

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE TRANSPORTES

**MODELAGEM DA GERAÇÃO E DISTRIBUIÇÃO DE VIAGENS
PARA ESCOLAS UTILIZANDO *CELLULAR AUTOMATA* E
AVALIAÇÃO MULTICRITÉRIO**

Eng^a VANESSA DA SILVA SANTOS

Orientador: Prof. Associado Antônio Néelson Rodrigues da Silva

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Engenharia Civil – Área de Concentração: Transportes.

São Carlos

2005

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

S237m Santos, Vanessa da Silva
Modelagem da geração e distribuição de viagens para
escolas utilizando *Cellular Automata* e avaliação
multicritério / Vanessa da Silva Santos. -- São
Carlos, 2005.

Dissertação (Mestrado) -- Escola de Engenharia de
São Carlos-Universidade de São Paulo, 2005.

Área: Transportes.

Orientador: Prof. Associado Antônio Néilson
Rodrigues da Silva.

1. *Cellular Automata*. 2. Avaliação multicritério.
3. Planejamento urbano. 4. Planejamento de
transportes. I. Título.

A meus pais, Maria do Rosário da Silva Santos e José Ribamar Santos por terem me incentivado e apoiado muito nos estudos.

A meus irmãos Leandro e Valéria pelo total apoio durante a pesquisa.

AGRADECIMENTOS

A Deus.

Ao Prof. Dr. Antônio Néelson Rodrigues da Silva, um excelente orientador.

Aos professores Osni, do Departamento de Geotecnia, e Ernesto Schaal pela boa vontade na cessão de dados para a pesquisa.

Aos Professores do Departamento de Transportes, que compartilharam seus conhecimentos de transportes e contribuíram para meu desenvolvimento profissional.

Aos professores José Leomar Fernandes Júnior e Kátia Livia Zambon pela valiosa colaboração no Exame de Qualificação.

Aos funcionários do departamento de Transportes, especialmente Heloisa, Beth, Lílian, Magaly e Suely.

À Marcela pela contribuição na pesquisa.

À Alexandra por estar sempre presente.

Aos especialistas Amilton Amorim, Cláudia Maria de Almeida, Flávia da Fonseca Feitosa, Ismael Ulysséa Neto, Lúcia Maria Capanema Álvares, Maurício Couto Polidori, Paulo de Martino Jannuzzi e Renato da Silva Lima, que dedicaram seu tempo ao preenchimento das matrizes de comparação par a par.

Aos demais colegas do STT pelo incentivo e pela agradável convivência.

À FAPESP, pelo apoio financeiro sem o qual este trabalho não seria realizado.

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

RESUMO

Em condições acentuadas de restrições de recursos para a construção e manutenção de infra-estruturas urbanas, torna-se importante caracterizar e localizar espacialmente sua demanda para auxiliar os planejadores no processo decisório que envolve sua implementação, ampliação e manutenção, de forma que os usuários sejam atendidos da melhor maneira possível e dentro das possibilidades financeiras das prefeituras. Este foi o ponto de partida para este trabalho, cujos objetivos são modelar a dinâmica populacional intra-urbana de uma cidade de médio porte brasileira através da combinação de modelos com *Cellular Automata* e Avaliação Multicritério e, a partir daí, a distribuição de viagens para uma infra-estrutura pontual específica, as Escolas Municipais de Educação Infantil (EMEI). Na aplicação prática da pesquisa inicialmente modela-se a dinâmica populacional intra-urbana na cidade de São Carlos como um todo com um modelo de *Cellular Automata* e um modelo demográfico de extrapolação de tendências. Baseando-se em valores de densidade populacional obtidos com dados dos censos do IBGE referentes aos anos de 1980, 1991 e 2000, obteve-se o cenário referente ao ano de 2010. O modelo urbano de *Cellular Automata* utilizado foi construído em três fases: quantificação da dispersão (cálculo da área total que deveria ser incorporada à mancha urbana), localização da dispersão (definição da localização das áreas que deveriam ser incorporadas à mancha urbana) e diferenciação da dispersão (cálculo das densidades demográficas na mancha urbana prevista). O modelo conseguiu apresentar bons resultados tanto na localização quanto na diferenciação da dispersão. Para caracterizar a demanda por EMEIs neste caso estabeleceu-se uma relação matemática entre a população total e a população na faixa etária que utiliza o serviço das EMEIs (4 a 6 anos). No caso do modelo demográfico, no qual a densidade populacional foi calculada a partir de uma curva de tendência linear, notou-se que devido à restrição da mancha urbana à área previamente ocupada, surgem valores de densidade populacional acima da faixa observada na série histórica. A demanda por EMEIs neste modelo também foi obtida através de uma curva de tendência linear, utilizando dados específicos da faixa etária de 4 a 6 anos. Após a caracterização da demanda por EMEIs foram criados cenários de distribuição de viagens para as mesmas nos anos de 2000 e 2010. Observou-se que os modelos apresentaram comportamentos distintos na caracterização da demanda, que foi expressivamente maior no modelo de *Cellular Automata* mas quando tiveram a demanda adaptada à oferta existente, através da multiplicação por uma taxa de atendimento, apresentaram resultados bastante semelhantes. No entanto, considerando um cenário muito provável de aumento da oferta, seja em escolas novas ou nas já existentes, o impacto sobre os transportes seria muito melhor caracterizado no caso dos modelos de CA, não só porque as estimativas de demanda foram significativamente maiores, mas também porque incorporaram a possibilidade de ocupação de novas áreas urbanas.

ABSTRACT

In conditions of extreme restrictions of resources for construction and maintenance of public facilities, the correct demand allocation is essential not only for the decision-making process, which involves their implementation and rational use, but also for reaching a significant share of the demand with a reasonable level of service and within the budget limits. That was the starting point for the definition of the objectives of this work, which are to model the population dynamics in a medium-sized city by using a *Cellular Automata* approach and Multicriteria Evaluation techniques and to simulate trip distribution patterns to a particular public facility, the EMEIs, i.e., schools for children between 4 and 6 years-old. The study starts with an application of two urban dynamics models in the city of São Carlos: a *Cellular Automata* model and a demographic model based on linear regression. A scenario of the year 2010 was built based on 1980, 1991, and 2000 census data. The *Cellular Automata* model was constructed in three phases: quantification of sprawl (definition of the total area added to the existing urban area), location of sprawl (allocation of areas of expansion), and differentiation of sprawl (definition of the population density in each cell of the new urban area). The model captured reasonably well the urban dynamics process in both location and differentiation of sprawl. The demand for EMEIs was then defined through the definition of a mathematical relationship between the total population and the target population. The population density predicted by the demographic model was calculated through a linear trend applied to historical data. It was significant in this case the occurrence of estimated values higher than the actual values mainly due to the restrictions of the urban area to its previous boundaries. The demand for EMEIs was also obtained using a linear trend, this time using specific data of the target population. Scenarios of trip distribution in the years 2000 and 2010 were created after the demand was modeled. The predictions of the models were very different, and the total demand estimation of the *Cellular Automata* model was higher than that obtained with the demographic model. When adjusting the demand to the existing supply in the year 2000, the trip distribution results were quite similar, although resulting in extremely different service rates. However, in a very likely scenario of supply growth, either through new or existing schools, the impacts on transportation would be better identified in the case of CA models, not only because they have produced higher demand estimates, but also because they have considered the inclusion of new areas into the modeled space.

LISTA DE FIGURAS

Figura 2.1 – Vizinhança quadrada de 3 x 3 células.....	5
Figura 2.2 – Estado das células no instante $t = 0$	6
Figura 2.3 – Número de células vizinhas vivas.....	6
Figura 2.4 – Células que passam do estado 0 para o estado 1.....	7
Figura 2.5 – Células que permanecem no estado 1.....	7
Figura 2.6 – Células de estado final 0.....	8
Figura 2.7 – Estado das células no instante $t = 1$	8
Figura 2.8 – Seis gerações de um modelo de <i>Cellular Automata</i>	8
Figura 2.9 – Obtenção do Mapa de condições únicas.....	20
Figura 2.10 – Cenários utilizados para o cálculo do coeficiente Fit (Costanza, 1989).....	27
Figura 2.11 – Cálculo do coeficiente F para três tamanhos de janelas.....	28
Figura 2.12 – Valor de F_1 para as 100 janelas de amostragem de tamanho 1 x 1.....	29
Figura 2.13 – Valor de F_2 para as 81 janelas de amostragem de tamanho 2 x 2.....	30
Figura 2.14 – Valor de F_w para os tamanhos de janelas de amostragem do exemplo.....	30
Figura 3.1 – Funções <i>Fuzzy</i> (Silva <i>et al.</i> , 2004).....	36
Figura 3.2 – Mancha Urbana da Cidade Fictícia.....	38
Figura 3.3 – Centro e Sistema Viário da Cidade Fictícia.....	38
Figura 3.4 – Exemplos de funções de normalização lineares.....	42
Figura 3.5 – Cenários gerados com funções crescentes.....	43
Figura 3.6 – Cenários gerados com funções decrescentes.....	43
Figura 3.7 – Cenários gerados com funções crescentes e decrescentes.....	43
Figura 5.1 – Dados Básicos utilizados na pesquisa.....	58
Figura 5.2 – Localização das EMEIs (2000).....	59
Figura 5.3 – Mapas obtidos com os dados básicos.....	60
Figura 5.4 – Malha de Células utilizadas, cobrindo a mancha urbana em 2000 e uma faixa a até 400 m da mesma.....	61
Figura 5.5 – Adequabilidade à ocupação (Adaptado de Aguiar, 1989).....	66
Figura 5.6 – Função de normalização do tipo 1 para o critério distância ao centro.....	69
Figura 5.7 – Função de normalização do tipo 1 para o critério distância à rodovia Washington Luiz.....	69
Figura 5.8 – Função de normalização do tipo 1 para o critério distância às vias urbanas principais.....	69
Figura 5.9 – Função de normalização do tipo 1 para o critério distância à mancha urbana.....	70
Figura 5.10 – Função de normalização do tipo 1 para o critério distância a pólos geradores de demanda.....	70
Figura 5.11 – Função de normalização do tipo 2 para o critério distância ao centro.....	71
Figura 5.12 – Função de normalização do tipo 2 para o critério distância à rodovia Washington Luiz.....	71
Figura 5.13 – Função de normalização do tipo 2 para o critério distância às vias urbanas principais.....	71
Figura 5.14 – Função de normalização do tipo 2 para o critério distância à mancha urbana.....	72
Figura 5.15 – Função de normalização do tipo 2 para o critério distância a pólos geradores de demanda.....	72
Figura 5.16 – Cenários de ocupação gerados pelos modelos e real para 1991.....	77
Figura 5.17 – Cenários de ocupação gerados pelos modelos e real para 2000.....	78
Figura 5.18 – Cenários de densidades populacionais gerados pelos modelos e real para 1991.....	80
Figura 5.19 – Cenários de densidades populacionais gerados pelos modelos e reais para 2000.....	81
Figura 5.20 – Cenários de densidades populacionais gerados pelo modelo de <i>Cellular Automata</i> para 2010.....	82
Figura 5.21 – Análise Espacial da População de 4 a 6 anos.....	83
Figura 5.22 – População de 4 a 6 anos prevista pelo modelo de Cellular Automata e agregada em setores de 500 metros de lado.....	84

Figura 5.23 – Cenários de densidades populacionais gerados pelo modelo demográfico para 2000.	87
Figura 5.24 – Cenários de densidades populacionais gerados pelo modelo demográfico para 2010.	87
Figura 5.25 – População de 4 a 6 anos prevista pelo modelo demográfico.	89
Figura 5.26 – Comparação entre fluxos do modelo de <i>Cellular Automata</i> e real em 2000.	92
Figura 5.27 – Comparação entre fluxos do modelo demográfico e real em 2000.	93
Figura 5.28 – Comparação entre fluxos do modelo demográfico e real em 2010.	93

LISTA DE TABELAS

Tabela 2.1 – Variáveis de entrada, condições de fronteira, classes de estado das células e vizinhança dos modelos.....	10
Tabela 2.2 – Regras de transição dos modelos.....	13
Tabela 2.3 – Atributos do mapa A.....	20
Tabela 2.4 – Atributos do mapa B.....	21
Tabela 2.5 – Atributos do mapa C.....	21
Tabela 2.6 – Atributos do mapa D (condições únicas).....	21
Tabela 2.7 – Valores reais de distribuição de áreas.....	21
Tabela 2.8 – Valores esperados de distribuição de áreas.....	22
Tabela 2.9 – Cálculo do chi-quadrado.....	22
Tabela 2.10 – Valores reais de distribuição de proporções de áreas.....	24
Tabela 2.11 – Valores de $p_{ij} \ln(p_{ij})$	24
Tabela 2.12 – Valores de F_i para diversos valores de k.....	31
Tabela 3.1 – Escala de comparação dos critérios.....	34
Tabela 3.2 – Escala de comparação de critérios.....	34
Tabela 3.3 – Matriz de comparação par a par do Avaliador 1.....	39
Tabela 3.4 – Matriz de comparação par a par do Avaliador 2.....	39
Tabela 3.5 – Matriz de comparação par a par do Avaliador 3.....	39
Tabela 3.6 – Índice de Aleatoriedade (RI) para $n = 1, 2, \dots, 15$	41
Tabela 3.7 – pontos de controle das funções de normalização.....	42
Tabela 4.1 – Matriz de transição de estados de t a $t + 10$	46
Tabela 4.2 – Matriz de transição de estados de $t + 10$ a $t + 20$	46
Tabela 5.1 – Matriz de transição de estados de 1980 a 1991.....	62
Tabela 5.2 – Matriz de transição de estados de 1991 a 2000.....	62
Tabela 5.3 – Matriz de probabilidades de transição de estados referente ao período de 1980 a 1991.....	62
Tabela 5.4 – Matriz de probabilidades de transição de estados referente ao período de 1991 a 2000.....	63
Tabela 5.5 – Matriz de probabilidades de transição de estados referente ao período de 1980 a 2000.....	63
Tabela 5.6 – Matriz de transição de estados de 2000 a 2010.....	63
Tabela 5.7 – Área da mancha urbana de São Carlos.....	64
Tabela 5.8 – Pesos atribuídos aos critérios em cada avaliação.....	67
Tabela 5.9 – Ponto de controle das funções de normalização exclusivamente crescentes ou decrescentes.....	67
Tabela 5.10 – Pontos de controle da função de normalização crescente e decrescente do tipo 1.....	68
Tabela 5.11 – Pontos de controle da função de normalização crescente e decrescente do tipo 2.....	70
Tabela 5.12 – Escore limite e número de conversões dos modelos.....	74
Tabela 5.13 – Valores do coeficiente Fit obtidos na fase de calibração (estimativa de valores para 1991).....	75
Tabela 5.14 – Valores do coeficiente Fit obtidos na fase de validação (estimativa de valores para 2000).....	76
Tabela 5.15 – Parâmetros e erros das redes neurais.....	79
Tabela 5.16 – Valores do coeficiente Fit obtidos para o ano de 1991.....	82
Tabela 5.17 – Valores do coeficiente Fit obtidos para o ano de 2000.....	82
Tabela 5.18 – Dados e estimativas de população total para a cidade de São Carlos.....	84
Tabela 5.19 – População total prevista para 2000 pelo modelo demográfico.....	86
Tabela 5.20 – População total prevista para 2010 pelo modelo demográfico.....	86
Tabela 5.21 – População de 4 a 6 anos prevista para 2000 pelo modelo demográfico.....	88
Tabela 5.22 – População de 4 a 6 anos prevista para 2010 pelo modelo demográfico.....	89
Tabela 5.23 – Proporção de pessoas de 4 a 6 anos.....	90
Tabela 5.24 – Proporção de pessoas de 4 a 6 anos.....	92

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Relevância do trabalho.....	3
1.2	Objetivo.....	3
1.3	Estrutura do Trabalho	3
2	CELLULAR AUTOMATA.....	5
2.1	Conceitos teóricos.....	5
2.2	Aplicações de Modelos de CA em Planejamento.....	9
2.3	Medidas de desempenho de modelos espaciais	17
2.3.1	Mapas de variáveis qualitativas.....	17
2.3.2	Mapas de variáveis quantitativas contínuas.....	31
3	AVALIAÇÃO MULTICRITÉRIO	33
	Exemplo de decisão entre alternativas com AMC.....	37
3.1	Avaliação de pesos para critérios	39
3.2	Normalização de critérios	41
3.3	Combinação de critérios.....	42
4	MÉTODO.....	44
4.1	Aspectos gerais	44
4.2	Modelagem da demanda por EMEIs com <i>Cellular Automata</i>	46
4.2.1	Quantificação da Dispersão	46
4.2.2	Localização da Dispersão	47
4.2.3	Diferenciação da Dispersão	51
4.2.4	Estimativa da demanda por EMEIs	52
4.3	Modelagem da demanda por EMEIs com modelo demográfico.....	53
4.4	Distribuição de viagens.....	53
4.5	Atividades Desenvolvidas	54
5	APRESENTAÇÃO E ANÁLISE DOS RESULTADOS.....	57
5.1	Base de dados	57
5.2	Modelagem da demanda por EMEIs com <i>Cellular Automata</i>	61
5.2.1	Quantificação da Dispersão	62

5.2.2	Localização da dispersão.....	64
5.2.3	Diferenciação da Dispersão	79
5.2.4	Estimativa da população de 4 a 6 anos	83
5.3	Modelagem da demanda por EMEIs com modelos demográficos.....	84
5.4	Distribuição de viagens	90
6	CONCLUSÕES E CONSIDERAÇÕES FINAIS.....	94
7	REFERÊNCIAS BIBLIOGRÁFICAS.....	97
	ANEXO 1 - Matrizes de comparação par a par	102

1 INTRODUÇÃO

Devido à expansão urbana e ao crescimento populacional de muitas cidades brasileiras, torna-se necessário conhecer a curto e médio prazo a demanda por determinados bens e serviços, tais como: escolas, postos de saúde e unidades do corpo de bombeiros. A demanda pela infra-estrutura necessária para o fornecimento dos mesmos, denominados equipamentos coletivos, está diretamente relacionada a dois fatores principais: o uso e a ocupação do solo urbano. Entende-se como uso o tipo de atividade predominante (residencial, comercial, industrial, etc.), e como ocupação a intensidade com que ocorre este uso (em geral medida pela quantidade de pessoas e/ou atividades por unidade de área).

Segundo Lima (2003) o planejamento de equipamentos coletivos é um problema de adequação entre demanda por determinados bens e serviços, da qual se conhece o valor inicial e perspectivas de evolução para um futuro próximo, e a oferta dos mesmos, de que se conhece a situação presente. Trata-se de um problema que não pode ser efetuado com sucesso sem os valores de projeto corretos da demanda nos diferentes pontos da região de intervenção. O estabelecimento desses valores envolve sempre a determinação da evolução da demanda ao longo do tempo. A evolução da demanda pode ser vista como a conjunção da evolução do universo a que o equipamento se destina a servir (a população, no todo ou em parte) com a evolução da taxa de utilização, por este universo, do serviço que o equipamento se destina a proporcionar.

No universo das cidades brasileiras, decidiu-se modelar a demanda por EMEIs nas cidades médias (cidades com população entre 100.000 e 500.000 habitantes) pelo fato de assumirem parcelas cada vez maiores da população total nos centros urbanos brasileiros. Em 1950 essa parcela era de 23,15 %, tendo aumentado para 27,24 % no ano de 2000, o que representa um aumento de

17,54 % na parcela da população residente nas cidades médias no período (ver Novak, 1998 e IBGE, 2000). No estado de São Paulo, esta parcela é ainda maior: 30,92 % em 2000 (IBGE 2000).

Segundo Novaes (1982) os modelos urbanos existentes classificam-se, de acordo com a abordagem, em: abordagem empírica, enfoque microanalítico ou comportamental, analogias e enfoque macroanalítico ou da física social.

Na *abordagem empírica* é feito um ajuste de relações matemáticas entre causa e efeito por um processo de tentativa e erro, usualmente apoiado na análise de regressão ou outro tipo de ferramental estatístico. Um exemplo seria uma função multiplicativa para obtenção das viagens produzidas entre duas zonas baseada na população e renda média per capita.

No *enfoque microanalítico ou comportamental* admite-se que a explicação de um determinado comportamento geral pode ser feita através da análise do comportamento do indivíduo.

Nas *analogias* busca-se relacionar um determinado fenômeno com outros, de natureza diversa, mas cujas leis são bastante conhecidas. Um exemplo clássico deste enfoque é o modelo gravitacional.

O *enfoque macroanalítico ou da física social* consiste em estabelecer as relações entre o comportamento típico de cada categoria de indivíduos e os níveis de demanda agregados, que são observados na realidade através de formulações probabilísticas.

Dentre os modelos existentes (ver Wegener, 1994), optou-se por pesquisar o *Cellular Automata* (CA), uma classe de modelos urbanos com comportamento análogo ao de sistemas matemáticos, que se destaca por permitir a expansão urbana durante a simulação, isto é, além de permitir a modelagem das características internas da mancha urbana existente, permite antecipar o aumento da mesma.

Um aspecto fundamental da abordagem aqui proposta é a integração, ao modelo de CA, de técnicas de Avaliação Multicritério (AMC) durante a identificação das regras de transição. Os modelos que integram AMC e CA podem ser usados em simulações para explorar como as formas urbanas evoluem sob diferentes padrões de desenvolvimento, representados por conjuntos de pesos relacionados a múltiplos critérios.

1.1 Relevância do trabalho

A relevância deste trabalho de pesquisa está na oportunidade de se desenvolver novos métodos que possibilitem associar modelos de *Cellular Automata* com técnicas de Avaliação Multicritério. O desenvolvimento de procedimentos que sejam capazes de aplicar de maneira efetiva o modelo elaborado, facilitaria aos planejadores e tomadores de decisão, a condução de análises das áreas que devem ser priorizadas na seleção de locais para implementar infra-estruturas urbanas, reduzindo os deslocamentos de seus usuários. Esses fatos motivam a pesquisa e o desenvolvimento de novos métodos que possam ser utilizados no planejamento urbano e de transportes.

1.2 Objetivo

Dadas as usuais restrições de recursos para a construção e manutenção de infra-estruturas urbanas, torna-se importante caracterizar e localizar espacialmente sua demanda para auxiliar os planejadores no processo decisório que envolve sua implementação, ampliação e manutenção, de forma que os usuários sejam atendidos da melhor forma possível e dentro das possibilidades financeiras das prefeituras.

Neste contexto, o objetivo principal desta pesquisa é:

- Modelar a geração e distribuição de viagens para uma infra-estrutura pontual específica (Escolas Municipais de Educação Infantil) utilizando *Cellular Automata* e Avaliação Multicritério.

1.3 Estrutura do Trabalho

Para atingir os objetivos propostos para o trabalho, dois eixos teóricos são fundamentais, por tratarem da base conceitual do modelo: os modelos de *Cellular Automata* e a Avaliação Multicritério. Estes conceitos e também aplicações são apresentados, respectivamente, nos Capítulos 2 e 3 e compreendem a revisão bibliográfica realizada no estudo. O quarto capítulo apresenta, por sua vez, o método empregado para o desenvolvimento do trabalho: a preparação do material básico para dar início à aplicação em questão e a definição dos procedimentos desenvolvidos na seqüência.

No quinto capítulo é apresentada a base de dados necessária para a condução da pesquisa, os resultados encontrados e, na seqüência, são conduzidas

algumas análises a eles relacionados. O sexto capítulo contém as conclusões mais importantes, assim como as considerações finais. Finalmente, no sétimo capítulo, encontram-se as referências bibliográficas citadas ao longo desta dissertação.

2 CELLULAR AUTOMATA

Este capítulo apresenta uma introdução sobre os conceitos básicos de *Cellular Automata* (CA) assim como os componentes básicos dos modelos e alguns exemplos de aplicação.

2.1 Conceitos teóricos

Os modelos de *Cellular Automata* (CA) procuram explicar alguns fenômenos através de regras simples, permitindo aos seus componentes interagir dinamicamente até o fenômeno macroscópico emergir (Taylor, 1992). Um dos modelos de CA mais simples é o Jogo da Vida, com o qual foi elaborado o exemplo abaixo. Os componentes do modelo são:

- Geometria da rede: quadrada e sem dimensão definida.
- Tamanho e formato da vizinhança: quadrada de tamanho 3 x 3 células (Figura 2.1).

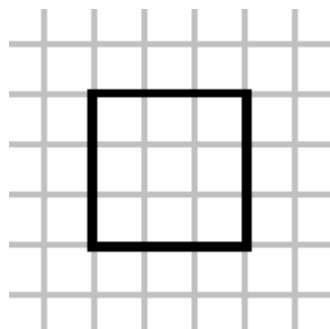


Figura 2.1 – Vizinhança quadrada de 3 x 3 células.

- Condições iniciais, apresentadas na Figura 2.2.

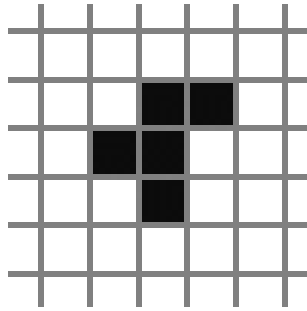


Figura 2.2 – Estado das células no instante $t = 0$.

- Classes de estado das células: 1 (preto, que neste caso representa vivo) ou 0 (branco, ou morto)
- Condições de fronteira: limite do conjunto células permitido no *software Life*.
- Regras de transição: o modelo adota regras de transição determinísticas, apresentadas a seguir.

Seja n o número de vizinhos pretos
1ª regra: Se (estado inicial = 0) e ($n = 3$) **Então** (estado final = 1)
2ª regra: Se (estado inicial = 1) e ($n = 2$ ou $n = 3$) **Então** (estado final = 1)
3ª regra: **Senão** (estado final = 0)

Definidos todos os componentes do modelo é possível determinar o estado das células no decorrer do tempo.

O primeiro passo a realizar é a contagem do número de células vizinhas vivas para cada célula, como apresentado na Figura 2.3.

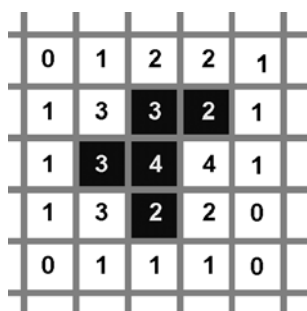


Figura 2.3 – Número de células vizinhas vivas.

