Tamura e Saitoh (2001); Wu e Webster (1998); Yeh e Li (2001a, 2001b e 2002).

Conforme Silva (2002) é importante observar que os CA apresentam algumas propriedades importantes. Em geral, pelo menos cinco características constituem um modelo CA, apresentadas a seguir:

- **Geometria da rede:** o território estudado é classificado totalmente em células;
- **Estados:** cada célula tem pelo menos um estado (por exemplo, urbano, não urbano) e o comportamento de cada célula é definido por um conjunto de regras de transição;

- **Vizinhança:** as relações de vizinhança têm um papel importante no CA. Conforme ocorre no mundo real, o espaço envolvente e a sua evolução ao longo do tempo explicam os comportamentos observados;
- **Regras de transição:** podem ser determinísticas ou estocásticas, associadas ao comportamento de cada célula e não há limite para o número de regras;
- **Seqüência de períodos temporais discretos:** quando ativado, o CA se comporta de modo iterativo, de um período temporal para o seguinte atualizando o espaço de maneira síncrona.

Santos (2005), de acordo com a Figura 4.1, apresenta um exemplo simples que permite compreender melhor algumas dessas características dos modelos de CA. Observa-se na Figura 4.1 que a **geometria da rede** é regular com células quadradas, a **vizinhança** considerada compreende oito células ao redor da célula central, as **regras de transição** adotadas são do tipo determinísticas e as células possuem apenas dois **estados** (branco ou preto). Por meio destas condições, pode-se observar o comportamento deste exemplo de CA após seis gerações do modelo.

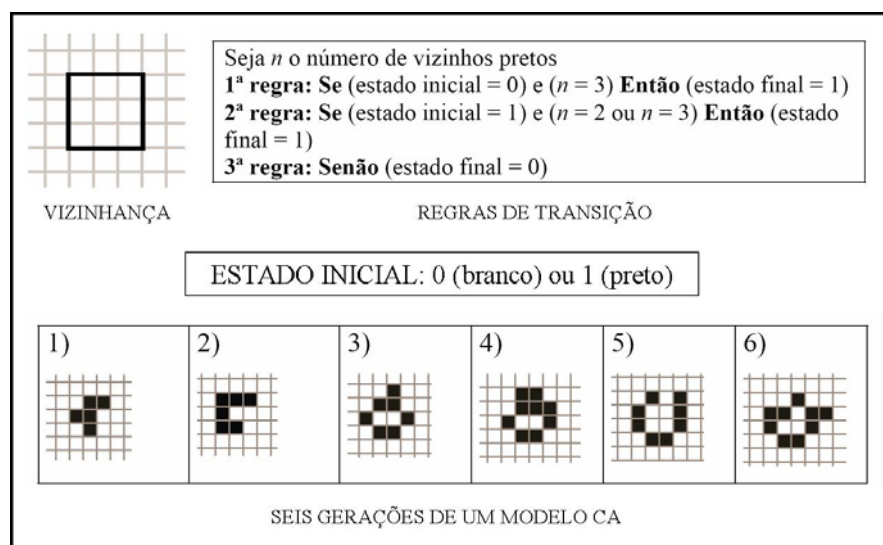


Figura 4.1: Exemplo de um CA simples (Fonte: adaptado de Santos, 2005)

Nesse sentido, conforme apresentado por Santos (2005) e Santos, Lima e Silva (2005), é relativamente fácil generalizar essas características para o estudo urbano, conforme se descreve a seguir.

A **geometria da rede**, que se refere à sua forma e dimensão, pode ser considerada como os territórios urbanos. De maneira geral é utilizada uma geometria regular com células quadradas, entretanto a utilização de redes irregulares contribuiria para o realismo dos modelos, uma vez que a maioria dos objetos urbanos não é regular.

O **estado** da célula pode representar os atributos do espaço territorial, como uso do solo e densidade populacional. Desta forma permite-se caracterizar e avaliar qualquer atributo desejado nos modelos de CA.

Com relação à **vizinhança**, esta se refere à própria célula e um conjunto de células adjacentes que podem interagir entre si. Os dois tipos de vizinhança mais utilizados são a circular e a de Moore (quando oito células formam um quadrado em torno da célula).

Finalmente, as **regras de transição** que são determinadas para refletir como ocorrem os fenômenos no mundo real, podem ser interpretadas como algoritmos na simulação, ou seja, definem as condições futuras da célula especificando o seu comportamento observado no tempo. Há diversas abordagens com relação às regras de transição, conforme se resume a seguir:

- **Determinística:** abordagem tradicional dos modelos de CA, destaca-se pela sua simplicidade e por permitir a reprodução plena das suas modelagens;
- **Estocástica:** abordagem em que é introduzida uma perturbação aleatória com o objetivo de reproduzir aspectos não explicáveis pelos modelos;
- **Obtida por redes neurais artificiais:** abordagem em que os valores dos parâmetros são determinados automaticamente a partir de treinamento da rede neural com os dados. Por um lado isso facilita a simulação, diminuindo a necessidade de dados, entretanto torna muito difícil conhecer as regras de transição adotadas. É importante observar que esta abordagem é determinística na adoção dos pesos das conexões entre os nós, entretanto o processo de treinamento incorpora, de alguma maneira, a distribuição estocástica dos dados.

Com relação à última abordagem, em que as regras de transição são obtidas por meio de redes neurais, esta tem sido muito utilizada devido ao seu alto desempenho no trato dos dados, se comparado aos modelos estatísticos convencionais. As redes neurais são robustas e tolerantes a erros, particularmente em dados que apresentam ruídos. Além disso, elas são especialmente adaptadas para o reconhecimento de padrões.

Em uma comparação, Ramos e Silva (2003a e 2003b) apresentam um estudo em que são identificadas para cada célula as transições ocorridas ao longo do tempo e, assim, se estabelecem regras de transição por meio de uma abordagem determinística utilizando uma estrutura do tipo “se... então... senão”. Em uma outra proposta apresentada (Ramos e Silva, 2007) os autores sugerem a utilização de redes

neurais artificiais para o estabelecimento das regras de transição, mostrando que os resultados obtidos foram superiores àqueles obtidos pela abordagem determinística.

No trabalho aqui desenvolvido, as regras de transição foram obtidas a partir de redes neurais artificiais. Portanto, no próximo item, apresentam-se alguns conceitos referentes a essa ferramenta.

4.3. Redes neurais artificiais (RNA)

Conforme apresentado por Silva *et al.* (2004), as redes neurais artificiais (RNA) são sistemas de equações, em geral não lineares e interligadas, cujo resultado de uma equação é o valor de entrada para várias outras redes. Constituem técnicas computacionais que apresentam um modelo matemático baseado nas funções das redes biológicas, que busca implementar seu comportamento e sua dinâmica.

As RNA derivam da simulação de cérebros humanos, os quais constituem em um grande número de neurônios interconectados. Conforme mencionado anteriormente, as RNA têm a capacidade de reconhecer e classificar padrões por meio de processos de treinamento ou aprendizagem.

A estrutura e processamento das RNA são relativamente simples. As unidades básicas de processamento em uma rede neural são chamadas de neurônios ou nós, os quais são organizados em camadas. Geralmente as redes têm uma camada de entrada e uma de saída e, entre essas, podem conter camadas intermediárias. Um neurônio da camada de entrada é conectado a todos os neurônios da camada intermediária, os

quais são também conectados ao(s) neurônio(s) da camada de saída. A Figura 4.2 ilustra uma estrutura básica de RNA.

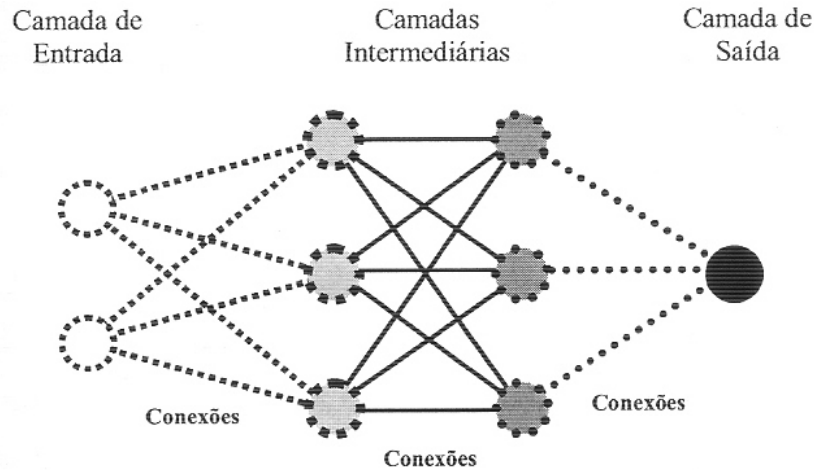


Figura 4.2: Estrutura básica de uma RNA (Fonte: Silva *et al.*, 2004)

As RNA possuem também alguns componentes principais, a saber: conjunto de elementos de processamento, estado de ativação, função de saída para cada elemento de processamento, ligações entre elementos de processamento, propagação e função de combinação, ativação, aprendizagem e ambiente onde o sistema opera. A seguir são apresentados alguns detalhes desses componentes:

- **Elementos de processamento:** são os principais elementos de uma RNA e estão relacionados à sua topologia, ou seja, a especificação do número de neurônios que constituirá a rede, bem como a sua organização nas camadas;
- **Estado de ativação:** cada neurônio em uma RNA tem um estado de ativação associado, o qual depende do que se procura representar com o modelo. Em geral, o estado de ativação pode ser representado por uma função contínua ou discreta;

- **Função de saída de cada elemento de processamento:** os neurônios produzem um sinal de saída que está relacionado com a sua ativação por meio de uma função de transferência (função de saída). Esta função tem o objetivo de controlar a intensidade da função de ativação para se obter o desempenho desejado na rede;
- **Ligações entre elementos de processamento:** as ligações entre os neurônios são representadas por uma matriz de pesos e definem o conhecimento do sistema. Pesos positivos são ligações que excitam gerando um reforço na ativação, e pesos negativos inibem a ativação do neurônio.
- **Propagação da função de combinação:** são regras com as quais os valores de entrada de um neurônio são combinados com os valores da respectiva matriz de pesos para produzir a entrada líquida;
- **Aprendizagem:** processo em que a rede extrai informações relevantes dos exemplos a ela apresentados, criando uma representação própria para o problema.

Para melhor compreensão do processo de utilização de RNA em aplicações diversas, é apresentado a seguir, um tutorial extraído da página de internet do Grupo de Sistemas Inteligentes do Departamento de Informática da Universidade Estadual de Maringá.

Passo 1: Coleta de dados e separação em conjuntos

Este primeiro passo do processo de desenvolvimento de redes neurais artificiais se refere à coleta de dados relativos ao problema e a sua separação em um conjunto de treinamento e um conjunto de testes. Esta tarefa requer uma análise

cuidadosa sobre o problema para minimizar ambigüidades e erros nos dados. Além disso, os dados coletados devem ser significativos e cobrir amplamente o domínio do problema. Não devem cobrir apenas as operações normais ou rotineiras, mas também as exceções e as condições nos limites do domínio do problema.

Normalmente, os dados coletados são separados em duas categorias: dados de treinamento, que serão utilizados para o treinamento da rede e dados de teste, que serão utilizados para verificar o seu desempenho sob condições reais de utilização. Além dessa divisão, pode-se usar também uma subdivisão do conjunto de treinamento, criando um conjunto de validação, utilizado para verificar a eficiência da rede quanto a sua capacidade de generalização durante o treinamento, e podendo ser empregado como critério de parada do treinamento.

Depois de determinados estes conjuntos, eles são geralmente colocados em ordem aleatória para prevenção de tendências associadas à ordem de apresentação dos dados. Além disso, pode ser necessário pré-processar estes dados, através de normalizações, escalonamentos e conversões de formato para torná-los mais apropriados à sua utilização na rede.

Passo 2: Configuração da rede

O segundo passo é a definição da configuração da rede, que pode ser dividido em três etapas: seleção do paradigma neural apropriado à aplicação, determinação da topologia da rede a ser utilizada (número de camadas, número de unidades em cada camada etc.) e determinação de parâmetros do algoritmo de treinamento e funções

de ativação. Este passo tem um grande impacto no desempenho do sistema resultante.

Existem metodologias, “dicas” e “truques” na condução destas tarefas. Normalmente estas escolhas são feitas de forma empírica. A definição da configuração de redes neurais é ainda considerada uma arte, que requer grande experiência dos projetistas.

Passo 3: Treinamento

O terceiro passo é o treinamento da rede. Nesta fase, seguindo o algoritmo de treinamento escolhido, serão ajustados os pesos das conexões. É importante considerar, nesta fase, alguns aspectos tais como a inicialização da rede, o modo de treinamento e o tempo de treinamento.

Uma boa escolha dos valores iniciais dos pesos da rede pode diminuir o tempo necessário para o treinamento. Normalmente, os valores iniciais dos pesos da rede são números aleatórios uniformemente distribuídos, em um intervalo definido. A escolha errada destes pesos pode levar a uma saturação prematura.

Quanto ao tempo de treinamento, vários fatores podem influenciar a sua duração, porém sempre será necessário utilizar algum critério de parada, como por exemplo, a taxa de erro médio por ciclo ou a capacidade de generalização da rede. Pode ocorrer que em um determinado instante do treinamento a generalização comece a degenerar, causando o problema de *over-training*, ou seja, a rede se

especializa no conjunto de dados do treinamento e perde a capacidade de generalização.

O treinamento deve ser interrompido quando a rede apresentar uma boa capacidade de generalização e quando a taxa de erro for suficientemente pequena, ou seja, menor que um erro admissível. Assim, deve-se encontrar um ponto ótimo de parada com erro mínimo e capacidade de generalização máxima.

Passo 4: Teste

O quarto passo é o teste da rede. Durante esta fase o conjunto de teste é utilizado para determinar o desempenho da rede com dados que não foram previamente utilizados. O desempenho da rede, medido nesta fase, é uma boa indicação de seu desempenho real.

Devem ser considerados ainda outros testes como análise do comportamento da rede utilizando entradas especiais e análise dos pesos atuais da rede, pois se existirem valores muito pequenos, as conexões associadas podem ser consideradas insignificantes e assim serem eliminadas. De modo inverso, valores substantivamente maiores que os outros poderiam indicar que houve *over-training* da rede.

Passo 5: Integração

Finalmente, com a rede treinada e avaliada, ela pode ser integrada em um sistema do ambiente operacional da aplicação. Para maior eficiência da solução, este sistema deverá conter facilidades de utilização como interface conveniente e

facilidades de aquisição de dados através de planilhas eletrônicas, interfaces com unidades de processamento de sinais, ou arquivos padronizados. Uma boa documentação do sistema e o treinamento de usuários são necessários para o sucesso do mesmo.

Além disso, o sistema deve periodicamente monitorar seu desempenho e fazer a manutenção da rede quando for necessário ou indicar aos projetistas a necessidade de re-treinamento. Outras melhorias poderão ainda ser sugeridas quando os usuários forem se tornando mais familiares com o sistema, estas sugestões poderão ser muito úteis em novas versões ou em novos produtos.

Após o comentário dos conceitos básicos envolvendo a modelagem espacial, é apresentado no próximo capítulo o método proposto no estudo de caso aqui desenvolvido.

5. Método

Este capítulo descreve o método adotado para o estudo da inter-relação entre aglomerados urbanos e oferta de transportes, discutindo-se primeiramente os principais aspectos relacionados à estatística espacial e à modelagem espacial referentes à aplicação no presente trabalho. Em seguida, esses aspectos são detalhados em um estudo de caso para o estado de São Paulo, apresentando a origem dos dados utilizados bem como o seu tratamento e formatação. Nesse capítulo também estão inseridas algumas características do estado referentes aos seus aspectos político-administrativos e econômicos, evidenciando-se as regiões metropolitanas existentes.

5.1. Técnicas empregadas

Uma das ferramentas utilizadas no estudo aqui desenvolvido foi a **ESTATÍSTICA ESPACIAL**, a qual engloba, dentre outras técnicas, a análise exploratória de dados espaciais (ESDA, em inglês). Os três elementos principais para a sua utilização, a matriz de proximidade espacial (W), o vetor de desvios (Z)

e o vetor de médias ponderadas (W_z), podem ser obtidos de diversas maneiras, de acordo com o pacote computacional utilizado.

Lopes (2005), por exemplo, apresenta os passos para a obtenção desses elementos por meio do *software* SPRING. Outra maneira é por meio do *software* TransCAD, a qual é apresentada ao final deste trabalho, no anexo. Uma outra forma possível ainda é utilizando-se o *software* ArcView, com a extensão *Spacestat* (Anselin e Bao, 1997; Anselin e Smirnov, 1998) incorporada ao programa. Entretanto, uma vez que essa extensão não é mais comercializada, esta última alternativa encontra-se restrita aos usuários que a possuem. Por outro lado, existe uma opção por meio do *software* GeoDa, o qual é distribuído gratuitamente pela Internet.

Os resultados obtidos com a aplicação da técnica de ESDA sobre qualquer variável espacial podem ser visualizados por meio de gráficos e mapas, como por exemplo, o gráfico e mapa de Moran. Em ambos os casos, os resultados são classificados em quatro categorias: quadrantes Q1, Q2, Q3 e Q4, conforme detalhado no Capítulo 3.

A justificativa de se utilizar essa técnica no presente caso se deve ao fato de que a representação em mapas temáticos simplesmente por meio do valor numérico obtido torna difícil a comparação direta entre as diversas variáveis envolvidas. Por meio da técnica de ESDA obtém-se uma padronização dos resultados através de um método que permite informar se existe ou não semelhança entre as variáveis ao considerar relações de vizinhança entre as entidades espaciais.

A **MODELAGEM ESPACIAL** é outra vertente da análise espacial que se integra à etapa de análise confirmatória de dados espaciais e que também se utiliza das ferramentas da estatística espacial.

Os modelos desenvolvidos no presente trabalho utilizam diferentes variáveis relacionadas aos resultados obtidos nas técnicas de ESDA para estabelecer as regras de transição. Este processo é semelhante ao adotado nos CA's em dois aspectos: na adoção de regras de transição e pela hipótese básica de que relações locais geram padrões globais. Por outro lado, o modelo desenvolvido é diferente dos "tradicionais" CA pela forma como subdivide o território. Ao contrário da habitual subdivisão em células iguais e regulares (como no modelo *raster* de SIG) é adotada uma subdivisão coincidente com os limites político-administrativos municipais (por meio de um modelo de SIG vetorial).

As regras de transição podem ser estabelecidas de diversas formas. Ramos e Silva (2003a e 2003b) apresentam um estudo em que são identificadas para cada célula as transições ocorridas ao longo do tempo (como por exemplo, mudanças de quadrantes do gráfico de Moran ao qual pertenciam os municípios) e, assim, se estabelecem regras de transição. Em uma outra proposta apresentada (Ramos e Silva, 2007), os autores sugerem a utilização de redes neurais artificiais para a sua determinação.

Em síntese, o método aqui proposto para alcançar os objetivos do presente trabalho pode ser compreendido por meio da Tabela 5.1, em que são apresentadas as etapas envolvidas no estudo.

Tabela 5.1: Método proposto para o estudo

Etapa	Descrição
Tratamento e formatação dos dados	Tabulação dos dados e obtenção dos indicadores populacional e de oferta de transportes
Análise exploratória dos dados espaciais	Identificação de regiões consideradas uniformes
Modelagem espacial	Construção e avaliação dos modelos referentes à caracterização de RUH's

A partir do conhecimento dos aspectos fundamentais que fazem parte das etapas metodológicas do presente estudo, é apresentado na seqüência um detalhamento das atividades para um estudo de caso desenvolvido para o estado de São Paulo.

5.2. Estudo de caso

O estado de São Paulo possui atualmente 645 municípios que apresentam algumas das maiores densidades demográficas verificadas no país. São Paulo conta também com a segunda maior rede de rodovias no país (pavimentadas e não pavimentadas), sendo a maior neste caso, a do estado de Minas Gerais. Entretanto, possui a maior rede viária pavimentada do Brasil, com mais de 25.000 km de extensão (GEIPOT, 2000).

O estado conta também com três regiões metropolitanas oficiais: a de São Paulo, a da Baixada Santista e a de Campinas, que juntas concentram cerca de 21,7

milhões de habitantes, ou seja, 58,6 % da população do estado e 12 % do total do país. Essas regiões metropolitanas ocupam uma área de aproximadamente 14.000 km², representando 5,6 % da área do estado de São Paulo. Apresentam também um Produto Interno Bruto (PIB) superior ao de muitos países, como Dinamarca e Noruega, e que corresponde a 63,2 % do PIB estadual e quase um quarto do nacional (EMPLASA, 2002).

As principais fontes de dados para o estudo aqui desenvolvido foram: o Instituto Brasileiro de Geografia e Estatística (IBGE), a Fundação Sistema Estadual de Análise de Dados (SEADE), o Instituto Geográfico e Cartográfico (IGC) e o Departamento de Estradas de Rodagem do Estado de São Paulo (DER-SP). Junto a essas instituições foram adquiridos os seguintes dados: a população em cada município do estado nos anos de 1960, 1970, 1980, 1991 e 2000; a base geográfica contendo as divisões administrativas municipais do estado de São Paulo em 1991 e 2000; os mapas rodoviários referentes aos anos de 1960, 1971, 1981 e 1990 e a malha rodoviária pavimentada geo-referenciada do estado, situação em 2000.

O estudo de caso aqui desenvolvido envolve três etapas principais, detalhadas nas seções a seguir: o tratamento e formatação dos dados, a caracterização de regiões urbanas homogêneas por meio da técnica de ESDA e a construção de modelos espaciais utilizando conceitos de CA com a ferramenta de redes neurais.

5.2.1. Tratamento e formatação dos dados

A etapa de tratamento e formatação dos dados tem como objetivo a adequação dos diferentes tipos e formatos dos dados adquiridos junto às instituições mencionadas, de forma a permitir a sua utilização em programas computacionais para diversas análises. O produto final desejado nesta etapa consistia em se determinar apenas duas variáveis principais: a densidade populacional para cada município e um indicador que representasse a oferta de transportes também para cada município. Em ambos os casos as variáveis foram consideradas como representando os anos de 1960, 1970, 1980, 1990 e 2000, dado que a defasagem, quando havia, era pequena (1991 ao invés de 1990, por exemplo). Assim, embora alguns dos dados adquiridos não correspondessem exatamente ao referido período, estes foram considerados equivalentes ao período estudado (regular de 10 em 10 anos).