A aplicação da primeira regra de transição analisa apenas as células de estado inicial igual a 0 (células mortas, representadas pela cor branca). Neste conjunto de células, as células com exatamente três vizinhos vivos passam do

estado 0 para estado 1. Na Figura 2.4, as células que passarão do estado 0 para o estado 1 aparecem em destaque.

0	1	2	2	1
1	3	3	2	1
1	3	4	4	1
1	3	2	2	0
0	1	1	1	0

Figura 2.4 – Células que passam do estado 0 para o estado 1.

A aplicação da segunda regra de transição analisa as células de estado inicial igual a 1 (vivas, de cor preta). Neste conjunto de células, as células com dois ou três vizinhos vivos permanecem no estado 1. Na Figura 2.5, as células que permanecerão no estado 1 aparecem em destaque.

0	1	2	2	1
1	3	3	2	1
1	3	4	4	1
1	3	2	2	0
0	1	1	1	0

Figura 2.5 – Células que permanecem no estado 1.

A aplicação da terceira regra de transição analisa as células que não satisfazem as condições anteriores. Neste conjunto de células, todas as de estado inicial 0 permanecem no mesmo estado e as de estado inicial 1 passam para o estado 0. Na Figura 2.6, as células que atendem a esta condição aparecem em destaque.

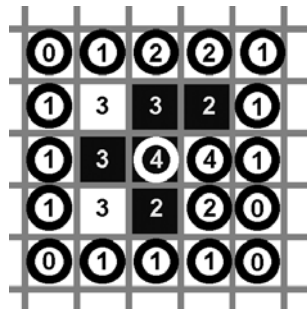


Figura 2.6 – Células de estado final 0.

A aplicação de todas as regras de transição do modelo à configuração inicial de estado de células apresentada na Figura 2.2 leva à configuração apresentada na Figura 2.7.

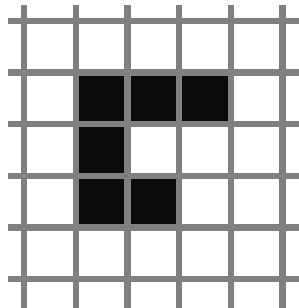


Figura 2.7 – Estado das células no instante $t = 1$.

A Figura 2.8 apresenta o resultado da simulação de cinco gerações deste modelo a partir do estado inicial definido na Figura 2.2:

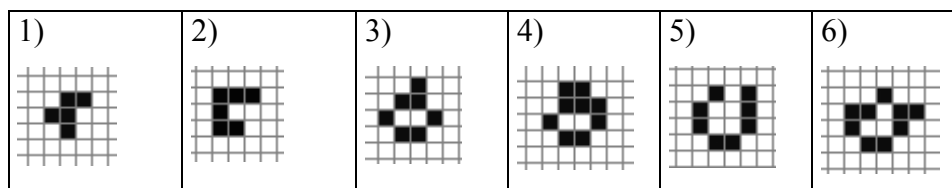


Figura 2.8 – Seis gerações de um modelo de *Cellular Automata*.

Segundo Ramos e Silva (2002), a construção de um modelo CA destinado a simular um problema específico real, tal como a dinâmica de crescimento populacional, deve obedecer a algumas escolhas. Dentre elas, as mais importantes são: a geometria da rede, o tamanho da vizinhança, condições de fronteira, condições iniciais, classe de estados e regras de transição, tal como observado no exemplo acima.

A *geometria da rede* consiste da dimensão e forma desta. Na maioria dos casos, utiliza-se uma rede quadrada, devido à facilidade de

representação e visualização (Viher *et al*, 1998). Entretanto, a maioria dos objetos das cidades não é regular, logo, a utilização de redes irregulares aumentaria o realismo dos modelos.

A *vizinhança* na qual as células podem interagir consiste na própria célula e de um conjunto de células adjacentes. Exemplos de vizinhança são a de Moore, na qual oito células formam um quadrado em torno da célula, e a de von Neumann, na qual quatro células adjacentes a uma célula formam uma cruz centrada nesta. As *condições de fronteira* permitem que o CA possa ser simulado com o auxílio do computador, pois a definição formal de CA exige que a rede seja infinita em todas as dimensões.

A *condição inicial* é o cenário de partida para a análise do problema (no exemplo, representada pela Figura 2.2). A *classe de estados* das células, de tamanho finito, pode representar qualquer característica a elas atribuída, como uso da terra (residencial ou comercial, por exemplo), densidade populacional, entre outras. As *regras de transição* são determinadas para refletir como o fenômeno real acontece, e podem ser interpretadas como algoritmos na simulação (Ramos e Silva, 2002). Elas especificam o comportamento das células com a evolução do tempo, definindo as futuras condições dessas células (Torrens, 2000).

2.2 Aplicações de Modelos de CA em Planejamento

É relativamente fácil generalizar as especificações básicas de CA para representar sistemas urbanos. O espaço da célula, na qual a automação celular opera, pode ser considerado equivalente aos territórios urbanos. A malha pode representar as estruturas espaciais e os estados da célula os atributos do espaço territorial como, por exemplo, densidade populacional (Torrens, 2000). Foram consultados, na revisão bibliográfica aqui apresentada, trabalhos que apresentam aplicações dos modelos de CA a várias cidades e regiões reais ou espaços teóricos.

A análise destas aplicações práticas baseou-se no levantamento dos principais componentes dos modelos empregados na simulação. A Tabela 2.1 apresenta, de forma resumida, as variáveis de entrada dos modelos, a geometria da rede, tamanho e formato da vizinhança, condições de fronteira e classe de estados, enquanto a Tabela 2.2 apresenta as principais características das regras de transição dos trabalhos consultados.

Tabela 2.1 – Variáveis de entrada, condições de fronteira, classes de estado das células e vizinhança dos modelos.

Trabalho	Variáveis de entrada	Condições de fronteira	Geometria da rede	Classes de estado da célula	Vizinhança
Almeida <i>et al.</i> (2002a), (2002b) e (2003)	Área coberta pela rede de atendimento de água, Faixa com densidade de ocupação média a alta (25 % a 40 %), presença de conjuntos habitacionais, distância a faixas de concentração de estabelecimentos comerciais definidas pelo estimador de Kernel, distância ao uso industrial, distância ao uso residencial, distâncias a loteamentos residenciais periféricos isolados da mancha urbana principal, distância a equipamentos institucionais periféricos e isolados da mancha urbana principal, distância à rede de principais vias existentes, distância ao eixo de serviços e corredor industrial, distância às principais vias de acesso planejadas e distância às vias periféricas que atravessam áreas não loteadas.	Limite da área de 457 x 649 células contendo a mancha urbana da cidade de Bauru (Brasil)	Quadrada, de 100 metros de lado	Usos do solo: residencial, comercial, industrial, institucional, serviços, uso misto, lazer/recreação e uso não urbano	3 x 3
Antoni (2002a e 2002b)	Uso do solo	Limite urbano da cidade de Belfort (França)	Quadrada, de 50 metros de lado	Uso do solo (áreas naturais, construídas) e redes (rodovias nacionais, interseções de rodovias, ferrovias, etc., subdivididas em 13 categorias)	Não definida, as células mais distantes possuem menor influência na definição do estado da célula

Tabela 2.1 – Variáveis de entrada, condições de fronteira, classes de estado das células e vizinhança dos modelos.

Trabalho	Variáveis de entrada	Condições de fronteira	Geometria da rede	Classes de estado da célula	Vizinhança
Batty <i>et al.</i> (1999)	Classificação da célula de acordo com tipo do uso do solo e estágio (inicial, consolidado ou em declínio) além do tempo em que a célula entrou no estágio.	Limite do conjunto de 257 x 350 pixels (espaço teórico)	Células retangulares 150 x 200 metros	Uso do solo	Depende do potencial da atividade de influenciar as outras: Ex. 3 x 3, 5 x 5, 8 x 8
Cecchini e Rizzi (2001)	Estado inicial da célula e de suas vizinhas, proximidade a áreas industriais e áreas verdes e ocorrência de eventos como a presença de novas áreas residenciais, presença de infra-estruturas e reclassificação ambiental.	Limite do conjunto de 156 x 116 células contendo centro e periferia da cidade de Alcalá de Henares (Espanha)	Quadrada de 50 metros de lado	Categorias de valor relativo de cada unidade de apartamento: classes variando de 1 (70000 a 100000 pesetas) até 8 (valor superior a 200000 pesetas)	Depende da regra de transição, analisa até 4 células de distância
Dewdney (1989)	Estado inicial das células vizinhas	Limite do conjunto de 100 x 100 células	Quadrada, de dimensão não definida	Não urbano (0) e Urbano (1)	3 x 3
Diappi <i>et al.</i> (2004)	Usos do solo da célula e de suas vizinhas	Limite urbano da cidade de Milão (Itália)	Quadrada de 500 metros de lado	Uso do solo	3 x 3
Haque e Okumura (2003)	Distância a rodovia e estação de trem, número de células em cada uso num buffer de 500 metros, proximidade à água, proximidade a área de desastres e planejamento regulatório	Limite urbano de Higashi-Hiroshima (Japão)	Quadrada, 100 metros de lado	Uso do solo	Todas as células num raio de até 5 células
Liu e Phinn (2003)	Limites inferiores e superiores de densidade populacional e tipo de padrão de ocupação	Limite do conjunto de 100 x 100 células	Quadrada, de dimensão não definida	O estado da célula assume um valor no intervalo entre 0 (não ocupado) e 1 (ocupado)	Circular de raio 1
Ohgai <i>et al.</i> (2001)	Zoneamento, distância do centro da cidade, acessibilidade a redes de transporte (ruas e estações) e intensidade da atividade urbana na vizinhança	Limite do conjunto de 22519 células contendo a área administrativa de Toyohashi (Japão)	Quadrada, de 100 metros de lado	População e número de empregos na célula	Todas as células num raio de 10 células

Tabela 2.1 – Variáveis de entrada, condições de fronteira, classes de estado das células e vizinhança dos modelos.

Trabalho	Variáveis de entrada	Condições de fronteira	Geometria da rede	Classes de estado da Célula	Vizinhança
Ramos e Silva (2003)	Densidade demográfica	Limite territorial de Portugal	Irregular, correspondente à divisão territorial em Concelhos	Quadrante em um gráfico obtido através do uso de estatística espacial	Regiões adjacentes
Roy e Snickars (1996)	Acessibilidade da célula a outras com o mesmo tipo de uso do solo	Limite urbano da cidade de Perth (Austrália)	Quadrada, de dimensão não definida	Área urbana e água	3 x 3
Sasaki <i>et al.</i> (2001)	Uso do solo e densidade populacional nas células vizinhas	Limite da Região metropolitana de Manila (Filipinas)	Quadrada, de 500 metros de lado	População	3 x 3
Wu e Webster (1998)	Custo de viagem ao centro da cidade, custo de viagem ao distrito industrial principal, acesso à estação de trem recentemente construída no subcentro a leste, acesso à rodovia, intensidade de ocupação na vizinhança e restrições topográficas e regulatórias à ocupação do solo	Limite urbano da cidade de Guangzhou (China)	Quadrada, de 200 metros de lado	Não urbano (0) e urbano (1)	3 x 3
Yeh e Li (2001a)	Número total de células ocupadas na vizinhança de raio 1, adequabilidade ao uso rural, distância de centros, distância de recursos protegidos (água potável, floresta, pântano) e consumo de solo.	Limite urbano da cidade de Dongguan (China)	Quadrada, de 50 metros de lado	O estado da célula assume um valor no intervalo entre 0 (não ocupado) e 1 (ocupado)	Circular de raio 1
Yeh e Li (2001b)	Distância à cidade principal, distância às cidades secundárias, distância à rodovia, via expressa ou ferrovia mais próxima, quantidade de vizinhos ocupados numa vizinhança 7 x 7, adequabilidade do solo ao uso agrícola	Delta do rio Pearl (China)	Quadrada, de 50 metros de lado	Probabilidade de transição	vizinhança 7 x 7

Tabela 2.1 – Variáveis de entrada, condições de fronteira, classes de estado das células e vizinhança dos modelos.

Trabalho	Variáveis de entrada	Condições de fronteira	Geometria da rede	Classes de estado da Célula	Vizinhança
Yeh e Li (2002)	Densidade máxima, distância ao centro principal e aos secundários e adequabilidade ambiental	Limite urbano da cidade de Dongguan (China)	Quadrada, de 50 metros de lado	O estado da célula assume um valor no intervalo entre 0 (não ocupado) e 1 (ocupado)	Circular de raio 1

Tabela 2.2 – Regras de transição dos modelos.

Trabalho	Regras de transição
Almeida <i>et al.</i> (2003) (2002a) e (2002b)	Estocásticas: implementadas pelo <i>software</i> DINAMICA.
Antoni (2002a)	Determinísticas: Sugere-se a divisão da transição em 3 etapas: quantificação da dispersão, localização da dispersão e diferenciação da dispersão. Na quantificação da dispersão é determinada a área construída entre períodos da série histórica de dados e estima-se o número de células pertencentes a cada estado no estado posterior com o uso de cadeias de Markov. Na localização, determina-se que células serão acrescentadas à mancha urbana através do uso de um modelo de potencial, enquanto que, na diferenciação, determina-se quais células pertencerão a cada uso.
Antoni (2002b)	Determinísticas: A transição é dividida em três etapas, das quais o artigo apresenta duas: quantificação da dispersão e localização da dispersão. Na quantificação da dispersão estima-se o número de células pertencentes a cada estado. Na localização, determina-se que células serão acrescentadas à mancha urbana, através do uso de um modelo de potencial $P_i = \sum \frac{m_j}{d_{ij}}$ onde P_i é o valor do potencial da célula i calculado num raio de 300 metros, m_j é o valor de massa para a célula j (determinado estatisticamente), que depende do uso da célula e d_{ij} é a distância de i a j .
Batty <i>et al.</i> (1999)	Estocásticas: implementada no <i>software</i> DUEM, que utiliza o parâmetro $A(t)$, sendo $A(t) = A_i(t) + A_m(t) + A_d(t)$, onde: A é uma atividade total no tempo t e i , m e d significam respectivamente inicial, consolidada (no original em inglês, <i>mature</i>) e em declínio um número aleatório n é gerado para cada atividade para classificá-la como inicial (se $n < \rho_j^i(t)$), consolidada (Se $\rho_j^i(t) \leq n \leq \rho_j^m(t)$), e em declínio (Se $n > \rho_j^m(t)$). Para cada $A_j^i(t)$ é computada a probabilidade da nova atividade $A_k(t)$ se localizar no local k . Um número aleatório entre 0 e 1 determina que células permanecem no conjunto de ocupações. Para cada par de atividades χ e ϕ há uma probabilidade da atividade χ dar lugar à atividade ϕ . O estudo não tinha, no entanto, implementado as transições de um uso do solo para outro, o que permitiria determinar que atividade ocuparia a célula.

Tabela 2.2 – Regras de transição dos modelos.

Trabalho	Regras de transição
Cecchini e Rizzi (2001)	<p>Regras determinísticas que se dividem em dois grupos: o primeiro com regras que aumentam ou decrescem o valor da célula quando próxima a células de valores maiores, regras de renovação ou decaimento e regras para áreas contíguas a áreas industriais ou áreas verdes; o segundo relacionado com a ocorrência de eventos como a presença de novas áreas residenciais e regulamento ambiental.</p> <p>Exemplo: Se (estado da célula = 2, 3 ou 4) e (há restauração de prédios históricos para uso público a até 3 células de distância) Então aumentar o valor das células em 1 ponto.</p>
Dewdney (1989)	<p>Determinísticas: Seja n o número de vizinhos vivos Se (estado inicial = 0) e ($n = 3$) então (estado final = 1) Se (estado inicial = 1) e ($n = 3$ ou 4) então (estado final = 1)</p>
Diappi <i>et al.</i> (2004)	<p>Não são explícitas, obtidas através de redes neurais artificiais (RNA)</p>
Haque e Okumura (2003)	<p>Modelo Logit calcula a probabilidade de transição, a modelagem segue em duas vertentes</p> <p>1) Determinística: O uso de solo de maior valor de probabilidade é associado à célula. 2) Estocástico: é gerado um número aleatório e o uso que tem probabilidade de transição mais próxima ao número aleatório é associado à célula.</p>
Liu e Phinn (2003)	<p>Determinística associada à lógica fuzzy: $S_{urban} = \{(x_{ij}, \mu_{S_{urban}}(x_{ij})) \mid x_{ij} \in X\}$ onde $\mu_{S_{urban}}(x_{ij})$ é uma função de ajuste baseada em lógica <i>fuzzy</i> e S_{urban} representa o estado da célula sob processo de desenvolvimento $0 \leq S_{urban} \leq 1$. Onde:</p> $\mu_{S_{urban}}(x_{ij}) = \begin{cases} 0 & \text{para } \rho_{x_{ij}} < \rho_0 \\ \frac{\rho_{x_{ij}} - \rho_0}{\rho_1 - \rho_0} & \text{para } \rho_0 \leq \rho_{x_{ij}} < \rho_1 (x_{ij} \in X) \\ 1 & \text{para } \rho_{x_{ij}} \geq \rho_1 \end{cases}$ <p>sendo ρ_0 e ρ_1 os limites inferior e superior de densidade populacional considerados na delimitação da área urbana.</p>
Ohgai <i>et al.</i> (2001)	<p>Estocástica: Para a população: Seja P a população em pessoas/ha e $\Delta S^P = 10$ pessoas/ha Se ($P^{+P} > P^P$) e ($0 < P^P < rand$) então $P^{t+1} = P^t + \Delta S^P$ Se ($P^{+P} < P^P$) e ($0 < P^P < rand$) então $P^{t+1} = P^t - \Delta S^P$ Para os empregos: Seja E o número de empregos/ha e $\Delta S^E = 5$ empregos/ha Se ($P^{+E} > P^E$) e ($0 < P^E < rand$) então $E^{t+1} = E^t + \Delta S^E$ Se ($P^{+E} < P^E$) e ($0 < P^E < rand$) então $E^{t+1} = E^t - \Delta S^E$ Sendo <i>rand</i> um número aleatório.</p>

Tabela 2.2 – Regras de transição dos modelos.

Trabalho	Regras de transição
Ramos e Silva (2003)	<p>Determinística: baseia-se na classificação dos vizinhos nos quadrantes de um gráfico de acordo com os valores de densidade populacional e na relação entre a densidade média nos vizinhos e na zona (usando técnicas de Análise Exploratória de Dados Espaciais). Por exemplo, as células pertencentes ao quadrante 2 poderiam passar aos quadrantes 1, 2, 3 ou 4 a partir da aplicação da seguinte regra.</p> <p>De 2 para 2 Se % viz em Q2 > 79% Senão de 2 para 4 Se Média da densidade de vizinhos/densidade populacional na zona < 1 Senão de 2 para 1 Se densidade média de vizinhos/densidade populacional da zona < 1.5 Senão de 2 para 3</p>
Roy e Snickars (1996)	<p>Determinísticas: Seja $a_{ik} = \sum_j \exp\left(-\mu(k) \times d_{ij}(k) \times \frac{x_j(k)}{N(k)}\right)$, onde: $\mu(k)$ é igual a 1, se a célula j é usada para a atividade do tipo k, e é igual a 0, caso contrário. $D_{ij}(k)$ é a distância da célula i à célula j para uma atividade do tipo k, $x_j(k)$ é o coeficiente de acessibilidade para a atividade do tipo k, $N(k)$ é o número de células contendo a atividade do tipo k, onde $\sum_j x_j(k) = N(k)$, a_{ik} é a acessibilidade da célula i às células contendo a atividade do tipo k, uma medida da atratividade da célula para a atividade k. Para cada uso é escolhida a célula mais atrativa, até que o número preestabelecido de células para o uso seja alocado. São implementadas pelo <i>software CityLife</i>.</p>
Sasaki <i>et al.</i> , (2001)	<p>Não são explícitas, obtidas através de redes neurais artificiais (RNA)</p>
Wu e Webster (1998)	<p>Estocástica: Seja $r_{ij} = \sum \beta_i F_i$ $p_{ij} = \exp\left[\alpha \left(\frac{r_{ij}}{r^{\max}} - 1\right)\right]$ onde r_{ij} é o escore da Avaliação Multicritério (AMC) β_i são os pesos obtidos da AMC F_i são os fatores de ocupação r^{\max} é o valor máximo de r_{ij} α é o parâmetro de dispersão p_{ij} é a probabilidade de ocupação Um processo de Monte Carlo escolhe aleatoriamente células do espaço celular Para cada célula selecionada Se $p_{ij} >$ número aleatório Então estado final = 1 As conversões são realizadas até que a demanda seja atendida.</p>
Yeh e Li (2001a)	<p>Estocástica: $G_{xy}^{t+1} = G_{xy}^t + \Delta G_{xy}^t$ Onde G é o “valor cinza” de ocupação no intervalo 0 (não ocupado) a 1 (ocupado), xy é a localização da célula e ΔG_{xy}^t é o ganho de “cinza” na iteração, dada pela equação: $\Delta G_{xy}^t = \left[1 + (-\ln \gamma)^\alpha\right] * f_{xy}(q, N) * \delta_{xy}(FORM) * \delta_{xy}(E_T) * \delta_{xy}(LAND)$ Onde: $[1 + (-\ln \gamma)^\alpha]$ é uma perturbação estocástica que representa erros desconhecidos durante a simulação, γ é um número aleatório no intervalo 0 a 1 e α é o parâmetro que controla o tamanho da perturbação estocástica; $f_{xy}(q, N)$ representa a ocupação na vizinhança, $\delta_{xy}(FORM)$ é a restrição a formas urbanas, $\delta_{xy}(ET)$ é a restrição que representa a adequabilidade ambiental e $\delta_{xy}(ET)$ é a restrição global que usa o consumo de solo para controlar a densidade de ocupação.</p>

Tabela 2.2 – Regras de transição dos modelos.

Trabalho	Regras de transição
Yeh e Li (2001b)	As RNA obtém a probabilidade de transição. A partir daí é seguida a abordagem determinística. Se a probabilidade de transição > limite pré-estabelecido Então estado final = ocupado
Yeh e Li (2002)	Determinística: $G_{xy}^{t+1} = G_{xy}^t + \Delta G_{xy}^t$ Onde G é o “valor cinza” de ocupação no intervalo 0 (não ocupado) a 1 (ocupado) e xy é a localização da célula e ΔG_{xy}^t é o ganho de “cinza” na iteração dado pela equação: $\Delta G_{xy}^t = \frac{\sum_{xy \in \Omega_N} D_{xy}}{D_{\max} \pi \xi^2} \delta_{xy}(FORM) * \delta_{xy}(E_T)$ Onde Dxy é uma função da densidade, ξ é o raio da vizinhança circular, Dmáx é o máximo valor de densidade de ocupação, $\delta_{xy}(FORM)$ é a restrição a formas urbanas, $\delta_{xy}(ET)$ é a restrição que representa a adequabilidade ambiental

Uma análise global dos trabalhos listados nas Tabelas 2.1 e 2.2 aponta que para o item *geometria da rede*, nas aplicações a uma área real, o tamanho da célula corresponde a uma parcela de tamanho definido no terreno e, nas aplicações teóricas, a uma parcela de dimensões não definidas do espaço teórico. Embora possam assumir diversos formatos regulares (quadrado, hexagonal, triangular, etc.) e até mesmo irregulares, predominou o formato de rede quadrada, de 50 a 500 metros de lado. Exceções são uma rede retangular de 150 x 200 metros (Batty *et al.*, 1999) e uma rede irregular correspondente à divisão de Portugal em *Concelhos* (Ramos e Silva, 2003).

As vizinhanças encontradas nos trabalhos se dividem em dois tipos: circular e vizinhança de Moore (em sua grande maioria 3 x 3, ou seja, formada pelas oito células contíguas à célula em análise).

As *condições de fronteira* são definidas pelo limite territorial da cidade ou região do estudo ou o limite do espaço teórico estudado.

Quanto às *classes de estados das células* dos modelos, três vertentes principais puderam ser identificadas: modelagem da *dispersão urbana* (urbano (1) ou não urbano (0)), do *uso do solo* (tipos de atividade desempenhadas na célula) e *ocupação do solo* (representada pela densidade populacional ou pela oferta de empregos).

Quanto às *regras de transição*, foram observadas três abordagens: determinística, estocástica e aquelas em que as redes neurais artificiais realizam toda a modelagem e estimam o estado final da célula. Estas abordagens podem ainda se associar a ferramentas complementares, como a lógica *fuzzy* (que usa

uma função contínua na definição da mancha urbana para representar áreas na zona de transição, onde se mesclam características de zona urbana e não urbana), a avaliação multicritério (que permite representar a interação entre efeitos de variáveis globais e locais através da adoção de pesos que representam sua importância no sistema) e as redes neurais artificiais (que neste caso estimam apenas a probabilidade de transição).

Os trabalhos que apresentam regras *determinísticas* são os que seguem a abordagem tradicional dos modelos de CA. Este tipo de regra de transição destaca-se por sua maior simplicidade e pela reprodutibilidade plena das modelagens. Na abordagem *estocástica* é introduzida, em pelo menos uma das fases da modelagem, uma perturbação aleatória que pretende reproduzir aspectos não explicáveis pelos modelos. Quando as regras de transição são representadas por *redes neurais artificiais*, os valores dos parâmetros são automaticamente obtidos do treinamento de rede neural com os dados, o que facilita a simulação e diminui a necessidade de dados, embora torne virtualmente impossível saber quais são as regras de transição adotadas.

2.3 Medidas de desempenho de modelos espaciais

Este item apresenta algumas das diversas formas de quantificar a semelhança entre mapas gerados por um modelo espacial e o mapa real.

2.3.1 Mapas de variáveis qualitativas

Este grupo de mapas apresenta dados associados a classes específicas. Mapas da ocupação urbana e de uso do solo são exemplos desta categoria. No primeiro caso as áreas são associadas às classes não urbana e urbana. No segundo caso as classes podem ser, por exemplo, urbana, residencial e comercial.

Segundo Bonham-Carter (1994), na maioria dos casos gerais de dois mapas com múltiplas classes, as medidas de correlação mais comuns são baseadas ou no teste do chi-quadrado ou em estatísticas entrópicas.