Primeiramente, com relação à determinação da densidade populacional, foi necessário compatibilizar os dados de população em cada período considerando-se uma divisão administrativa municipal equivalente à do ano de 2000. No Brasil, os municípios do estado de São Paulo, como em qualquer outro estado, sofreram diversos desmembramentos ao longo do tempo dando origem a novos municípios, os quais receberam parcelas da população do município inicialmente existente. Uma vez determinada a situação de cada município em termos populacionais, essa proporção da população foi então aplicada para os períodos anteriores ao desmembramento, “atualizando-se” desta forma, os dados tanto do município criado (emancipado) como daquele que o originou.

A Figura 5.1 ilustra um esquema do processo de atualização dos dados de população dos municípios. Por exemplo, em 1990 o município possuía 1000 habitantes e sofreu um desmembramento em dois municípios, registrando-se em 2000, um município com 500 habitantes e o outro, com 700 habitantes. Se considerado como um único município em 2000, este contaria com uma população de 1200 habitantes. Então, a partir das proporções $500/1200$ e $700/1200$, obtém-se as parcelas para cada um dos municípios que deverão ser aplicadas em 1990 para se determinar as populações compatibilizadas, ou seja, que cada subdivisão conteria.

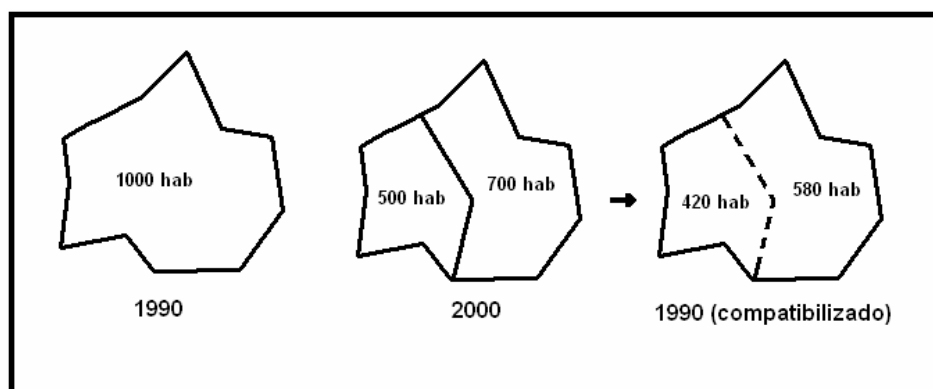


Figura 5.1: Esquema ilustrativo do processo de atualização dos dados para municípios que sofreram desmembramento ao longo do tempo

Completado esse processo, os dados de população ficaram atualizados segundo uma divisão administrativa comum (referente a 2000), possibilitando uma análise coerente da densidade populacional para cada um dos municípios no período de 1960 a 2000.

Quanto à representação da oferta de transportes, é importante observar ainda como foram obtidas as malhas rodoviárias para o período de 1960 a 2000.

Conforme mencionado anteriormente, obteve-se junto ao DER-SP apenas a malha rodoviária geo-referenciada (formato *shapefile*) referente ao ano de 2000.

A partir dessa base geográfica e de mapas analógicos referentes aos anos anteriores (relativos às décadas de 1960, 1970, 1980 e 1990) efetuou-se um processo de ajuste da base geográfica atual aos períodos anteriores, apagando-se os registros de trechos de rodovias que não existiam nos referidos períodos. Após esse processo, foram obtidas cinco bases geográficas referentes às malhas rodoviárias no período de 1960 a 2000 (apresentadas na Figura 5.2), com as quais foi possível o cálculo de um indicador que representasse a oferta de transportes para cada município.

Finalizada essa etapa de tratamento e formatação dos dados, procedeu-se à análise exploratória dos dados espaciais sobre as duas variáveis: a densidade populacional e um indicador de oferta de transportes, conforme se apresenta no próximo item.

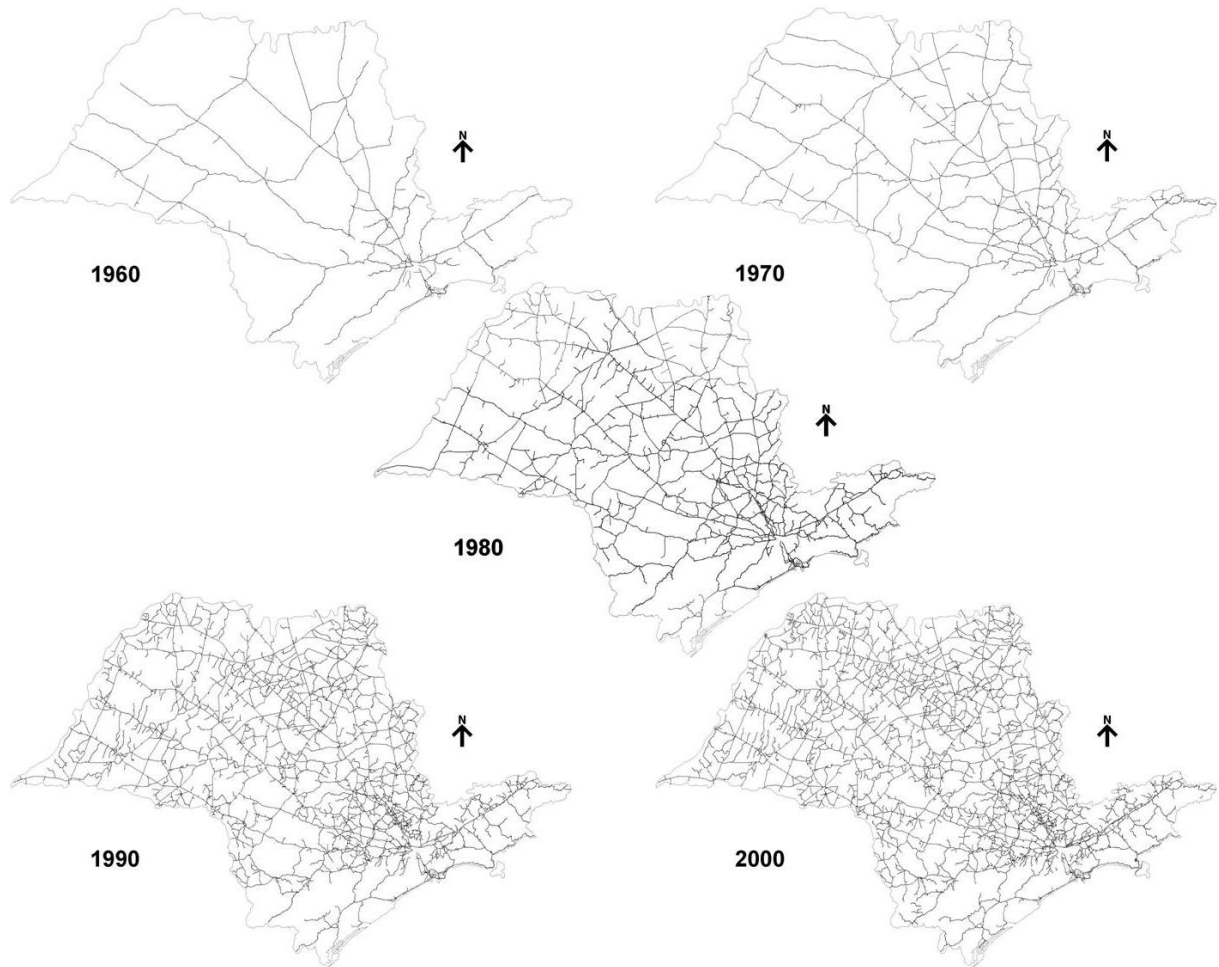


Figura 5.2: Bases geográficas referentes às malhas rodoviárias do estado de São Paulo no período de 1960 a 2000

5.2.2. Análise exploratória dos dados espaciais

Conforme discutido no item 5.1, há várias formas de se determinar os elementos necessários para as técnicas de ESDA. No estudo aqui realizado utilizou-se o *software* ArcView, com a extensão *Spacestat* (Anselin e Bao, 1997; Anselin e Smirnov, 1998) incorporada ao programa para a determinação da matriz de proximidade espacial (W). As informações dessa matriz, utilizada em conjunto com uma planilha eletrônica previamente programada, calculam os outros elementos, Z e W_z , necessários na análise. Conforme mencionado no item 5.1, o

procedimento realizado por esse *software* é restrito aos usuários que o possuem. No entanto, para o entendimento da parte operacional que envolve o procedimento de obtenção dos elementos necessários, pode-se verificar no anexo um tutorial de como determiná-los por meio do *software* TransCAD (apesar do mesmo não ter sido utilizado no presente trabalho).

O objetivo desta etapa é a caracterização de regiões urbanas homogêneas por meio da densidade populacional. Embora este não seja o critério ideal, conforme discutido no Capítulo 2, alguns estudos (Ramos e Silva, 2003a, 2003b e 2007 e Ramos, Silva e Miranda, 2004) apresentam bons resultados quando esta variável é analisada com a técnica de ESDA.

Além da aplicação da técnica de ESDA sobre a densidade populacional para a caracterização de regiões urbanas homogêneas, Manzato, Soares e Silva (2006), em uma tentativa de se testar a viabilidade de utilização de indicadores de oferta de transportes com o mesmo objetivo, apresentam um estudo com a análise de diversos indicadores de oferta de transportes comparados entre si e com a distribuição populacional. Embora restrito aos resultados do ano 2000, os autores sugerem uma combinação dos indicadores populacionais e de oferta viária para a caracterização de regiões urbanas homogêneas. Dentre os indicadores analisados, o índice de cobertura espacial de estrutura de circulação viária, proposto por Magalhães, Teixeira e Yamashita (2004) e dado pela Equação 5.1, foi o que apresentou os melhores resultados. Este indicador diz respeito à área influenciada por um sistema de transporte e é composto por uma faixa ou região ao redor deste sistema (bandas ou *buffers*).

$$IC_{RODO,x} = \frac{\sum^n \gamma(i) \cdot A_i}{A_x}$$

Equação 5.1: Índice de cobertura de estrutura de circulação viária

Em que,

γ_i : função que determina o peso atribuído à i -ésima faixa de cobertura

(*buffer*) tal que $\gamma_i \in [0,1]$;

A_i : área da i -ésima faixa (*buffer*) pertence a X;

A_x : área da região de estudo X;

n : número de faixas.

É importante observar que a função que determina o peso atribuído à i -ésima faixa de cobertura (γ_i) pode ser do tipo linear, polinomial de grau n , dentre outras. No caso específico deste estudo, foram testadas diversas funções e adotou-se uma função do tipo polinomial de grau 5. Além disso, foram consideradas 10 (dez) faixas de cobertura, eqüidistantes 4 km.

A hipótese de combinação dos indicadores foi analisada em Manzato e Silva (2006), ainda com algumas restrições relacionadas aos períodos utilizados (1991, 1996 e 2000, para a densidade populacional e apenas 2000 para a oferta viária). Os autores apresentam uma metodologia comparando a densidade populacional com o indicador de oferta viária proposto por Magalhães, Teixeira e Yamashita (2004), por meio da técnica de ESDA, para a caracterização de regiões urbanas homogêneas. No presente estudo, uma vez que estavam disponíveis tanto os dados sobre a oferta viária como os de densidade populacional no período de 1960 a

2000, procurou-se ampliar a análise de forma semelhante, na tentativa de se identificar regiões urbanas homogêneas a partir da metodologia proposta por Manzato e Silva (2006).

A partir dos resultados obtidos pelo indicador populacional (densidade) e pelo indicador de oferta de infra-estrutura de transportes e, de acordo com os conceitos de ESDA discutidos neste capítulo, foram gerados mapas temáticos, classificados segundo os quadrantes do gráfico de Moran. Esses quadrantes representam a ocorrência das variáveis sobre o território, conforme pode ser constatado no Capítulo 6.

Adicionalmente à metodologia apresentada por Manzato e Silva (2006), uma outra forma de se combinar o indicador populacional com o indicador de oferta viária é por meio de modelos espaciais. Esta alternativa foi analisada no presente estudo e é apresentada no próximo item.

5.2.3. Modelos espaciais

Conforme discutido no item 5.1, há diversas formas de se construir modelos espaciais, principalmente no que se refere ao estabelecimento de regras de transição e à subdivisão do território adotada. Nos modelos espaciais aqui desenvolvidos, as regras de transição foram estabelecidas por meio de redes neurais artificiais, utilizando-se para isso o *software* EasyNN-Plus e uma subdivisão do território coincidente com a divisão administrativa municipal do estado de São Paulo.

Basicamente, dois tipos de modelos foram testados no presente estudo. O primeiro, denominado “3 IN 1 OUT”, foi construído com os 645 registros referentes aos municípios do estado de São Paulo, em que os valores de entrada do modelo eram variáveis de três períodos subseqüentes (como por exemplo: 1960, 1970 e 1980) e os valores de saída, referentes a um período (como por exemplo, 1990). O segundo tipo de modelo, denominado “T, T+10”, apresenta apenas um período de entrada e um de saída. Foi construído de forma que os 645 registros fossem alinhados verticalmente, de acordo com o número de períodos desejados. Ao invés de se ter um modelo com apenas 645 registros, era obtido um modelo com 1290 (645 x 2) registros ou 1935 (645 x 3) registros. A Figura 5.3 ilustra um esquema desses tipos de modelos.

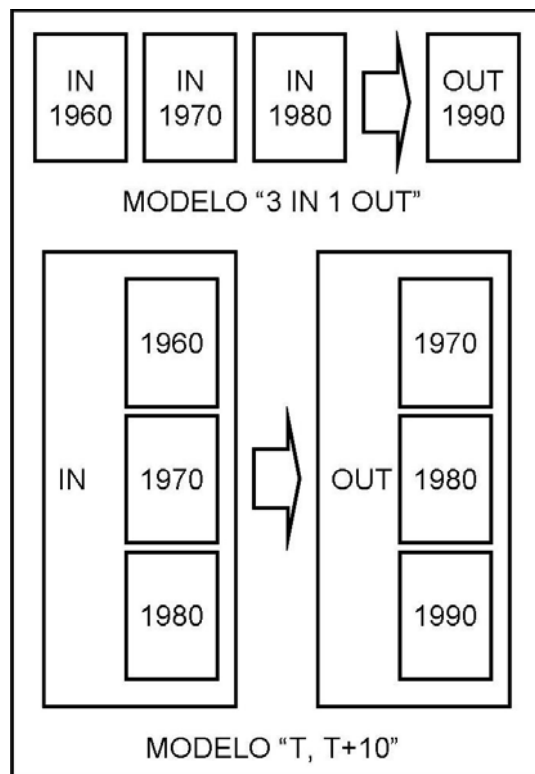


Figura 5.3: Esquema ilustrativo dos tipos de modelos desenvolvidos

Embora a idéia inicial seja analisar a oferta de transportes em conjunto com a distribuição populacional, procurou-se avaliar o desempenho dos modelos espaciais utilizando-se diversas combinações de variáveis. Quanto às variáveis utilizadas, os modelos foram subdivididos em três grupos, conforme é detalhado a seguir:

1. Modelos desenvolvidos apenas com informações sobre a **DENSIDADE POPULACIONAL**. Estes modelos possuem as seguintes variáveis de entrada para cada município: densidade populacional (DP_i), quadrante do gráfico de Moran ao qual pertencem (Q_i), densidade populacional média dos vizinhos adjacentes (DP_j) e número de vizinhos em cada quadrante Q_1 , Q_2 , Q_3 e Q_4 (N° Viz Q_1 , N° Viz Q_2 , N° Viz Q_3 e N° Viz Q_4). A variável de saída se refere à densidade populacional (DP_i).
2. Modelos desenvolvidos com informações sobre a **DENSIDADE POPULACIONAL** e com dados sobre a **OFERTA DE TRANSPORTES** incorporada. Estes modelos possuem as mesmas variáveis de entrada do grupo anterior e a oferta de transportes é adicionada por meio do indicador, não o seu valor simplesmente, mas na forma de variação ao longo do tempo. Essa variação é ainda normalizada entre 0,1 e 0,9 (Δ_{OFERTA}). Aqui a variável de saída também se refere à densidade populacional (DP_i).
3. Modelos desenvolvidos com informações categóricas sobre os **QUADRANTES** do gráfico de Moran. Estes modelos possuem apenas as informações sobre os quadrantes ao qual pertencem, tanto do **INDICADOR POPULACIONAL** ($Q_i DP$) como do **INDICADOR DE OFERTA VIÁRIA** ($Q_i Oferta$)

como variável de entrada. A variável de saída se refere ao quadrante do gráfico de Moran da densidade populacional (Q_i DP).

A partir da combinação dessas informações, sete modelos foram desenvolvidos no presente estudo, apresentados na Tabela 5.2 e detalhados a seguir:

- Modelo I: do tipo “3 IN 1 OUT” apenas com as variáveis sobre densidade populacional;
- Modelo II: do tipo “T, T+10” apenas com as variáveis sobre densidade populacional;
- Modelo III: do tipo “3 IN 1 OUT” com as variáveis sobre densidade populacional e variação da oferta de transportes;
- Modelo IV: do tipo “T, T+10” com as variáveis sobre densidade populacional e variação da oferta de transportes;
- Modelo V: do tipo “3 IN 1 OUT” com variáveis categóricas sobre os quadrantes referentes à densidade populacional e ao indicador de oferta viária;
- Modelo VI: do tipo “T, T+10” com variáveis categóricas sobre os quadrantes referentes à densidade populacional e ao indicador de oferta viária;
- Modelo VII: do tipo “T, T+10” com variáveis categóricas sobre os quadrantes referentes à densidade populacional e ao indicador de oferta viária, mas construídos apenas com os registros (municípios) que sofreram alguma transição relacionada à distribuição populacional entre 1960 e 2000. Uma vez que a maior parte dos registros não sofre alterações no

quadrante referente à densidade populacional durante o período considerado, este modelo procura avaliar o desempenho de um conjunto de dados que sofreram transições entre 1960 e 2000, na tentativa de se identificar algum comportamento diferente na fase de treinamento.

Tabela 5.2: Modelos espaciais desenvolvidos

MODELOS		Variável	
(I)	3 in 1 out	Densidade Populacional	Entrada: $DP_i + Q_i + DP_j + N^{\circ} \text{Viz Q1} + N^{\circ} \text{Viz Q2} + N^{\circ} \text{Viz Q3} + N^{\circ} \text{Viz Q4}$ Saída: DP_i
(II)	T, T+10		
(III)	3 in 1 out	Densidade Populacional + Oferta viária	Entrada: $DP_i + Q_i + DP_j + N^{\circ} \text{Viz Q1} + N^{\circ} \text{Viz Q2} + N^{\circ} \text{Viz Q3} + N^{\circ} \text{Viz Q4} + \Delta \text{oferta}$ Saída: DP_i
(IV)	T, T+10		
(V)	3 in 1 out	<u>Quadrantes</u>	Entrada: $Q_i DP + Q_i \text{ Oferta}$ Saída: $Q_i DP$
(VI)	T, T+10	Densidade Populacional + Oferta viária	
(VII)	T, T+10 (seleção)		

Após a apresentação dos aspectos gerais dos modelos aqui desenvolvidos, é importante discutir alguns detalhes desta etapa de modelagem. Conforme visto no Capítulo 4, de maneira geral a utilização de redes neurais artificiais envolve três fases: treinamento, validação e previsão. No presente estudo, na fase de treinamento da rede, os dados foram selecionados aleatoriamente em 25 % para validação, 25 % para teste e 50 % para treinamento. Após essa fase, nas fases subsequentes (validação e previsão) foram utilizados 100 % dos dados, nos respectivos períodos ao resultado desejado. A Tabela 5.3 apresenta cada uma dessas fases juntamente com as variáveis envolvidas em cada modelo aqui desenvolvido.

Tabela 5.3: Variáveis utilizadas em cada modelo e em cada fase da modelagem por RNA

Modelo	Fase de Treinamento		Fase de Validação		Fase de Previsão	
	Entrada	Saída	Entrada	Saída	Entrada	Saída
I	Variáveis relacionadas à densidade populacional em 1960, 1970 e 1980	Densidade populacional em 1990	Variáveis relacionadas à densidade populacional em 1970, 1980 e 1990	Densidade populacional em 2000	Variáveis relacionadas à densidade populacional em 1980, 1990 e 2000	Densidade populacional em 2010
II	Variáveis relacionadas à densidade populacional em 1960, 1970 e 1980	Densidade populacional em 1970, 1980 e 1990	Variáveis relacionadas à densidade populacional em 1990	Densidade populacional em 2000	Variáveis relacionadas à densidade populacional em 2000	Densidade populacional em 2010
III	Variáveis relacionadas à densidade populacional em 1960, 1970 e 1980 Variação da oferta de transportes entre 1960 e 1980	Densidade populacional em 1990	Variáveis relacionadas à densidade populacional em 1970, 1980 e 1990 Variação da oferta de transportes entre 1970 e 1990	Densidade populacional em 2000	Variáveis relacionadas à densidade populacional em 1980, 1990 e 2000 Variação da oferta de transportes entre 1980 e 2000	Densidade populacional em 2010
IV	Variáveis relacionadas à densidade populacional em 1970 e 1980 Variação da oferta de transportes entre 1960 e 1970 e entre 1970 e 1980	Densidade em 1980 e 1990	Variáveis relacionadas à densidade populacional em 1990 Variação da oferta de transportes entre 1980 e 1990	Densidade populacional em 2000	Variáveis relacionadas à densidade populacional em 2000 Variação da oferta de transportes entre 1990 e 2000	Densidade populacional em 2010
V	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 1960, 1970 e 1980	Quadrante referente à densidade populacional em 1990	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 1970, 1980 e 1990	Quadrante referente à densidade populacional em 2000	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 1980, 1990 e 2000	Quadrante referente à densidade populacional em 2010
VI	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 1960, 1970 e 1980	Quadrante referente à densidade populacional em 1970, 1980 e 1990	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 1990	Quadrante referente à densidade populacional em 2000	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 2000	Quadrante referente à densidade populacional em 2010
VII	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 1960, 1970 e 1980 (seleção daqueles registros que sofreram alteração relacionada à densidade populacional)	Quadrante referente à densidade populacional em 1970, 1980 e 1990	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 1990	Quadrante referente à densidade populacional em 2000	Quadrantes referentes à densidade populacional e ao indicador de oferta viária em 2000	Quadrante referente à densidade populacional em 2010

Com esses modelos foram estimadas as densidades populacionais para o ano de 2000, as quais foram representadas por meio de mapas temáticos classificados segundo os quadrantes do gráfico de Moran e comparadas com os dados reais. Além disso, o desempenho de cada modelo foi medido por meio do cálculo do número de acertos em cada quadrante e o seu acerto total. Finalmente, a partir dos modelos obtidos foi feita uma extrapolação para o ano de 2010, verificando-se a distribuição populacional neste período. Os resultados desta etapa e das outras discutidas anteriormente podem ser observados no Capítulo 6.

Adicionalmente, é apresentada no próximo item uma síntese das atividades desenvolvidas no presente estudo.

5.3. Síntese das atividades desenvolvidas

Com o objetivo de facilitar o entendimento das etapas envolvidas na presente pesquisa, apresenta-se a seguir uma síntese das atividades desenvolvidas. Para as etapas de tratamento e formatação dos dados e análise exploratória dos dados espaciais, os passos são:

1. Aquisição dos dados de população nos períodos desejados e de uma base geográfica com divisão administrativa municipal. Obtenção de informações sobre os processos de desmembramentos ocorridos, como por exemplo, o ano em que ocorreu o desmembramento e de quais municípios foram originados;
2. Ajuste dos dados de população para o cálculo da densidade populacional segundo uma divisão administrativa comum entre os períodos considerados;
3. Aquisição de informações sobre a oferta viária: base de dados georeferenciada atual e mapas dos períodos anteriores;
4. Caso não haja as bases geográficas nos períodos anteriores, obtê-las por meio de um ajuste, apagando-se os registros que não existiam anteriormente;
5. Cálculo do indicador de oferta viária, segundo a Equação 5.1;

6. Determinação dos elementos W , Z e W_Z , referentes à análise exploratória dos dados espaciais, para as duas variáveis: a densidade populacional e o índice de cobertura de estrutura de circulação viária;
7. Determinação dos quadrantes a que pertence cada um dos municípios, para as duas variáveis. Representação desses quadrantes em forma de mapas temáticos, classificados segundo os quadrantes Q1, Q2, Q3 e Q4.

Com relação aos modelos espaciais, os passos necessários são:

1. Além da utilização da densidade populacional e do quadrante a que pertence cada município (os quais já foram determinados), é necessário o cálculo das seguintes variáveis adicionais: densidade populacional média dos vizinhos de cada município e número de vizinhos de cada município que pertence aos quadrantes Q1, Q2, Q3 e Q4;
2. Normalização dos valores do indicador de oferta viária entre 0,1 e 0,9 e cálculo de sua variação nos períodos desejados (como por exemplo, variação de 1960 a 1970, ou 1960 a 1980 etc.);
3. Construção dos modelos espaciais, de acordo com as informações das Tabelas 5.2 e 5.3;
4. Para a fase de treinamento da rede, selecionar aleatoriamente as proporções de registros para validação, teste e treinamento. O tempo destinado a esta fase depende do desempenho da rede, o qual é fornecido pelo próprio *software*, por meio do erro médio;
5. Para a fase de validação, construir mapas temáticos (*Box maps*) com os resultados gerados pela rede, para cada um dos modelos. Comparar a

porcentagem de acertos de quadrantes obtidos pelos modelos com os dados reais;

6. Para a fase de previsão, construir mapas temáticos (*Box maps*) com os resultados gerados pela rede, para cada um dos modelos.

6. Análise dos resultados

Este capítulo apresenta a descrição das análises dos resultados obtidos no estudo de caso do presente trabalho. São apresentados primeiramente os resultados da análise exploratória de dados espaciais (ESDA) aplicadas sobre as duas variáveis: a densidade populacional e o indicador de cobertura espacial de estrutura de circulação viária. Em seguida são apresentados os resultados dos modelos espaciais desenvolvidos para o ano de 2000, discutindo-se os principais pontos a respeito do desempenho de cada um deles e, ao final, apresentam-se os resultados dos modelos aplicados em uma extrapolação dos dados para o ano de 2010.

6.1. Análise exploratória dos dados de densidade populacional

As Figuras 6.1, 6.2, 6.3, 6.4 e 6.5 ilustram os resultados obtidos por meio da análise exploratória dos dados de densidade populacional nos anos de 1960, 1970, 1980, 1990 e 2000, respectivamente. Estas figuras, que representam mapas temáticos classificados segundo os quadrantes do gráfico de Moran (*Box map*),

apresentam, em cor clara, municípios pertencentes ao quadrante Q1, os quais possuem valores altos de densidade populacional cercados por vizinhos que também possuem valores altos. É importante observar que a maioria dos municípios dessa região homogênea coincide com os limites das regiões metropolitanas oficiais (Campinas, São Paulo e Baixada Santista), mostradas nos mapas temáticos.

Observam-se também, em cinza escuro, os municípios pertencentes ao quadrante Q3, os quais possuem valores baixos para a densidade populacional enquanto que os seus vizinhos possuem valores altos. Essa condição especial torna esses municípios potenciais ingressantes à região homogênea composta por aqueles inseridos em Q1, dada a proximidade existente entre os municípios em Q3 e em Q1. Além disso, constata-se por meio da análise das Figuras 6.1, 6.2, 6.3, 6.4 e 6.5 alguns municípios pertencentes ao quadrante Q4, os quais possuem um alto valor de densidade populacional e seus vizinhos possuem valores baixos. Esses municípios podem ter uma expressiva importância local por concentrarem serviços que atendam os municípios vizinhos. Os índices de Moran obtidos para os respectivos períodos foram iguais a 0,2581; 0,3821; 0,4943; 0,5428 e 0,5702; o que indica um razoável (e crescente) grau de autocorrelação espacial nos cinco casos.

Como um dos objetivos propostos no presente estudo é a identificação de regiões urbanas homogêneas utilizando técnicas de análise espacial, os resultados obtidos até esta etapa permitem inferir esta possibilidade por meio da densidade populacional, conforme já apresentado em Ramos e Silva (2003a, 2003b e 2007) e Ramos, Silva e Miranda (2004). Essa identificação pode ser feita, primeiramente, por meio dos municípios em Q1 e, adicionalmente, observando-se a existência de municípios em Q3 ao redor.

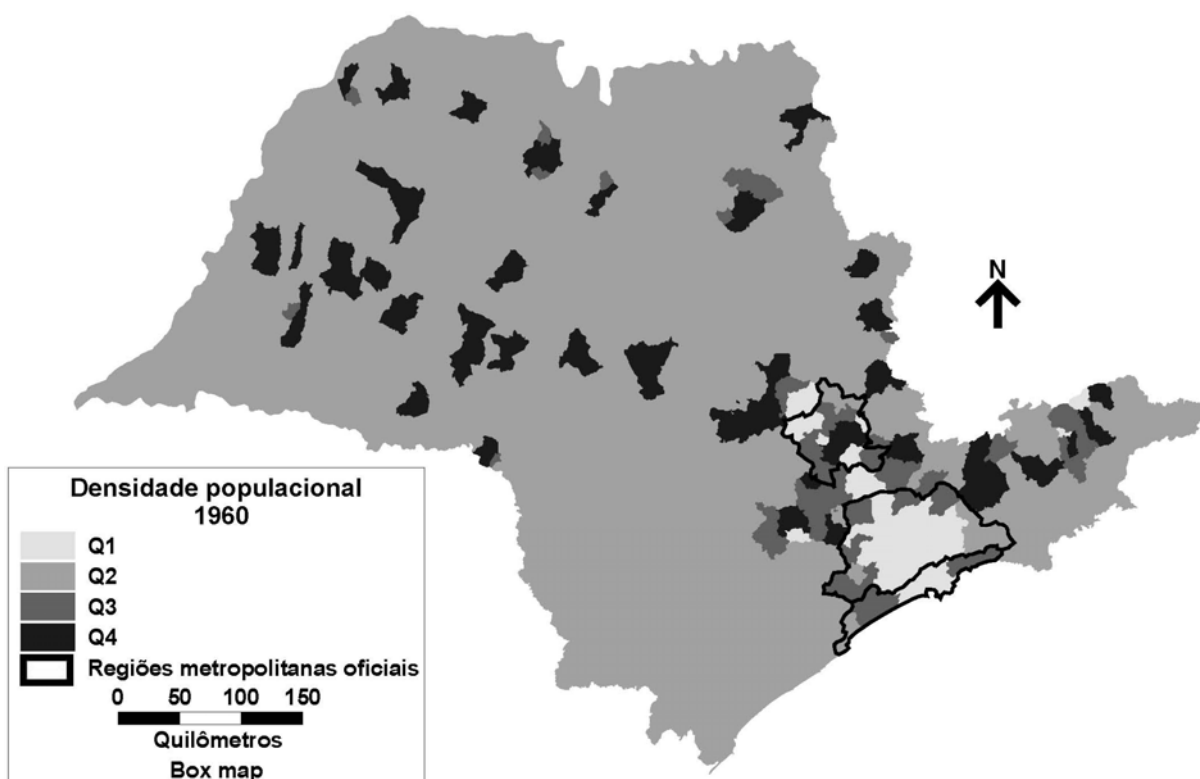


Figura 6.1: *Box map* representando a densidade populacional em 1960

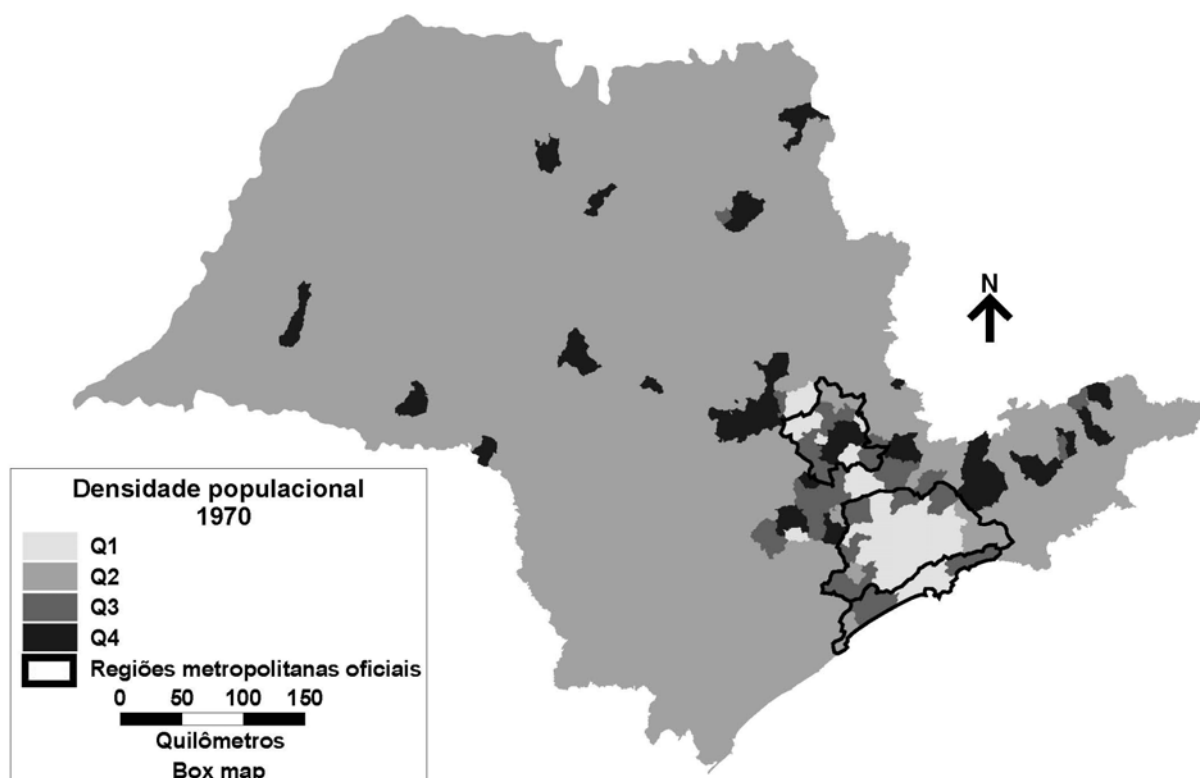


Figura 6.2: *Box map* representando a densidade populacional em 1970

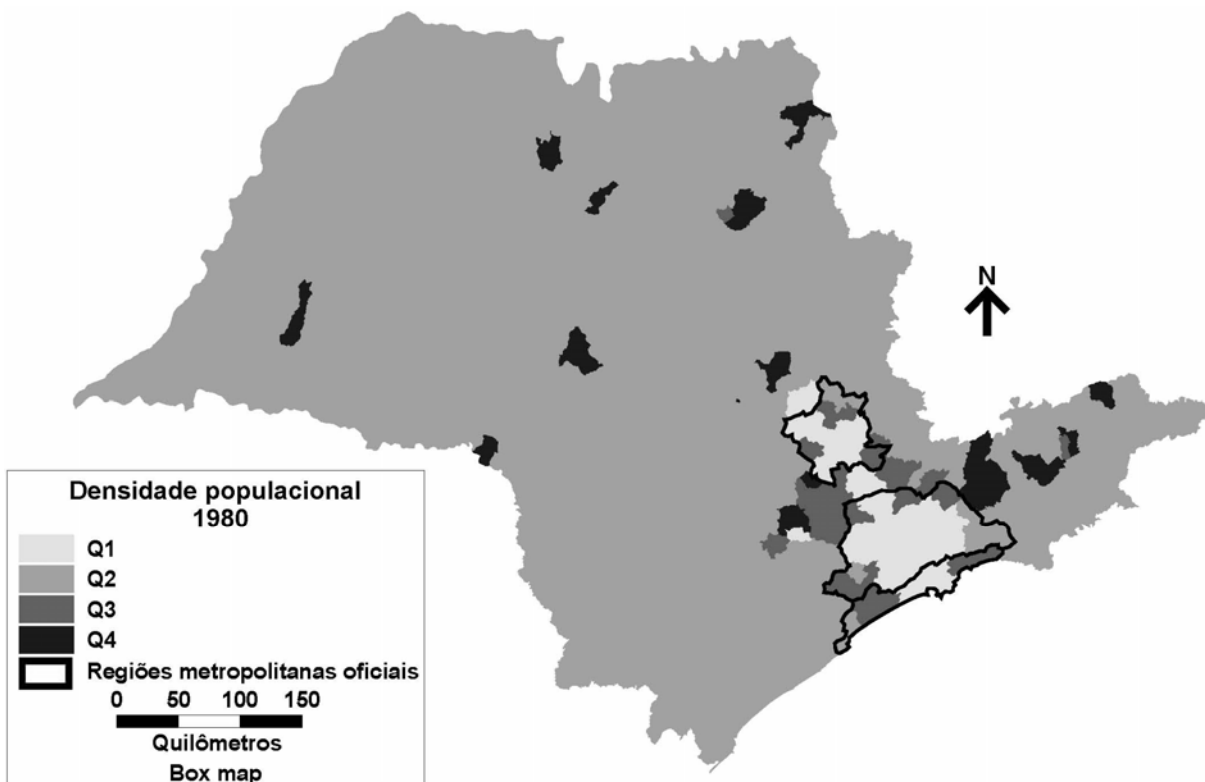


Figura 6.3: *Box map* representando a densidade populacional em 1980

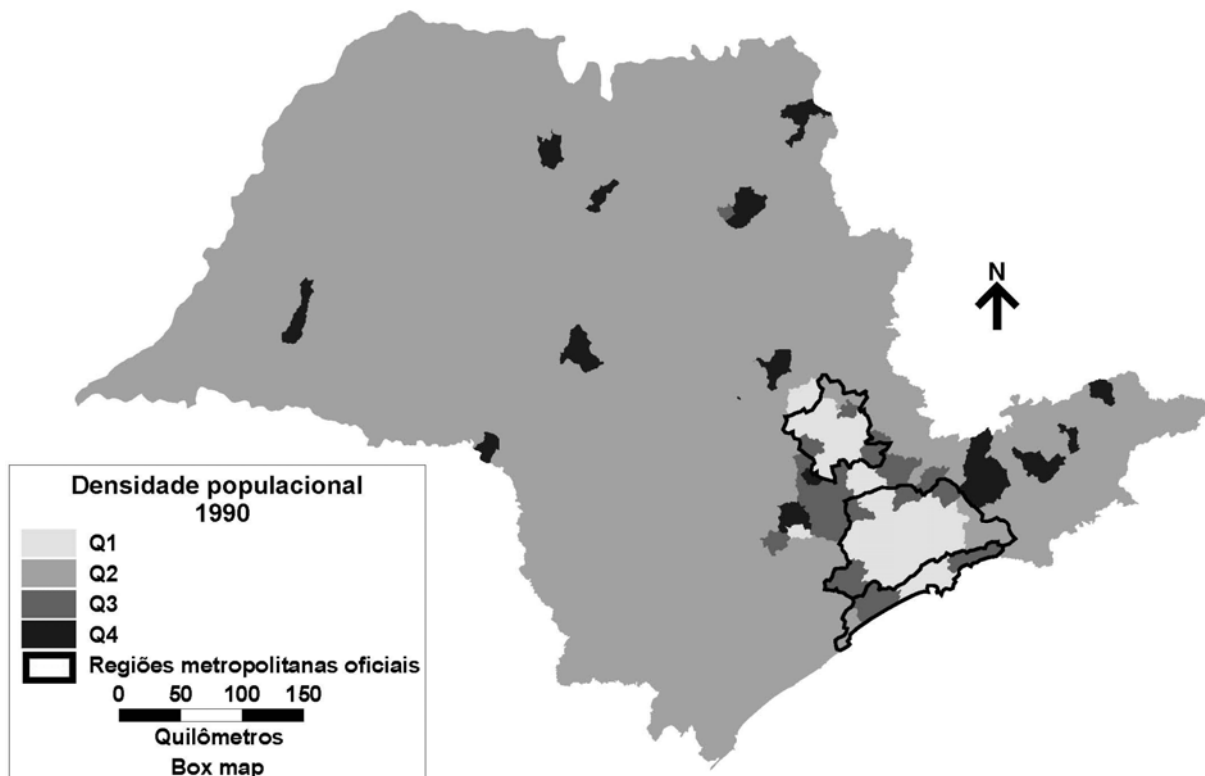


Figura 6.4: *Box map* representando a densidade populacional em 1990

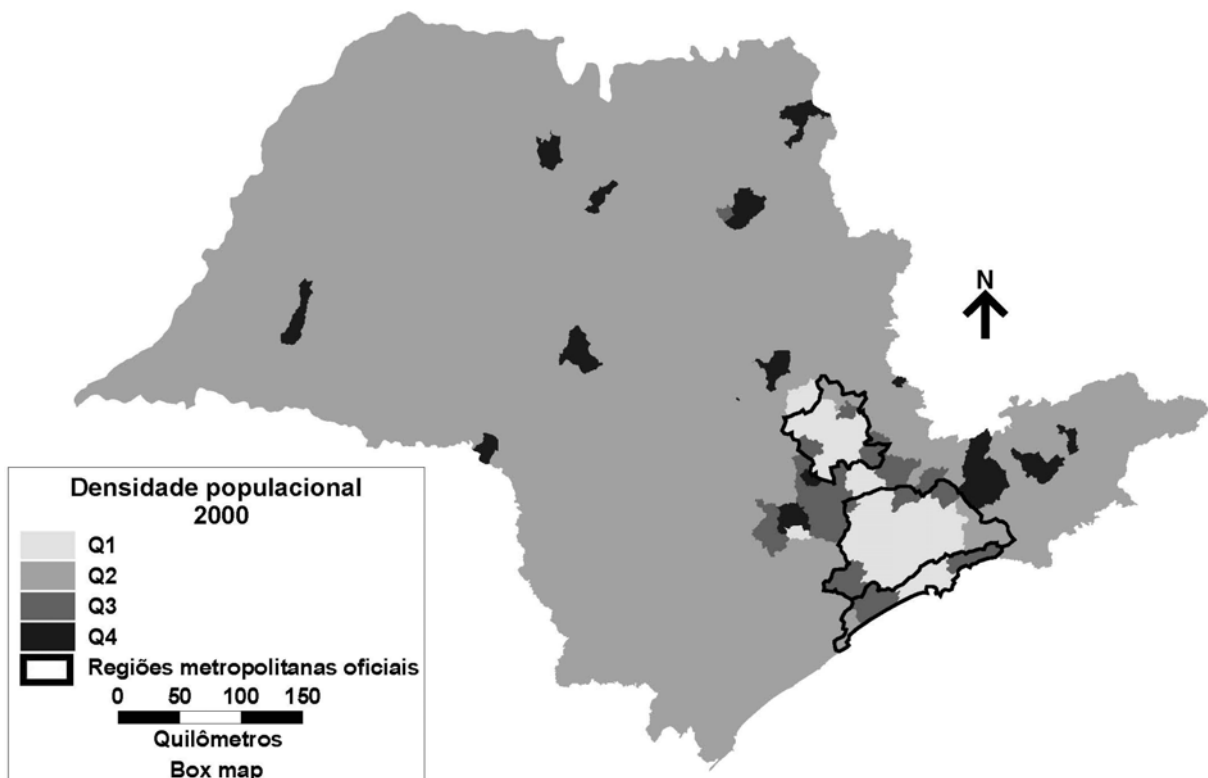


Figura 6.5: Box map representando a densidade populacional em 2000

6.2. Análise exploratória dos dados de infra-estrutura

As Figuras 6.6, 6.7, 6.8, 6.9 e 6.10 ilustram os resultados obtidos por meio da análise exploratória de dados espaciais aplicada sobre o indicador de cobertura de estrutura de circulação viária nos anos de 1960, 1970, 1980, 1990 e 2000, respectivamente. Estas figuras, que representam mapas temáticos classificados segundo os quadrantes do gráfico de Moran (*Box map*), apresentam, em cor clara, municípios pertencentes ao quadrante Q1, os quais possuem valores altos para o indicador de oferta viária cercados por vizinhos que também possuem valores altos. Isto significa uma região composta por municípios com boa dotação de infra-estrutura rodoviária.

Observam-se também, em cinza escuro, os municípios pertencentes ao quadrante Q3, os quais possuem valores baixos para o indicador de oferta viária, enquanto que os seus vizinhos possuem valores altos. Além disso, constata-se por meio da análise das Figuras 6.6, 6.7, 6.8, 6.9 e 6.10 alguns municípios pertencentes ao quadrante Q4, os quais possuem um alto valor para o indicador de oferta viária e seus vizinhos possuem valores baixos.

Ao analisar estas figuras não se percebe um padrão segundo a evolução da oferta de transportes ao longo do período 1960 a 2000, ao contrário do que se observa com a densidade populacional. Isto pode ser confirmado pelos valores dos índices de Moran, respectivamente iguais a: 0,5698; 0,4910; 0,3955; 0,4680 e 0,4618.

Desta forma torna-se difícil qualquer caracterização de regiões urbanas homogêneas por meio do indicador de oferta de infra-estrutura viária. Além disso, ao contrário do que foi sugerido por Manzato, Soares e Silva (2006) e Manzato e Silva (2006), a combinação direta de indicadores populacionais e de infra-estrutura de transportes com técnica de ESDA (como por exemplo, *Box maps*) para a identificação de regiões urbanas homogêneas não é tão evidente. Portanto, conforme é apresentado no próximo item e já discutido na metodologia, a combinação da oferta de transportes com a distribuição populacional pode ser feita por meio dos modelos espaciais, em que são testadas diferentes maneiras de se obter um resultado que expresse de maneira razoável a inter-relação entre uso do solo e oferta de transportes.