Este item apresenta coeficientes baseados no número total de acertos, no teste do chi-quadrado e em estatísticas entrópicas, além do coeficiente Fit sugerido por Costanza (1989).

2.3.1.1 Coeficientes baseados no número de acertos

Este grupo é formado por coeficientes que fazem uma análise global dos mapas. Destacam-se o coeficiente de correspondência simples (C_A) e o coeficiente de Jaccard (C_J).

O coeficiente de correspondência simples é a medida de ajuste mais simples e intuitiva. No caso de um espaço celular é o cálculo da proporção de células que estão corretamente classificadas pelo modelo, isto é, numa comparação feita célula a célula apresentam a mesma classificação no cenário real e no cenário gerado pelo modelo. É calculado segundo a Eq. (1).

$$C_A = \frac{c}{N} \quad (1)$$

Sendo c o número de células corretamente classificadas e N o número total de células no espaço celular avaliado.

O coeficiente de Jaccard (C_J) serve apenas para comparar mapas binários (mapas cujas células assumem valor 0 ou 1). Pode ser aplicado, por exemplo à modelagem da mancha urbana de uma cidade, associando-se 0 às áreas não ocupadas e 1 às áreas ocupadas. Varia de 0 (nenhum acerto) a 1 (acerto total) e é calculado conforme a Eq. (2).

$$C_J = \frac{T_{11}}{T_{11} + T_{21} + T_{12}} \quad (2)$$

Onde T_{11} é a área de acerto (área em que os dois mapas apresentam valor 1) e T_{12} e T_{21} são as áreas em que houve erro (área em que apenas um dos mapas apresenta valor 1).

Existe também a área T_{22} (área em que os dois mapas apresentam o valor 0) que não influencia o cálculo do índice. Esta é uma característica favorável ao índice, pois seu valor não é influenciado por células de valor 0 acrescentadas à periferia, uma situação necessária aos modelos urbanos de CA para que se possa realizar a dispersão da mancha urbana.

2.3.1.2 Coeficientes baseados no teste do chi-quadrado

O teste do chi-quadrado testa a relação entre duas variáveis nominais utilizando a matriz de tabulação cruzada de áreas. Denominando a tabela de áreas entre os mapas A e B de matriz T , com elementos T_{ij} , onde $i = 1, 2..n$ classes do mapa B (linhas da tabela) e $j = 1, 2..m$ classes do mapa A (colunas da tabela). Os totais marginais de T são definidos como T_{iT} , para a soma da i -ésima

linha, T_{jT} para a soma da j -ésima coluna e T_{TT} para a soma dos elementos de todas as colunas e linhas. Se os dois mapas forem independentes entre si, com nenhuma correlação entre os dois, então a área esperada em cada categoria é dada pelo produto dos totais marginais dividido pelo total geral. Então a área esperada T_{ij}^* para a i -ésima linha e j -ésima coluna é:

$$T_{ij}^* = \frac{T_{iT} T_{jT}}{T_{TT}} \quad (3)$$

E o teste do chi-quadrado é definido como

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{(T_{ij} - T_{ij}^*)^2}{T_{ij}^*} \quad (4)$$

A expressão $\frac{(\text{observado} - \text{esperado})^2}{\text{esperado}}$ tem limite inferior de 0

quando as áreas observadas são exatamente iguais às esperadas e os dois mapas são completamente independentes. Conforme as áreas observadas se diferenciam mais das áreas esperadas, o valor de chi-quadrado aumenta em magnitude e tem um limite superior variável. Além disso é fortemente dependente das unidades de medida, sendo proporcional ao tamanho da unidade de área. Assim, se a unidade linear mudar com um fator de 100 de metros para centímetros, então a unidade de área aumenta com um fator de 10.000 e o chi-quadrado também.

Segundo Bonham-Carter (1994) dois coeficientes de associação baseados nos valores de chi-quadrado são o coeficiente de Cramers, V , e o coeficiente de Contingência C . O primeiro é definido como

$$V = \sqrt{\frac{\chi^2}{T_{TT} M}} \quad (5)$$

e o último como

$$C = \sqrt{\frac{\chi^2}{T_{TT} + \chi^2}} \quad (6)$$

Onde M é o mínimo de $(n - 1, m - 1)$, sendo n e m o número de classes de cada mapa. A magnitude dos dois coeficientes é independente das unidades de medida, e varia entre 0 (indicando nenhuma correlação entre os mapas) e um valor máximo menor que 1 (indicando total correlação entre os mapas).

Exemplo de cálculo dos coeficientes de Cramers e de Contingência

O exemplo a seguir ilustra o cálculo dos coeficientes de Cramers e de Contingência entre dois mapas A e B fictícios. Para isso é necessário primeiramente construir a matriz de tabulação cruzada de áreas. A Figura 2.9 apresenta os passos para obter o mapa de condições únicas. A partir dos mapas A e B obtém-se o mapa C, que retrata todas as combinações de condições entre os mapas A e B. No Mapa D as áreas são agregadas e as que apresentam a mesma condição recebem o mesmo código identificador da condição. 20

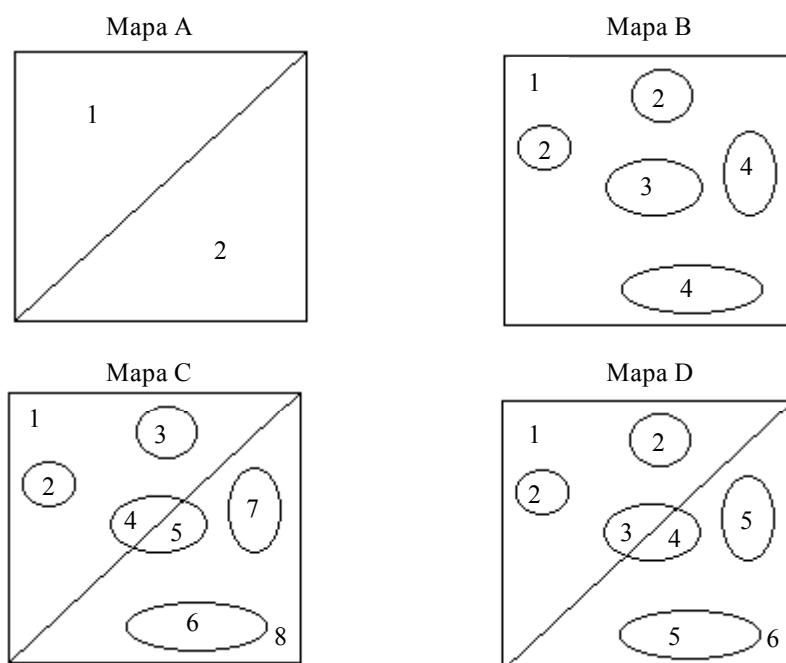


Figura 2.9 – Obtenção do Mapa de condições únicas

A Tabela 2.3 discrimina a área correspondente e cada atributo do Mapa A, enquanto a Tabela 2.4 discrimina os atributos do mapa B.

Tabela 2.3 – Atributos do mapa A.

Condição	Área
1	50
2	50

Tabela 2.4 – Atributos do mapa B.

Condição	Área
1	68
2	3
2	4
3	8
4	12
4	5

A Tabela 2.5 apresenta os atributos dos elementos do mapa C, destacando as áreas em que as condições se repetem. Estas áreas são agregadas, recebendo o mesmo identificador, resultando na Tabela 2.6, que contém somente condições únicas.

Tabela 2.5 – Atributos do mapa C.

Área	Mapa A	Mapa B	Área
1	1	1	39
2	1	2	3
3	1	2	4
4	1	3	4
5	2	3	4
6	2	4	12
7	2	4	5
8	2	1	29

Tabela 2.6 - Atributos do mapa D (condições únicas).

Condição	Mapa A	Mapa B	Área
1	1	1	39
2	1	2	7
3	1	3	4
4	2	3	4
5	2	4	17
6	2	1	29

Para calcular o chi-quadrado, elabora-se uma tabela que apresenta a distribuição da área de acordo com os valores dos atributos dos dois mapas. A Tabela 2.7 apresenta a matriz com distribuição de áreas para o exemplo.

Tabela 2.7 – Valores reais de distribuição de áreas.

Mapa B	Mapa A		Total
	1	2	
1	39	29	68
2	7	0	7
3	4	4	8
4	0	17	17
Total	50	50	100

O valor esperado das áreas, calculado com a Eq. (3), é apresentado na Tabela 2.8.

Tabela 2.8 – Valores esperados de distribuição de áreas.

Mapa B	Mapa A		Total
	1	2	
1	34	34	68
2	3,5	3,5	7
3	4	4	8
4	8,5	8,5	17
Total	50	50	100

A Tabela 2.9 apresenta as parcelas do cálculo do chi-quadrado, efetuadas com a Eq. (4).

Tabela 2.9 – Cálculo do chi-quadrado.

Mapa A	1	2	Total
1	0,7353	0,7353	1,4706
2	3,5000	3,5000	7,0000
3	0,0000	0,0000	0,0000
4	8,5000	8,5000	17,0000
Total	12,7353	12,7353	25,4706

O valor do chi-quadrado é o total das parcelas apresentadas na Tabela 2.9:

$$\chi^2 = 25,4706$$

Para calcular o índice de Cramers (V), utiliza-se a Eq. (5):

$$V = \sqrt{\frac{\chi^2}{T_{TT}M}} = \sqrt{\frac{25,4706}{100*1}} = 0,5047$$

O coeficiente de Contingências (C) é obtido com a Eq. (6):

$$C = \sqrt{\frac{\chi^2}{T_{TT} + \chi^2}} = \sqrt{\frac{25,4706}{100 + 25,4706}} = 0,4506$$

2.3.1.3 Coeficiente baseado em entropia

Segundo Bonham-Carter (1994), medidas entrópicas baseadas na matriz de tabulação cruzada podem também ser usadas para medir associações entre mapas. Supondo que os valores de T_{ij} sejam transformados em proporções de área, p_{ij} , através da divisão de cada elemento de área por T_{TT} . Assim:

$$p_{ij} = \frac{T_{ij}}{T_{TT}} \quad (7)$$

e as proporções marginais são definidas como

$$p_{iT} = \frac{T_{iT}}{T_{TT}} \quad (8)$$

$$p_{jT} = \frac{T_{jT}}{T_{TT}} \quad (9)$$

As medidas de entropia, também conhecidas como estatísticas de informação, podem ser definidas usando as proporções de área como estimativas da probabilidade. As proporções são adimensionais, logo as medidas de entropia têm a vantagem, em relação às medidas do chi-quadrado, de não serem afetadas pelas unidades de medidas.

Assumindo que uma matriz de proporções de área para o mapa A e para o mapa B tenha sido determinada a partir da matriz T, então as entropias de A e B são definidas como:

$$H(A) = -\sum_{j=1}^m p_{jT} \ln p_{jT} \quad (10)$$

e

$$H(B) = -\sum_{i=1}^m p_{iT} \ln p_{iT} \quad (11)$$

onde \ln é o logaritmo natural. A entropia conjunta da combinação, $H(A, B)$, é simplesmente:

$$H(A, B) = -\sum_{i=1}^n \sum_{j=1}^m p_{ij} \ln p_{ij} \quad (12)$$

Então o coeficiente de incerteza da informação comum (do inglês *joint information uncertainty*), $U(A, B)$, pode ser usado como uma medida de associação, sendo definido como:

$$U(A, B) = 2 \times \left[\frac{H(A) + H(B) - H(A, B)}{H(A) + H(B)} \right] \quad (13)$$

e variando entre 0 e 1. Quando os dois mapas são completamente independentes, então $H(A, B) = H(A) + H(B)$ e $U(A, B)$ é 0. Esta medida pode ser interpretada como uma combinação simétrica de duas medidas de incerteza, sendo uma a incerteza com que B prevê A e a outra a incerteza com que A prevê B. Uma discussão dessas quantidades e seus cálculos pode ser encontrada em Press *et al.* (1986).

Exemplo de cálculo do coeficiente de incerteza da informação comum

Utilizando os dados do exemplo de cálculo anterior é possível construir a Tabela 2.10, aplicando a Eq. (7).

Tabela 2.10 – Valores reais de distribuição de proporções de áreas.

Mapa A	1	2	Total
1	0,39	0,29	0,68
2	0,07	0	0,07
3	0,04	0,04	0,08
4	0,00	0,17	0,17
Total	0,50	0,50	1,00

Os valores de $H(A)$ e $H(B)$ são calculados substituindo os valores nas Eq. (10) e (11):

$$H(A) = -\sum_{j=1}^4 p_{jT} \ln(p_{jT}) = -(0,68 \times \ln(0,68) + 0,07 \times \ln(0,07) + 0,08 \times \ln(0,08) + 0,17 \times \ln(0,17))$$

$$H(A) = 0,9517$$

$$H(B) = -\sum_{i=1}^2 p_{iT} \ln(p_{iT}) = -(0,5 \times \ln(0,5) + 0,5 \times \ln(0,5))$$

$$H(B) = 0,6932$$

Os valores de $p_{ij} \ln(p_{ij})$, da Eq. (12) utilizados para calcular $H(A, B)$ são apresentados na Tabela 2.11.

Tabela 2.11 – Valores de $p_{ij} \ln(p_{ij})$.

Mapa A	1	2
1	-0,3672	-0,3590
2	-0,1861	0,0000
3	-0,1287	-0,1288
4	0,0000	-0,3012

O valor de $H(A, B)$ é a soma dos itens da Tabela.

$$H(A, B) = -\sum_{i=1}^n \sum_{j=1}^m p_{ij} \ln p_{ij} = 1,4711$$

Para calcular o coeficiente de incerteza da informação comum, substituem-se os os valores de $H(A)$, $H(B)$ e $H(A, B)$

$$U(A, B) = 2 \times \left[\frac{H(A) + H(B) - H(A, B)}{H(A) + H(B)} \right] = 2 \times \left[\frac{0,9517 + 0,6932 - 1,4711}{0,9517 + 0,6932} \right]$$

$$U(A, B) = 0,2112$$

2.3.1.4 Coeficiente FIT

Segundo Costanza (1989) a maioria dos testes estatísticos tende a mascarar ou ignorar os padrões temporais e espaciais dos dados. O teste do chi-quadrado, por exemplo, pode ser usado para medir o ajuste entre dois mapas de

uso do solo, mas ignora o padrão espacial das variáveis do uso do solo. Ele verifica apenas se o número total de *pixels* em cada categoria do uso do solo é significativamente diferentes entre dois mapas. No entanto, os mapas podem ter o mesmo número total de *pixels* em cada categoria do uso do solo e ser radicalmente diferentes no arranjo espacial dos mesmos. No caso do coeficiente de correspondência simples, a limitação é que esta medida desconsidera completamente todas as células que foram erroneamente classificadas, mesmo nos erros muito próximo dos valores corretos. Dois mapas com a mesma percentagem de correção podem exibir padrões muito diferentes em seus resíduos (erros) e este padrão é desconsiderado no cálculo deste índice.

O método Fit considera que é necessário medir o fenômeno em diferentes resoluções e interpretar o modo como os resultados mudam com as mudanças na resolução para chegar a uma descrição significativa do ajuste. Por isso dá algum peso às células pertencentes a cada categoria de localização próxima à localização real, assim como aos acertos, e calcula o ajuste total do padrão. O método aplica um procedimento padrão de ajuste a várias resoluções espaciais (representadas por janelas de amostragem de diferentes tamanhos) e usa o modo como o ajuste se modifica com a resolução para interpretar a performance do modelo.

Há diversas alternativas de medidas de ajuste numa janela de um tamanho particular, dependendo da natureza dos dados e outras considerações. A abordagem de resoluções múltiplas não depende de qual medida de ajuste é utilizada, desde que ela seja consistentemente aplicada a todas resoluções.

Para uma janela de tamanho w , uma medida conveniente de ajuste para dados categóricos é a proporção de células corretamente definidas. O ajuste para cada tamanho de janela pode ser estimado como 1 menos a proporção de células que teriam de mudar de categoria para que a janela de amostra tenha o mesmo número de células em cada categoria do cenário real, desconsiderando o arranjo espacial. A imagem inteira é analisada através de janelas de amostragem e o ajuste para o cenário total é a média do ajuste de todas as janelas de amostragem daquele tamanho, calculada pela Eq. (14):

$$F_w = \frac{\sum_{s=1}^{tw} \left[1 - \frac{\sum_{i=1}^p |a_{1i} - a_{2i}|}{2w^2} \right]_s}{t_w} \quad (14)$$

Onde F_w é o ajuste para o tamanho de janela de amostragem w , w é a dimensão de um dos lados da janela de amostragem (quadrada), a_{ki} é o número de células da categoria i no cenário k na janela de amostragem, p é o número de diferentes categorias (por exemplo, tipos de habitat) nas janelas de amostragem no cenário para o tamanho de janela de amostragem w .

Para usar estas medidas para determinar um grau geral de ajuste entre dois mapas, a informação na relação das janelas de amostragem com o ajuste, representada no gráfico, deve ser resumida. Uma média ponderada do ajuste em diferentes tamanhos de janelas é uma maneira possível de resumir o ajuste geral que permite dar mais peso às menores janelas sem ignorar totalmente as janelas maiores. Para este propósito pode ser utilizada a seguinte fórmula:

$$F_t = \frac{\sum_{w=1}^n F_w e^{-k(w-1)}}{\sum_{w=1}^n e^{-k(w-1)}} \quad (15)$$

Onde F_t é uma média ponderada dos ajustes em todos os tamanhos de janelas, F_w é o ajuste para as janelas de amostragem da dimensão linear w , k é uma constante e w é uma dimensão linear da janela de amostragem. Esta fórmula dá exponencialmente menos peso ao ajuste para resoluções mais baixas. O valor de k determina quanto peso deve ser dado a pequenas janelas de amostragem em relação às maiores. Se $k = 0$, o peso para todos os tamanhos de janelas é igual. Para $k = 1$, apenas os menores tamanhos de janelas terão importância.

A importância relativa de ajuste de padrões nas janelas de diferentes tamanhos deve ser determinada no contexto dos objetivos do modelo e da qualidade dos dados. Segundo Costanza (1989), a partir de estudos anteriores, sabe-se que modelos simples podem sempre ser de algum modo incorretos, mas eles podem fazer um bom trabalho de ajuste do padrão geral, especialmente quando há alta variância e incerteza nos dados. Para o propósito de uso do solo, $k \approx 0,10$ dá um ajuste adequado às janelas maiores.

Exemplo de cálculo do coeficiente Fit

Este exemplo, adaptado de Costanza (1989), calcula o coeficiente de ajuste Fit entre dois cenários que possuem 100 células em 4 classes de estado, apresentados na Figura 2.10.

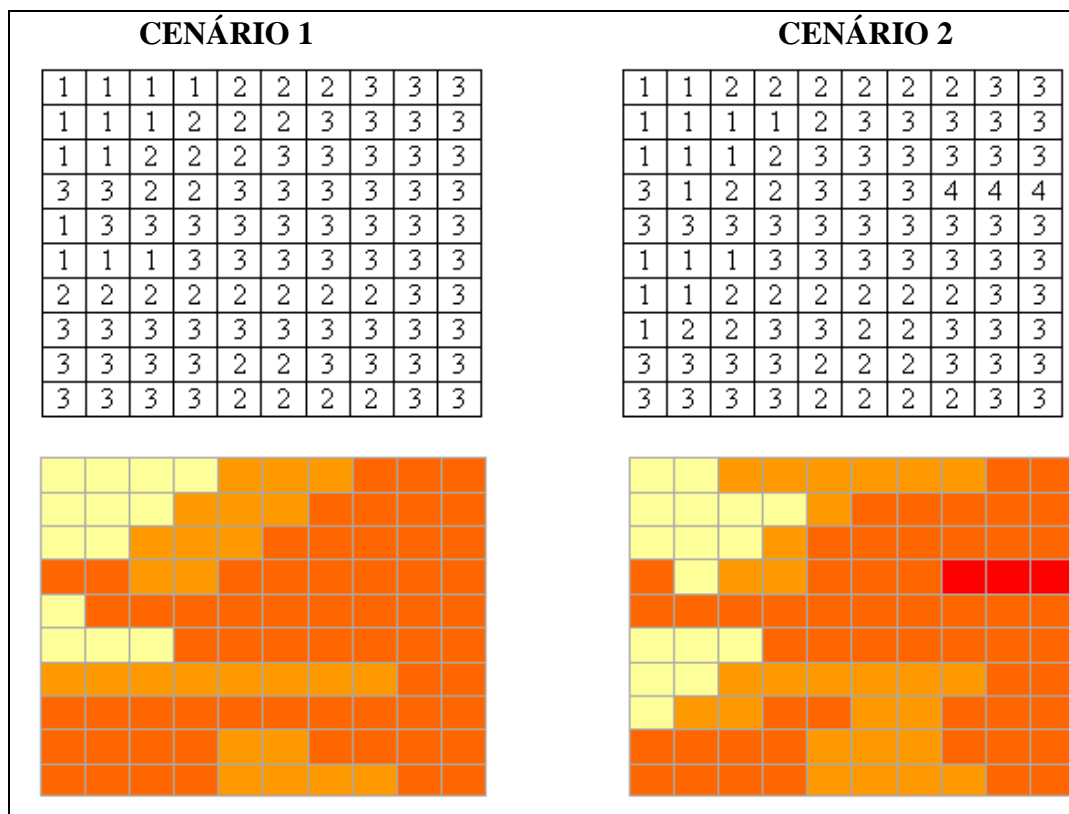


Figura 2.10 – Cenários utilizados para o cálculo do coeficiente Fit (Costanza, 1989)

O coeficiente de ajuste entre os dois cenários é determinado através de análise em janelas de amostragem de dimensão w . O coeficiente de cada janela de amostragem é calculado como 1 menos a proporção de células que teriam de mudar de categoria para que a janela de amostra tenha o mesmo número de células em cada categoria do cenário real (desconsiderando o arranjo espacial) e o valor de F_w é calculado como o valor médio do coeficiente obtido para as janelas.

A Figura 2.11 apresenta os resultados do cálculo do coeficiente Fit para três tamanhos de janelas. No caso das janelas de $w = 1$ e $w = 2$, o exemplo é detalhado a seguir.

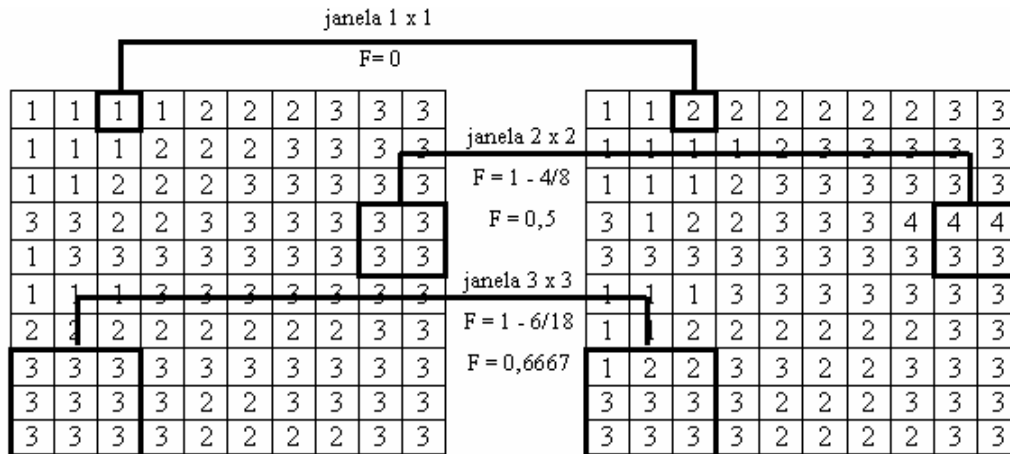


Figura 2.11 – Cálculo do coeficiente F para três tamanhos de janelas.

No cálculo da parcela $\sum_{i=1}^p |a_{1i} - a_{2i}|$ da Eq. (14) tem-se que p é igual a 4 (há 4 estados de células possíveis nos cenários).

No caso da janela de dimensão $w = 1$ destacada na Figura 2.11 é realizado o cálculo das parcelas $|a_{1i} - a_{2i}|$ apresentado a seguir:

$$|a_{11} - a_{21}| = |1 - 0| = 1,$$

$$|a_{12} - a_{22}| = |0 - 1| = 1$$

$$|a_{13} - a_{23}| = |0 - 0| = 0$$

$$|a_{14} - a_{24}| = |0 - 0| = 0$$

portanto tem-se que

$$\sum_{i=1}^p |a_{1i} - a_{2i}| = 2$$

e a proporção de células que teriam de mudar de categoria para que a janela de amostra tenha o mesmo número de células em cada categoria do cenário real é:

$$\frac{\sum_{i=1}^p |a_{1i} - a_{2i}|}{2w^2} = \frac{2}{2 * 1^2} = 1$$

O valor do coeficiente F_1 para esta janela de amostragem é:

$$F_1 = 1 - \frac{\sum_{i=1}^p |a_{1i} - a_{2i}|}{2w^2} = 1 - \frac{2}{2 * 1^2} = 0$$

A Figura 2.12 apresenta o valor de F_1 para todas as janelas de amostragem de $w = 1$.

1	1	0	0	1	1	1	0	1	1
1	1	1	0	1	0	1	1	1	1
1	1	0	1	0	1	1	1	1	1
1	0	1	1	1	1	1	0	0	0
0	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
0	0	1	1	1	1	1	1	1	1
0	0	0	1	1	0	0	1	1	1
1	1	1	1	1	1	0	1	1	1
1	1	1	1	1	1	1	1	1	1

Figura 2.12 – Valor de F_I para as 100 janelas de amostragem de tamanho 1 x 1.