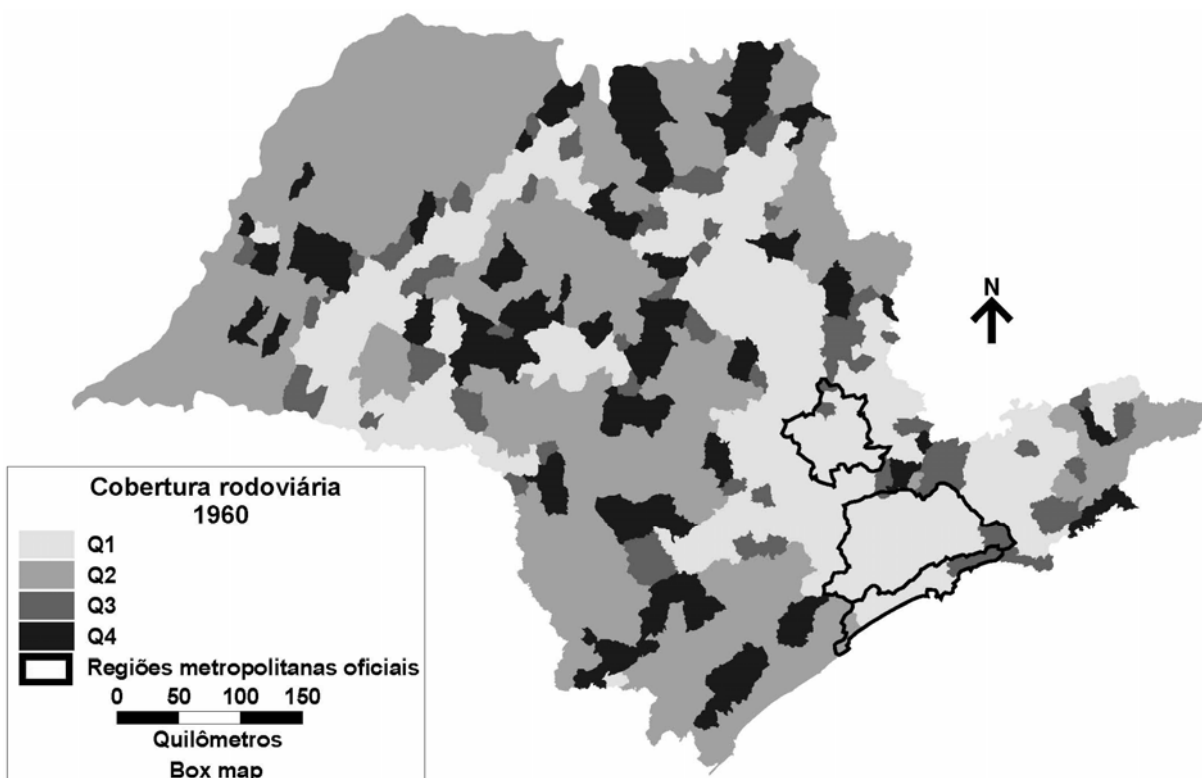


Figura 6.6: *Box map* representando o indicador de cobertura espacial de estrutura de circulação viária em 1960

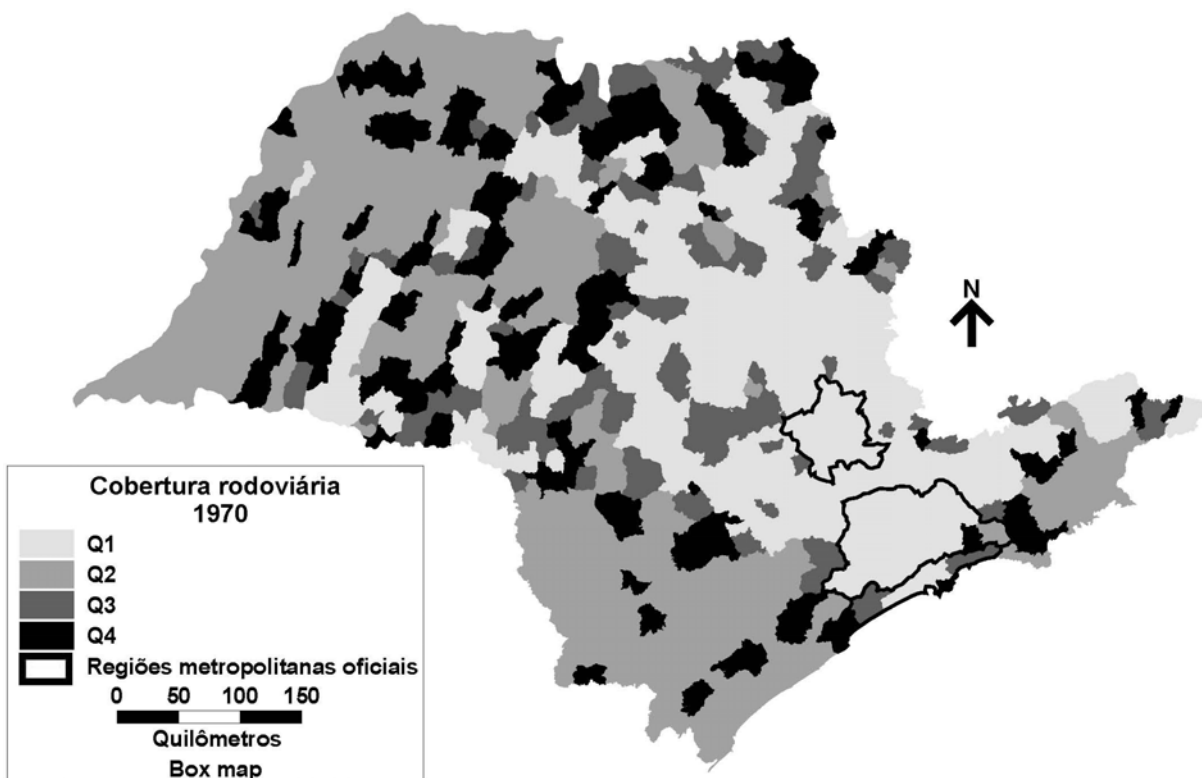


Figura 6.7: *Box map* representando o indicador de cobertura espacial de estrutura de circulação viária em 1970

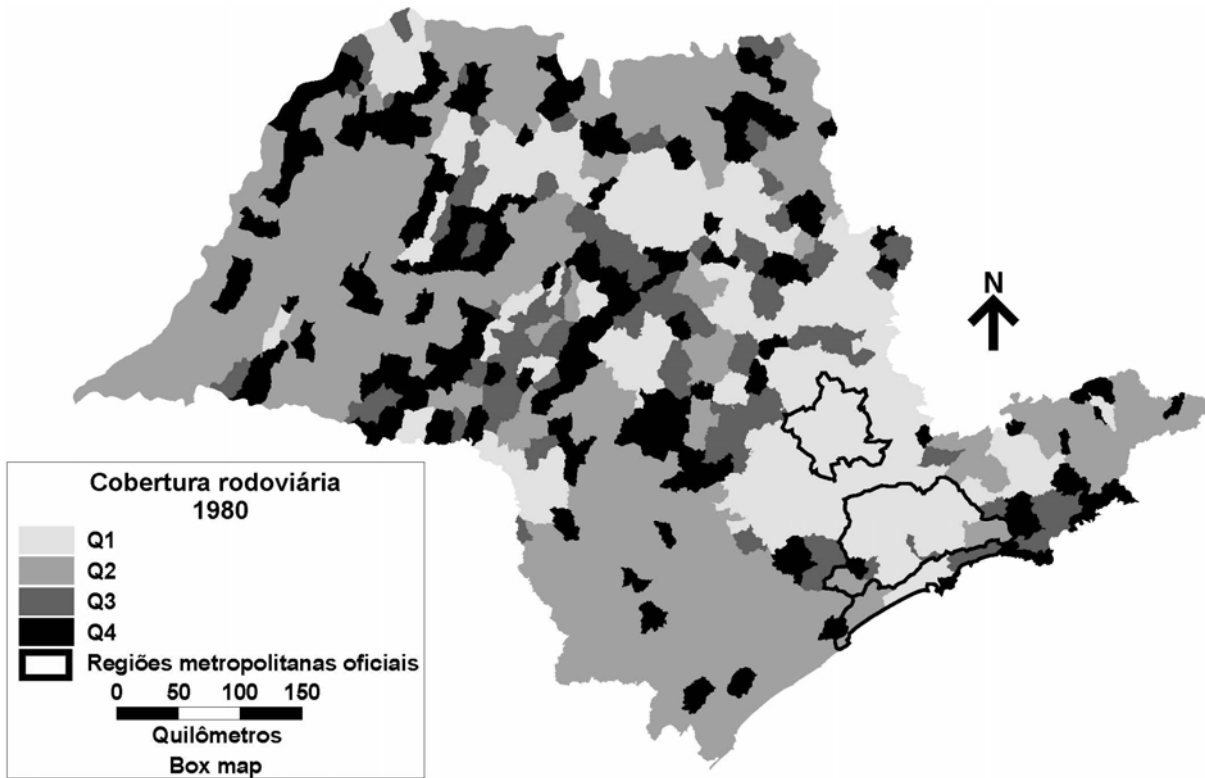


Figura 6.8: *Box map* representando o indicador de cobertura espacial de estrutura de circulação viária em 1980

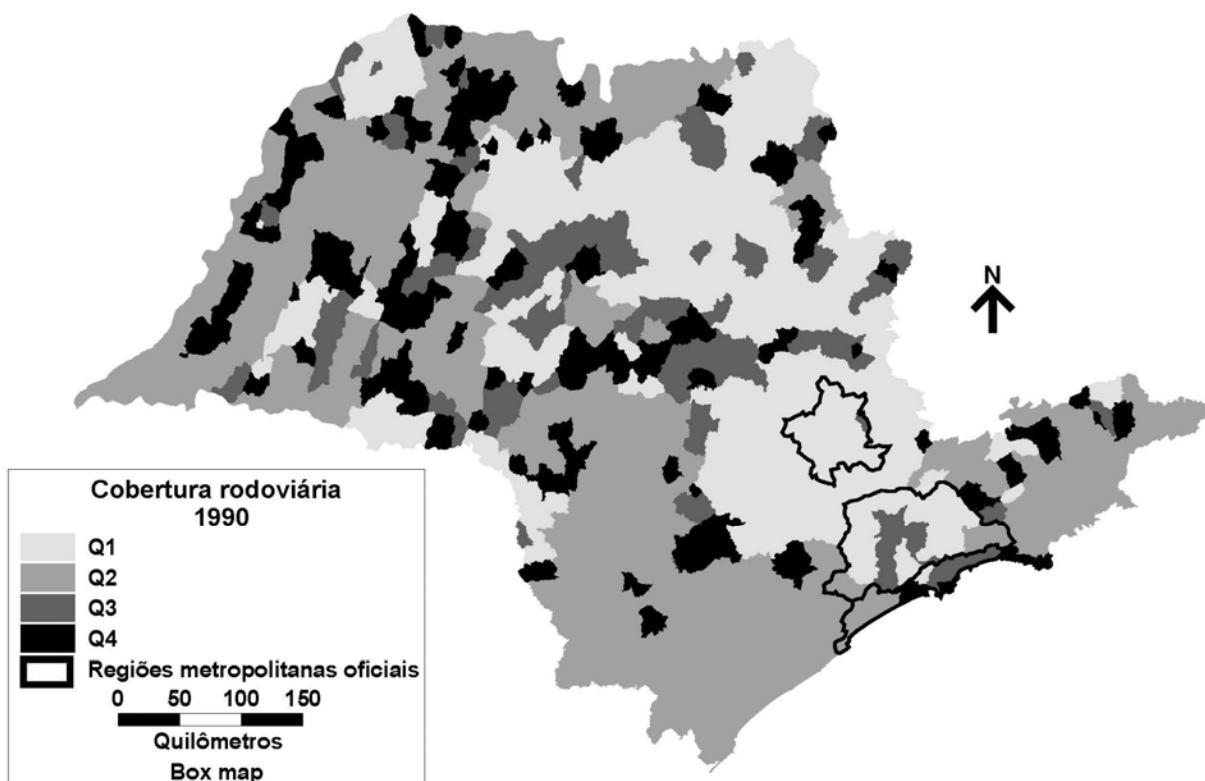


Figura 6.9: *Box map* representando o indicador de cobertura espacial de estrutura de circulação viária em 1990

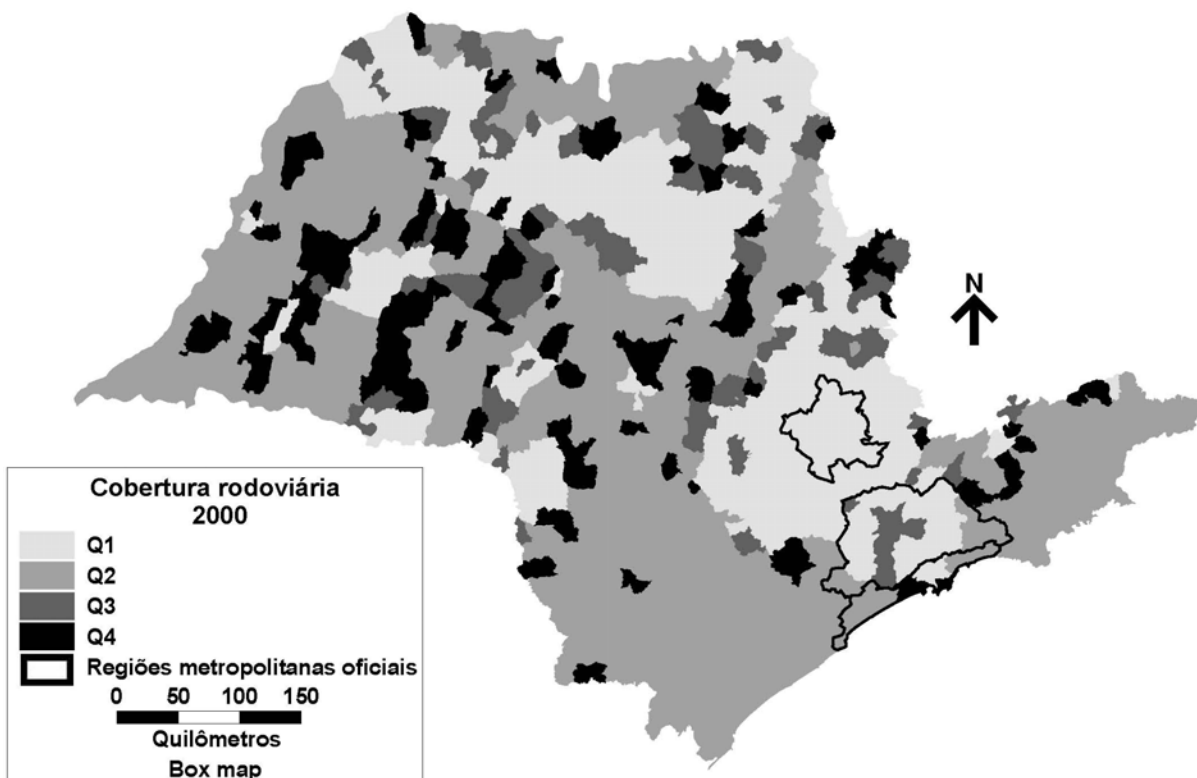


Figura 6.10: Box map representando o indicador de cobertura espacial de estrutura de circulação viária em 2000

6.3. Modelos espaciais obtidos para o ano de 2000

Os resultados dos modelos espaciais desenvolvidos no presente estudo são apresentados nas Figuras 6.11, 6.12, 6.13, 6.14, 6.15, 6.16 e 6.17. Essas figuras ilustram, em mapas temáticos classificados segundo os quadrantes do gráfico de Moran, a representação da densidade populacional estimada para o ano de 2000 por meio dos modelos I, II, III, IV, V, VI e VII, respectivamente. Além disso, para facilitar a comparação, é fornecida a cada uma das figuras uma imagem reduzida (sem escala) da Figura 6.5, a qual corresponde aos quadrantes obtidos com os dados reais referentes ao ano de 2000. É importante observar que os resultados apresentados pelas Figuras 6.11, 6.12, 6.13, 6.14, 6.15, 6.16 e 6.17 se referem à fase da validação dos modelos (100 % dos dados). Adicionalmente, apresenta-se na Tabela 6.1 um resumo

com as proporções de acertos parciais (por quadrante) e totais de cada um dos modelos, além do número absoluto de acertos sobre o total de municípios em cada classe (quadrante).

Os modelos I e II são construídos com base apenas nas informações de distribuição populacional. Conforme se verifica na Tabela 6.1, apesar de apresentarem uma porcentagem de acertos relativamente alta referente aos quadrantes (quando comparados aos quadrantes obtidos com os dados reais), a visualização sobre o território dos resultados estimados pelos modelos apresenta algumas inconsistências, como pode ser observado comparando-se as Figuras 6.11 e 6.12 com a Figura 6.5. Este fato poderia justificar a hipótese inicial do presente estudo, sobre a existência de uma inter-relação entre uso do solo e oferta de transportes.

Nesse sentido, como um dos objetivos do presente trabalho era a determinação de um modelo que expressasse de forma conjunta a inter-relação entre uso do solo e oferta de transportes, foram obtidos os modelos III, IV, V, VI e VII. Esses modelos apresentam a combinação da oferta de transportes com a densidade populacional.

Como pode ser observado pelos resultados contidos na Tabela 6.1, os modelos que incorporam a oferta de transportes apresentam um desempenho ligeiramente superior àqueles obtidos pela modelagem da densidade populacional apenas. Ao compará-los visualmente por meio das Figuras 6.11, 6.12, 6.13, 6.14, 6.15, 6.16 e 6.17 e ainda com a Figura 6.5, observa-se que a inserção da oferta de transportes como uma variável melhora bastante o desempenho dos modelos.

Outro ponto dessa análise a ser observado é o fato de que aqueles modelos estimados de forma categórica (Modelos V, VI e VII), apenas com as informações sobre os quadrantes ao qual pertencia cada município, apresentam desempenho ainda melhor, conforme mostra a Tabela 6.1. No entanto, estes modelos têm a desvantagem de não permitirem o conhecimento do valor da densidade populacional estimada. Ao contrário, os modelos que estimam essa variável (Modelos III e IV) permitem revelar o seu valor para cada município, embora necessitem de variáveis adicionais para a sua determinação, conforme se descreve na metodologia.

Finalmente, um último aspecto na análise dos resultados obtidos pelos modelos que incorporam a oferta de transportes é o resultado obtido pelo Modelo VII. Este modelo foi construído com base somente nas transições referentes aos quadrantes obtidos pelo indicador populacional observadas durante o período de 1960 a 2000. É importante destacar que este modelo apresenta um desempenho muito bom, conforme se verifica na Tabela 6.1, entretanto o modelo “erra” nos resultados obtidos para a classe referente ao quadrante Q4, classificando-os como Q3.

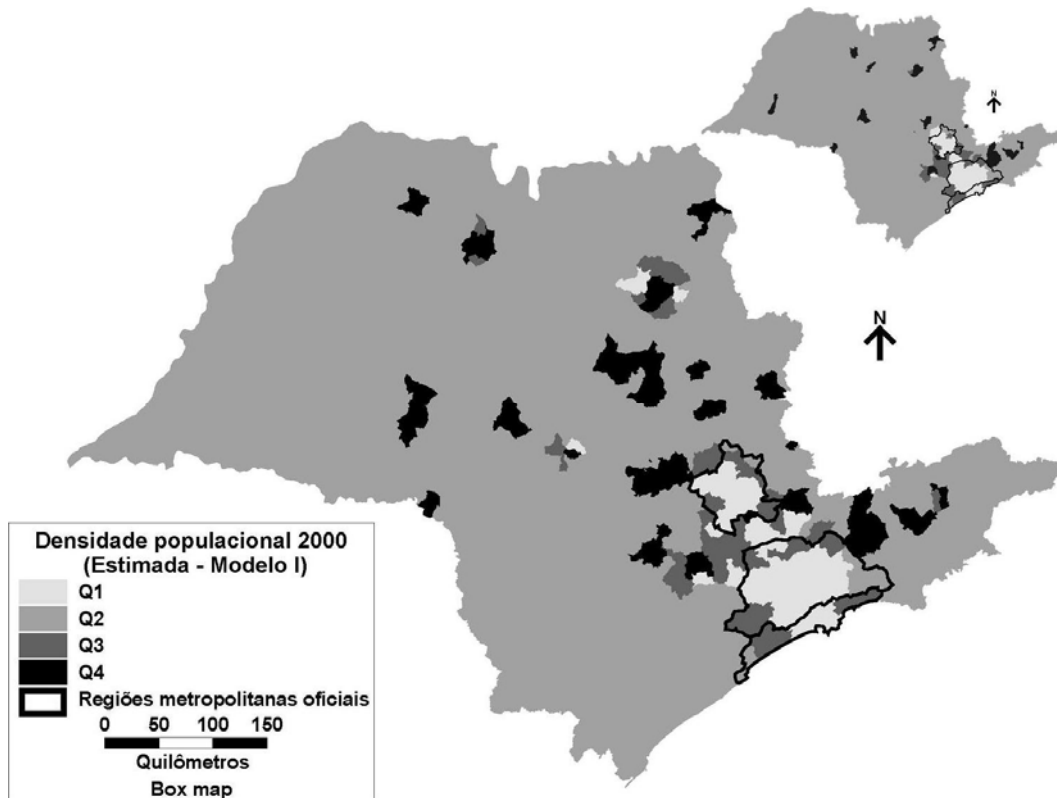


Figura 6.11: *Box map* representando a densidade populacional estimada pelo Modelo I em 2000

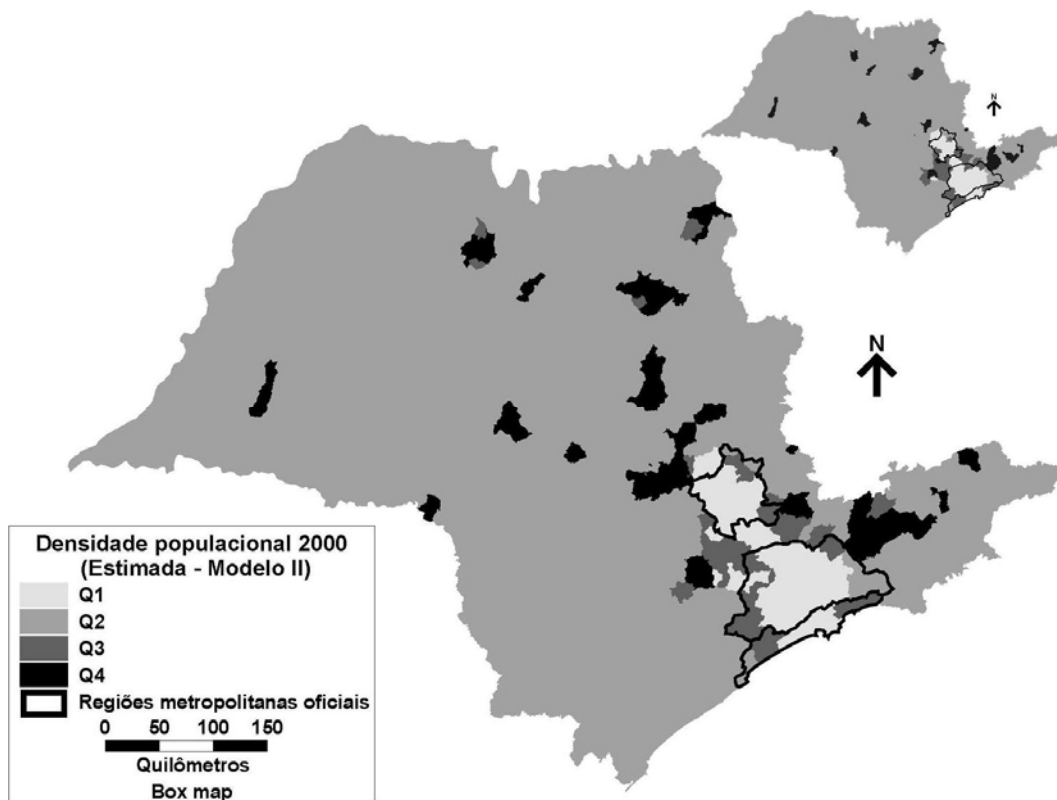


Figura 6.12: *Box map* representando a densidade populacional estimada pelo Modelo II em 2000

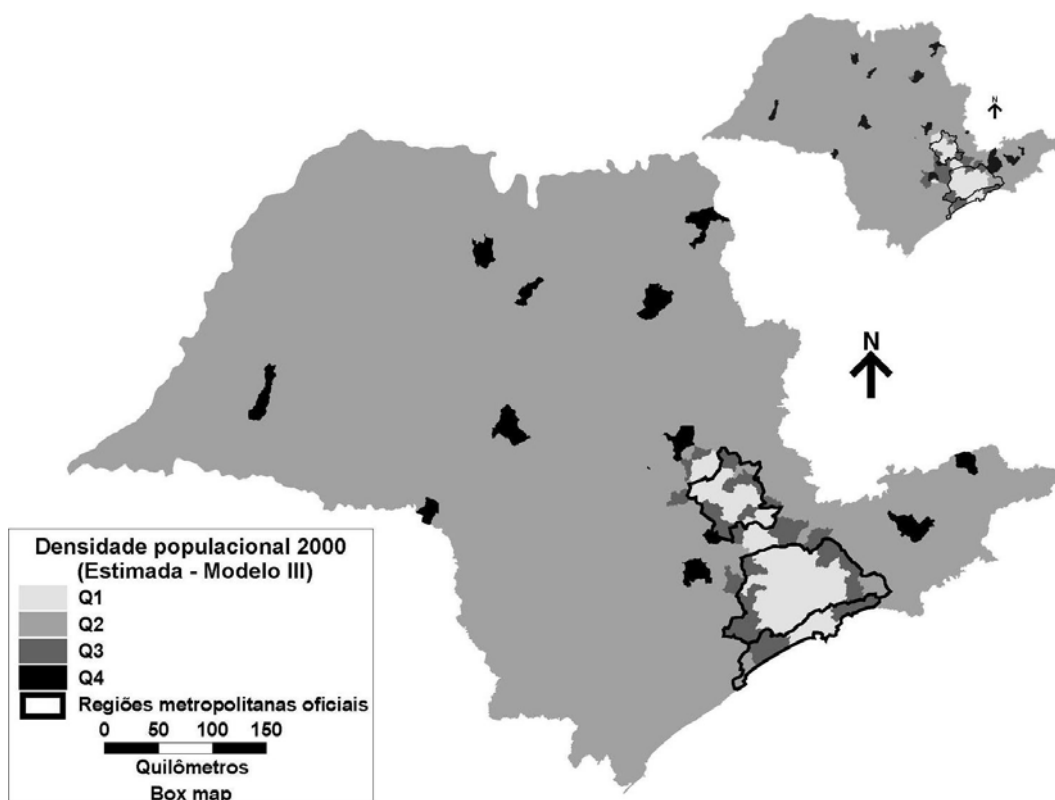


Figura 6.13: *Box map* representando a densidade populacional estimada pelo Modelo III em 2000

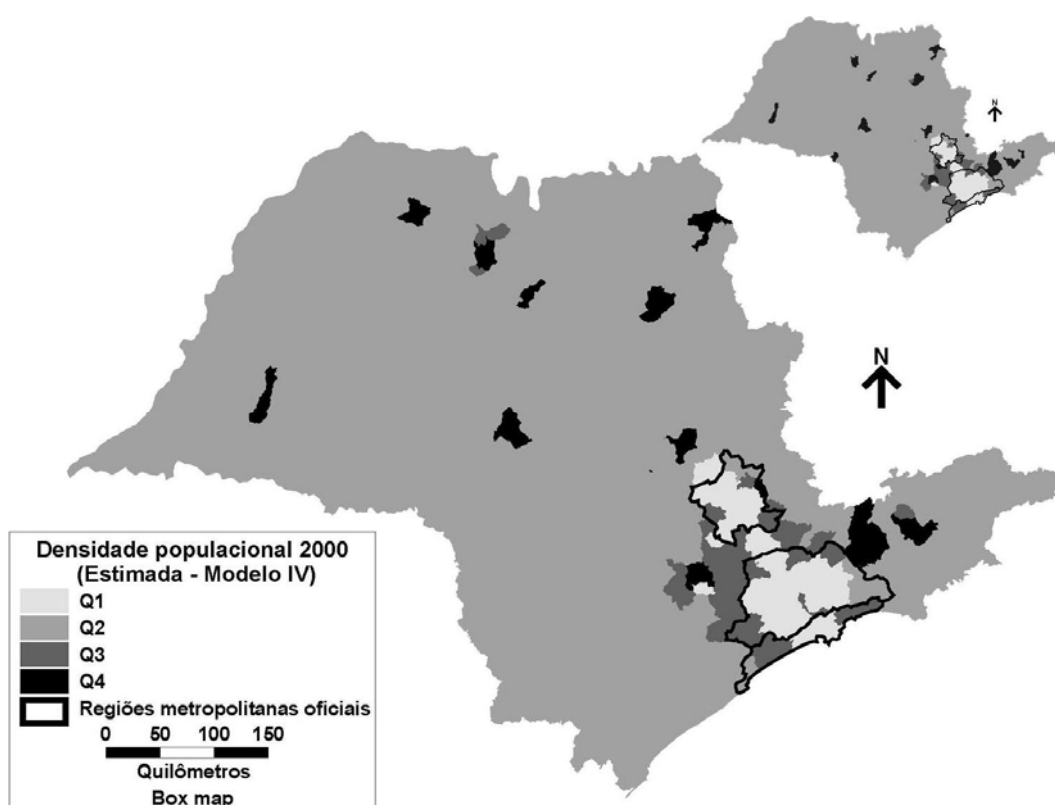


Figura 6.14: *Box map* representando a densidade populacional estimada pelo Modelo IV em 2000

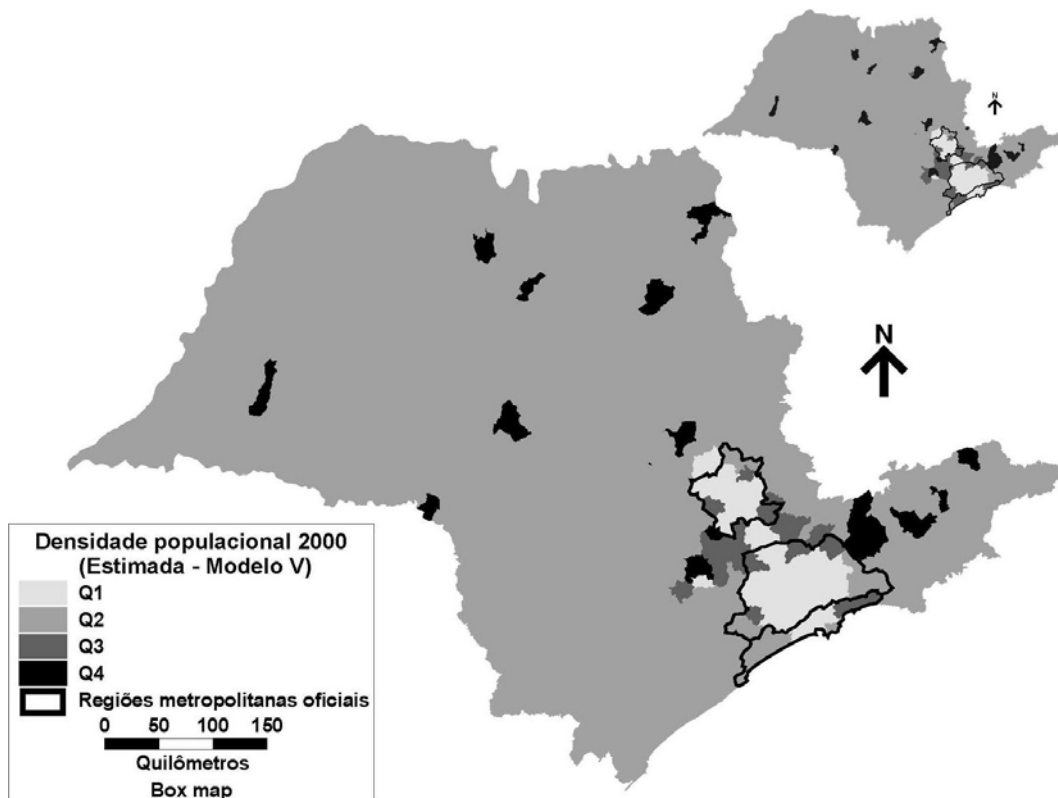


Figura 6.15: *Box map* representando a densidade populacional estimada pelo Modelo V em 2000

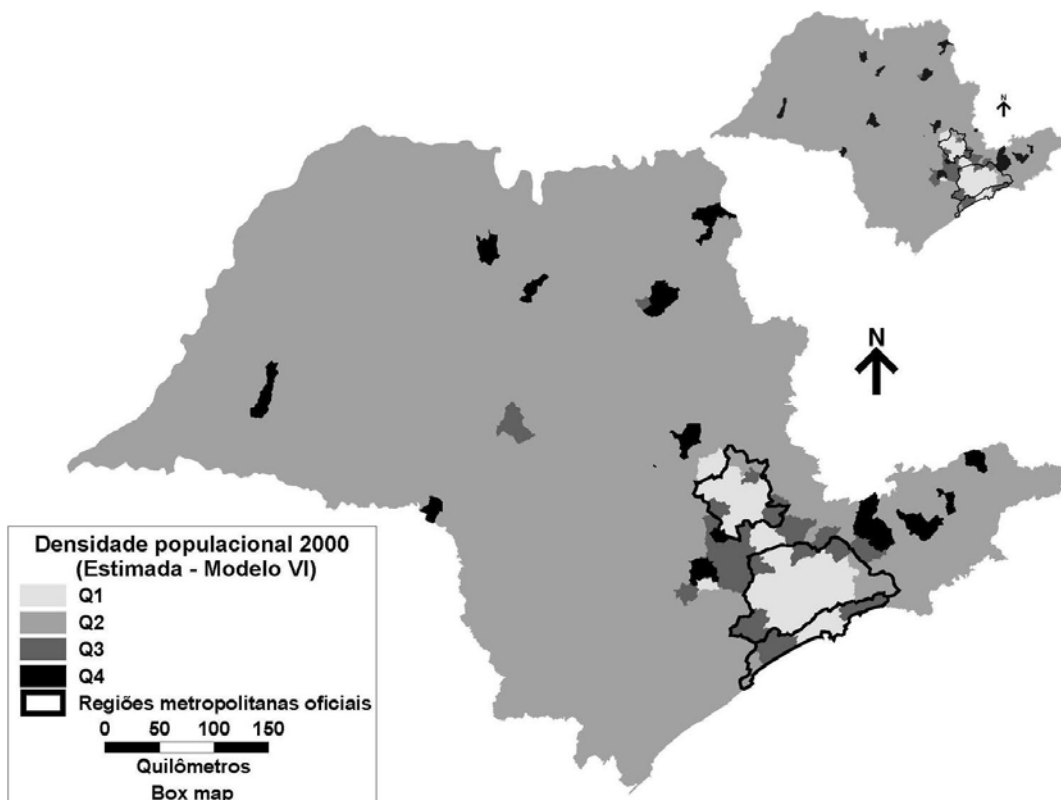


Figura 6.16: *Box map* representando a densidade populacional estimada pelo Modelo VI em 2000

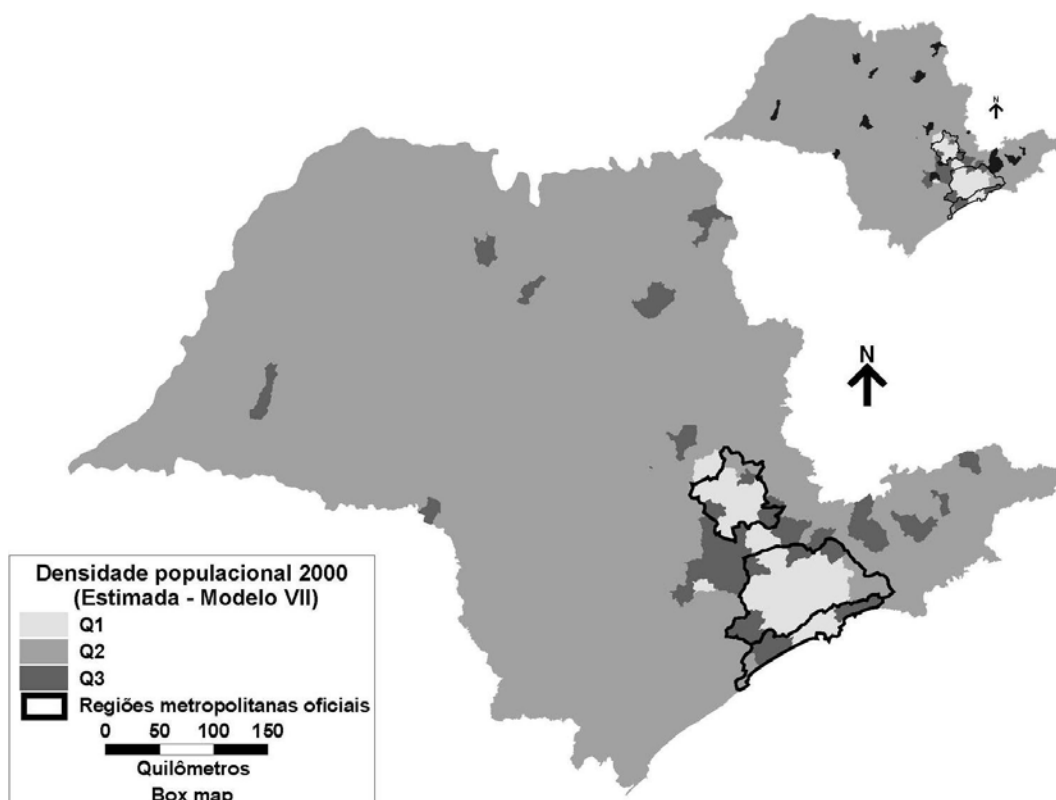


Figura 6.17: Box map representando a densidade populacional estimada pelo Modelo VII em 2000

Tabela 6.1: Proporções de acertos parciais (por quadrante) e totais de cada modelo

MODELOS		Variável	Q1	Q2	Q3	Q4	TOTAL
(I)	3 in 1 out	Densidade Populacional	88,89% (48/54)	95,05% (519/546)	92,59% (25/27)	76,47% (13/17)	93,94%
(II)	T, T+10		92,59% (50/54)	96,70% (528/546)	70,37% (19/27)	94,11% (16/17)	95,19%
(III)	3 in 1 out	Densidade Populacional + Oferta viária	85,18% (46/54)	98,53% (538/546)	55,55% (15/27)	70,58% (12/17)	94,88%
(IV)	T, T+10		94,44% (51/54)	98,90% (540/546)	96,29% (26/27)	70,58% (12/17)	97,67%
(V)	3 in 1 out	Quadrantes	96,29% (52/54)	99,81% (545/546)	70,37% (19/27)	94,11% (16/17)	98,14%
(VI)	T, T+10		Densidade Populacional	98,14% (53/54)	99,81% (545/546)	92,59% (25/27)	82,35% (14/17)
(VII)	T, T+10 (seleção)	Densidade Populacional + Oferta viária	98,14% (53/54)	99,81% (545/546)	92,59% (25/27)	0% (0/17)	96,74%

6.4. Modelos espaciais para a previsão no ano 2010

Em uma tentativa de se visualizar o cenário para um período de projeção no futuro, aplicou-se cada um dos modelos espaciais determinados anteriormente para se estimar os resultados em 2010. As Figuras 6.18, 6.19, 6.20, 6.21, 6.22, 6.23 e 6.24 ilustram mapas temáticos classificados segundo os quadrantes do gráfico de Moran para cada um dos modelos I, II, III, IV, V, VI e VII, respectivamente, com a estimativa da densidade populacional para 2010.

Exceto as Figuras 6.18 e 6.19, que se referem aos modelos I e II, os quais já não apresentaram um desempenho desejável, ao se comparar as Figuras 6.20, 6.21, 6.22, 6.23 e 6.24 entre si pode-se observar que a estimativa da densidade populacional para o período de projeção considerado em cada um dos modelos não sofre sensíveis diferenças. Todavia percebe-se que aqueles modelos que estimam a densidade populacional (Modelos III e IV, Figuras 6.20 e 6.21) apresentam um aspecto comum entre eles: a ampliação dos municípios pertencentes ao quadrante Q3. Além disso, observam-se no Modelo IV (Figura 6.21) municípios que passaram a pertencer ao quadrante Q4, os quais apresentam uma expressiva importância local no contexto em que estão inseridos.

Em síntese, pode-se inferir que é possível desenvolver um modelo espacial para a previsão de período de projeção no futuro conforme um dos objetivos propostos no presente trabalho. A única recomendação quanto ao uso de um ou outro tipo de modelo se refere ao produto desejado, uma vez que os modelos construídos com base somente nas informações referentes aos quadrantes não permitem

conhecer a densidade populacional estimada, ao contrário daqueles que são desenvolvidos a partir da densidade populacional e suas variáveis auxiliares. No próximo capítulo são apresentadas as conclusões do presente trabalho com algumas recomendações e inferências a respeito dos resultados aqui obtidos.

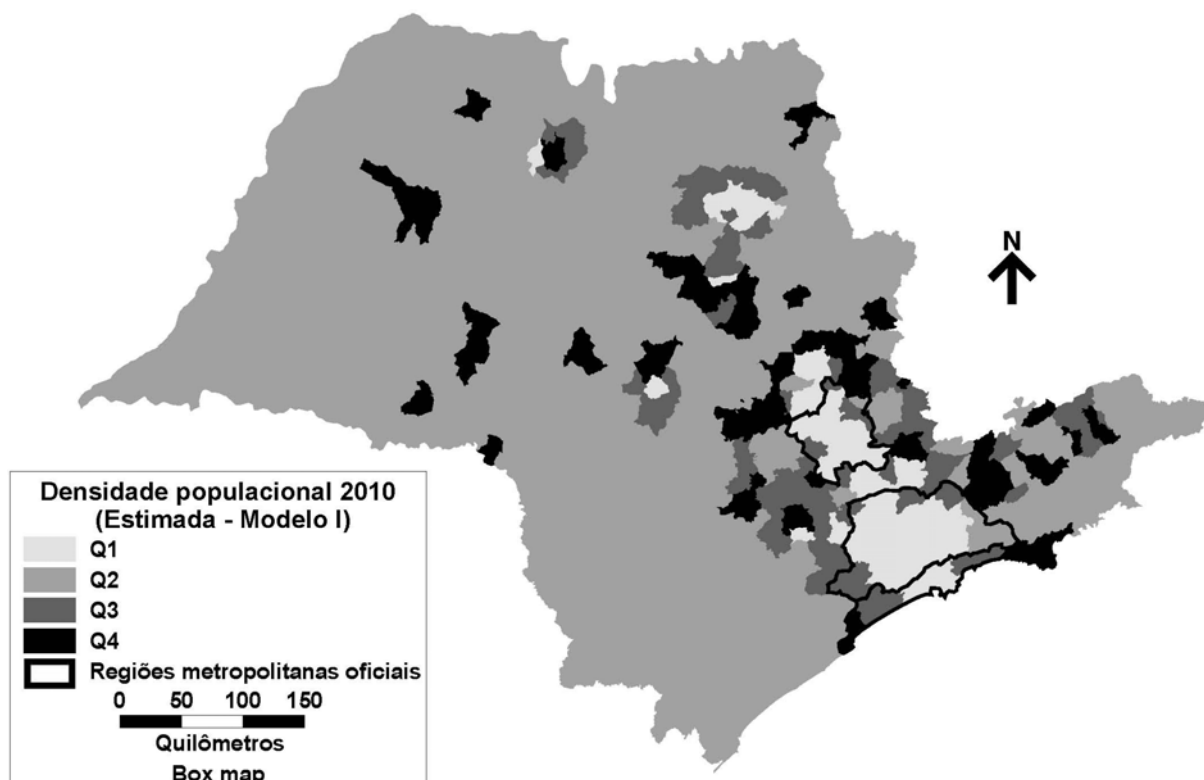


Figura 6.18: *Box map* representando a densidade populacional estimada pelo Modelo I em 2010