Substituindo os valores na Eq. (14), tem-se que:

$$F_2 = \frac{\sum_{s=1}^{100} \left[1 - \frac{\sum_{i=1}^4 |a_{1i} - a_{2i}|}{2 * 1^2} \right]}{100} = \frac{80}{100} = 0,80$$

No caso da janela de dimensão $w = 2$ destacada na Figura 2.11 é realizado o cálculo das parcelas $|a_{1i} - a_{2i}|$ apresentado a seguir:

$$|a_{11} - a_{21}| = |0 - 0| = 0,$$

$$|a_{12} - a_{22}| = |0 - 0| = 0$$

$$|a_{13} - a_{23}| = |4 - 2| = 2$$

$$|a_{14} - a_{24}| = |0 - 2| = 2$$

portanto tem-se que

$$\sum_{i=1}^p |a_{1i} - a_{2i}| = 4$$

e a proporção de células que teriam de mudar para fazer com que a janela de amostra tenha o mesmo número de células em cada categoria do cenário real é:

$$\frac{\sum_{i=1}^p |a_{1i} - a_{2i}|}{2w^2} = \frac{4}{2 * 2^2} = 0,5$$

O valor do coeficiente F_I para esta janela de amostragem é:

$$F_1 = 1 - \frac{\sum_{i=1}^p |a_{1i} - a_{2i}|}{2w^2} = 1 - \frac{4}{2 * 2^2} = 0,5$$

A Figura 2.13 apresenta o valor de F_2 para todas as janelas de amostragem de $w = 2$.

1,00	0,75	0,75	1,00	0,75	0,75	0,75	0,75	1,00
1,00	0,75	0,50	0,50	0,50	0,75	1,00	1,00	1,00
0,75	0,50	0,75	0,75	0,75	1,00	0,75	0,50	0,50
1,00	0,75	1,00	1,00	1,00	1,00	0,75	0,50	0,50
0,75	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,50	0,75	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,25	0,50	0,75	1,00	0,75	0,50	0,75	1,00	1,00
0,50	0,50	0,75	1,00	0,75	0,25	0,50	1,00	1,00
1,00	1,00	1,00	1,00	1,00	0,75	0,75	1,00	1,00

Figura 2.13 – Valor de F_2 para as 81 janelas de amostragem de tamanho 2 x 2.

Substituindo os valores na Eq. (14) tem-se que:

$$F_1 = \frac{\sum_{s=1}^{81} \left[1 - \frac{\sum_{i=1}^4 |a_{1i} - a_{2i}|}{2 * 2^2} \right]}{81} = \frac{66,5}{81} = 0,82$$

As janelas de amostragem devem analisar toda a área dos cenários. Para os cenários do exemplo, para o tamanho 2 são necessárias 81 janelas de amostragem para cobrir toda a área, enquanto para o tamanho 10 é necessária apenas uma janela. A Figura 2.14 apresenta o gráfico com os valores do coeficiente Fit médio obtido para os 10 tamanhos de janela de amostragem (w) utilizados.

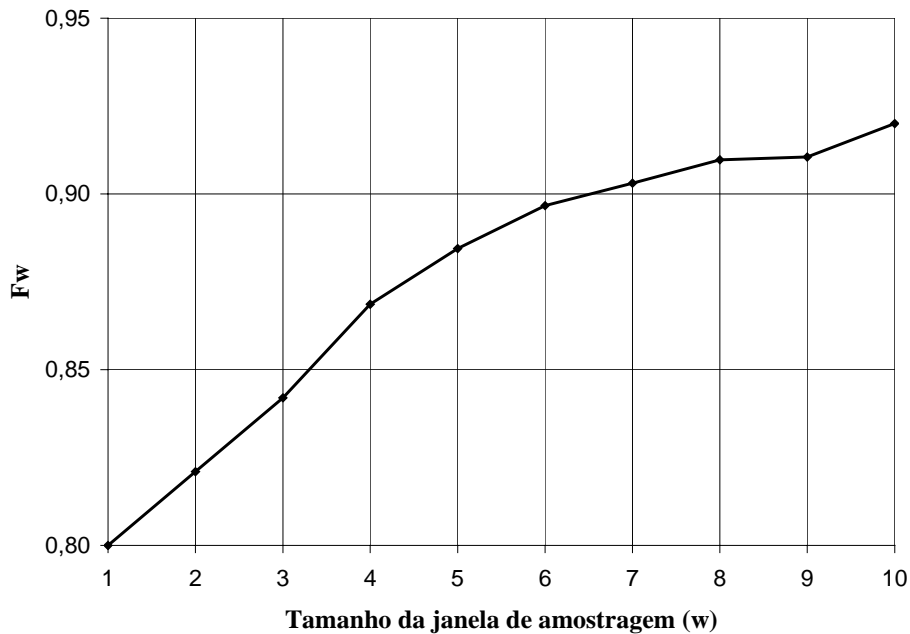


Figura 2.14 – Valor de F_w para os tamanhos de janelas de amostragem do exemplo.

A Tabela 2.12 apresenta os valores de F_t resultantes para vários valores de k .

Tabela 2.12 – Valores de F_t para diversos valores de k .

k	F_t
0	0,875601
0,1	0,864576
0,2	0,853826
0,3	0,844206
0,5	0,829576
1	0,812282

Para o exemplo na Figura 2.14, $k = 0,1$ leva a $F_t = 0,86$.

O uso de outros métodos de agregação ou ajuste produziria valores similares (mas não idênticos) de resultados, mas a idéia fundamental permaneceria. Se o gráfico se comporta como no exemplo (aumenta rapidamente no início) então o padrão de ajuste entre os dois cenários é alto, ainda que o ajuste inicial para o tamanho de janela de amostragem $w = 1$ seja baixo. Isto ocorreria se os padrões fossem similares quando observados globalmente mas apresentassem um ligeiro deslocamento na máxima resolução. De maneira inversa, se o gráfico inicia com um ajuste relativamente baixo e é horizontal, então o padrão espacial não é bem ajustado, mesmo se o ajuste inicial for maior que o padrão para os cenários do exemplo. Se os cenários fossem exatamente idênticos, ocorreria um ajuste perfeito em todos os tamanhos de janela.

2.3.2 Mapas de variáveis quantitativas contínuas

Variáveis quantitativas contínuas são variáveis cujos elementos assumem valores dentro de um intervalo contínuo. São exemplos mapas de densidade populacional e de elevações topográficas. De acordo com Bonham-Carter (1994), variáveis deste tipo podem ser comparadas com o coeficiente de correlação do produto do momento (do inglês *product moment correlation coefficient*), r :

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (16)$$

Onde x e y são os valores nos dois mapas, \bar{x} e \bar{y} são suas respectivas médias e há $i = 1, 2..n$ objetos espaciais. O coeficiente r varia entre 1 (correlação perfeita) e -1 (perfeita correlação negativa) passando por 0 (nenhuma

correlação ou independência). Os objetos espaciais podem ser pontos, *pixels* ou polígonos irregulares. Para correlações baseadas em polígonos, é desejável utilizar uma forma ponderada pela área, para que grandes polígonos tenham maior influência no resultado que pequenos polígonos. A fórmula é:

$$r_w = \frac{\sum_{i=1}^n T_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n T_i (x_i - \bar{x})^2 \sum_{i=1}^n T_i (y_i - \bar{y})^2}} \quad (17)$$

onde T_i é a área do i -ésimo polígono e x e y são médias ponderadas pela área.

2.3.2.1 Coeficiente Fit

No caso de mapas de variáveis quantitativas contínuas, pode-se aplicar o coeficiente Fit utilizando um coeficiente de ajuste adequado a este tipo de dado. Como os dados neste tipo de mapa são numéricos, o coeficiente de correlação R^2 é uma das medidas que podem ser utilizadas como coeficiente de ajuste F , calculado para as diversas dimensões de janelas de amostragem, de maneira análoga à apresentada no item 2.3.1.

3 AVALIAÇÃO MULTICRITÉRIO

Segundo Rosenthal (1985) uma análise da extensa bibliografia existente sobre o tema da tomada de decisão nos domínios da Gestão, da Pesquisa Operacional e do Planejamento do Território revela uma grande heterogeneidade e divergência no uso da terminologia referente à Avaliação Multicritério. Silva *et al.* (2004) determinam um conjunto das definições consensuais em matéria de Teoria da Decisão, apresentados a seguir:

- *Decisão*: escolha entre alternativas. Por exemplo, a escolha entre três classificações de aptidão (Alta, Média, Baixa) para um determinado uso do solo em relação a uma determinada parcela de terreno (também designada *indivíduo*) é um ato que se designa por *decisão*.
- *Alternativas*: podem representar diferentes localizações, diferentes planos, diferentes classificações, diferentes hipóteses sobre um fenômeno, etc.
- *Critério*: representa uma condição que se pode quantificar ou avaliar e que serve de base para a tomada de decisão. Os critérios podem ser de dois tipos: *exclusões* (critérios que limitam as alternativas em consideração na análise) ou *fatores* (critérios que acentuam ou diminuem a aptidão de uma determinada alternativa para o objetivo do planejamento).
- *Regra de decisão*: procedimento através do qual os critérios são combinados para chegar a uma determinada avaliação, incluindo a própria comparação entre avaliações no sentido de produzir decisões.

Silva *et al.* (2004) apresentam a regra de decisão em 3 etapas: avaliação de pesos para critérios, normalização de critérios e combinação de critérios.

Na etapa de *avaliação de pesos para critérios* é definida a importância relativa de cada critério no processo decisório, o que é feito normalmente atribuindo um peso a cada critério. Existem várias propostas de

métodos para a definição de pesos na literatura: baseados no ordenamento de critérios, em escalas de pontos, na distribuição de pontos e o método AHP, baseado em matrizes de comparações par a par (para 4 detalhes consultar Voogd, 1983; von Winterfeldt e Edwards, 1986; Malczewski, 1999; Saaty, 1977, 1980, 1987 e Silva *et al.*, 2004)

O método AHP, desenvolvido por Thomas Saaty, baseia-se numa matriz quadrada de dimensão $n \times n$, de comparação entre os n critérios, onde as linhas e as colunas correspondem aos critérios (na mesma ordem ao longo das linhas e das colunas). Os elementos a_{ij} da matriz representam a importância do critério na linha i em relação ao da coluna j e seguem uma escala de comparação apresentada nas Tabelas 3.1 e 3.2.

Tabela 3.1 – Escala de comparação dos critérios.

1/9	1/7	1/5	1/3	1	3	5	7	9
Extremamente menos importante	Bastante	Muito	Pouco	Igual	Pouco	Muito	Bastante	Extremamente mais importante

Tabela 3.2 – Escala de comparação de critérios.

Valor	Definição e explicação
1	Igual importância – os dois critérios contribuem de uma forma idêntica para o objetivo.
3	Pouco mais importante – um critério é um pouco mais importante que o outro.
5	Muito mais importante – um critério é claramente mais importante que outro.
7	Bastante mais importante – um dos critérios é predominante para o objetivo.
9	Extremamente mais importante – um dos critérios é absolutamente predominante para o objetivo.
2, 4, 6, 8	Valores intermediários – também podem ser utilizados.
Valores recíprocos aos anteriores	Se um critério i possui um dos valores anteriores quando comparado com o critério j , então o critério j possui o valor recíproco quando comparado com o critério i .

Fonte: Adaptado de Saaty (1980)

A matriz de comparação é recíproca e possui diagonal (elementos que comparam um critério a ele mesmo) igual a 1, conforme as Eqs. (18) e (19)

$$a_{ij} = \frac{1}{a_{ji}} \quad (18)$$

$$a_{ii} = 1 \quad (19)$$

O processo de determinação dos pesos desenvolve-se ao longo de sete etapas:

- Etapa 1: Construção da matriz de comparação par a par;
- Etapa 2: Cálculo do autovetor (no original, *eigenvector*) principal;
- Etapa 3: Cálculo do máximo autovalor (no original, *eigenvalue*);
- Etapa 4: Cálculo do Índice de Consistência (CI – *Consistency Index*);

- Etapa 5: Cálculo do Índice de Aleatoriedade (RI – *Random Index*);
- Etapa 6: Cálculo do Grau de Consistência (CR – *Consistency Ratio*);
- Etapa 7: Eventual reavaliação da matriz de comparação se CR superior a 0,1.

Segundo Silva *et al.* (2004) normalmente os valores de diferentes critérios não são comparáveis entre si, o que inviabiliza a sua agregação imediata. Para resolver este problema é necessário normalizar, para a mesma escala de valores, a avaliação dos diferentes critérios. Existem várias funções que podem ser utilizadas para reger a variação entre o ponto mínimo, a partir do qual os valores de escore do critério começam a contribuir para a decisão, e o valor máximo, a partir do qual escores mais elevados não trazem contribuição adicional para a decisão.

A maioria dos processos de *normalização de critérios* utiliza os valores máximo e mínimo para a definição de uma escala. A forma mais simples é uma variação linear, Outras formas de normalização são o *z-score* e o uso de funções *fuzzy* como a Sigmoidal, *J-Shaped*, Linear e Complexa (ver Bossard, 1999; Zadeh, 1965; Eastman, 1997 e Mendes, 2000). A Figura 3.1 apresenta alguns exemplos de funções de normalização crescente, isto é, aquelas cujo escore normalizado cresce com o escore original.

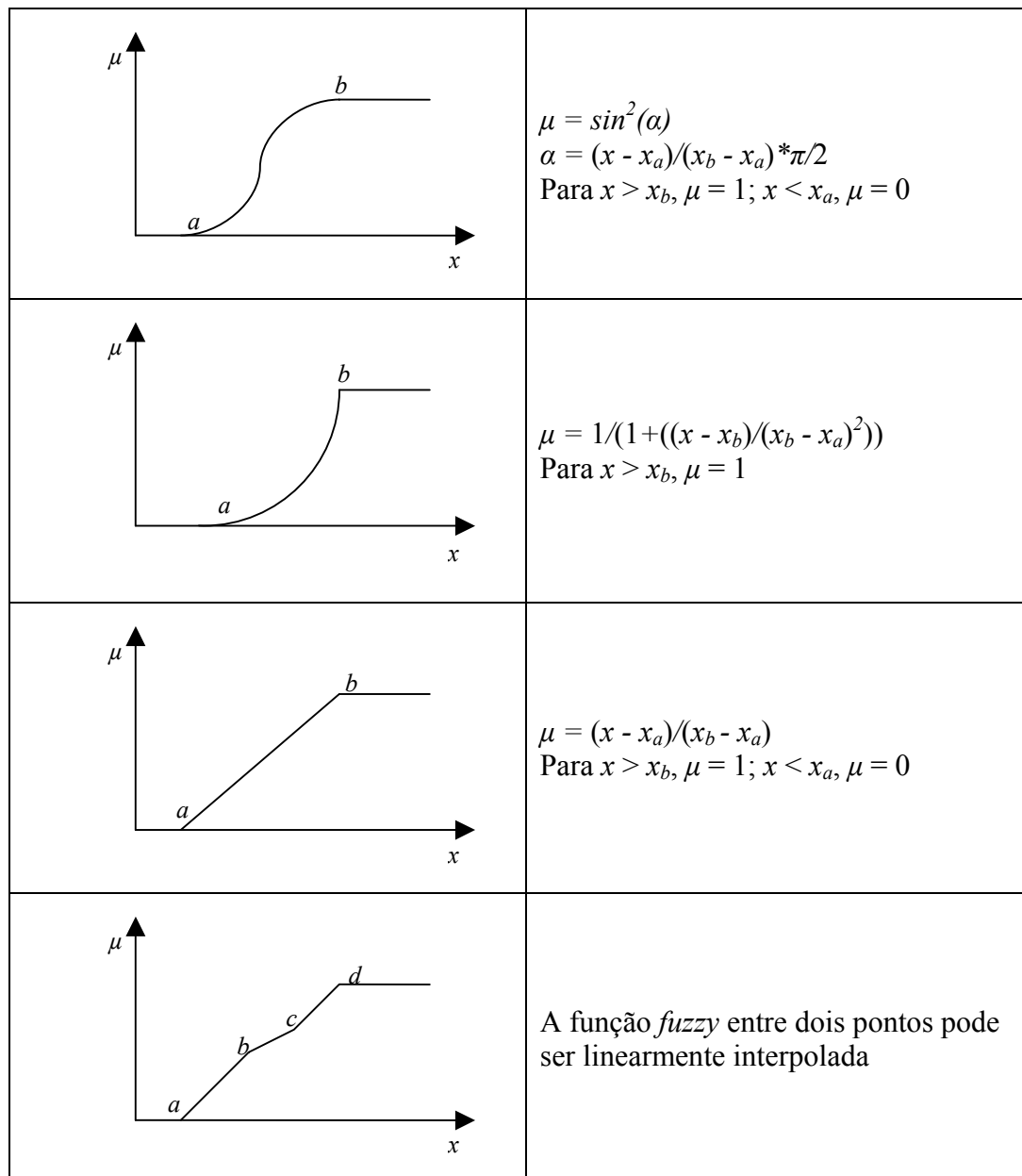


Figura 3.1 – Funções *Fuzzy* (Silva *et al.*, 2004).

Segundo Silva *et al.* (2004) quando se verifica a situação inversa, isto é, quando a uma variação positiva dos escores originais corresponde uma variação negativa dos escores normalizados, a curva é decrescente e a respectiva equação deverá ser ajustada. É o caso, por exemplo, de um critério relativo à acessibilidade, onde maiores distâncias a uma estrada correspondem a áreas menos aptas para um objetivo, logo de escore normalizado mais baixo.

A escolha da função *fuzzy* adequada depende da natureza do critério, sendo que a função sigmoideal é a mais utilizada. A seleção dos pontos de controle é um dos aspectos críticos no processo de normalização (*fuzzification*), já que de certa forma calibra a função para critérios e realidades particulares (Silva *et al.*, 2004)

Existem diversas classes de operadores para a *combinação de critérios* (ver Malczewski, 1999; Voogd, 1983; Yager, 1988; Eastman e Jiang, 1996; Eastman, Jiang e Toledano, 1998). Os dois procedimentos mais relevantes são a Combinação Linear Ponderada (WLC, do inglês *Weighted Linear Combination*) e a Média Ponderada Ordenada (OWA, do inglês *Ordered Weighted Average*).

A combinação Linear Ponderada combina os fatores através de uma média ponderada, dada pela Eq. (20). Dado que o somatório dos pesos é a unidade, o escore final é obtido na mesma escala dos escores normalizados dos fatores.

$$S = \sum_i w_i x_i \quad (20)$$

Sendo S o valor final do escore, w o peso do fator i e x_i o valor normalizado para o mesmo fator.

Nos casos em que, além dos fatores (que se expressam em escalas de escore contínuas), se aplicam também exclusões (que se expressam em escala binária 0/1), o procedimento pode ser alterado multiplicando o escore calculado com base nos fatores pelo produto das exclusões, como na Eq. (21).

$$S = \sum_i w_i x_i \times \prod_j c_j \quad (21)$$

Exemplo de decisão entre alternativas com AMC

Este exemplo apresenta um problema de decisão das áreas que apresentam maior adequabilidade para ocupação por uso residencial. O objetivo do exemplo é demonstrar o processo decisório com Avaliação Multicritério e enfatizar o importante papel das funções de normalização na modelagem. As alternativas possíveis são 276 células quadradas de 100 m de lado que representam a periferia não ocupada de uma cidade fictícia formada por 328 células. Dentre estas alternativas devem ser escolhidas 7 células que representam área da expansão territorial da mancha urbana da cidade. Os critérios utilizados são distância ao centro, distância à rodovia próxima à cidade, distância às vias urbanas principais e distância à área ocupada (mancha urbana) da cidade. A Figura 3.2 apresenta a mancha urbana da cidade e a Figura 3.3 apresenta características da cidade, utilizadas como critérios na Avaliação Multicritério.

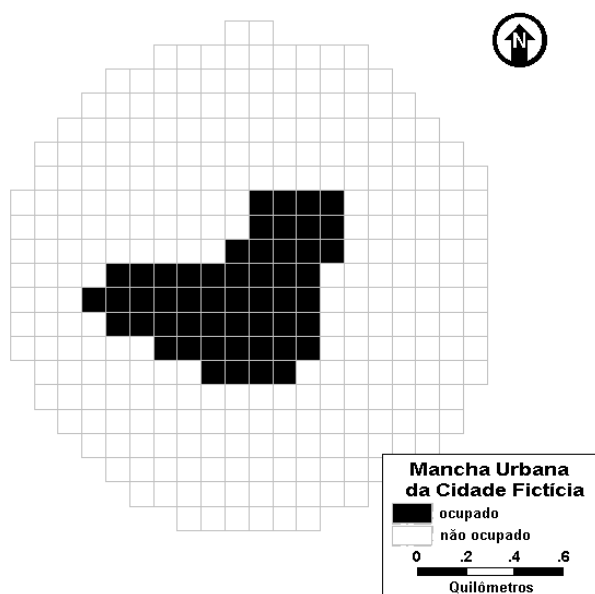


Figura 3.2 – Mancha Urbana da Cidade Fictícia.

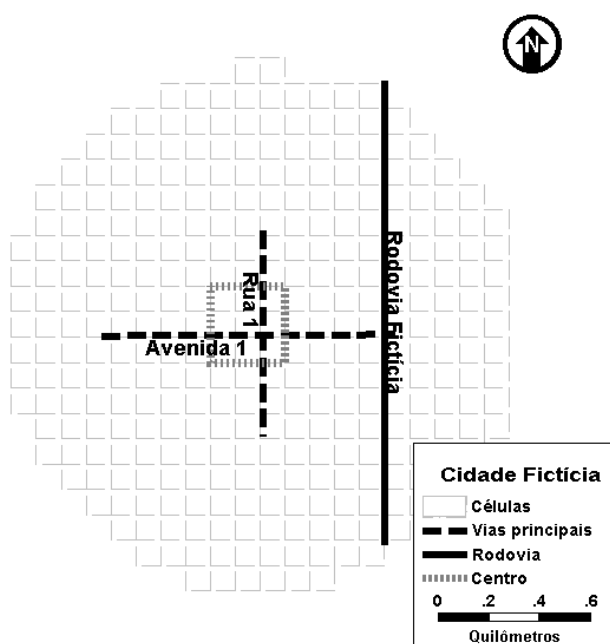


Figura 3.3 – Centro e Sistema Viário da Cidade Fictícia.

A regra de decisão é realizada em três etapas: avaliação de pesos para critérios, normalização de critérios e combinação de critérios.

A avaliação de pesos para critérios utiliza o método AHP para determinar o peso de cada critério por cada um dos três avaliadores. A normalização é realizada através de três tipos de funções *fuzzy* lineares. Na fase de

combinação de critérios é utilizada a combinação linear ponderada. Finalmente são apresentados os nove cenários resultantes.

A seguir é apresentada a realização destas etapas.

3.1 Avaliação de pesos para critérios

Através de entrevistas realizadas com cada membro do painel de especialistas, que de acordo com a escala de nove níveis anteriormente referida, atribuiu grau à relação entre cada par de critérios, foram definidas as três matrizes apresentadas nas Tabelas 3.3, 3.4 e 3.5.

Tabela 3.3 – Matriz de comparação par a par do Avaliador 1.

	distância ao centro	distância à rodovia	distância às vias principais	distância à mancha urbana
distância ao centro	1	7	1	3
distância à rodovia	1/7	1	1/7	1/5
distância às vias principais	1	7	1	7
distância à mancha urbana	1/3	5	1/7	1

Tabela 3.4 – Matriz de comparação par a par do Avaliador 2.

	distância ao centro	distância à rodovia	distância às vias principais	distância à mancha urbana
distância ao centro	1	9	3	5
distância à rodovia	1/9	1	1/7	1/5
distância às vias principais	1/3	7	1	3
distância à mancha urbana	1/5	5	1/3	1

Tabela 3.5 – Matriz de comparação par a par do Avaliador 3.

	distância ao centro	distância à rodovia	distância às vias principais	distância à mancha urbana
distância ao centro	1	1/5	1	3
distância à rodovia	5	1	5	7
distância às vias principais	1	1/5	1	3
distância à mancha urbana	1/3	1/7	1/3	1

Os itens a seguir apresentam as etapas necessárias à obtenção dos pesos que refletem as avaliações dos três especialistas.

3.1.1.1 Cálculo do autovetor principal

Para as matrizes de comparação par a par dos Avaliadores, pode-se calcular o vetor w_k pela resolução do sistema da Eq. (22):

$$A_w = \lambda_{\max} \times w_k \quad (22)$$

Sendo A_w a matriz de comparação par a par, w_k o vetor de pesos pretendidos pelo avaliador k e $\lambda_{\text{máx}}$ o máximo autovalor da matriz A .

Saaty (1980) mostrou que o autovetor resultante do máximo autovalor da matriz A traduz a prioridade dos fatores e preserva a preferência ordinal entre as alternativas. Os valores do vetor w podem então ser obtidos pela Eq. (23).

$$w_k = \frac{\sum_{j=1}^n \frac{a_{ij}}{\sum_{i=1}^n a_{ij}}}{n} \quad (23)$$

O que produz os resultados abaixo no exemplo:

$$w_1 = \begin{bmatrix} 0,3648 \\ 0,0470 \\ 0,4541 \\ 0,1341 \end{bmatrix}, w_2 = \begin{bmatrix} 0,5577 \\ 0,0417 \\ 0,2676 \\ 0,1330 \end{bmatrix} \text{ e } w_3 = \begin{bmatrix} 0,5577 \\ 0,0417 \\ 0,2676 \\ 0,1330 \end{bmatrix}$$

Este vetor apresenta uma estimativa dos pesos que estão sendo avaliados na comparação par a par.

3.1.1.2 Cálculo do máximo autovalor

O máximo autovalor ($\lambda_{\text{máx}}$) é dado pela Eq. (24).

$$\lambda_{\text{máx}} = \frac{1}{n} \left(\frac{w'_1}{w_1} + \frac{w'_2}{w_2} + \dots + \frac{w'_3}{w_3} \right) \quad (24)$$

Onde w' é um vetor obtido da forma apresentada na Eq. (25):

$$w' = A \times w \quad (25)$$

Sendo A a matriz de comparação par a par e w o vetor de pesos encontrado na etapa anterior.

Para o exemplo apresentado obtém-se $\lambda_{\text{máx}} = 4,2618$ para o Avaliador 1, $\lambda_{\text{máx}} = 4,1747$ para o Avaliador 2 e $\lambda_{\text{máx}} = 4,0739$.

3.1.1.3 Cálculo do Índice de Consistência (CI ou Consistency Index)

O valor de CI é obtido pela Eq. (26):

$$CI = \frac{\lambda_{\text{máx}} - n}{n - 1} \quad (26)$$

Para o exemplo apresentado obtém-se $CI = 0,0873$ para o Avaliador 1, $CI = 0,0582$ para o Avaliador 2 e $CI = 0,0246$ para o Avaliador 3.