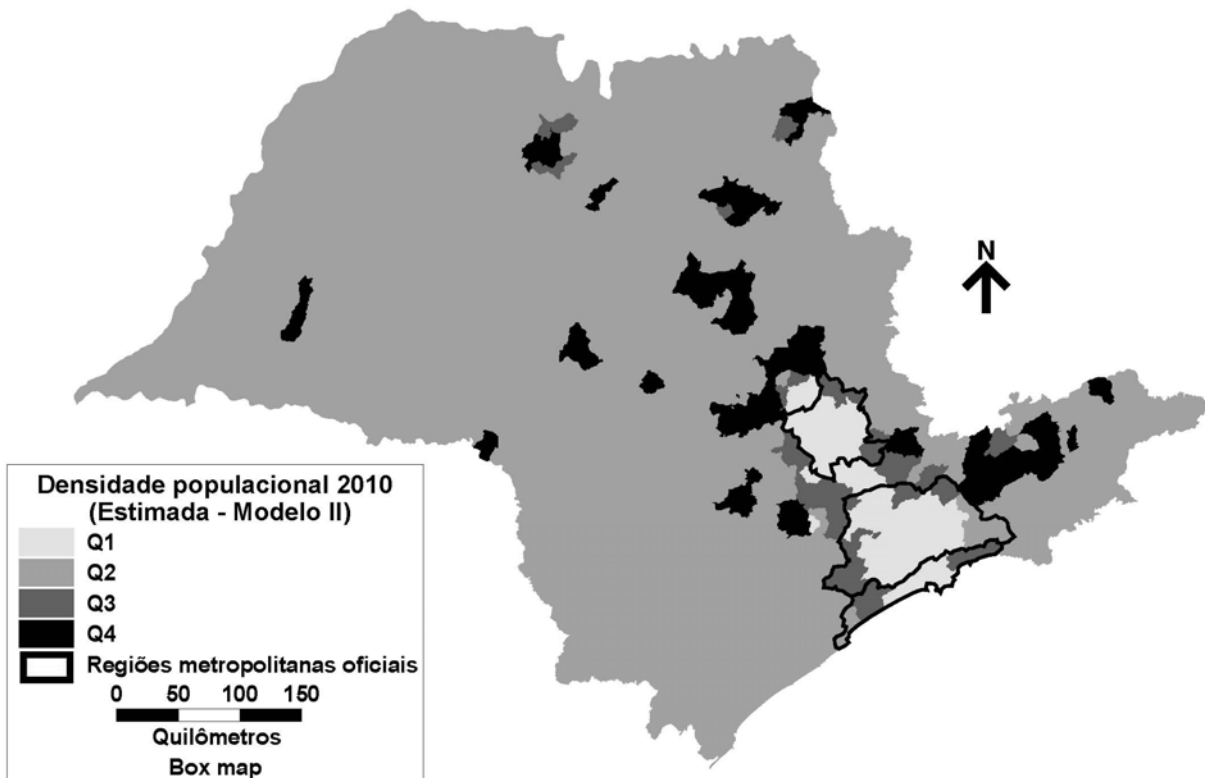


Figura 6.19: *Box map* representando a densidade populacional estimada pelo Modelo II em 2010

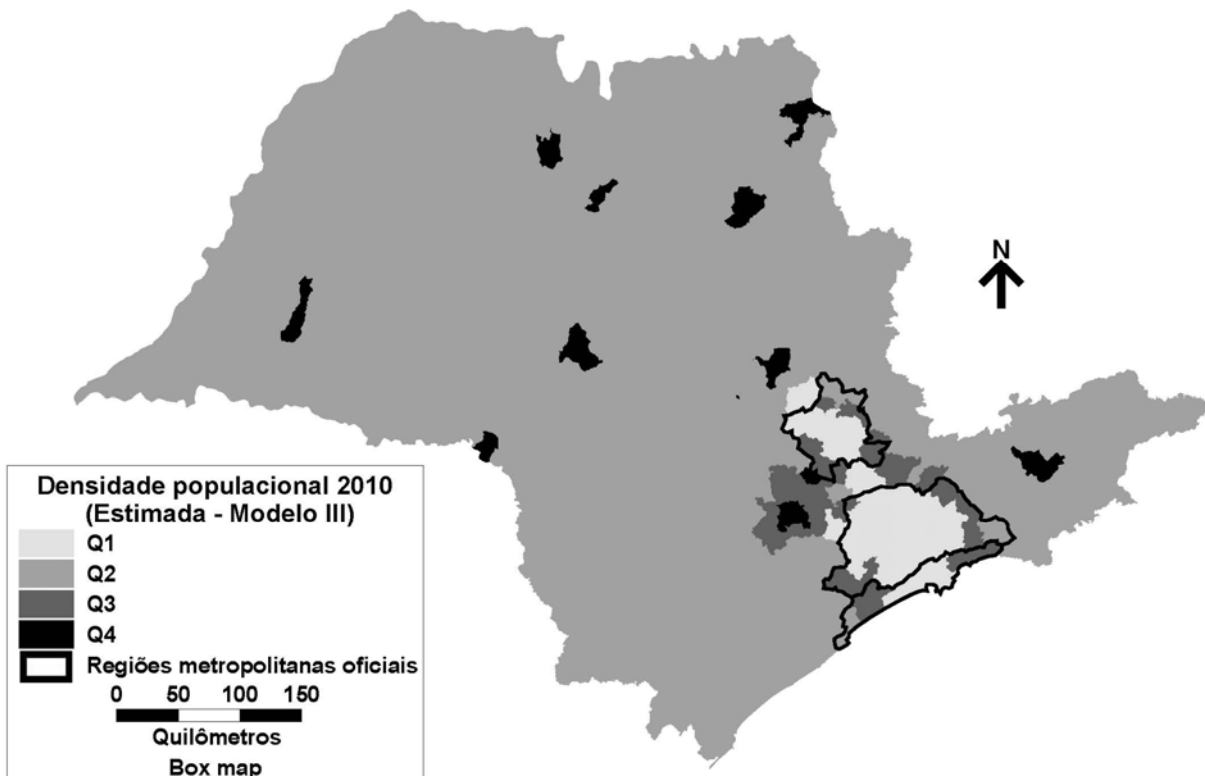


Figura 6.20: *Box map* representando a densidade populacional estimada pelo Modelo III em 2010

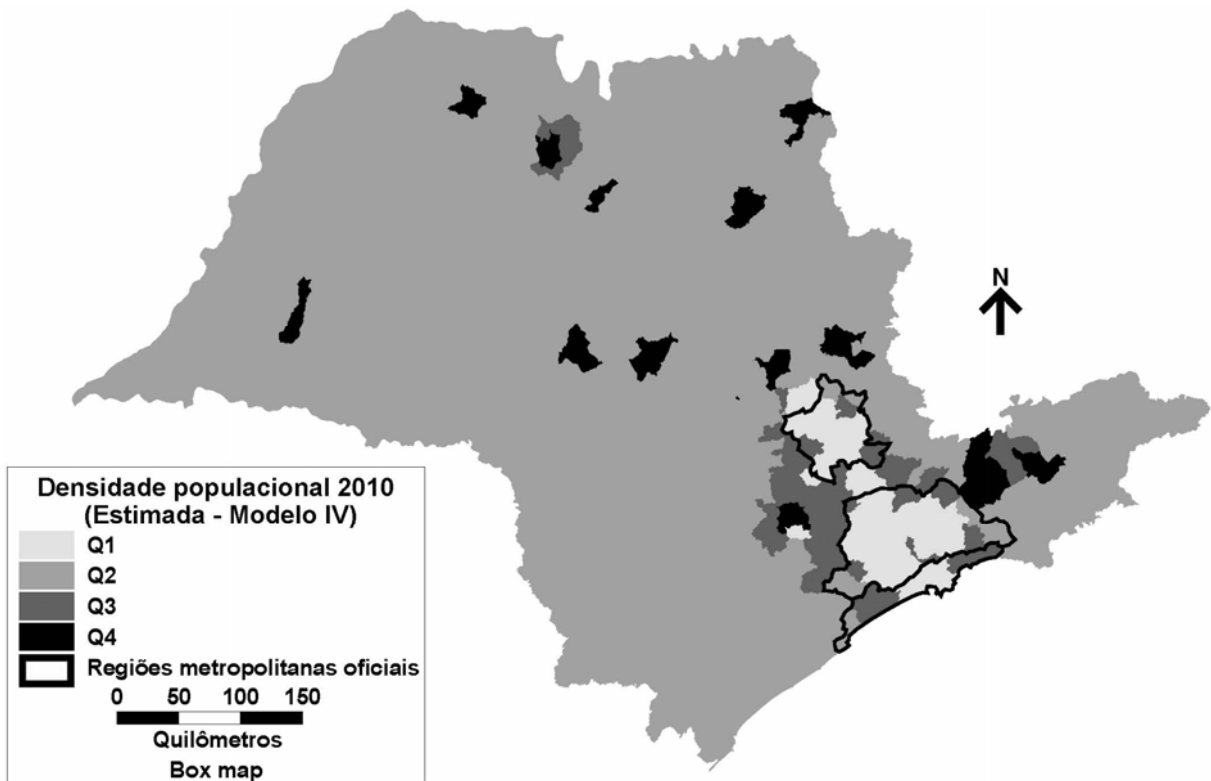


Figura 6.21: *Box map* representando a densidade populacional estimada pelo Modelo IV em 2010

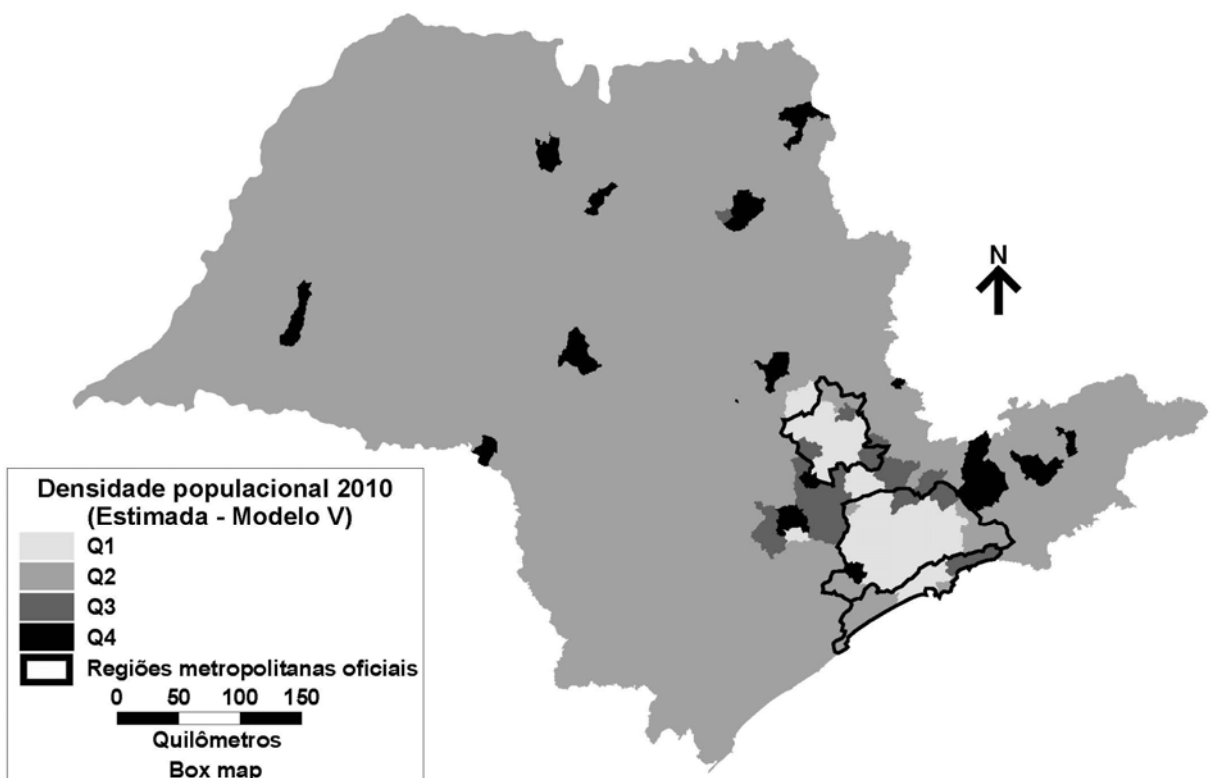


Figura 6.22: *Box map* representando a densidade populacional estimada pelo Modelo V em 2010

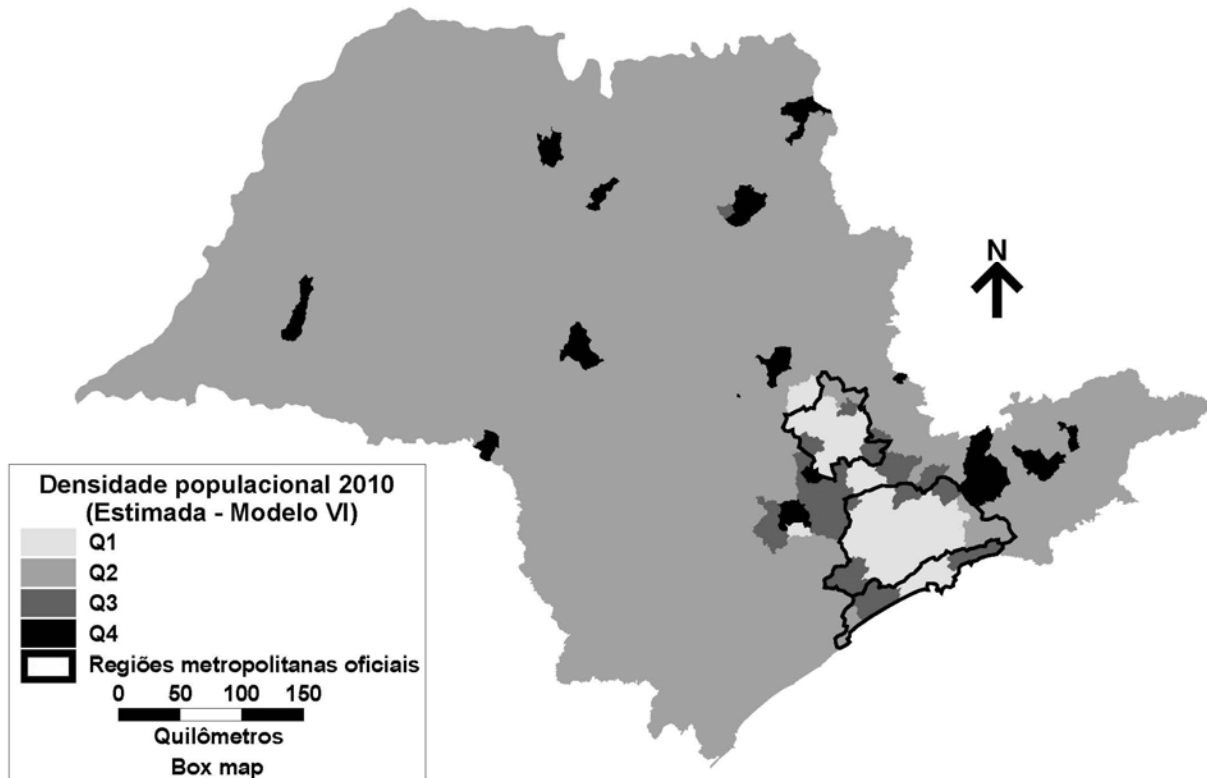


Figura 6.23: *Box map* representando a densidade populacional estimada pelo Modelo VI em 2010

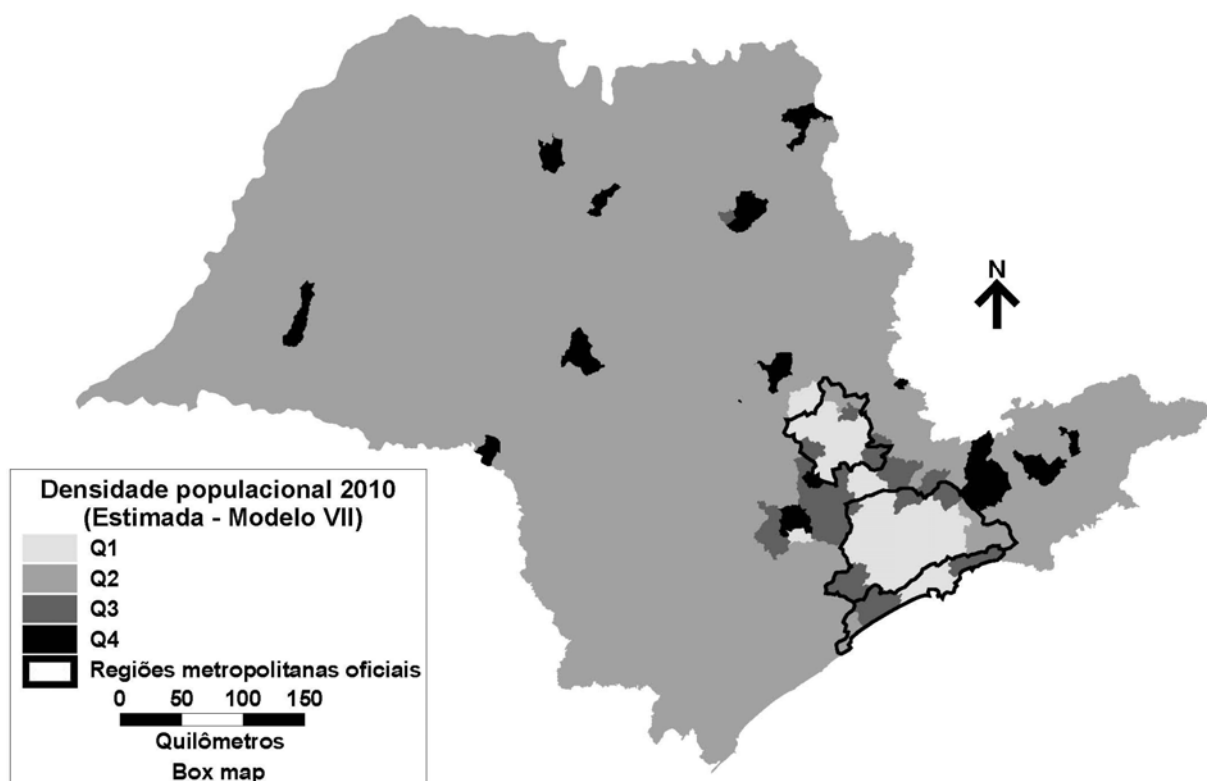


Figura 6.24: *Box map* representando a densidade populacional estimada pelo Modelo VII em 2010

7. Conclusões e recomendações

Um dos objetivos do presente trabalho era a identificação de regiões urbanas homogêneas por meio de técnicas de análise espacial. Conforme observado em estudos sobre a delimitação de espaços metropolitanos (RAMOS; SILVA, 2003a; 2003b; 2007; RAMOS; SILVA; MIRANDA, 2004), a utilização de técnicas de análise exploratória de dados espaciais parece ser bastante promissora para o planejamento urbano e regional, razão pela qual investiu-se na pesquisa aqui apresentada.

Inicialmente, analisando-se as distribuições espaciais dos valores de densidade populacional por meio da sua representação em mapas temáticos classificados segundo os quadrantes do gráfico de Moran (ou *Box Map*), observa-se que é possível identificar razoavelmente bem as regiões urbanas homogêneas existentes, incluindo uma boa parcela dos municípios pertencentes às regiões metropolitanas oficiais, ao longo do período considerado de 1960 a 2000.

Por outro lado, conforme apresentado em Manzato, Soares e Silva (2006), em uma tentativa de se identificar regiões urbanas homogêneas por meio de indicadores de oferta de infra-estrutura viária utilizando-se os conceitos de ESDA, verificou-se

que elas nem sempre coincidem com a distribuição espacial da população e tampouco coincidem de forma exata com as regiões metropolitanas oficiais.

Ao ampliar a investigação no presente estudo, utilizando-se os dados no período de 1960 a 2000, a análise das distribuições espaciais do indicador de oferta de infra-estrutura viária também confirmou que não há uma coincidência com a distribuição espacial da população nem com as regiões metropolitanas oficiais. Além disso, ao contrário do que ocorre com a distribuição populacional, a representação do indicador de oferta de infra-estrutura por meio de mapas temáticos classificados segundo o gráfico de Moran não apresenta um padrão com relação a sua evolução no período considerado. Este fato impossibilita qualquer forma de comparação visual que permita estabelecer um critério de combinação entre as duas variáveis, como sugerido em Manzato e Silva (2006).

No entanto, dada a natureza de forte influência recíproca entre a ocupação do solo e a oferta de infra-estrutura de transportes, é de se esperar que haja alguma forma de avaliação conjunta dessas variáveis. Isso vem ao encontro do segundo objetivo do presente trabalho, o qual é o desenvolvimento de um modelo espacial para a previsão de regiões urbanas homogêneas para um período futuro.

Os primeiros modelos espaciais aqui desenvolvidos (Modelos I e II) levavam em consideração apenas a distribuição da população e, apesar de apresentarem um bom desempenho, ao se analisar os resultados obtidos sobre o território (conforme mostra a Tabela 6.1), verificam-se algumas inconsistências nos modelos gerados. Essas inconsistências se referem a uma estimativa discrepante da densidade populacional para alguns municípios, a qual pode ser observada nos mapas temáticos

classificados segundo os quadrantes do gráfico de Moran, comparando-se as Figuras 6.11 e 6.12 com a Figura 6.5.

Nos modelos subseqüentes, em que se incorporou a oferta de transportes, os resultados obtidos permitiram observar a forte influência dessa variável. Os modelos apresentaram um desempenho ainda superior, tanto no caso em que se estimava a densidade populacional (Modelos III e IV) como nos casos em que os resultados eram obtidos de forma categórica a partir dos quadrantes (Modelos V, VI e VII). Este fato corrobora a hipótese inicial do trabalho, a qual se refere à inter-relação entre uso e ocupação do solo e oferta de transportes.

Quanto ao grupo de modelos que levavam em consideração a oferta de transportes, pôde-se observar ainda que dois daqueles modelos desenvolvidos a partir dos quadrantes (Modelos V e VI) apresentaram uma proporção de acertos totais maior do que aqueles modelos em que se estimava a densidade populacional (Modelos III e IV). A partir desse fato pode-se inferir que talvez não seja necessária a construção de um modelo composto por diferentes variáveis, uma vez que um modelo desenvolvido a partir apenas das informações sobre quadrantes ao qual pertence cada município apresentou um desempenho ligeiramente superior. Entretanto, ao contrário dos Modelos III e IV, os Modelos V e VI apresentam a limitação de não permitirem o conhecimento do valor da densidade populacional estimada. A avaliação do produto desejado em análises deste tipo é fundamental para outras aplicações em que se utilize a mesma metodologia.

É importante observar os resultados obtidos a partir do Modelo VII. Este modelo também foi desenvolvido a partir das informações dos quadrantes dos

municípios, mas somente daqueles que sofreram alguma transição entre 1960 e 2000, nos quadrantes referentes à densidade populacional ao qual pertenciam. Apesar de este modelo apresentar uma proporção de acertos relativamente alta, o modelo não foi capaz de acertar os municípios pertencentes ao quadrante Q4, classificando-os como Q3.

Quanto à extrapolação para o ano de 2010, obtida por meio dos modelos espaciais aqui desenvolvidos, pode-se observar certa semelhança entre eles (exceto nos Modelos I e II). É possível verificar que os Modelos III, IV, V, VI e VII caracterizam bem as regiões homogêneas no quadrante Q1, embora não apontem para uma expansão ou apresentem novas regiões pertencentes a esse quadrante. Entretanto, os modelos sugerem a transição de municípios pertencentes ao Q2 para o Q3, principalmente nas regiões próximas às regiões metropolitanas oficiais, ampliando-se a região existente de municípios no Q3. É de se esperar que essa região venha a se desenvolver ainda mais, observando-se um padrão de transição dos municípios da seguinte forma: Q2-Q3-Q1. Os modelos sugerem também a transição de alguns municípios no Q2 para o Q4, surgindo desta forma, municípios com elevada expressão local no contexto em que estão inseridos.

No caso deste padrão se confirmar, espera-se em médio prazo uma ampliação da região metropolitana de Campinas e de São Paulo, envolvendo alguns municípios pertencentes ao Q1 que atualmente não fazem parte das regiões metropolitanas oficiais. É conveniente observar também que, nesse caso, diversos municípios no Q2 e no Q3 também seriam englobados na ampliação das regiões metropolitanas existentes, além de alguns municípios no Q4 (como por exemplo, Rio Claro, Sorocaba e São José dos Campos). Além disso, devido à proximidade entre as regiões urbanas

apresentadas, poder-se-ia antever uma junção das regiões metropolitanas existentes, formando uma grande RUH. É evidente que a extinção das regiões metropolitanas atuais, dando origem a uma única RUH, é um tema muito complexo por envolver diversos fatores político-administrativos, econômicos e sociais. Entretanto, em tese, é o que se pode observar a partir dos resultados apresentados.

Em termos gerais, a natureza das análises aqui desenvolvidas pode ser útil ao planejamento urbano e regional dada a forte inter-relação entre uso e ocupação do solo e oferta de transportes. Uma vez que o desenvolvimento urbano e regional depende necessariamente da infra-estrutura, é possível induzir o crescimento de áreas ainda pouco desenvolvidas por meio da oferta de infra-estrutura de transportes. Apesar de se ter utilizado apenas as informações sobre a infra-estrutura rodoviária no estudo aqui desenvolvido, é importante ressaltar que a incorporação de dados da oferta de outras infra-estruturas (como aeroportos, ferrovias, hidrovias etc.) pode apresentar resultados robustos e com maior fidelidade no contexto estudado.

Além disso, a continuidade dessa linha de pesquisa poderia ter como novas investigações a avaliação de outros indicadores de oferta de infra-estrutura de transportes, como por exemplo, indicadores de acessibilidade. Uma outra avaliação possível se refere ao cálculo do indicador populacional, que para o presente estudo considerou-se a área total do município. Alternativamente poder-se-ia considerar algumas variações para o cálculo da densidade populacional, como por exemplo, a utilização da área urbana ao invés da área total.

Finalmente, poderia ser realizada uma ampliação do presente estudo baseado na metodologia apresentada para outros estados ou regiões. Por meio de uma

avaliação dessa natureza seria possível obter informações em termos nacionais que certamente contribuiriam para o planejamento urbano e regional do Brasil.

Referências bibliográficas

ALMEIDA, C. M.; MONTEIRO, A. M. V.; CÂMARA, G.; SOARES-FILHO, B. S.;
CERQUEIRA, G. C.; PENNACHIN, C. L.; BATTY, M. Modelos espaço-temporais em
planejamento urbano: desafios para a nova geração de SIG. In: *Anais eletrônicos do GIS
Brasil 2002*, Curitiba, Brasil, 2002a.

ALMEIDA, C. M.; MONTEIRO; A. M. V.; CÂMARA, G.; SOARES-FILHO, B. S.;
CERQUEIRA, G. C.; PENNACHIN, C. L. E BATTY, M. *Empiricism and Stochastics in
Cellular Automaton Modeling of Urban Land Use Dynamics*. Paper 42, Centre for
Advanced Spatial Analysis (CASA), University College Londres, Inglaterra, 2002b.
(Disponível em: www.casa.ucl.ac.uk/how_ca_work.pdf)

ALMEIDA, C. M.; MONTEIRO, A. M. V.; CÂMARA, G.; SOARES-FILHO, B. S.;
CERQUEIRA, G. C.; PENNACHIN, C. L. E BATTY, M. Modelos celulares de dinâmicas
espaço-temporais urbanas. In: CÂMARA, G.; MONTEIRO, A. M. V. (Eds.) *Introdução à
Modelagem Dinâmica Espacial*. Simpósio Brasileiro de Sensoriamento Remoto, capítulo 3,
Belo Horizonte, 2003.

ANSELIN, L. *Spatial Data Analysis with GIS: An Introduction to Application in the Social Sciences*. Technical Report 92-10, National Center for Geographic Information and Analysis, University of California, California, EUA, 1992. (Disponível em: www.ncgia.ucsb.edu/Publications/Tech_Reports/92/92-10.PDF. Acesso em: 11 de set. de 2004)

ANSELIN, L. Local indicators of spatial association - LISA. *Geographical Analysis*, v. 27, p. 91–114, 1995.

ANSELIN, L. The Moran scatter plot as an ESDA tool to assess local instability in spatial association. In: FISCHER, M.; SCHOLTEN, H.; UNWIN, D. (Eds.). *Spatial Analytical Perspectives on GIS*. Londres: Taylor & Francis, 1996.

ANSELIN, L. Exploratory spatial data analysis in a geocomputational environment. In: LONGLEY, P.; BROOKS, S.; MACMILLAN B. E MCDONNELL R. (Ed.). *GeoComputation, a Primer*. New York: Wiley, 1998a.

ANSELIN, L. Interactive techniques and exploratory spatial data analysis. In: LONGLEY, P.; GOODCHILD, M.; MAGUIRE, D. E RHIND, D. (Ed.). *Geographical Information Systems: Principles, Techniques, Management and Applications*. Cambridge: Geoinformation International, 1998b.

ANSELIN, L.; BAO, S. Exploratory spatial data analysis linking SpaceStat and ArcView. In: FISHER M.; GETIS, A. (Eds.) *Recent Developments in Spatial Analysis*. Berlin: Springer-Verlag, 1997.

ANSELIN, L.; SMIRNOV, O. *The SpaceStat Extension for ArcView 3.0*. Regional Research Institute, West Virginia University, Morgantown, 1998.

ANTONI, J. P. Urban sprawl modelling: combining models to make decisions. In: TIMMERMANS, H. (Ed.) *Design and Decision Support Systems in Urban Planning*. Architecture and Urban Planning, Ellecom, Holanda, 2002a.

ANTONI, J. P. Urban sprawl modelling: a methodological approach. *Cybergeo 2002*, 2002b. (Disponível em: www.cybergeo.presse.fr/ectqg12/antoni/antoni.htm)

ARAÚJO, M. P. *Infra-estrutura de Transporte e Desenvolvimento Regional: uma abordagem de equilíbrio geral inter-regional*. Tese (Doutorado) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2006.

BAILEY, T. C. A review of statistical spatial analysis in geographical information systems. In: FOTHERINGHAM, S.; ROGERSON, P. (Eds.) *Spatial Analysis and GIS*, Londres: Taylor & Francis Ltd, 1994.

BAILEY, T.; GATTREL, A. C. *Interactive Spatial Data Analysis*. Londres: Longman, 1995.

BARREDO, J. I.; KASANKO, M.; MCCORMICK, N.; LAVALLE, C. Modelling dynamic spatial processes: simulation of urban future scenarios through Cellular Automata. *Landscape and Urban Planning*, v. 64, n. 3, p. 145–160, 2003.

BATTY, M.; XIE, Y.; SUN, Z. Modeling urban dynamics through GIS-based Cellular Automata. *Computers, Environments and Urban Systems*, v. 23, p. 205-233, 1999.

CÂMARA, G.; MONTEIRO, A. M. V. *Introdução ao Geoprocessamento: Aula 1 - Conceitos Básicos*. 2001. (Disponível em: www.dpi.inpe.br/cursos/ser300/aulas/aula1-definicoes. Acesso em: 11 de set. de 2004).

CÂMARA, G.; CARVALHO, M. S.; CRUZ, O. G.; CORREA, V. Análise de dados de áreas. In: FUKS, S. D.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. (Eds.). *Análise Espacial de Dados Geográficos*. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil, 2002a. (Disponível em: www.dpi.inpe.br/gilberto/livro/analise. Acesso em: 11 de set. de 2004).

CÂMARA, G.; MONTEIRO, A. M.; CARVALHO, M. S. Análise espacial e geoprocessamento. In: FUKS, S. D.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. (Eds.) *Análise Espacial de Dados Geográficos*. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil, 2002b. (Disponível em: www.dpi.inpe.br/gilberto/livro/analise. Acesso em: 11 de set. de 2004).

CARVALHO, M. S. *Aplicação de Métodos de Análise Espacial na Caracterização de Áreas de Risco à Saúde*. Tese (Doutorado) – COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1997.

CECCHINI, A.; RIZZI, P. The reasons why Cellular Automata are a useful tool in the working-kit for the millennium urban planner in governing the territory. In: *Anais eletrônicos do VII International Conference on Computers in Urban Planning and Urban Management, CUPUM*, Honolulu, EUA, 2001.

CRESSIE, N. *Statistics for Spatial Data*. New York: Wiley, 1991.

DEWDNEY, A. K. The turing omnibus. *Computer Science Press*, Rockville, EUA, 1989.

DIAPPI, L.; BOLCHI, P; BUSCEMA, M. Improved understanding of urban sprawl using Neural Networks. In: VAN LEEUWEN, J. P.; TIMMERMANS, H. J. P. (Eds.) *Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning*, p. 33-49, Sint-Michielsgestel, Holanda, 2004.

EMPLASA Empresa Paulista de Planejamento Metropolitano, 2002. (Disponível em: www.emplasa.sp.gov.br, Acesso em 7 de abril de 2006).

FERRÃO, J.; RODRIGUES, D.; VALA, F. *As regiões metropolitanas portuguesas no contexto ibérico*. Lisboa: DGOTDU, 2002.

FERRÃO, J.; VALA, F. Delimitação das aglomerações metropolitanas de Lisboa e Porto com base no critério de continuidade de espaço construído. *Revista de Estudos Regionais - Região Lisboa e Vale do Tejo*, n. 2, p. 7-35, 2001.

FERREIRA, M. J.; ROSADO, A. R. As grandes áreas urbanas portuguesas. Conceitos e delimitação espacial. In: PEREIRA, M. (Ed.). *Grandes Áreas Urbanas. Reorganização institucional e territorial. O caso da Área Metropolitana de Lisboa*, p. 38-50, Universidade Nova de Lisboa: Centro de Estudos de Geografia e Planeamento, 1999.

GEIPOT – Empresa Brasileira de Planejamento de Transportes, *Anuário Estatístico dos Transportes*, Ministério dos Transportes, Brasília, 2000. (Disponível em: www.geipot.gov.br/IndexG.htm, Acesso em 7 de abril de 2006).

GOUVÊA, R. G. *A Questão Metropolitana no Brasil*. Primeira ed. Rio de Janeiro: Editora FGV, 2005.

GRIGOLON, A. B.; MANZATO, G. G.; SILVA, A. N. R. Spatial growth patterns and transportation impacts of major trip generators – the case of Campinas, Brazil. *Artigo submetido para a World Conference on Transport Research Society*, a ser realizada em Berkeley, Califórnia, EUA, 2007.

HAQUE, S. M.; OKUMURA, M. Simulation of land use changes for strategic urban management with a GIS based statistical model. In: *Anais eletrônicos do VIII International Conference on Computers in Urban Planning and Urban Management, CUPUM*, Sendai, Japão, 2003.

KABOS, S.; CSILLAG, F. The analysis of spatial association on a regular lattice by join-count statistics without the assumption of first-order homogeneity. *Computers & Geosciences*, n. 2, p. 7-35, 2002.

KAMPEL, S. A.; CÂMARA, G.; QUINTANILHA, J. A. Análise exploratória das relações espaciais do desflorestamento na Amazônia legal brasileira. In: *Anais do GIS Brasil 2000*, 2000. (Disponível em: www.dpi.inpe.br/gilberto/papers/silvana_gisbrasil2000.pdf. Acesso em: 11 de set. de 2004).

KREMPI, A. P. *Explorando Recursos de Estatística Espacial para Análise da Acessibilidade da Cidade de Bauru*. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2004.

LACOUR, C.; PUISSANT, S. *La métropolisation: croissance, diversité, fractures*. Paris: Anthropos, 1999.

LEMOS, D. S. C. P. S; SANTOS, M. P. S.; PORTUGAL, L. S. Análise da relação entre o sistema de transporte e a exclusão social na cidade do Rio de Janeiro. *Revista Engevista*, v. 6, n. 3, p. 36-53, 2004.

LEMOS, D. S. C. P. S; SANTOS, M. P. S.; PORTUGAL, L. S. Análise das relações existentes entre o uso do solo e o sistema de transportes, na cidade do Rio de Janeiro, Brasil. In: *Anais eletrônicos do 1º Congresso Luso-Brasileiro para o Planejamento Urbano Regional Integrado Sustentável, PLURIS 2005*, São Carlos, Brasil, 2005

LEVINE, N. Spatial statistics and GIS: software tools to quantify spatial patterns. *Journal of the American Planning Association*, v. 62, n. 3, p. 381–392, 1996.

LI, L.; SATO, Y.; ZHU, H. Simulating spatial urban expansion based on a physical process. *Landscape and Urban Planning*, v. 64, p. 67–76, 2003.

LIU, Y.; PHINN, S. R. Modelling urban development with Cellular Automata incorporating fuzzy-set approaches. *Computers, Environment and Urban Systems*, v. 27, n. 6, p. 637–658, 2003.

LOIBL, W.; TOETZER, T. Modeling growth and densification processes in suburban regions - simulation of landscape transition with spatial agents. *Environmental Modelling & Software*, v. 18, n. 6, p. 553–563, 2003.

LONGLEY, P. B. M. *Spatial Analysis: Modelling in a GIS Environment*, New York: John Wiley & Sons Inc., 1996.

LOPES, S. B. *Efeitos da Dependência Espacial em Modelos de Previsão de Demanda por Transporte*. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2005.

MAGALHÃES, M. T. Q.; TEIXEIRA, G. L.; YAMASHITA, Y. Indicadores de cobertura espacial para diagnóstico da dotação de estrutura de circulação rodoviária brasileira. In: *Anais eletrônicos do XVIII Congresso de Pesquisa e Ensino em Transportes, ANPET*, Florianópolis, Brasil, 2004.

MANZATO, G. G.; SILVA, A. N. R. Incorporando um indicador de oferta de infra-estrutura de transportes na definição de regiões metropolitanas. In: *Anais eletrônicos do XX Congresso de Pesquisa e Ensino em Transportes, ANPET*, Brasília, Brasil, 2006.

MANZATO, G. G.; SOARES, A. J.; SILVA A. N. R. Aglomerações urbanas e oferta de transportes no estado de São Paulo, Brasil. In: *Anais eletrônicos do 2º. Congresso Luso-Brasileiro para o Planejamento Urbano Regional Integrado Sustentável, PLURIS 2006*, Braga, Portugal, 2006.

METROPOLITAN AREA STANDARDS REVIEW COMMITTEE. Final report and recommendations from the Metropolitan Area Standards Review Committee to the Office of Management and Budget concerning changes to the standards for defining metropolitan areas. *Federal Register*, v. 65, n. 163, 2000.

MONTEIRO, J. L. Caracterização dos espaços urbanos na RLVT: o contributo da análise discriminante. *Revista de Estudos Regionais - Região Lisboa e Vale do Tejo*, n. 1, p. 21–46, 2000.

NUREC - NETWORK ON URBAN RESEARCH IN THE EUROPEAN UNION. *Atlas of agglomerations in the European Union*. Duisburg, 1994.

NEVES, M. C.; RAMOS, F. R.; CAMARGO, E. C. G.; CÂMARA, G.; MONTEIRO, A. M. Análise exploratória espacial de dados sócio-econômicos de São Paulo. In: *Anais do GIS Brasil 2000*, 2000. Disponível em: www.dpi.inpe.br/gilberto/papers/marcos_gisbrasil2000.pdf. Acesso em: 11 de set. de 2004).

OFFICE OF MANAGEMENT AND BUDGET. Alternative approaches to defining metropolitan and non-metropolitan areas. *Federal Register*, v. 63, n. 244, 1998.

OFFICE OF MANAGEMENT AND BUDGET. Recommendations from the Metropolitan Area Standards Review Committee to the Office of Management and Budget concerning changes to the standards for defining metropolitan areas. *Federal Register*, v. 64, n. 202, 1999.

OFFICE OF MANAGEMENT AND BUDGET. Standards for defining metropolitan and micropolitan statistical areas. *Federal Register*, v. 65, n. 249, 2000.

OHGAI, A.; IGARASHI, M.; MA, S. R.; IKARUGA, S. A Cellular Automata model to investigate the influence of land use policies on future distributions of population and employees. In: *Anais eletrônicos do VII International Conference on Computers in Urban Planning and Urban Management, CUPUM*, Honolulu, EUA, 2001.

OLIVEIRA, C.; RODRIGUES, D. Mobilidade e território da região de Lisboa e Vale do Tejo: pistas para uma análise integrada. *Revista de Estudos Regionais - Região Lisboa e Vale do Tejo*, n. 3, p. 43–65, 2001.

ORFORD, S. Identifying and comparing changes in the spatial concentrations of urban poverty and affluence: a case study of inner London. *Computers, Environment and Urban Systems*, v. 28, n. 6, p. 701–717, 2004.

O'SULLIVAN, D.; TORRENS, P. M. Cellular models of urban systems. In: BANDINI, S.; WORSCH, T. (Ed.). *Theoretical and Practical Issues on Cellular Automata*. London: Springer-Verlag, 2000.

PEDROSA, B. M.; CÂMARA, G. Modelagem dinâmica e geoprocessamento. In: FUKS, S. D.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. (Eds.) *Análise Espacial de Dados Geográficos*, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil, 2002. (Disponível em: www.dpi.inpe.br/gilberto/livro/analise/. Acesso em: 11 de set. de 2004).

PING, J. L.; GREEN, C. J.; ZARTMAN, R. E.; BRONSON, K. F. Exploring spatial dependence of cotton yield using global and local autocorrelation statistics. *Field Crops Research*, v. 89, p. 219–236, 2004.

PREMO, L. S. Local spatial autocorrelation statistics quantify multi-scale patterns in distributional data: an example from the Maya Lowlands. *Journal of Archaeological Science*, v. 31, n. 7, p. 855–866, 2004.

QUEIROZ, M. P. *Análise Espacial dos Acidentes de Trânsito do Município de Fortaleza*. Dissertação (Mestrado) – Universidade Federal do Ceará, Fortaleza, 2003.

RAMOS, F. R. *Análise Espacial de Estruturas Intra-urbanas: o caso de São Paulo*. Dissertação (Mestrado) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2002.

RAMOS, R. A. R.; SILVA, A. N. R. A data-driven approach for the definition of metropolitan regions. In: *Anais eletrônicos do VIII International Conference on Computers in Urban Planning and Urban Management, CUPUM*, Sendai, Japão, 2003a.

RAMOS, R. A. R.; SILVA, A. N. R. Um contributo para a delimitação da área metropolitana do noroeste de Portugal. *Revista Portuguesa de Estudos Regionais*, p. 61–82, 2003b.

RAMOS, R. A. R.; SILVA, A. N. R. A spatial analysis approach for the definition of metropolitan regions – the case of Portugal. *Environment and Planning B*, v. 34, n. 1, p. 171–185, 2007.

RAMOS, R. A. R.; SILVA, A. N. R.; MIRANDA, V. P. A comparison of two methods for the definition of regional metropolitan areas through an application in the north of Portugal. In: *Anais eletrônicos do 44th European Congress of the European Regional Science Association*, Porto, Portugal, 2004

RODRIGUES, D.; VALA, F.; MONTEIRO, J. Áreas de influência das cidades de Lisboa e Vale do Tejo. *Revista de Estudos Regionais - Região Lisboa e Vale do Tejo*, n. 5, p. 43–64, 2002.

ROY, G. G.; SNICKARS, F. Citylife: a study of Cellular Automata in urban dynamics. In: FISHER, M.; SCHOLTEN, H. J.; UNWIN, D. *Spatial Analytical Perspectives on GIS*, p. 213–218, Londres: Taylor e Francis, 1996.

SANTOS, V. S. *Modelagem da Dinâmica Populacional Intra-Urbana com Cellular Automata e Avaliação Multicritério para Antecipar Necessidades de Transportes*. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2005.

SANTOS, V. S.; LIMA, R. S.; SILVA, A. N. R. Modelagem da dinâmica populacional intra-urbana com Cellular Automata e avaliação multicritério. In: SILVA, A. N. R.; SOUZA, L. C. L.; MENDES, J. F. G. (Eds.) *Planejamento Urbano, Regional, Integrado e Sustentável. Desenvolvimentos recentes no Brasil e em Portugal*, São Carlos: Edição dos autores, 2005.

SASAKI, K.; TAMURA, T.; SAITOH, K. Application of neural Cellular Automata to land use-population structure – In case of Metro Manila. In: *IX World Conference on Transport Research, WCTR*, Seul, Coréia, 2001.

SILVA, E. A. Cenários da expansão urbana na área metropolitana de Lisboa. *Revista de Estudos Regionais - Região Lisboa e Vale do Tejo*, n. 5, p. 23–41, 2002.

SILVA, M. A. S. *Mapas Auto-Organizáveis na Análise Exploratória de Dados Geoespaciais Multivariados*. Dissertação (Mestrado) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos 2004.

SILVA, A. N. R.; RAMOS, R. A. R.; SOUZA, L. C. L.; RODRIGUES, D. S.; MENDES, J. F. G. *SIG – Uma Plataforma para Introdução de Técnicas Emergentes no Planejamento Urbano, Regional e de Transportes: Uma Ferramenta 3D para Análise Ambiental Urbana, Avaliação Multicritério, Redes Neurais Artificiais*. São Carlos: Edição dos autores, 2004.

SILVA, A. A.; VALA, F. Acessibilidades e construção na área metropolitana de Lisboa, 1991-2001. *Revista de Estudos Regionais - Região Lisboa e Vale do Tejo*, n. 3, p. 25–40, 2001.

TAYLOR, C. E. “Fleshing Out” Artificial Life II. In: *Artificial Life*, Santa Fe, New Mexico, p. 25-38, 1992.