3.1.1.4 Cálculo do Índice de Aleatoriedade (RI ou Random Index)

Saaty (1980) propôs valores para RI através do cálculo do valor médio de CI obtido para matrizes recíprocas geradas aleatoriamente, apresentados na Tabela 3.6.

Tabela 3.6 – Índice de Aleatoriedade (RI) para $n = 1, 2, \dots, 15$.

n	RI	n	RI	n	RI
1	0,00	6	1,24	11	1,51
2	0,00	7	1,32	12	1,48
3	0,58	8	1,41	13	1,56
4	0,90	9	1,45	14	1,57
5	1,12	10	1,49	15	1,59

Fonte: Adaptado de Ramos (2000).

No exemplo, tem-se $RI = 0,90$.

3.1.1.5 Cálculo do Grau de Consistência (CR – Consistency Ratio)

O Grau de Consistência para matrizes de mesma dimensão é dado pela Eq. (27):

$$CR = \frac{CI}{RI} \quad (27)$$

Para o exemplo tem-se $CR = 0,0970$ para o Avaliador 1, $CR = 0,06470$ para o Avaliador 2 e $CR = 0,02735$ para o Avaliador 3.

3.1.1.6 Eventual reavaliação da matriz de comparação se CR superior a 0,1

Segundo Saaty e Vargas (1991), com base na aplicação do método AHP em vários contextos envolvendo problemas complexos, é razoável aceitar os valores obtidos para os pesos sempre que se obtenha para CR um valor inferior a 0,1, sendo necessário reavaliar as opções adotadas na avaliação feita quando este valor é ultrapassado. Todas as matrizes do exemplo têm CR inferior a 0,1, portanto nenhuma matriz necessita ser refeita.

3.2 Normalização de critérios

Neste exemplo serão analisados três tipos de função *fuzzy* lineares: decrescente, crescente e crescente num intervalo e decrescente em outro intervalo, apresentadas na Figura 3.4.

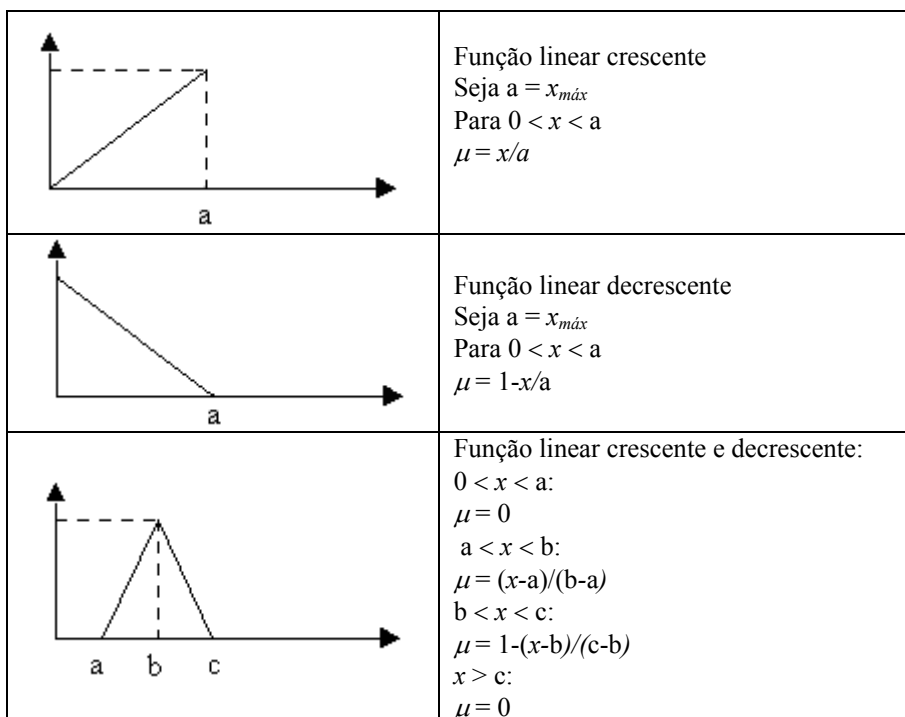


Figura 3.4 – Exemplos de funções de normalização lineares.

A Tabela 3.7 apresenta os valores dos pontos de controle utilizados para a normalização dos critérios segundo cada função.

Tabela 3.7 – pontos de controle das funções de normalização.

Critério	Funções crescentes	Funções decrescentes	Funções crescentes e decrescentes		
	a (km)	a (km)	a (km)	b (km)	c (km)
Distância ao centro	0,87	0,87	0,20	0,30	0,40
Distância à rodovia	1,41	1,41	0,80	0,90	1,00
Distância às vias principais	0,72	0,72	0,30	0,40	0,50
Distância à mancha urbana	0,70	0,70	0,10	0,40	1,30

3.3 Combinação de critérios

A combinação de critérios é realizada através da combinação linear ponderada, que calcula o escore de cada célula com a Eq. (20).

O escore das células varia de acordo com o avaliador e com a função de normalização adotada. Como há 3 Avaliadores que atribuíram diferentes pesos aos critérios e três funções de normalização para cada variável é possível gerar até nove cenários diferentes para este exemplo. Os cenários do exemplo são apresentados nas Figuras 3.5 a 3.7.

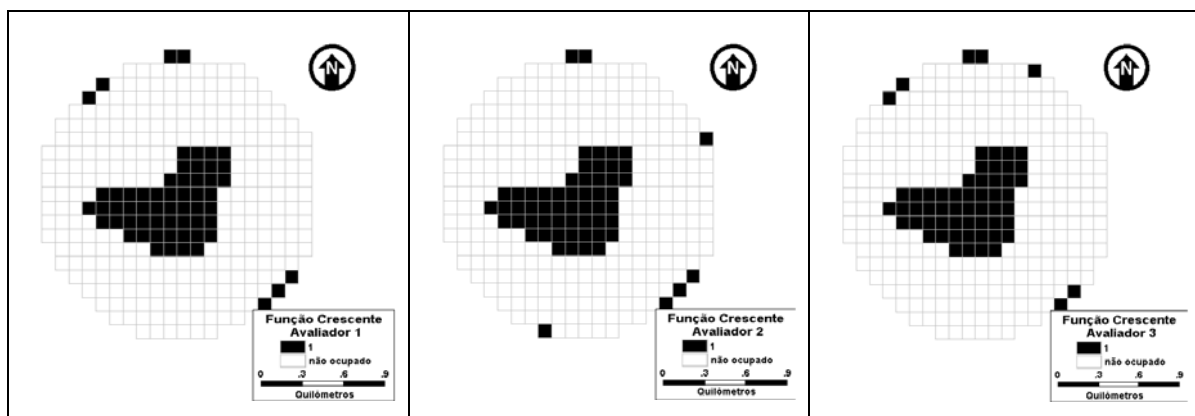


Figura 3.5 – Cenários gerados com funções crescentes.

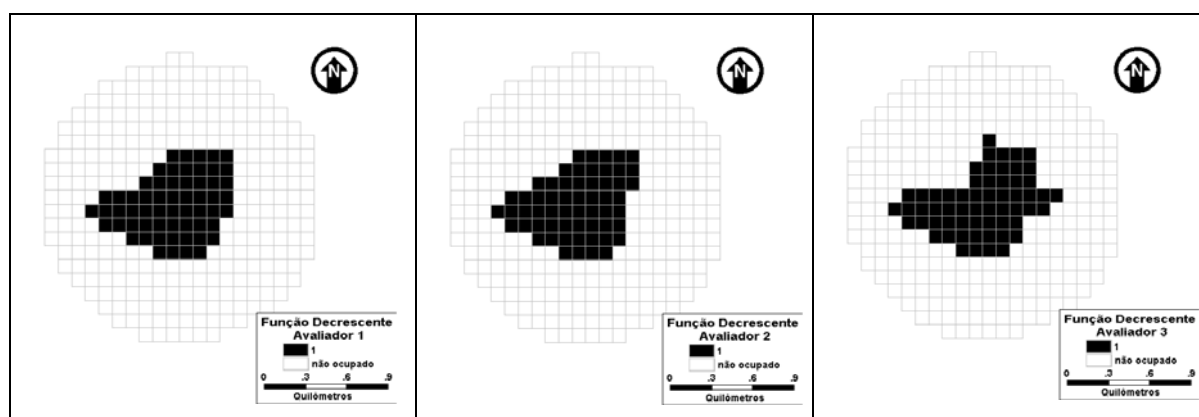


Figura 3.6 – Cenários gerados com funções decrescentes.

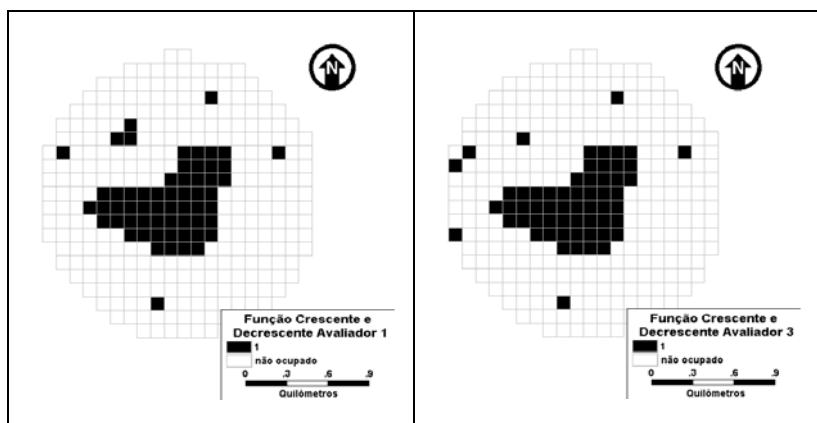


Figura 3.7 – Cenários gerados com funções crescentes e decrescentes.

É possível observar que os cenários gerados com as funções de normalização crescentes são dispersos (são escolhidas as células mais distantes da mancha urbana), os cenários gerados com as funções decrescentes são compactos e os gerados pelas funções crescentes e decrescentes são intermediários. O cenário criado com função crescente e decrescente e pesos atribuídos aos critérios pelo Avaliador 2 não é apresentado por ser idêntico ao gerado pelos critérios do Avaliador 1.

4 MÉTODO

Apresenta-se neste capítulo um resumo da abordagem adotada para o desenvolvimento de uma aplicação em transportes de um modelo baseado em *Cellular Automata* e em Avaliação Multicritério através da modelagem da demanda e distribuição de viagens para Escolas Municipais de Educação Infantil (EMEI).

4.1 Aspectos gerais

Durante a pesquisa realizou-se a modelagem da dinâmica populacional intra-urbana na cidade de São Carlos com o uso de *Cellular Automata* e Avaliação Multicritério. Estimou-se a demanda por EMEIs a partir dos resultados obtidos e realizou-se a distribuição de viagens para as mesmas com o uso de redes neurais artificiais.

O desenvolvimento da modelagem da demanda baseia-se no método definido por Antoni (2002a) e nos estudos realizados por Wu e Webster (1998) e por Sasaki *et al.* (2001). Antoni (2002a) modelou o uso do solo na cidade de Belfort (França) com um modelo de *Cellular Automata*, Wu e Webster (1998) modelaram a mancha urbana de Guangzhou (China) utilizando a Avaliação Multicritério para definir as áreas que seriam ocupadas e Sasaki *et al.* (2001) modelaram a densidade populacional na região metropolitana de Manila com redes neurais artificiais. A modelagem foi realizada em três fases: quantificação da dispersão (cálculo da área total que deveria ser incorporada à mancha urbana), localização da dispersão (definição da localização das áreas que deveriam ser incorporadas à mancha urbana) e diferenciação da dispersão (cálculo das densidades demográficas na mancha urbana prevista) e utilizou dois SIGs (TransCAD[®] e ArcView[®]), um *software* de redes neurais artificiais (EasyNNPlus[®]) e uma planilha eletrônica.

Em duas das fases desta etapa da pesquisa foram realizadas avaliações das estimativas produzidas pelos modelos. A primeira, aplicada durante a fase de localização da dispersão, avalia a forma como o modelo aloca a expansão da mancha urbana da cidade, comparando com dados reais de localização. Para realizar esta avaliação é determinada a correlação entre o mapa da mancha urbana do cenário real e o do cenário resultante da modelagem. A segunda fase da avaliação é aplicada durante a fase de diferenciação da dispersão e avalia como o modelo distribui a população na mancha urbana, através da comparação entre a densidade populacional real e aquela determinada pelo modelo.

Entre as medidas de avaliação que podem ser empregadas para identificar os modelos de melhor desempenho na etapa de localização da dispersão estão:

- Coeficiente de correspondência simples;
- Coeficiente de Jaccard;
- Índice de Cramers;
- Coeficiente de Contingência;
- Coeficiente da incerteza da informação comum;
- Fit.

As que podem ser utilizadas para avaliar os modelos na fase de diferenciação da dispersão são:

- Correlação do produto do momento;
- Fit.

Optou-se por utilizar o coeficiente Fit nas duas fases da avaliação. Com base nesta medida foram selecionados os melhores modelos para realizar a modelagem urbana com *Cellular Automata*.

A obtenção da demanda por EMEIs no modelo de CA baseou-se na relação entre a população na faixa etária que frequenta este tipo de estabelecimento, a população no setor censitário e sua distância ao centro. Após sua determinação foram utilizadas redes neurais artificiais para realizar a distribuição das viagens.

Nos itens a seguir são descritos todos os procedimentos realizados durante a pesquisa. Finalmente é apresentado um resumo das atividades desenvolvidas, para auxiliar a elaboração de trabalhos futuros.

4.2 Modelagem da demanda por EMEIs com *Cellular Automata*

Este item apresenta a modelagem da dinâmica populacional intra-urbana da cidade de São Carlos com modelos de *Cellular Automata* e o estudo da relação entre a distribuição da população total e a da faixa etária que utiliza o serviço das EMEIs.

4.2.1 Quantificação da Dispersão

A etapa de quantificação da dispersão foi realizada através do processo de cadeias de Markov, obtendo-se a matriz de probabilidade de transição baseando-se nas matrizes de transição de estado entre anos distintos, determinadas através da análise da ocupação das células.

Tabela 4.1 – Matriz de transição de estados de t a $t + 10$.

Área (ha)		$t + 10$	
		0	1
t	0	A00	A01
	1	A10	A11

Tabela 4.2 – Matriz de transição de estados de $t + 10$ a $t + 20$.

Área (ha)		$t + 20$	
		0	1
$t + 10$	0	A00	A01
	1	A10	A11

Os elementos da matriz de probabilidades de transição de estados referente ao período de t a $t + 20$ foram calculados como apresentado nas Eqs. de (27) a (30):

$$P_{11} = \frac{A00_{t-t+10} + A00_{t+10-t+20}}{A00_{t-t+10} + A00_{t+10-t+20} + A01_{t-t+10} + A01_{t+10-t+20}} \quad (27)$$

$$P_{12} = \frac{A01_{t-t+10} + A01_{t+10-t+20}}{A00_{t-t+10} + A00_{t+10-t+20} + A01_{t-t+10} + A01_{t+10-t+20}} \quad (28)$$

$$P_{21} = \frac{A10_{t-t+10} + A10_{t+10-t+20}}{A10_{t-t+10} + A10_{t+10-t+20} + A11_{t-t+10} + A11_{t+10-t+20}} \quad (29)$$

$$P_{22} = \frac{A11_{t-t+10} + A11_{t+10-t+20}}{A10_{t-t+10} + A10_{t+10-t+20} + A11_{t-t+10} + A11_{t+10-t+20}} \quad (30)$$

Sendo P_{ij} o elemento da linha i e coluna j da matriz de probabilidades de transição e Aij_{a-b} a área ocupada por células de estado i no instante a e j no instante b .

De posse das probabilidades de transição de estados são obtidos os elementos da matriz de transição de estados de $t + 20$ a $t + 30$, calculados conforme as Eq. de (31) a (34) apresentadas a seguir:

$$A00_{t+20-t+30} = P_{11} * T0_{t+20} \quad (31)$$

$$A01_{t+20-t+30} = P_{12} * T0_{t+20} \quad (32)$$

$$A10_{t+20-t+30} = P_{21} * T1_{t+20} \quad (33)$$

$$A11_{t+20-t+30} = P_{22} * T1_{t+20} \quad (34)$$

Sendo Ti_a a área total de células no estado i no tempo a .

Na aplicação conduzida para esta pesquisa, a transição do estado 1 para o estado 0 foi desprezada e considerou-se como dispersão da mancha urbana a variação da área ocupada (área no estado 1) entre os períodos considerados, calculada conforme a Eq. (35):

$$D_{a-b} = A10_{a-b} + A11_{a-b} \quad (35)$$

Sendo D_{a-b} a área que será incorporada à mancha urbana entre os instantes a e b .

4.2.2 Localização da Dispersão

Os passos utilizados para a localização da dispersão foram: estabelecimento dos critérios da Avaliação Multicritério, avaliação de pesos para critérios, normalização de critérios, combinação de critérios, aplicação das regras de transição e escolha dos melhores modelos.

4.2.2.1 Estabelecimento dos Critérios

A partir da análise das variáveis dos modelos estudados na literatura apresentados nas Tabelas 2.1 e 2.2 foram identificados critérios que pudessem influenciar na ocupação urbana, levando-se em consideração a possibilidade efetiva de cálculo de seus valores nas células. Foram utilizados tanto critérios do tipo fator quanto critérios do tipo exclusão.

4.2.2.2 Avaliação de Pesos para Critérios

Na pesquisa, diversos valores dos pesos foram obtidos a partir da aplicação de diferentes matrizes de comparação par a par, com o objetivo de refletir o comportamento do maior número possível de decisores. A seleção dos especialistas que atribuiriam os valores dos pesos associados aos critérios foi

realizada através de pesquisa na Plataforma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico (disponível no *site* <http://lattes.cnpq.br/pl>).

4.2.2.3 Normalização de Critérios

A fase da normalização pretende transformar os valores de um critério a um mesmo intervalo conhecido para que possam ser combinados. Esta fase foi realizada com quatro funções de normalização para cada variável: uma exclusivamente crescente, outra exclusivamente decrescente e duas crescentes e decrescentes. Para todas as funções adotou-se o formato linear, com equações apresentadas na Figura 3.4.

4.2.2.4 Combinação de Critérios

Para a combinação de critérios foi utilizado o procedimento Combinação Linear Ponderada, no qual o escore de cada célula, que representa a adequabilidade da célula para o uso residencial foi obtido segundo a Eq. (37):

$$r_{ij} = \left(\sum_{k=1}^m F_{ijk} W_k \right) \prod_{k=m+1}^n F_{ijk} \quad (37)$$

Sendo F_{ijk} o escore normalizado do critério k da célula de localização ij , W_k o peso do critério k , $1 \leq k \leq m$ os critérios do tipo fator; $m < k \leq n$ os critérios do tipo exclusão.

4.2.2.5 Aplicação das Regras de Transição

Após realizar o cálculo dos valores de escore resultantes de cada avaliação foi realizada a conversão de uso segundo a regra de transição apresentada a seguir:

Seja r_{ij} o escore obtido na AMC pela célula não ocupada no instante t , n_s o número simulado de células que devem pertencer à mancha urbana e $n(t+10)$ o número de células que devem pertencer à mancha urbana segundo a quantificação da dispersão no instante $t+10$

Repetir

$r_{lim} = 0$

Para todas as células

Se $r_{ij} \geq r_{lim}$

Estado $(t+10) = \text{ocupado}$

$n_s(t+10) = n(t)+1$

Se $n_s(t+10) > n(t+10)$

Aumentar r_{lim}

até $n_s(t+10) = n(t+10)$

Em cada simulação adotou-se um valor de escore limite para que o número de conversões realizado pelo modelo fosse o mais próximo possível do número de conversões observados nos instantes utilizados na modelagem.

A normalização, combinação e a aplicação das regras de transição foram realizadas em planilhas eletrônicas, de maneira conjunta. Os dados necessários para sua implementação são:

- Pesos atribuídos aos critérios pelos avaliadores;
- Pontos de controle das funções de normalização;
- Dados das células: coluna e linha (índices de localização determinados no TransCAD para possibilitar a análise de vizinhança na planilha eletrônica), valores dos critérios;
- Número de células ocupadas nos anos de t , $t + 10$, $t + 20$ e previstas para $t + 30$ na fase de quantificação da dispersão.

Foram criadas colunas para calcular e armazenar os valores normalizados dos critérios, segundo as quatro funções de normalização, o valor do escore das células segundo os valores das funções de normalização e os pesos dos avaliadores e os cenários de ocupação (criados em função dos escores das células, do valor do escore limite definido pelo número de células previstas para cada ano e do critério de restrição ambiental). No total foram criados, para cada avaliador, 4 cenários de ocupação (um para cada tipo de função de normalização) para cada ano modelado: $t + 10$, $t + 20$ e $t + 30$.

Foram gerados cenários de ocupação para $t + 10$, utilizando dados de t , e para $t + 20$, utilizando dados de $t + 10$. Os cenários gerados foram comparados aos cenários de ocupação reais em $t + 10$ e $t + 20$ através do índice Fit e, para os melhores modelos, foi realizada a fase de diferenciação da dispersão.

4.2.2.6 Escolha dos melhores modelos de localização da dispersão

Para avaliar os cenários de ocupação gerados criou-se uma planilha eletrônica para comparar cenários qualitativos, que calcula o coeficiente Fit para células pertencentes a duas classes de estados: ocupado e não ocupado.

Os dados necessários para o cálculo são:

- Coluna e linha das células;
- Ocupação real;
- Ocupação simulada;
- Valor do coeficiente k para o cálculo do coeficiente Fit global.

A planilha eletrônica de cálculo é composta pelas seguintes pastas:

- Dados: contém os dados exportados do TransCAD (identificadores de coluna e linha, ocupação estimada e real);
- Resultados: contém o cálculo do coeficiente Fit médio para cada tamanho de janela e do coeficiente Fit total;
- TD ocupr: contém uma tabela dinâmica que informa, para cada linha e coluna do cenário, a ocupação real da célula;
- TD ocups: contém uma tabela dinâmica que informa, para cada linha e coluna do cenário, a ocupação simulada da célula;
- Fw ($w = 1, 2, \dots, 10$): contém o cálculo do coeficiente Fit para as janelas de tamanho w .

As fórmulas utilizadas no cálculo seguem o formato a seguir:

```
SE(NÃO(
E((SOMA("TD ocupr"! intervalo w x w)=0);
(SOMA("TD ocups"! intervalo w x w)=0)));
1-(ABS(CONT.SE("TD ocupr"! intervalo w x w;"=1") - CONT.SE("TD ocups"!
intervalo w x w;"=1"))
+ABS(CONT.SE("TD ocupr"! intervalo w x w;"=0")-
CONT.SE("TD ocups"! intervalo w x w;"=0")))/(2*w^2);
FALSO)
```

Onde o intervalo $w \times w$ é o intervalo de valores de ocupações reais e estimadas utilizados para o cálculo do coeficiente Fit para cada janela de lado w .

Com a atribuição do valor FALSO às janelas em que não há nenhuma célula ocupada nos cenários real e estimado, a função média desconsidera o valor do coeficiente Fit das janelas vazias, o que reduz a alteração do valor médio do coeficiente Fit pelas células das áreas periféricas.

O coeficiente é calculado nos tamanhos de janelas de 1 a 10 conforme a Eq. (14). Em seguida calcula-se o valor médio F_w (que despreza os valores FALSO) para cada tamanho de janela. O coeficiente global F_t é calculado através da combinação dos valores de F_w , segundo a Eq. (15), utilizando o valor de k escolhido.

Após a realização da avaliação de todos os modelos criados, foram escolhidos os quatro modelos que tiveram melhor desempenho, representados por maiores valores do coeficiente Fit, para serem utilizados na fase de diferenciação da dispersão.

4.2.3 Diferenciação da Dispersão

A fase de diferenciação da dispersão foi realizada com o uso de redes neurais artificiais e foi implementada no *software* de redes neurais artificiais EasyNNPlus. Os dados necessários são: densidades populacionais nas células nos anos t , $t + 10$ e $t + 20$, segundo dados do IBGE, e previsão de população para o ano $t + 30$. Foram calibradas redes com os dados das células ocupadas em $t + 10$, utilizando como parâmetros de entrada os dados referentes ao ano t e de saída os referentes a $t + 10$, ambos normalizados entre 0 e 1. Os dados foram divididos em três grupos: o primeiro grupo, destinado ao treinamento, continha 50 % dos dados, os grupos de validação e teste correspondiam a 25 % cada.

Calculou-se o erro de cada uma das redes com a Eq. (38)

$$Erro = média \left(\frac{dens_{ij}^s - dens_{ij}^r}{máximo(dens_{ij}^r)} \right) \quad (38)$$

Onde $dens_{ij}^s$ é a densidade populacional simulada na rede e $dens_{ij}^r$ é a densidade populacional real.

Após a calibração do modelo, a rede foi alimentada com dados de entrada referentes a t (apenas os dados das células escolhidas pelos melhores modelos para ser ocupadas) e forneceu como saída valores normalizados referentes ao ano de $t + 10$. Estes valores normalizados foram multiplicados por uma constante, de tal maneira que a população total nas células fosse a população referente ao ano de $t + 10$. Às células não selecionadas na fase de localização da dispersão era atribuído valor 0 de densidade populacional.

Este procedimento foi repetido com os dados relativos ao ano $t + 10$. Foram obtidos oito cenários para os quatro melhores modelos, que foram avaliados novamente para obter o melhor modelo.

4.2.3.1 Escolha do melhor modelo global de diferenciação

Para avaliar os cenários de ocupação gerados criou-se uma planilha de cálculo análoga à planilha de avaliação qualitativa, que calcula o coeficiente Fit para células associadas a valores de densidade populacional.

Os dados necessários para o cálculo são:

- Coluna e linha das células (índices de localização determinados no TransCAD para possibilitar a análise de vizinhança na planilha eletrônica);
- Densidade populacional segundo censo do IBGE;

- Densidade populacional simulada;
- População real e simulada na célula;
- Valor do coeficiente k para o cálculo do coeficiente Fit global;
- A planilha eletrônica de cálculo é composta pelas mesmas pastas da planilha de avaliação qualitativa, alterando apenas a maneira como o coeficiente é determinado.