TEIXEIRA, G. L. *Uso de Dados Censitários para Identificação de Zonas Homogêneas para Planejamento de Transportes Utilizando Estatística Espacial*. Dissertação (Mestrado) – Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2003.

TORRENS, P. M. *How cellular models of urban systems work - WP-28*. University College London, England: Centre for Advanced Spatial Analysis (CASA), 2000.

TORRENS, P. M.; O’SULLIVAN, D. Cellular Automata and urban simulation: where do we go from here? *Environment and Planning B: Planning and Design*, v. 28, p. 163–168, 2001.

UNWIN, A.; UNWIN, D. Exploratory spatial data analysis with local statistics. *Journal of the Royal Statistical Society, series D, The Statistician*, v. 47, p. 415–422, 1998.

WARD, D. P.; MURRAY, A. T.; PHINN, S. R. A stochastically constrained cellular model of urban growth. *Environment and Urban Systems*, v. 24, p. 539–558, 2000.

WEBSTER, C.; WU, F. Coase, spatial pricing and self-organising cities. *Urban Studies*, v. 38, n. 11, p. 2037-2054, 2001

WU, F.; WEBSTER, C. J. Simulation of land development through the integration of Cellular Automata and multicriteria evaluation. *Environment and Planning B*, v. 25, p. 103-126, 1998.

YEH, A. G.; LI, X. A constrained CA model for the simulation and planning of sustainable urban forms by using GIS. *Environment and Planning B*, v. 28, p. 733-753, 2001a.

YEH, A. G. E LI, X. Neural network CA simulation urban forms. In: *Anais eletrônicos do VII International Conference on Computers in Urban Planning and Urban Management, CUPUM*, Havai, EUA, 2001b.

YEH, A. G.; LI, X. A Cellular Automata model to simulate development density for urban planning. *Environment and Planning B*, v. 29, p. 431-450, 2002.

Anexo: ESDA por meio do *software* TransCAD

Esse anexo apresenta uma forma de se utilizar a técnica de ESDA por meio do *software* TransCAD. Em síntese, essa técnica envolve a determinação dos elementos W , Z e W_Z , com a sua representação por meio do gráfico e do mapa de espalhamento de Moran. A seguir é detalhado como podem ser obtidos esses elementos para uma base de áreas.

A.1. Matriz de adjacência (W)

A matriz de adjacência é obtida por meio da ferramenta *Statistics > Adjacency Matrix*. Ao clicar nesta ferramenta, dê um rótulo para a matriz e, em seguida, salve-a no computador dando um nome para o arquivo.

Geralmente o TransCAD fornece, além da matriz de adjacência, duas matrizes adicionais: uma com o nível da adjacência e outra com o comprimento da fronteira comum entre as áreas. Entretanto, o que interessa no processo apresentado é somente a matriz de adjacência.

Ao acessar esta matriz pode-se observar que ela é preenchida com o valor “1,0” quando há adjacência entre uma área e outra. Quando não há adjacência, a célula fica vazia.

Para a utilização dessa matriz na técnica de ESDA, ela deve ser normalizada, ou seja, a soma dos valores de cada linha deve ser igual a 1,0. Como o TransCAD não faz isso automaticamente, deve-se proceder da seguinte maneira:

1. Ativar os marginais da matriz referente à soma, em *Matrix > Settings*, conforme mostra a Figura A.1:



Figura A. 1: Ativação dos marginais da matriz

2. Exportar os valores da coluna marginal para uma nova tabela, em “*Matrix > Export Rows or Columns*” (será salvo um arquivo no formato DBF). Nenhuma célula deverá estar selecionada neste procedimento. É importante observar que ao desejar extrair informações referentes à coluna marginal, deve-se selecionar “*Row Marginals*” e vice-versa. A Figura A.2 ilustra este procedimento:

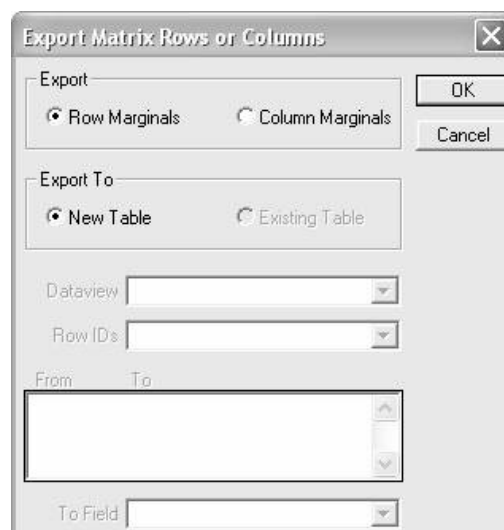


Figura A. 2: Exportação da coluna marginal para uma tabela

3. Adicionar uma nova matriz em “*Matrix > Contents*” denominando-a de “SomaMarg”, por exemplo. Acessá-la e, selecionando qualquer coluna desta matriz, importar os valores da coluna marginal extraída no passo anterior. Para isso, em “*Matrix > Import Rows or Columns*”, selecionar a tabela em que estão os dados, clicar em “*All Columns*” (os valores deverão ser importados para todas as colunas) e selecionar o campo “*SUM*”. A Figura A.3 ilustra esse procedimento e a Figura A.4 mostra como deve ficar a matriz “SomaMarg”:

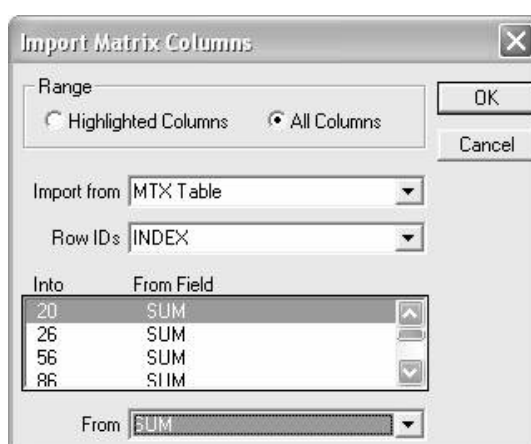


Figura A. 3: Importação da coluna marginal para a matriz

	20	26	56	86	92	104	110	134	176	Sum
20	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	1062.00
26	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	1062.00
56	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	885.00
86	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	1062.00
92	8.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00	1416.00
104	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	1062.00
110	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	1239.00
134	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	531.00
176	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00	1062.00
224	8.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00	1416.00
254	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	9.00	1593.00
Sum	948.00	948.00	948.00	948.00	948.00	948.00	948.00	948.00	948.00	167796.00

Figura A. 4: Matriz com os valores importados da coluna marginal

4. Novamente, adicionar outra matriz denominando-a de, por exemplo, “Normaliz” para se obter a matriz normalizada. Após adicioná-la, selecionar qualquer célula e, em “*Matrix > Fill*”, preenchê-la com a divisão da matriz de adjacência pela matriz com a soma marginal (Figura A.5). Assim, obtém-se a matriz normalizada como mostra a Figura A.6, em que os valores da coluna marginal são iguais a 1,0 agora.

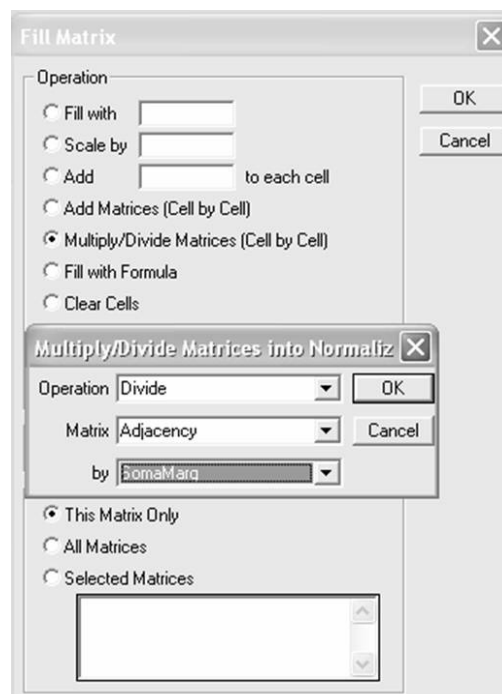


Figura A. 5: Janela para o cálculo da matriz de adjacência normalizada

	20	26	56	86	92	104	110	134	176	Sum
20	1.00	--	--	--	--	--	--	--	--	1.00
26	--	1.00	--	--	--	--	--	--	--	1.00
56	--	--	1.00	--	--	--	--	--	--	1.00
86	--	--	0.17	1.00	--	--	--	--	--	1.00
92	--	--	--	--	1.00	--	--	--	--	1.00
104	--	--	--	--	--	1.00	--	--	--	1.00
110	--	--	--	--	--	--	1.00	--	--	1.00
134	--	--	--	--	--	--	--	1.00	--	1.00
176	--	--	--	--	--	--	--	--	1.00	1.00
224	--	--	--	--	--	--	--	--	--	1.00
254	--	--	--	--	--	--	--	--	--	1.00
Sum	1.55	1.19	0.88	1.43	1.42	1.09	1.64	0.39	1.30	77.00

Figura A. 6: Matriz de adjacência normalizada

Após esse processo obtém-se a matriz de adjacência normalizada permitindo-se a sua utilização na técnica ESDA. O próximo elemento a ser determinado é o vetor de desvios, conforme é apresentado a seguir.

A.2. Vetor de desvios (Z)

Por definição, o vetor de desvios é calculado pela diferença entre o valor do atributo considerado na análise e o valor da média geral obtida no conjunto dos dados. No *dataview* há uma ferramenta que fornece a média de todos os campos da base de dados. Após abri-lo, clique em “*Dataview > Statistics*” e salve a tabela (arquivo DBF) no computador. Localize o valor da média do atributo em estudo e anote-o separadamente.

Retorne ao *dataview* e adicione uma coluna (“*Dataview > Modify Table*”) denominando-a, por exemplo, de Z. Após a sua criação, preencha-a com a fórmula (“*Edit > Fill > Formula*”): ATRIBUTO – MÉDIA.

Retorne à matriz e adicione uma nova, com o nome “Z”, por exemplo. Nesta matriz deverão ser importados os valores de Z calculados no passo anterior. Para isso, primeiro certifique-se que a coluna correspondente a Z é a última coluna à direita do *dataview*. Segundo, selecione nesta matriz qualquer linha e importe os valores de Z (que estão no *dataview*) em “*Matrix > Import Rows or Columns*”. Deve-se selecionar “*All Rows*”, verificar de onde os dados estão sendo extraídos (“*Import from*”) e selecionar o campo referente a “Z”.

A Figura A.7 ilustra esse procedimento e a Figura A.8 mostra como deverá ficar a matriz ‘Z’ após o processo de importação dos valores. É conveniente observar que este processo pode ser demorado. Para uma base de aproximadamente 10.000 registros, por exemplo, o processo pode levar mais de 24 horas.

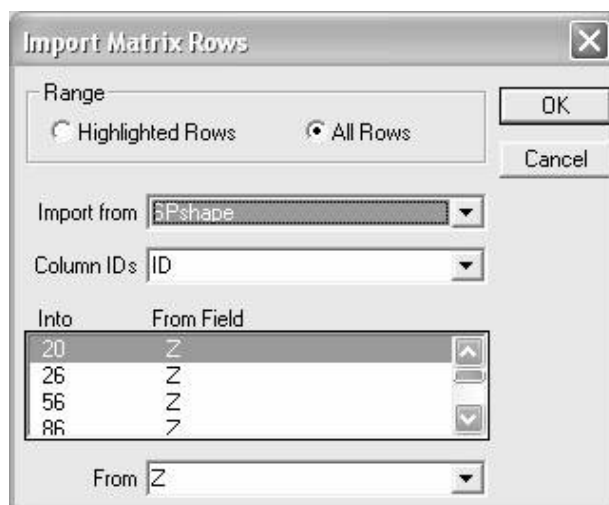


Figura A. 7: Importação dos valores de Z do *dataview* na matriz

	20	26	56	86	92	104	110	134	176		Sum
20	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
26	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
56	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
86	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
92	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
104	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
110	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
134	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
176	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
224	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
254	122761.13	-79552.87	-70005.87	-74455.87	28119.13	46794.13	29332.13	883398.13	-80619.87	-63!	0.01
Sum	21728720.01	4080857.99	2391038.99	3178688.99	4977086.01	8282561.01	5191787.01	56361469.01	4269716.99	1315!	1.77

Figura A. 8: Matriz com os valores de Z

Após esse procedimento, obtém-se uma matriz com os valores de Z , a qual será utilizada para o cálculo do terceiro elemento, W_Z , conforme é mostrado no próximo item.

A.3. Vetor W_Z

O vetor W_Z é a diferença entre a média móvel espacial e a média geral do conjunto de dados, que também pode ser obtido pela multiplicação de W por Z . O TransCAD não faz operações de matrizes com vetores, razão pela qual no passo anterior já foi obtido o vetor de desvios na forma de uma matriz (observe na Figura A.8 que todas as linhas são iguais).

Além disso, qualquer operação entre matrizes feita pelo *software* é sempre célula a célula e não seguindo as regras da matemática. Portanto, esse é um outro detalhe a ser observado na Figura A.8, em que os valores do vetor de desvios já estão transpostos, na forma adequada para se fazer a operação de multiplicação de W por Z , pelo TransCAD.

Para isso, uma nova matriz deve ser adicionada, denominando-a, por exemplo, de “ W_Z ”. Após a sua criação, acesse-a e clique em qualquer célula. Preencha esta matriz com a multiplicação da matriz “ W ” (normalizada) pela matriz “ Z ”. Após esse passo, o valor de W_Z desejado é a coluna marginal da matriz. As Figuras A.9 e A.10 ilustram, respectivamente, o processo para o cálculo de “ W_Z ” e a matriz propriamente dita já preenchida.

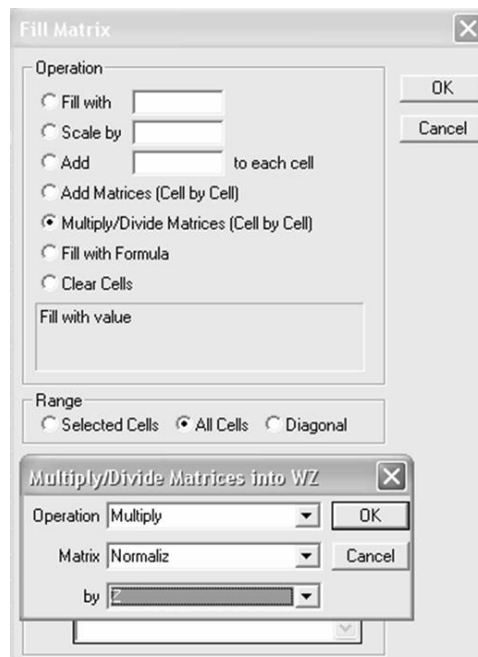


Figura A. 9: Janela para o cálculo de W_Z

	20	26	56	86	92	104	110	134	176	224	254	Sum
20		101702.96
26	-74845.54
56		-14891.17	-71802.87
86	-11667.64		-70744.70
92	-63035.24
104	8134.46
110	202394.56
134	260726.46
176	-77695.37
224	-72641.12
254		-72933.87
266	-79947.54
272	-78852.37
296	-70265.87
380	-36838.67
392	-72564.87
404	69.80
410	-4518.37
422	-47436.12
446	5623.83	-51074.87
470	-68931.37
2598	-67835.27
2610	-24818.62	-71583.20
2622	-77241.87
Sum	189987.46	-94800.50	-61921.86	-106720.08	39913.54	51195.01	48235.06	343225.03	-104805.83	-89577.92	-144983.59	-3311.925797.35

Figura A. 10: Matriz com os valores de W_Z

Exporte essa coluna para uma tabela (em formato DBF). A Figura A.11 reforça o processo de exportação da coluna marginal (que neste caso, a soma representa W_Z) indicando a seleção de “*Row Marginals*”.

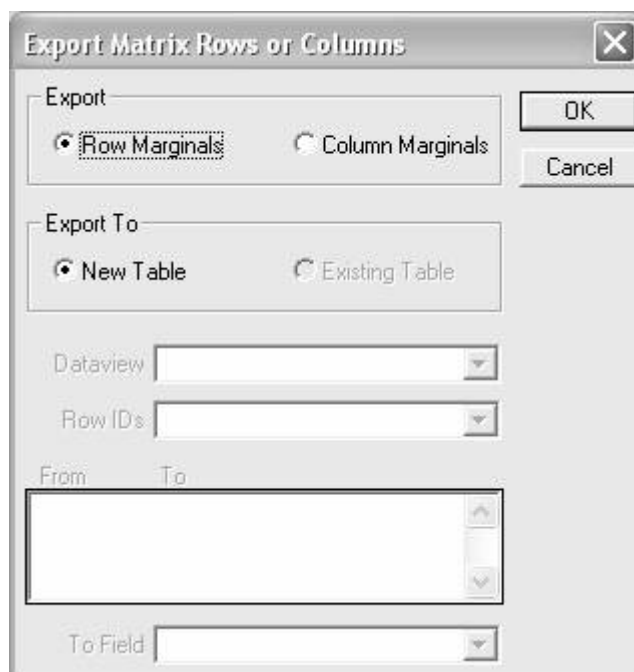


Figura A. 11: Exportação da coluna marginal com o valor de W_Z

Após esses passos, os três elementos da técnica de ESDA estão determinados, faltando apenas normalizá-los segundo o desvio-padrão, permitindo a construção do gráfico e do mapa (*Box map*) de espalhamento de Moran. O gráfico pode ser construído exportando-se os valores de Z e W_Z (normalizados) para uma planilha eletrônica, por exemplo, que ofereça alguma ferramenta de construção de gráficos de dispersão. O procedimento para a construção do *Box map* é apresentado no próximo item.

A.4. *Box map*

O procedimento para a construção do *Box map* no TransCAD é relativamente simples, uma vez calculados os elementos necessários: Z e W_Z . No item anterior o

valor de W_Z foi armazenado em uma tabela no formato DBF após a sua exportação a partir da matriz.

Agora, basta inserir esses valores no *dataview* por meio do comando “*Join*”. Acrescente mais duas colunas no *dataview*: uma para armazenar os valores de W_Z (é recomendável que se copiem os valores para uma coluna nova, pois ao fechar o programa perdem-se as informações que foram juntadas) e outra para armazenar as informações sobre os quadrantes.

O cálculo dos quadrantes é efetuado por uma fórmula com uma estrutura do tipo *IF, THEN, ELSE*. A sintaxe utilizada pelo TransCAD para essa estrutura é a seguinte (em que Z e W_Z são os nomes dados aos campos referentes à esses elementos):

```
IF (Z > 0 AND WZ > 0) THEN 1 ELSE IF (Z < 0 AND WZ < 0) THEN 2 ELSE  
IF (Z < 0 AND WZ > 0) THEN 3 ELSE IF (Z > 0 AND WZ < 0) THEN 4 ELSE  
0
```

A coluna com o campo relacionado ao quadrante deve ser preenchida com essa fórmula e, após esse procedimento, deve-se gerar um mapa temático para a sua representação sobre o território. As Figuras A.12 e A.13 ilustram, respectivamente, um gráfico e um mapa de espalhamento de Moran utilizando o procedimento aqui descrito.

Em síntese, o procedimento envolve a criação de várias matrizes para o cálculo e armazenamento dos três elementos básicos da técnica de ESDA. Para a

determinação dos quadrantes, esses elementos devem estar disponíveis no *dataview*, para o cálculo dos mesmos segundo uma fórmula com estrutura do tipo *IF, THEN, ELSE*, permitindo a sua representação por meio de gráfico ou de mapa.

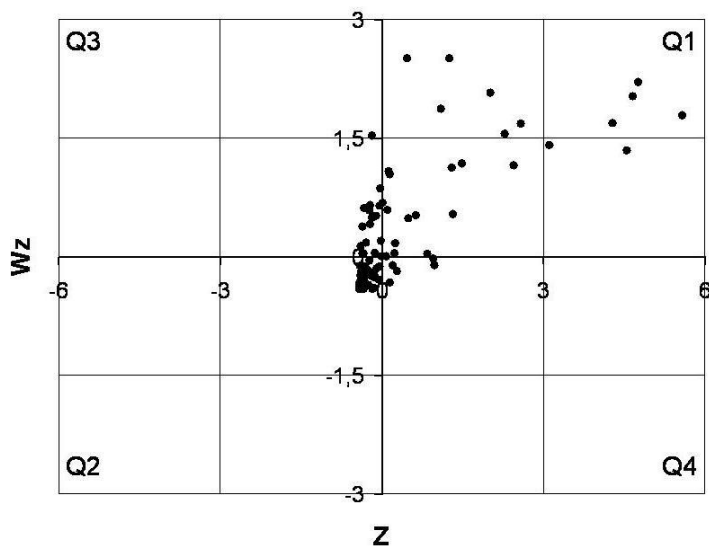


Figura A. 12: Gráfico de espalhamento de Moran

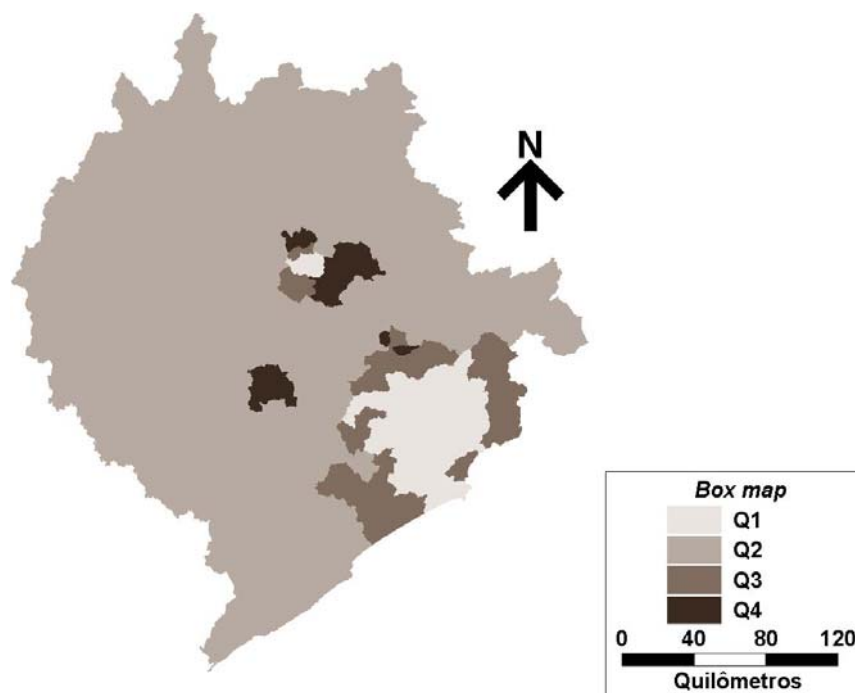


Figura A. 13: Mapa de espalhamento de Moran