O coeficiente quantitativo é calculado com a diferença da população total dentro de cada janela. A diferença é dividida pela diferença de população entre uma janela ocupada com densidade populacional real máxima e uma janela de densidade populacional nula. O coeficiente Fit é calculado como 1 (um) menos essa diferença normalizada, o que resulta num valor 1 (um) para densidades iguais (diferença entre populações nula) e 0 (zero) para a diferença de população entre uma janela ocupada com a densidade populacional máxima e outra janela não ocupada. As fórmulas utilizadas no cálculo seguem o formato a seguir:

```
SE(
SOMA("TD pop_r"!intervalo w x w;"TD pop_s"! intervalo w x w)=0;
FALSO;
1-ABS(SOMA("TD pop_r"! intervalo w x w)-SOMA("TD pop_s"! intervalo
w x w))/(máx(Dados!intervalo pop real)*w^2))
```

Onde o intervalo $w \times w$ é o intervalo de valores de populações reais e estimadas utilizados para o cálculo do coeficiente Fit para cada janela de lado w , e $\text{máx}(\text{Dados!intervalo pop real}) * w^2$ é a diferença de população entre uma janela de ocupação com densidade populacional real máxima e uma janela de densidade populacional nula.

4.2.4 Estimativa da demanda por EMEIs

Na fase de estimativa da demanda por EMEIs foram utilizados dados dos setores censitários ajustados. Foi calculada a distância dos centróides dos setores até o centro da cidade, assim como a proporção entre a população na faixa etária de 4 a 6 anos e a população total.

Em seguida estabeleceu-se uma função linear relacionando as duas variáveis calculadas e agregou-se a distribuição da população obtida pelo modelo obtido na fase anterior em setores quadrados de 500 metros de lado. Os setores que apresentaram população maior que 2000 habitantes (número máximo de habitantes usualmente adotado como limite por setor censitário do IBGE) foram divididos em dois.

De posse da função e da população por setor foi realizado o cálculo da população na faixa etária de 4 a 6 anos nos setores, que considera a multiplicação da demanda por uma taxa de atendimento.

4.3 Modelagem da demanda por EMEIs com modelo demográfico

Além do modelo de *Cellular Automata* foi construído um modelo demográfico de extrapolação de tendências para fins comparativos. Neste modelo, a densidade populacional de $t + 30$ foi estimada a partir de uma curva de tendência linear construída com os dados de densidade populacional dos anos de t , $t + 10$ e $t + 20$, agregados na divisão de setores censitários ajustados de $t + 30$. Após o cálculo da população prevista pelo modelo, comparou-se o valor de população total previsto por Lima (2003), determinando-se um índice de ajuste global para as densidades, de modo que o valor da população total coincidissem. Para modelar a faixa etária dos 4 a 6 anos, o procedimento de obtenção de densidade a partir de uma curva de ajuste linear foi repetido, substituindo-se as densidades populacionais totais pelas densidades populacionais da faixa etária de 4 a 6 anos. As densidades obtidas foram multiplicadas pelo fator de ajuste obtido no modelo global e os valores resultantes foram multiplicados pela taxa de atendimento para obter a demanda por EMEIs.

4.4 Distribuição de viagens

A fase de distribuição das viagens baseou-se no estudo de Akamine e Silva (2004), que modelaram, utilizando redes neurais artificiais, o fluxo de alunos na cidade de São Carlos, ou seja, o total de alunos que reside em um determinado setor e se desloca para uma determinada escola.

A rede que apresentou melhor desempenho na predição de fluxos do estudo teve seu funcionamento reproduzido em uma planilha, utilizada nesta pesquisa. Para utilizá-la são necessários os seguintes dados:

- Total de viagens produzidas nos pontos de origem;
- Total de viagens atraídas nos pontos de destino (capacidade da escola);
- Distância entre os pontos de origem-destino.

De posse destes dados é possível determinar o fluxo entre todas as combinações entre os setores e escolas.

4.5 Atividades Desenvolvidas

A condução da pesquisa requer a realização de várias atividades e é necessário o planejamento de sua seqüência com a finalidade de direcionar o estudo para alcançar os objetivos propostos. Estão descritas a seguir essas atividades, agrupadas em quatro grandes etapas para uma melhor compreensão do desenvolvimento dos procedimentos detalhados no próximo capítulo.

1. Preparação dos dados

- 1.1. Obtenção dos limites dos setores censitários da cidade em estudo para no mínimo três períodos;
- 1.2. Obtenção de dados de população total por setor censitário, referentes aos períodos estudados;
- 1.3. Obtenção de uma malha de células de 100 x 100 metros de lado;
- 1.4. Obtenção da densidade populacional média por setor censitário, referentes aos períodos estudados;
- 1.5. Associação à malha de células de 100 x 100 células dos valores de densidade populacional média nos anos de t , $t + 10$ e $t + 20$;
- 1.6. Seleção dos critérios a serem considerados na pesquisa considerando sua possibilidade de cálculo;
- 1.7. Cálculo dos valores dos critérios e associação ao arquivo geográfico de células;
- 1.8. Exportação dos dados do arquivo geográfico de células para o SIG ArcView;
- 1.9. Geração de arquivos de extensão *.gal através do uso da extensão SpaceStat do SIG ArcView para os campos ID e densidade populacional média nos anos de t , $t + 10$ e $t + 20$;
- 1.10. Processamento dos arquivos de extensão *.gal em planilha eletrônica para calcular os valores da variável densidade populacional média nas células vizinhas nos anos de t , $t + 10$ e $t + 20$;
- 1.11. Exportação dos valores da variável densidade populacional média nas células vizinhas nos anos de t , $t + 10$ e $t + 20$ para o SIG TransCAD;
- 1.12. Criação de um arquivo geográfico de linhas com uma linha horizontal acima das células, cálculo da distância de todas as células à linha criada e

obtenção de um índice de localização linha para todas as células, calculado dividindo-se a distância à linha criada pela dimensão da célula;

1.13. Criação de um arquivo geográfico de linhas com uma linha vertical à

criada e obtenção de um índice de localização coluna para todas as células, calculado dividindo-se a distância à linha criada pela dimensão da célula;

1.14. Construção de planilha eletrônica para avaliação de variáveis qualitativas e quantitativas;

1.15. Criação de setores para agregar os resultados de demanda obtidos na modelagem;

1.16. Cálculo da distância de cada setor criado a cada EMEI;

Observação: os passos 1.8, 1.9 e 1.10 podem ser substituídos pelos passos a seguir:

1.17. Exportação dos valores de densidades para planilha eletrônica;

1.18. Cálculo em planilha eletrônica da densidade média nas células vizinhas.

2. Modelagem da demanda com modelo de *Cellular Automata*

2.1. Exportação de todos os critérios e dos índices de localização para planilha eletrônica;

2.2. Criação de planilha para entrevistas com especialistas, com o cálculo dos pesos dos critérios analisados através de comparação par a par;

2.3. Escolha dos especialistas em Planejamento e Modelagem Urbana;

2.4. Realização de entrevistas para determinação dos pesos dos critérios;

2.5. Implementação em planilha eletrônica de modelos baseados em Avaliação Multicritério para realizar a conversão de estado não urbano para urbano;

2.6. Análise dos modelos utilizando os dados reais de $t + 10$ e $t + 20$ na planilha de avaliação de variável qualitativa e seleção do melhor modelo de CA para estimar a dispersão da mancha urbana;

2.7. Implementação em *software* de redes neurais artificiais de modelos para realizar a diferenciação da dispersão (determinação da densidade populacional das células);

2.8. Análise dos modelos utilizando os dados reais de $t + 10$ e $t + 20$ na planilha de avaliação de variável quantitativa e seleção do melhor modelo de CA para estimar a diferenciação da dispersão;

- 2.9. Aplicação do modelo para estimar o cenário de $t + 30$;
- 2.10. Agregação dos dados de distribuição de população em $t + 30$ nos setores criados;
- 2.11. Estabelecimento de uma função relacionando a população na faixa etária de 4 a 6 anos, a população total do setor e a distância ao centro;
- 2.12. Cálculo da população na faixa etária de 4 a 6 anos em $t + 30$ nos setores criados;

3. Modelagem da demanda com modelo demográfico

- 3.1. Cálculo dos valores da densidade populacional em $t + 20$ e $t + 30$ na divisão de setores de $t + 30$ para o modelo demográfico, com uma curva de tendência linear;
- 3.2. Cálculo da população em $t + 20$ e $t + 30$ com os valores de densidade populacional estimados pelo modelo demográfico;
- 3.3. Cálculo do fator de ajuste, utilizando a população em $t + 20$ e $t + 30$ obtida pelo modelo demográfico e os valores da população dos respectivos anos obtidos dos de $t + 20$ e de previsão realizada analisando a tendência demográfica global;
- 3.4. Cálculo da densidade populacional de 4 a 6 anos em $t + 30$ na divisão de setores de $t + 20$, com uma curva de tendência linear;
- 3.5. Multiplicação dos valores obtidos pelo fator de ajuste global;
- 3.6. Cálculo da população de 4 a 6 anos em $t + 30$ com os valores de densidade populacional estimados;
- 3.7. Associação dos dados do modelo demográfico à divisão espacial dos setores criados;

4. Distribuição de viagens

- 4.1. Preenchimento da coluna geração da planilha eletrônica de distribuição de viagens com os dados de demanda dos modelos
- 4.2. Preenchimento da coluna atração com a capacidade das EMEIs
- 4.3. Preenchimento da coluna distância com as distâncias calculadas de cada setor criado para cada EMEI
- 4.4. Cálculo e comparação dos resultados da distribuição de viagens realizada a partir dos dados obtidos com os dois modelos.

5 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Neste capítulo são apresentados a base de dados utilizada na pesquisa, os resultados e análise dos resultados obtidos.

5.1 Base de dados

A pesquisa estuda a cidade de São Carlos, uma cidade de médio porte com 210841 habitantes (IBGE, 2004) situada na região central do estado de São Paulo. O material básico necessário para dar início ao desenvolvimento deste trabalho inclui: o mapa digitalizado do sistema viário da cidade de São Carlos (Figura 5.1a) os limites dos setores censitários dos recenseamentos de 1980, 1991 e 2000, nos quais foram introduzidos limites de vazios urbanos conhecidos (Figura 5.1b, c e d, respectivamente), localização e capacidade das EMEIs (Figura 5.2) e dados da população total e na faixa etária de 4 a 6 anos por setor censitário. Com exceção da localização e capacidade das EMEIs, obtidas por Akamine e Silva (2004), os dados utilizados na pesquisa foram os mesmos empregados no trabalho de Lima (2003).

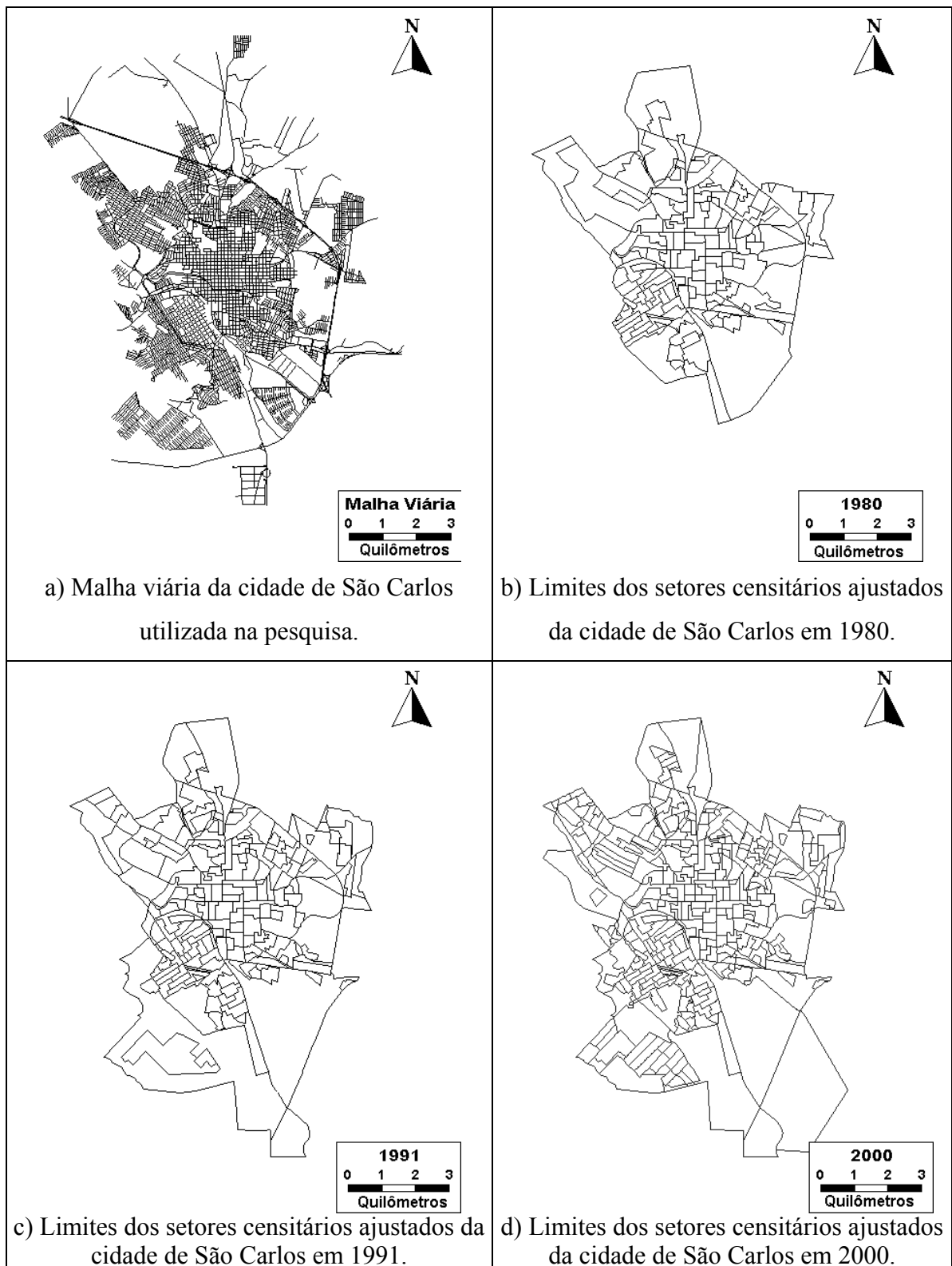


Figura 5.1 – Dados Básicos utilizados na pesquisa.

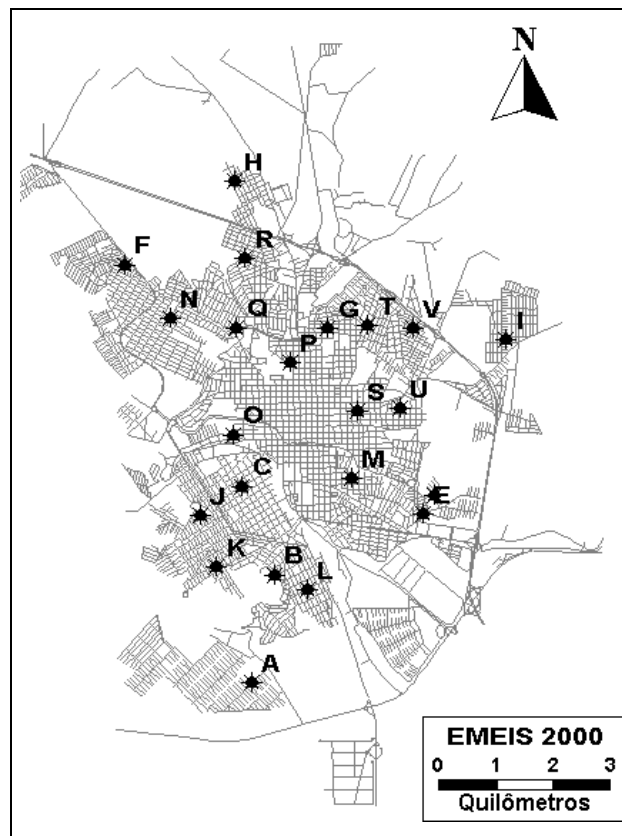


Figura 5.2 – Localização das EMEIs (2000).

Com estes dados básicos foram obtidos os mapas que delimitam o centro da cidade de São Carlos, as vias urbanas principais, a localização da rodovia Washington Luiz, que liga a cidade à São Paulo, e as zonas geradoras de alta demanda por viagens (compostas por duas universidades e um *shopping center*), apresentados na Figura 5.3.

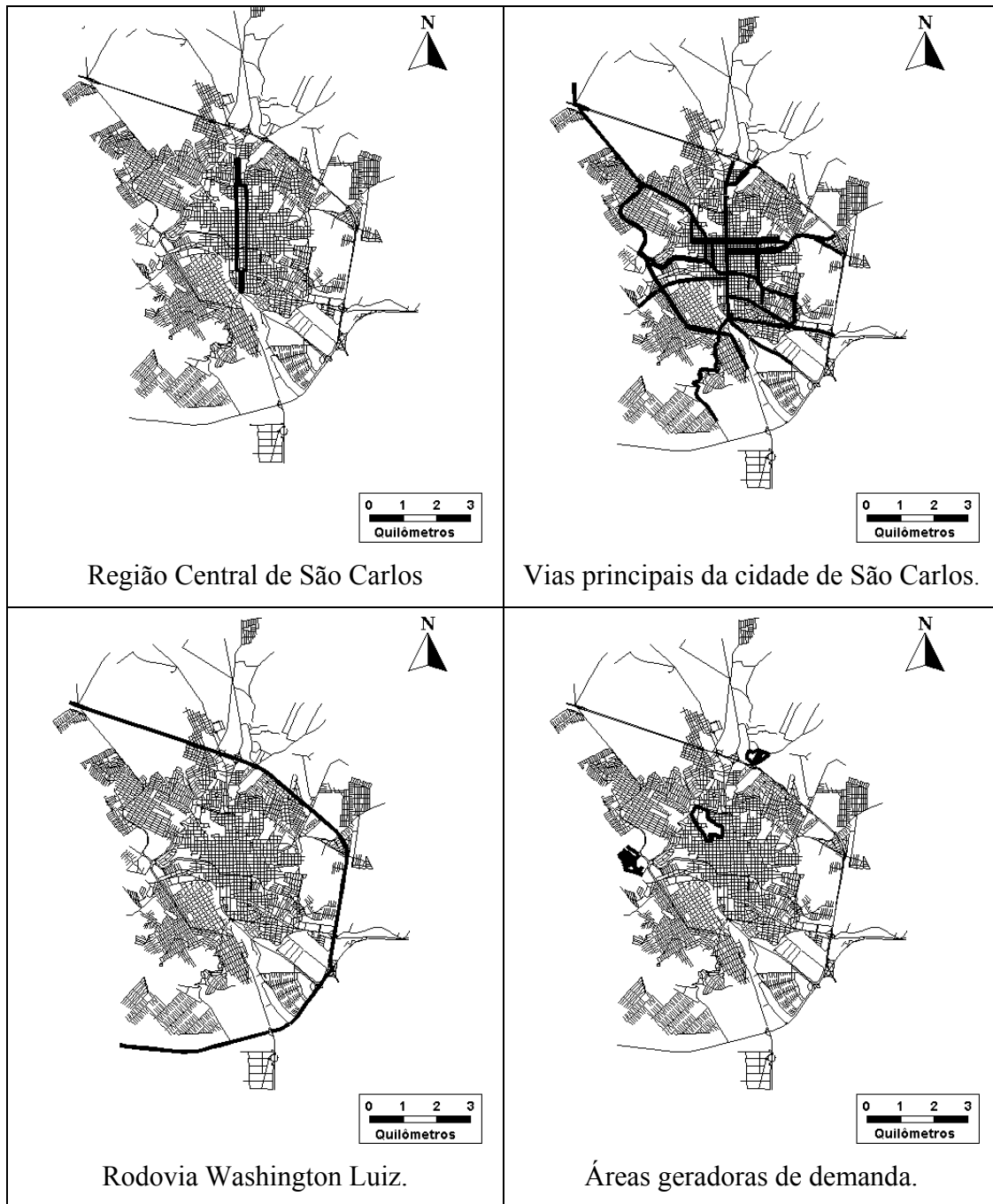


Figura 5.3 – Mapas obtidos com os dados básicos.

Foi definida também uma malha de células de 100 x 100 metros de lado (Figura 5.4) sobre toda a mancha urbana da cidade de São Carlos no ano 2000 e sobre a área periférica de até 400 metros de distância da mesma.

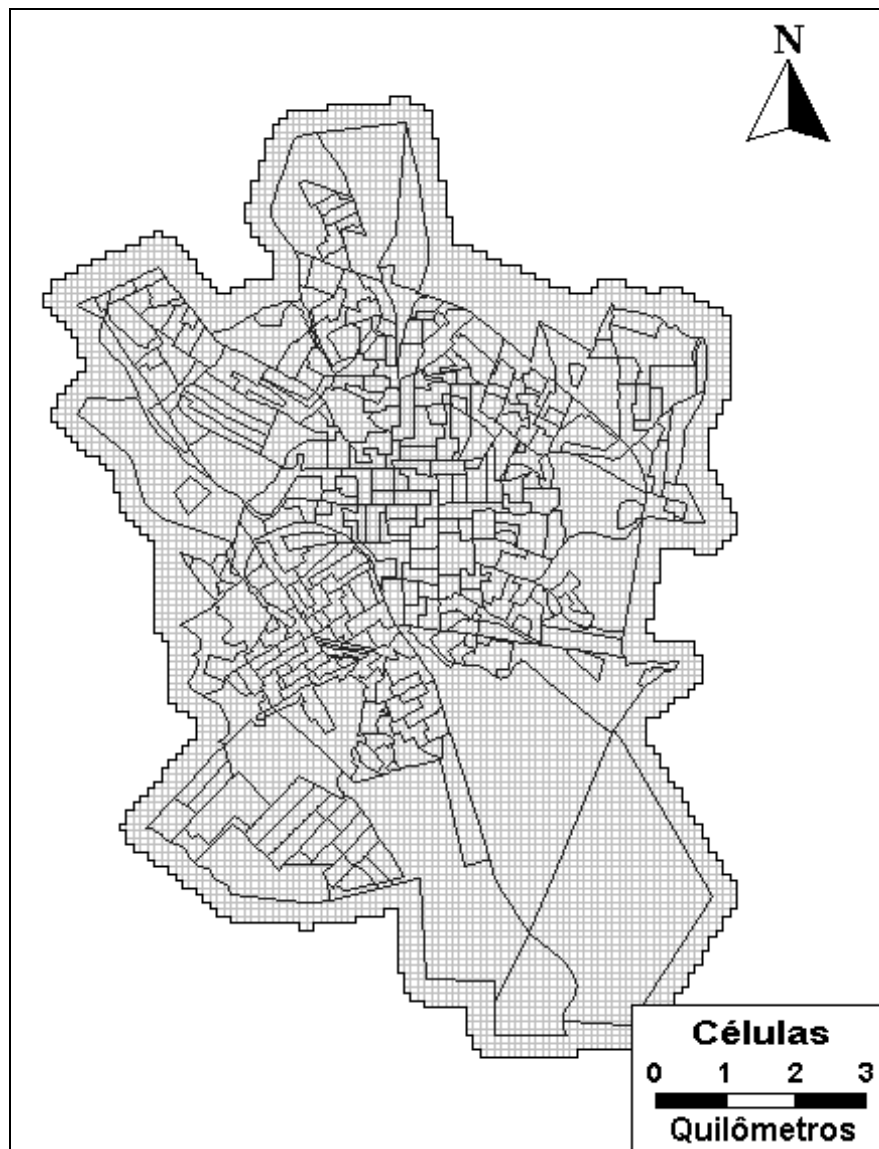


Figura 5.4 – Malha de Células utilizadas, cobrindo a mancha urbana em 2000 e uma faixa a até 400 m da mesma.

Após a obtenção dos dados básicos foram realizadas as etapas do método descrito no capítulo anterior para a estimativa da demanda por EMEIs. Os resultados obtidos são apresentados no item a seguir.

5.2 Modelagem da demanda por EMEIs com *Cellular Automata*

A modelagem da demanda por EMEIs foi realizada através da construção de um modelo da dinâmica populacional intra-urbana da cidade de São Carlos como um todo, que é posteriormente utilizado para gerar cenários da distribuição da população e da faixa etária específica de 4 a 6 anos (faixa etária que utiliza as EMEIs) para o ano de 2010.

Os componentes escolhidos para o modelo global foram:

- Geometria da rede: quadrada de dimensão 100 m x 100 m.
- Tamanho e formato da vizinhança: quadrada de tamanho 3 x 3 células.
- Condições iniciais: Distribuição espacial da densidade populacional em 1980, obtida através da divisão da população nos setores censitários por sua área (retirando os vazios urbanos).
- Estado das células: densidade populacional (hab/ha).
- Condições de fronteira: limite do conjunto de células que contém a área urbanizada de São Carlos.

Regras de transição: determinísticas e associadas à Avaliação Multicritério.

As etapas a seguir descrevem a obtenção da distribuição da população total e da demanda por EMEIs.

5.2.1 Quantificação da Dispersão

Obteve-se a matriz de probabilidade de transição baseando-se nas matrizes de transição de estado entre 1980 e 1991 (Tabela 5.1) e entre 1991 e 2000 (Tabela 5.2), através do processo de cadeias de Markov.

Tabela 5.1 – Matriz de transição de estados de 1980 a 1991.

Área (ha)	1991			
	0	1	Total	
1980	0	5448	381	5829
	1	14	3023	3037
Total		5462	3404	8866

Tabela 5.2 – Matriz de transição de estados de 1991 a 2000.

Área (ha)	2000			
	0	1	Total	
1991	0	4919	543	5462
	1	13	3391	3404
Total		4932	3934	8866

A matriz de probabilidade referente ao período é obtida através das duas matrizes de transição de estados e apresenta valores de probabilidade de transição próximos aos das matrizes de probabilidades de transição referentes a cada período considerado na pesquisa, apresentadas nas Tabelas 5.3 e 5.4.

Tabela 5.3 – Matriz de probabilidades de transição de estados referente ao período de 1980 a 1991.

	0	1
0	0,9346	0,0654
1	0,0046	0,9954

Tabela 5.4 – Matriz de probabilidades de transição de estados referente ao período de 1991 a 2000.

	0	1
0	0,9006	0,0994
1	0,0038	0,9962

Os elementos da matriz de probabilidades de transição de estados referente ao período, apresentada na Tabela 5.5, foram calculados conforme as Eq. (27) a (30):

$$P_{11} = \frac{A00_{1980-1991} + A00_{1991-2000}}{A00_{1980-1991} + A00_{1991-2000} + A01_{1980-1991} + A01_{1991-2000}} = 0,9170$$

$$P_{12} = \frac{A01_{1980-1991} + A01_{1991-2000}}{A00_{1980-1991} + A00_{1991-2000} + A01_{1980-1991} + A01_{1991-2000}} = 0,0830$$

$$P_{21} = \frac{A10_{1980-1991} + A10_{1991-2000}}{A10_{1980-1991} + A10_{1991-2000} + A11_{1980-1991} + A11_{1991-2000}} = 0,0042$$

$$P_{22} = \frac{A11_{1980-1991} + A11_{1991-2000}}{A10_{1980-1991} + A10_{1991-2000} + A11_{1980-1991} + A11_{1991-2000}} = 0,9958$$

Tabela 5.5 – Matriz de probabilidades de transição de estados referente ao período de 1980 a 2000.

	0	1
0	0,9170	0,0830
1	0,0042	0,9958

De posse das probabilidades de transição de estados são obtidos os elementos da matriz de transição de estados de 2000 a 2010, apresentados na Tabela 5.6 e calculados conforme as Eq. (31) a (34):

$$A00_{2000-2010} = P_{11} * T0_{2000} = 0,9170 * 4932$$

$$A01_{2000-2010} = P_{12} * T0_{2000} = 0,0830 * 4932$$

$$A10_{2000-2010} = P_{21} * T1_{2000} = 0,0042 * 3934$$

$$A11_{2000-2010} = P_{22} * T1_{2000} = 0,9958 * 3934$$

Tabela 5.6 – Matriz de transição de estados de 2000 a 2010.

Área (ha)	2010			Total
	0	1		
2000	0	4528	404	4932
	1	17	3917	3934
Total		4545	4321	8866

Nesta pesquisa, a transição do estado 1 (um) para o estado 0 (zero) foi desprezada e considerou-se como dispersão a variação da área ocupada (área

no estado 1) entre os períodos considerados na pesquisa, apresentada na Tabela 5.7.

Tabela 5.7 – Área da mancha urbana de São Carlos.

Ano	Área Ocupada (ha)	Variação de área (ha)
1980	3037	-
1991	3404	367
2000	3934	530
2010	4321	387

5.2.2 Localização da dispersão

Os itens a seguir descrevem os passos realizados na fase de localização da dispersão.

5.2.2.1 Estabelecimento dos critérios

Os critérios utilizados na Avaliação Multicritério foram:

- Distância ao centro da cidade (dc).
- Densidade populacional média nas células vizinhas (mv).
- Distância à rodovia Washington Luiz (dw).
- Distância às vias principais (dv).
- Distância à mancha urbana (do).
- Distância a pólos geradores de demanda por viagens, como *shopping centers* e universidades (dd).
- Adequabilidade à ocupação.

Os critérios acima estabelecidos foram selecionados levando-se em conta a possibilidade efetiva de cálculo de seus valores nas células. São critérios do tipo fator, com exceção do critério de adequabilidade à ocupação, que é um critério de exclusão.

A partir dos dados de população, dos valores da área das células e dos mapas com o sistema viário e localização do centro da cidade e de zonas geradoras de alta demanda, utilizando-se de funções próprias do SIG foram calculados os valores dos critérios.

O cálculo do valor do critério densidade média populacional nas células vizinhas foi realizado por Lima (2003) com os dados de densidade demográfica, calculados dividindo-se a população em cada setor do IBGE pela sua área e retirando-se a área de alguns vazios urbanos conhecidos. Os valores, obtidos no SIG-T TransCAD, foram transferidos para o SIG Arcview, onde,

através da extensão SpaceStat, foram gerados arquivos que listavam o valor da densidade de cada célula e de suas vizinhas. Estes arquivos foram processados em planilha eletrônica, onde foi calculada a densidade média nas vizinhas de cada células. Os valores obtidos, associados a cada célula, foram novamente transferidos para armazenagem no SIG-T TransCAD.

As outras variáveis foram obtidas através de ferramentas de cálculo de distância do SIG-T TransCAD, utilizando como dados a localização do centro da cidade, o traçado geométrico das vias principais e da rodovia e os limites da área dos setores censitários, dos quais foram excluídos os vazios urbanos.

O critério de adequabilidade à ocupação baseou-se no estudo de Aguiar (1989), que sugere áreas geotecnicamente homogêneas, indicando aquelas potenciais ou problemáticas para ocupação. É realizado para este fim um estudo em escala 1:25000 dos aspectos geomorfológicos, pedológicos e climatológicos, da vegetação e uso do solo, legislação ambiental, geologia e hidrologia da região. É produzida, neste estudo, entre outros resultados (cartas de erodibilidade, escavabilidade, para obras viárias, etc.), uma carta para ser utilizada no planejamento urbano que define áreas propícias para ocupação residencial, industrial e rural (agricultura, pecuária, extrativismo vegetal, etc.) e delimita áreas de proteção ambiental ou inadequadas para ocupação. Estas últimas foram utilizadas no estudo para retirar as áreas não adequadas à ocupação das opções na Avaliação Multicritério através da associação ao número 0, destacadas na Figura 5.5.

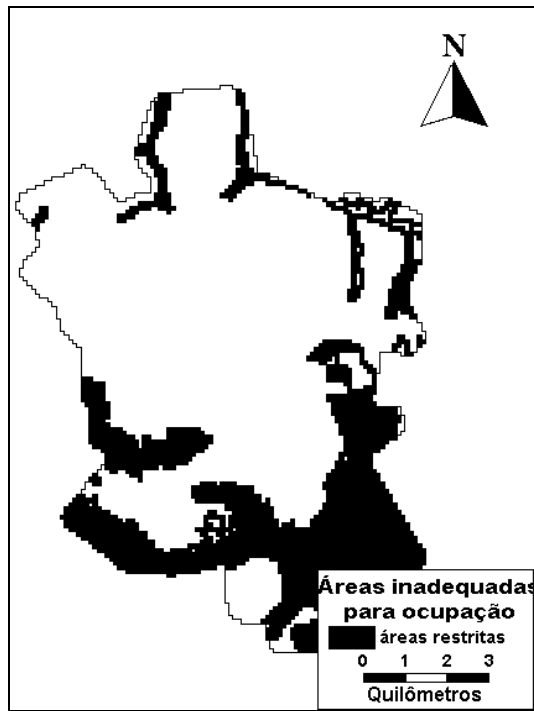


Figura 5.5 – Adequabilidade à ocupação (Adaptado de Aguiar, 1989).

5.2.2.2 Avaliação de pesos para critérios

Foi realizada uma busca na Plataforma Lattes por pesquisadores da área de Planejamento Urbano e Regional que abordassem o assunto modelagem urbana. Dos nomes resultantes foram selecionados 27 pesquisadores que tinham qualificação na área, os quais foram convidados a participar da pesquisa através de mensagens de correio eletrônico.

Para a obtenção de pesos foi adotado o método de comparações par a par, realizadas através de matrizes em planilha eletrônica. Dos 27 avaliadores, 8 realizaram as avaliações, o que resultou na criação de 8 matrizes distintas. Cada matriz só poderia ser utilizada na obtenção de pesos se apresentasse índice de consistência inferior a 0,1, critério que foi obedecido em todas as avaliações.

Após a seleção dos especialistas, matrizes de comparação par a par, (apresentadas no Anexo 1), preenchidas pelos mesmos foram utilizadas para obter os valores de pesos e graus de consistência apresentados na Tabela 5.8.

Tabela 5.8 – Pesos atribuídos aos critérios em cada avaliação.

Avaliador	CR	w_{dc}	w_{mv}	w_{dw}	w_{dv}	w_{do}	w_{dd}
1	0,094	0,077	0,222	0,085	0,375	0,180	0,063
2	0,073	0,240	0,036	0,175	0,093	0,060	0,396
3	0,098	0,490	0,042	0,053	0,132	0,118	0,164
4	0,081	0,490	0,234	0,030	0,072	0,074	0,100
5	0,028	0,039	0,039	0,084	0,165	0,337	0,337
6	0,098	0,353	0,228	0,085	0,165	0,120	0,049
7	0,069	0,140	0,454	0,029	0,053	0,246	0,077
8	0,081	0,137	0,083	0,045	0,457	0,251	0,028

Não se nota um padrão nos pesos, o que gerou modelos com grande diversidade de cenários, apesar do número reduzido de especialistas.

5.2.2.3 Normalização de Critérios

Na etapa seguinte foi realizada a normalização dos critérios para o mesmo intervalo com o objetivo de combiná-los posteriormente. Foi realizada a calibração do modelo com dados dos censos gerais do IBGE referentes aos anos de 1980 e 1991, sendo este último utilizado posteriormente para a projeção de cenários, que foram avaliados com os dados reais do censo de 2000.

Como descrito no método, na fase da normalização foram testadas quatro funções de normalização linear para cada variável: uma exclusivamente crescente, outra exclusivamente decrescente e duas crescentes e decrescentes, cujas equações são apresentadas na Figura 3.4.

Para as funções exclusivamente crescentes e exclusivamente decrescentes foram utilizados os pontos de controle apresentados na Tabela 5.9, que representam os valores máximos observados em cada variável.

Tabela 5.9 – Ponto de controle das funções de normalização exclusivamente crescentes ou decrescentes.

Critério	a
dc	6,79 km
mv	130,00 hab/ha
dw	2,94 km
dv	4,67 km
do	4,83 km
dd	9,03 km

A fase de normalização com funções crescentes e decrescentes teve início com a observação do número de ocorrências da transição não ocupado para ocupado em relação aos diferentes valores das variáveis.

A análise se revela particularmente importante para variáveis de distância (como distância ao centro, por exemplo) para as quais se deseja determinar a distância mínima e máxima observando-se a transição. No caso deste

tipo de variável, se é adotada uma função normalizadora exclusivamente crescente, observa-se que são selecionadas as alternativas mais distantes da mancha urbana (e essa distância é dependente do número de células acrescentadas na periferia da mancha urbana). Se, por outro lado, é adotada uma função normalizadora exclusivamente decrescente, o cenário obtido escolhe as alternativas mais próximas, o que nem sempre ocorre no espalhamento urbano real das cidades.

A análise das variáveis para normalização, realizada em planilha eletrônica, teve início com a seleção, nos dados, daquelas células que sofreram a transição de não ocupado para ocupado no período de 1980 a 1991. Nesta parcela dos dados foi observado o número de transições por faixa de valores de cada variável.

Observou-se, para cada variável, três valores: o valor mínimo e máximo da variável que delimitavam o intervalo em que ocorria a maioria das transições não-ocupado para ocupado e o valor da variável que apresentava maior frequência, à qual foi atribuído valor normalizado 1. Estes valores, obtidos através de um ajuste visual, foram adotados como os pontos de controle *a*, *b* e *c* para as funções de normalização crescentes e decrescentes do tipo 1, conforme equação apresentada na Figura 3.4.

Para as funções crescentes e decrescentes observou-se o número de ocorrências da transição não ocupado para ocupado em relação aos diferentes valores das variáveis, normalizando-os posteriormente entre zero e um, através da sua divisão pelo valor máximo observado na variável.

Os valores mínimo e máximo da variável que delimitavam o intervalo em que ocorria o maior número significativo de transições não-ocupado para ocupado e o valor da variável que apresentava maior frequência foram adotados como os pontos de controle *a*, *b* e *c* para as funções de normalização crescentes e decrescentes do tipo 1, e são apresentados na Tabela 5.10.

Tabela 5.10 – Pontos de controle da função de normalização crescente e decrescente do tipo 1.

Critério	a	b	c	Número de ocorrências máximo
dc	1,7	3,5	4,5	41
dw	0	0,6	2,1	59
dv	0	1,3	2,5	38
do	0	0,1	2,2	34
dd	0,8	4,2	5,3	25

As Figuras 5.6 a 5.10 apresentam os gráficos das funções de normalização do tipo 1.

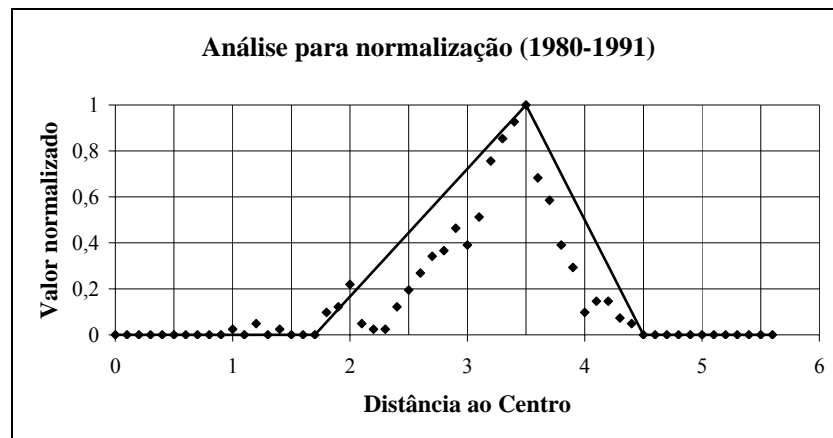


Figura 5.6 – Função de normalização do tipo 1 para o critério distância ao centro.

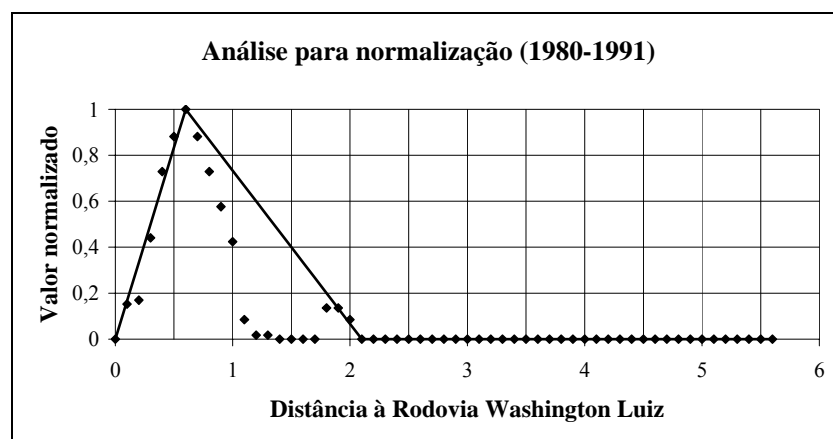


Figura 5.7 – Função de normalização do tipo 1 para o critério distância à rodovia Washington Luiz.

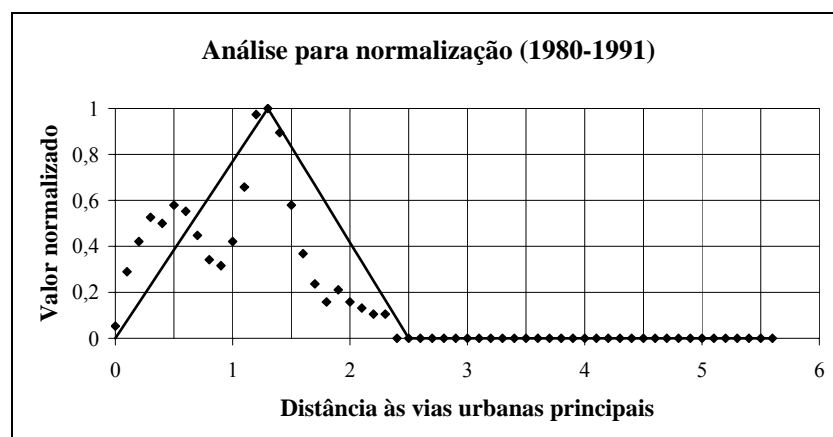


Figura 5.8 – Função de normalização do tipo 1 para o critério distância às vias urbanas principais.

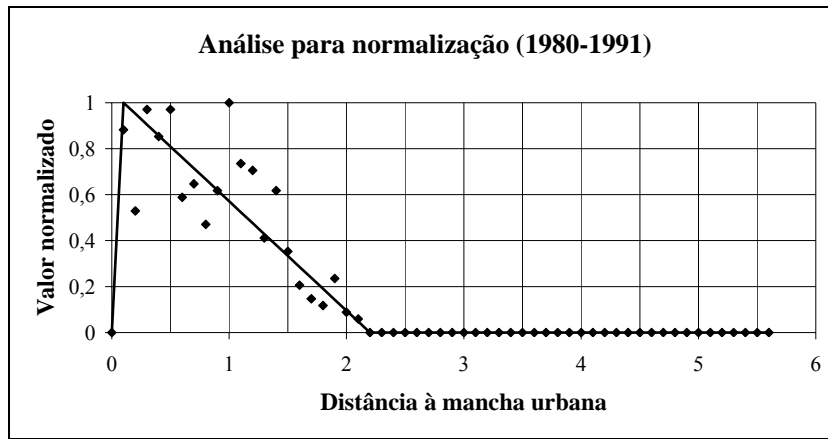


Figura 5.9 – Função de normalização do tipo 1 para o critério distância à mancha urbana.

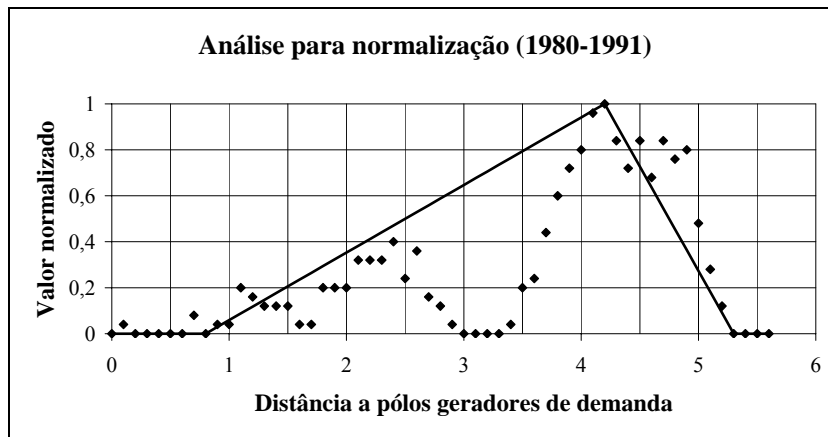


Figura 5.10 – Função de normalização do tipo 1 para o critério distância a pólos geradores de demanda.

Para as funções de normalização crescentes e decrescentes do tipo 2 foram utilizados como pontos de controle os valores máximo e mínimo dos critérios, além do ponto de controle b , idêntico ao das funções de normalização do tipo 1. Os pontos de controle são apresentados na Tabela 5.11 a seguir.

Tabela 5.11 – Pontos de controle da função de normalização crescente e decrescente do tipo 2.

Critério	a	b	c
dc	0	3,5	6,79
dw	0	0,6	2,94
dv	0	1,3	4,67
do	0	0,1	4,83
dd	0	4,2	9,03

As Figuras 5.11 a 5.15 apresentam os gráficos das funções de normalização do tipo 2.

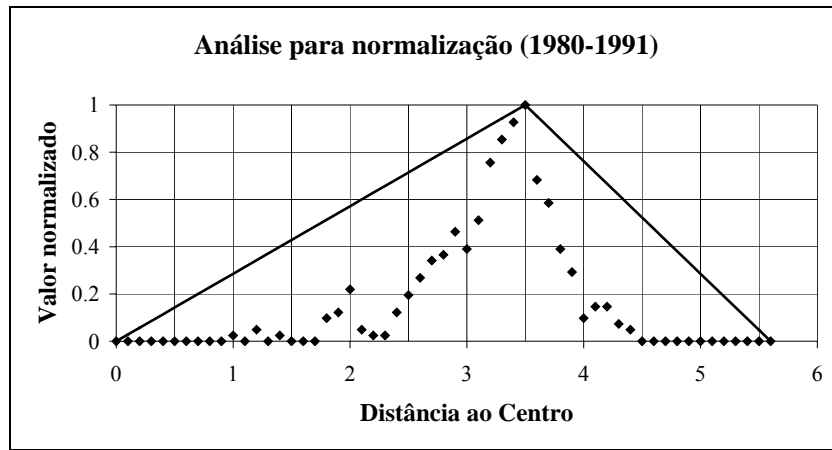


Figura 5.11 – Função de normalização do tipo 2 para o critério distância ao centro.

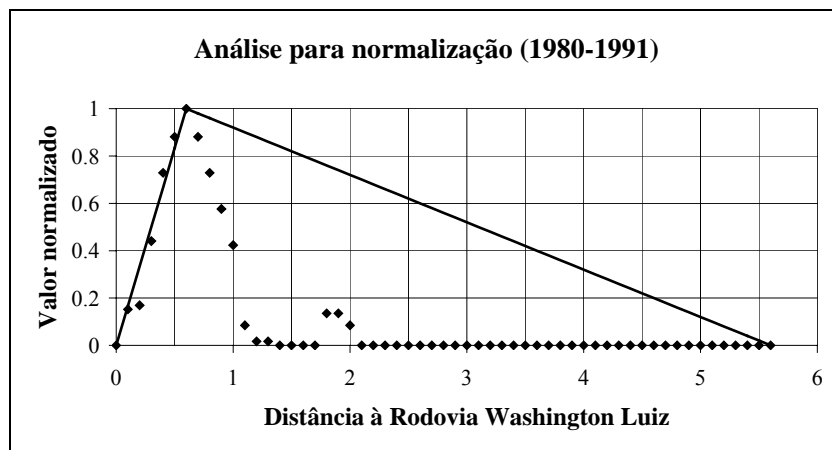


Figura 5.12 – Função de normalização do tipo 2 para o critério distância à rodovia Washington Luiz.

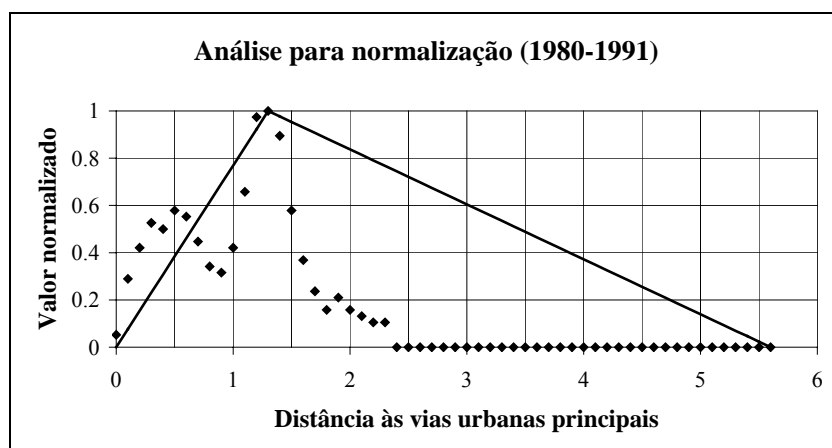


Figura 5.13 – Função de normalização do tipo 2 para o critério distância às vias urbanas principais.

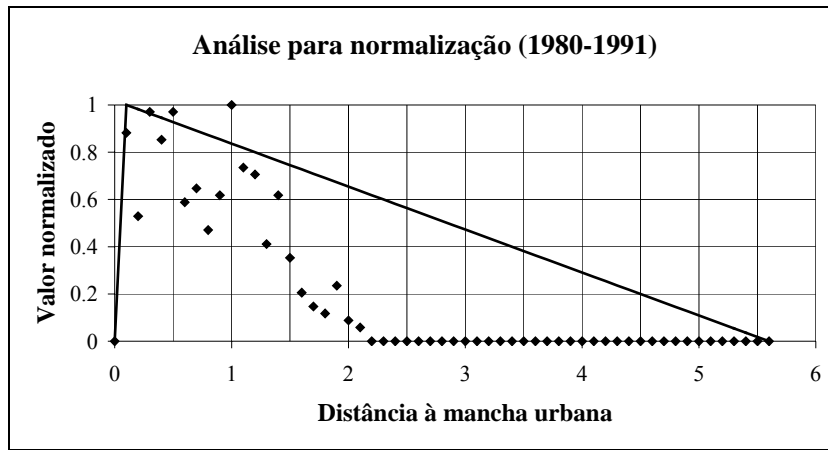


Figura 5.14 – Função de normalização do tipo 2 para o critério distância à mancha urbana.

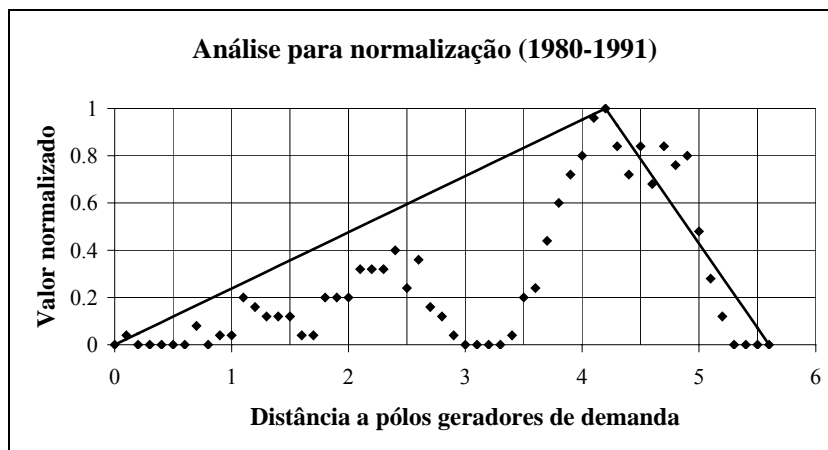


Figura 5.15 – Função de normalização do tipo 2 para o critério distância a pólos geradores de demanda.

Excepcionalmente para a variável densidade populacional média nas células vizinhas, devido à grande ocorrência de células do valor 0 (92,38 % do total), adotou-se uma função linear exclusivamente decrescente, de valor máximo 1 e valor mínimo 0 calculada conforme a Eq. (39):

$$n(mv) = 1 - \frac{mv_{xy}}{máx(mv_{xy})} \quad (39)$$

Onde $n(mv)$ é o valor normalizado da variável densidade média dos vizinhos, mv_{xy} é o valor da variável na célula e $máx(mv_{xy})$ é o máximo valor de mv_{xy} encontrado no conjunto de dados de 1980. Para os dados em que $mv_{xy} > m_{máx}$ adotou-se $n(mv) = 0$. O valor de $máx(mv_{xy})$ utilizado foi de 130,00 hab/ha.

5.2.2.4 Combinação de Critérios

Para a combinação de critérios foi utilizado o procedimento Combinação Linear Ponderada, no qual o escore de cada célula, que representa a adequabilidade da célula para o uso residencial foi obtido segundo a Eq. (39):

$$r_{ij} = \pi_{ad} (w_{dc} \times n(dc) + w_{mv} \times n(mv) + w_{dw} \times n(dw) + w_{dv} \times n(dv) + w_{do} \times n(do) + w_{dd} \times n(dd)) \quad (39)$$

Sendo w_{dc} , w_{mv} , w_{dw} , w_{dv} , w_{do} e w_{dd} os pesos obtidos pelos critérios to tipo fator na Avaliação Multicritério de cada avaliador; $n(dc)$, $n(mv)$, $n(dw)$, $n(dv)$, $n(do)$ e $n(dd)$, seus valores normalizados e π_{ad} o valor do critério de adequabilidade à ocupação para a célula.

5.2.2.5 Aplicação das Regras de Transição

Após realizar o cálculo dos valores de r_{ij} resultantes de cada avaliação foi realizada a conversão de uso não ocupado para ocupado para as células cujo escore fosse maior que o escore limite (r_{lim}), determinado de tal modo que o número de células resultante fosse o mais próximo possível do valor n , determinado na fase de quantificação da dispersão. Os valores dos escores limites e do número de conversões são apresentados na Tabela 5.12, na qual não se consegue observar um padrão no valor limite do escore, o que possibilitaria mais uma alternativa para a realização da quantificação da dispersão.

Os modelos são representados por um código que indica o tipo de função de normalização utilizada (F1 para função crescente e decrescente do tipo 1, F2 para funções crescentes e decrescentes do tipo 2, F3 para funções exclusivamente crescentes e F4 para funções exclusivamente decrescente) e o avaliador que atribuiu os pesos utilizados no modelo (de 1 a 8).

Tabela 5.12 – Escore limite e número de conversões dos modelos.

modelo	r_{lim} 1991	n 1991	r_{lim} 2000	n 2000	r_{lim} 2010	n 2010
real		3404		3934		
F1A1	0,7971	3403	0,7616	3934	0,7584	4321
F1A2	0,8118	3404	0,7500	3934		
F1A3	0,8490	3404	0,8206	3934		
F1A4	0,8529	3404	0,8033	3934		
F1A5	0,7517	3404	0,6318	3934		
F1A6	0,7971	3403	0,7615	3934	0,7584	4321
F1A7	0,7830	3404	0,7299	3934		
F1A8	0,7829	3403	0,7503	3934		
F2A1	0,8850	3404	0,8607	3934		
F2A2	0,8779	3404	0,8261	3934		
F2A3	0,9032	3404	0,8771	3934		
F2A4	0,9153	3404	0,8891	3934	0,8889	4321
F2A5	0,8357	3405	0,8010	3934		
F2A6	0,8850	3404	0,8606	3934		
F2A7	0,8821	3404	0,8550	3934	0,8518	4321
F2A8	0,8798	3404	0,8564	3934		
F3A1	0,4535	3404	0,3395	3934		
F3A2	0,5336	3405	0,2760	3934		
F3A3	0,6813	3405	0,1728	3934		
F3A4	0,7206	3404	0,2160	3934		
F3A5	0,3861	3404	0,2341	3934		
F3A6	0,6310	3404	0,2565	3934		
F3A7	0,5286	3404	0,3531	3934		
F3A8	0,4630	3404	0,2815	3934		
F4A1	0,8650	3404	0,8458	3934		
F4A2	0,8670	3404	0,8449	3934		
F4A3	0,9227	3404	0,9107	3934		
F4A4	0,8921	3404	0,8789	3934		
F4A5	0,8882	3404	0,8626	3934		
F4A6	0,8793	3404	0,8655	3934		
F4A7	0,8418	3404	0,8153	3934		
F4A8	0,9197	3404	0,9025	3934		

Seguindo o método, cenários de localização da dispersão foram determinados e comparados aos cenários reais de ocupação. As Tabelas 5.13 e 5.14 apresentam os valores do coeficiente Fit obtidos pelos modelos nas fases de calibração e validação, para os tamanhos de janela de 1 a 5, além do coeficiente F_t , utilizado para comparar os modelos. Os valores foram ordenados de acordo com o valor médio do coeficiente Fit combinado (F_i). Observando as tabelas, é possível notar que os modelos que utilizam funções exclusivamente decrescentes (F4) são os que apresentam os piores resultados, seguidos pelos modelos de funções exclusivamente crescentes (F3). Os modelos de funções crescentes e decrescentes apresentaram os melhores resultados, alternando-se funções com pontos de contorno nos valores extremos das variáveis (F2) e em intervalos menores calibrados com os dados de 1980 a 1991 (F1).

Tabela 5.13 – Valores do coeficiente Fit obtidos na fase de calibração (estimativa de valores para 1991).

Modelo	Calibração					F_t	F_t médio
	Fit 1	Fit 2	Fit 3	Fit 4	Fit 5		
F2A7	0,8905	0,9072	0,9410	0,9422	0,9438	0,9325	0,9141
F2A4	0,8822	0,8997	0,9389	0,9386	0,9391	0,9268	0,9120
F1A1	0,8866	0,9031	0,9382	0,9378	0,9387	0,9279	0,9112
F1A6	0,8866	0,9031	0,9382	0,9378	0,9387	0,9279	0,9112
F1A4	0,8854	0,9013	0,9367	0,9366	0,9376	0,9265	0,9109
F2A1	0,8795	0,8986	0,9374	0,9390	0,9409	0,9276	0,9086
F2A6	0,8795	0,8986	0,9374	0,9390	0,9409	0,9276	0,9086
F2A3	0,8800	0,8981	0,9369	0,9392	0,9417	0,9281	0,9045
F1A7	0,8667	0,8855	0,9290	0,9283	0,9283	0,9149	0,9045
F1A3	0,8757	0,8944	0,9350	0,9354	0,9360	0,9225	0,9015
F2A5	0,8738	0,8926	0,9357	0,9371	0,9384	0,9237	0,8988
F2A2	0,8787	0,8964	0,9388	0,9420	0,9447	0,9290	0,8982
F2A8	0,8536	0,8772	0,9281	0,9309	0,9332	0,9148	0,8980
F1A8	0,8539	0,8773	0,9272	0,9293	0,9309	0,9131	0,8966
F1A2	0,8770	0,8946	0,9355	0,9373	0,9393	0,9248	0,8958
F1A5	0,8583	0,8801	0,9266	0,9264	0,9268	0,9113	0,8913
F3A4	0,8080	0,8378	0,9109	0,9177	0,9231	0,8961	0,8853
F3A3	0,8075	0,8365	0,9116	0,9183	0,9236	0,8961	0,8840
F3A5	0,8119	0,8391	0,9132	0,9199	0,9255	0,8985	0,8832
F3A7	0,8079	0,8372	0,9112	0,9181	0,9236	0,8962	0,8823
F3A1	0,8041	0,8334	0,9089	0,9155	0,9207	0,8931	0,8821
F3A8	0,8044	0,8330	0,9091	0,9157	0,9209	0,8931	0,8816
F3A6	0,8030	0,8330	0,9090	0,9150	0,9200	0,8930	0,8816
F3A2	0,8027	0,8316	0,9088	0,9156	0,9211	0,8930	0,8811
F4A3	0,8020	0,8243	0,8738	0,8733	0,8748	0,8592	0,8558
F4A2	0,8087	0,8290	0,8769	0,8772	0,8792	0,8641	0,8542
F4A4	0,8054	0,8264	0,8711	0,8716	0,8739	0,8595	0,8489
F4A8	0,8025	0,8222	0,8545	0,8570	0,8614	0,8509	0,8475
F4A6	0,8011	0,8224	0,8647	0,8649	0,8677	0,8546	0,8474
F4A5	0,8025	0,8219	0,8644	0,8658	0,8688	0,8551	0,8467
F4A1	0,8011	0,8187	0,8536	0,8563	0,8608	0,8494	0,8445
F4A7	0,8025	0,8203	0,8585	0,8615	0,8656	0,8527	0,8427

Tabela 5.14 – Valores do coeficiente Fit obtidos na fase de validação (estimativa de valores para 2000).

Modelo	Validação					F_t	F_t médio
	Fit 1	Fit 2	Fit 3	Fit 4	Fit 5		
F2A7	0,8190	0,8492	0,9083	0,9127	0,9176	0,8957	0,9141
F2A4	0,8325	0,8597	0,9075	0,9104	0,9141	0,8971	0,9120
F1A1	0,8253	0,8532	0,9064	0,9097	0,9131	0,8944	0,9112
F1A6	0,8253	0,8532	0,9064	0,9097	0,9131	0,8944	0,9112
F1A4	0,8306	0,8559	0,9073	0,9100	0,9128	0,8952	0,9109
F2A1	0,8135	0,8452	0,9039	0,9071	0,9110	0,8896	0,9086
F2A6	0,8135	0,8452	0,9039	0,9071	0,9110	0,8896	0,9086
F2A3	0,7991	0,8320	0,8978	0,9009	0,9044	0,8809	0,9045
F1A7	0,8319	0,8569	0,9049	0,9069	0,9100	0,8940	0,9045
F1A3	0,8111	0,8390	0,8969	0,8989	0,9008	0,8805	0,9015
F2A5	0,7971	0,8280	0,8914	0,8935	0,8961	0,8738	0,8988
F2A2	0,7893	0,8186	0,8865	0,8891	0,8914	0,8674	0,8982
F2A8	0,7963	0,8298	0,8975	0,9021	0,9061	0,8811	0,8980
F1A8	0,8030	0,8334	0,8951	0,8987	0,9018	0,8800	0,8966
F1A2	0,7929	0,8218	0,8840	0,8859	0,8883	0,8668	0,8958
F1A5	0,8084	0,8338	0,8873	0,8880	0,8896	0,8713	0,8913
F3A4	0,7722	0,8083	0,8898	0,8979	0,9046	0,8745	0,8853
F3A3	0,7685	0,8029	0,8872	0,8958	0,9029	0,8718	0,8840
F3A5	0,7607	0,7947	0,8874	0,8954	0,9018	0,8678	0,8832
F3A7	0,7611	0,7983	0,8868	0,8947	0,9010	0,8684	0,8823
F3A1	0,7638	0,8006	0,8895	0,8973	0,9038	0,8711	0,8821
F3A8	0,7634	0,8002	0,8888	0,8965	0,9031	0,8701	0,8816
F3A6	0,7622	0,7986	0,8885	0,8966	0,9032	0,8701	0,8816
F3A2	0,7662	0,8011	0,8860	0,8939	0,9004	0,8691	0,8811
F4A3	0,7910	0,8149	0,8655	0,8645	0,8661	0,8524	0,8558
F4A2	0,7769	0,8015	0,8606	0,8608	0,8619	0,8442	0,8542
F4A4	0,7737	0,7977	0,8519	0,8519	0,8539	0,8383	0,8489
F4A8	0,7894	0,8093	0,8486	0,8506	0,8546	0,8441	0,8475
F4A6	0,7785	0,8017	0,8526	0,8533	0,8561	0,8401	0,8474
F4A5	0,7737	0,7977	0,8519	0,8519	0,8539	0,8383	0,8467
F4A1	0,7833	0,8036	0,8462	0,8484	0,8525	0,8396	0,8445
F4A7	0,7737	0,7945	0,8416	0,8435	0,8474	0,8326	0,8427

5.2.2.6 Escolha dos melhores modelos globais de localização

Foram escolhidos os três modelos que apresentaram os maiores valores do coeficiente F_t médio, para serem utilizados na fase de diferenciação da dispersão. Os modelos escolhidos (F1A1, F1A6, F2A4 e F2A7, sendo que F1A1 e F1A6 geraram cenários idênticos). Todos os modelos escolhidos apresentaram altos valores do coeficiente Fit, acima de 0,9100.

Os cenários gerados pelos modelos são apresentados nas Figuras 5.16 e 5.17.

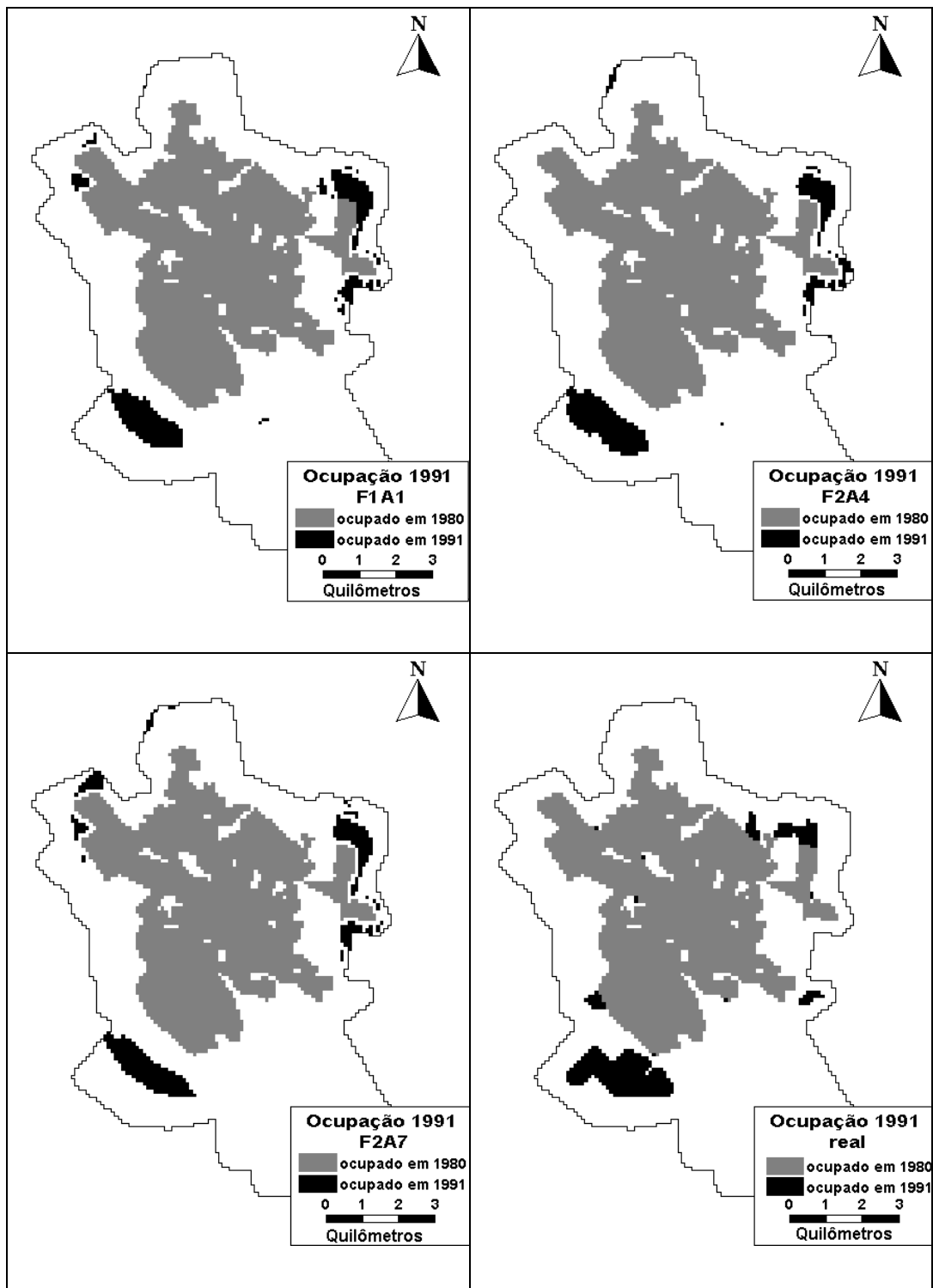


Figura 5.16 – Cenários de ocupação gerados pelos modelos e real para 1991.

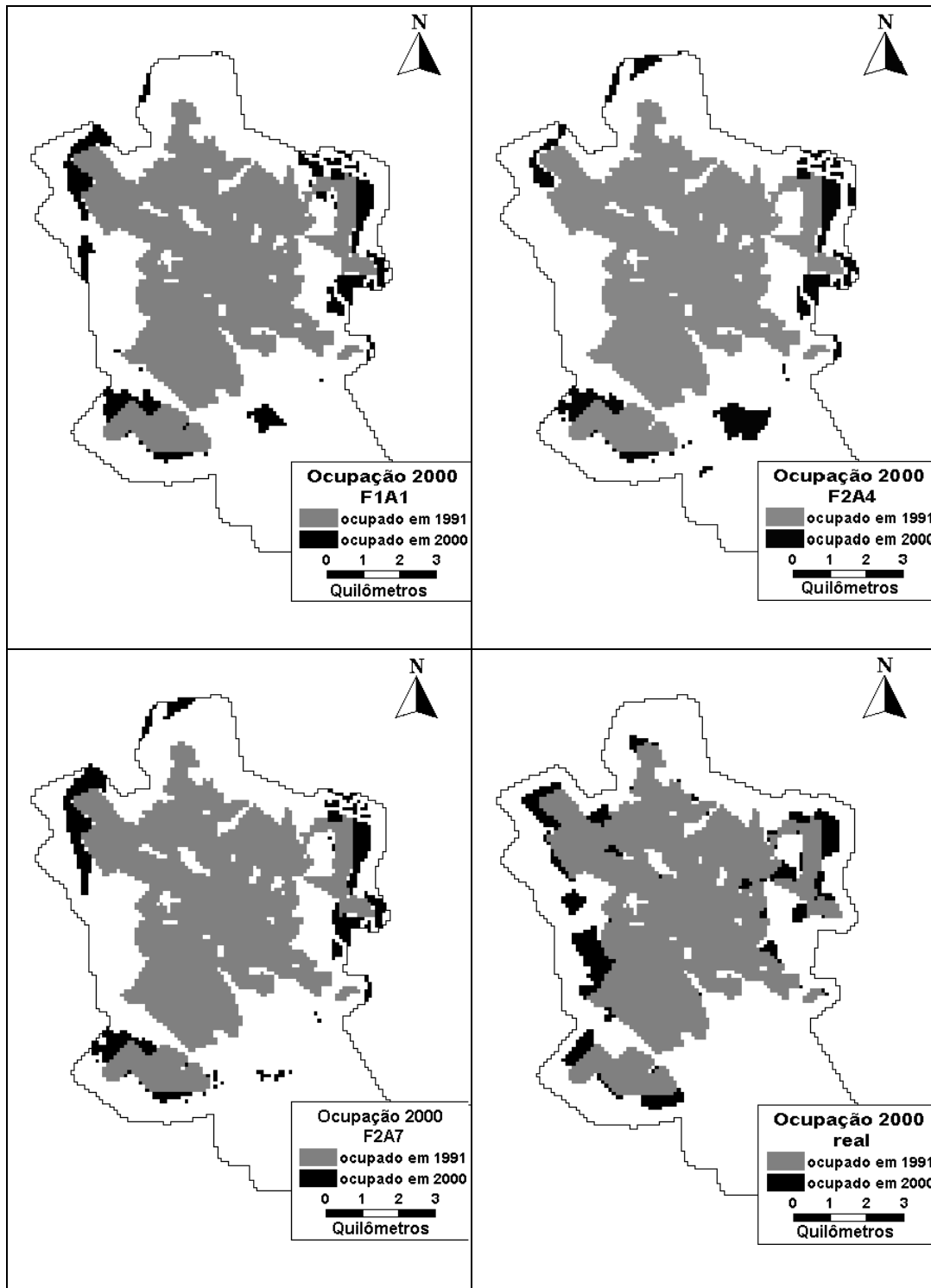


Figura 5.17 – Cenários de ocupação gerados pelos modelos e real para 2000.

É possível notar que os modelos conseguem identificar bem as regiões de ocupação a nordeste e sudoeste da cidade em 1991 e todos os modelos detectaram a existência das novas ocupações a noroeste da cidade. Nota-se ainda que algumas das células escolhidas para ocupação estão no limite do conjunto de células, o que indica que o conjunto de células vazias na periferia pode ser pequeno para a modelagem. Observa-se também que em alguns casos os modelos

escolhem células próximas à áreas ocupadas, porém não contíguas às mesmas decorrente do uso de uma função crescente e decrescente no critério distância à mancha urbana existente. Provavelmente o uso de uma função exclusivamente decrescente na normalização deste critério específico, a exemplo do realizado com o critério densidade média das células vizinhas, eliminaria esta característica dos modelos.

5.2.3 Diferenciação da Dispersão

No *software* de redes neurais artificiais EasyNNPlus foram calibradas quatro redes neurais artificiais, utilizando como dados de entrada os valores dos critérios densidade populacional média no instante inicial, densidade populacional média dos vizinhos no instante inicial, distância ao centro, à rodovia Washington Luiz, às vias principais, à mancha urbana e às zonas geradoras de demanda, e como dado de saída a densidade populacional média no instante final. Todos os valores foram normalizados entre 0 e 1 dividindo-os pelo seu valor máximo. Os parâmetros das redes utilizadas, assim como o erro associado a cada uma, calculado conforme a Eq. (38) são apresentados na Tabela 5.15.

Tabela 5.15 – Parâmetros e erros das redes neurais.

Rede	Taxa de aprendizado	Momentum	Número de ciclos	Erro 1991	Erro 2000	Erro médio
1	0,6	0,4	300	0,1044	0,1348	0,1196
2	0,4	0,4	300	0,1069	0,1345	0,1207
3	0,6	0,4	500	0,1051	0,1263	0,1157
4	0,4	0,4	500	0,0903	0,1345	0,1124

Utilizou-se a rede 4 para obtenção dos cenários para os melhores modelos de localização da dispersão por apresentar o menor valor médio do erro.

Após a seleção da rede, esta foi utilizada com dados de entrada referentes a 1980 (apenas os dados das células escolhidas pelos melhores modelos para ser ocupadas) e forneceu como saída valores normalizados referentes ao ano de 1991. Estes valores normalizados foram multiplicados por uma constante, de tal maneira que a população total nas células fosse a população referente ao ano de 1991. Às células não selecionadas na fase de localização da dispersão era atribuído valor 0 de densidade populacional. Este procedimento foi repetido com os dados relativos ao ano 2000. Foram obtidos oito cenários para os melhores modelos, que foram avaliados novamente para obtenção do melhor de todos. Os cenários de densidades populacionais obtidos pelos modelos, assim como os reais são apresentados nas Figuras 5.18 e 5.19.

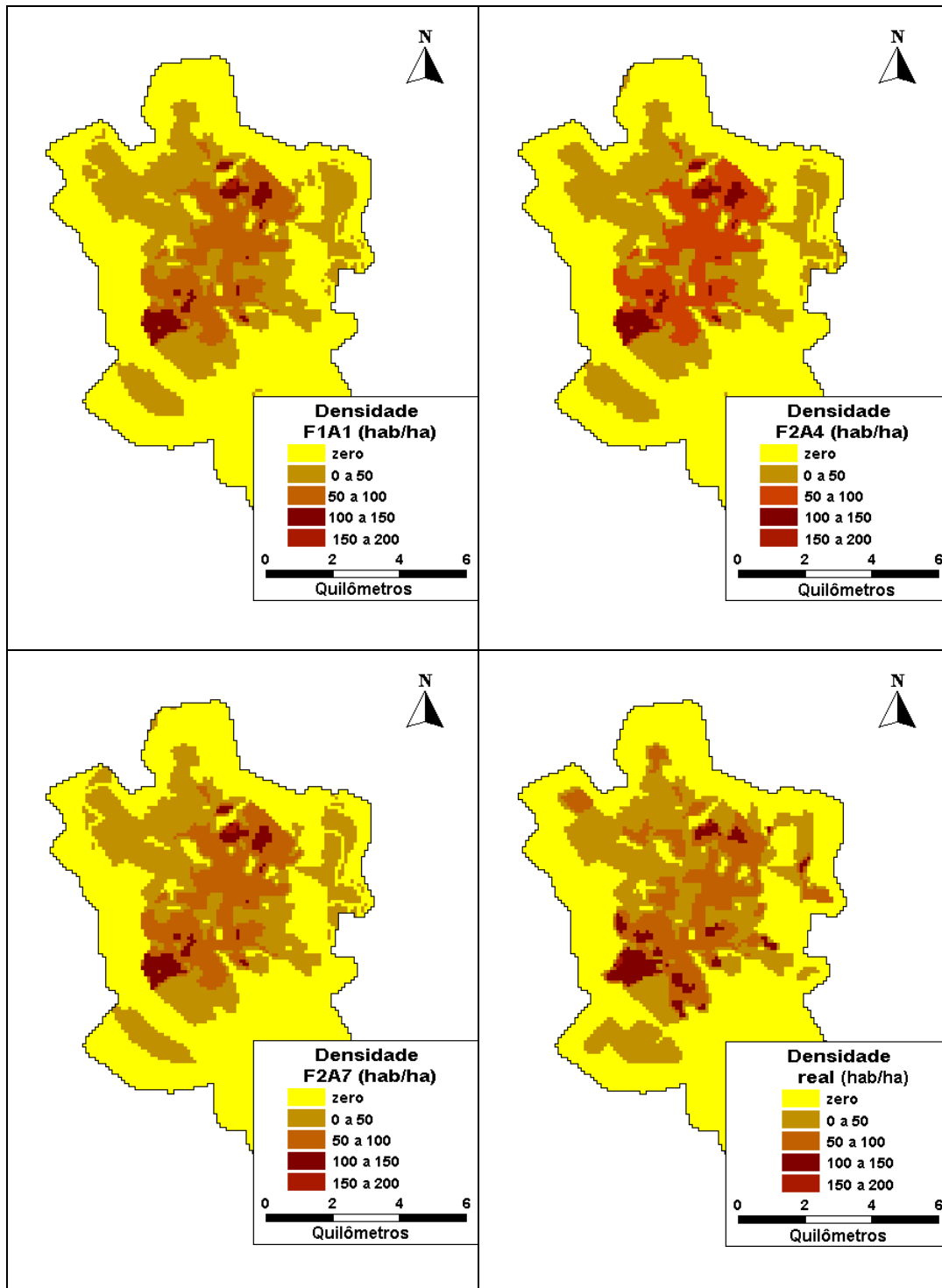


Figura 5.18 – Cenários de densidades populacionais gerados pelos modelos e real para 1991.

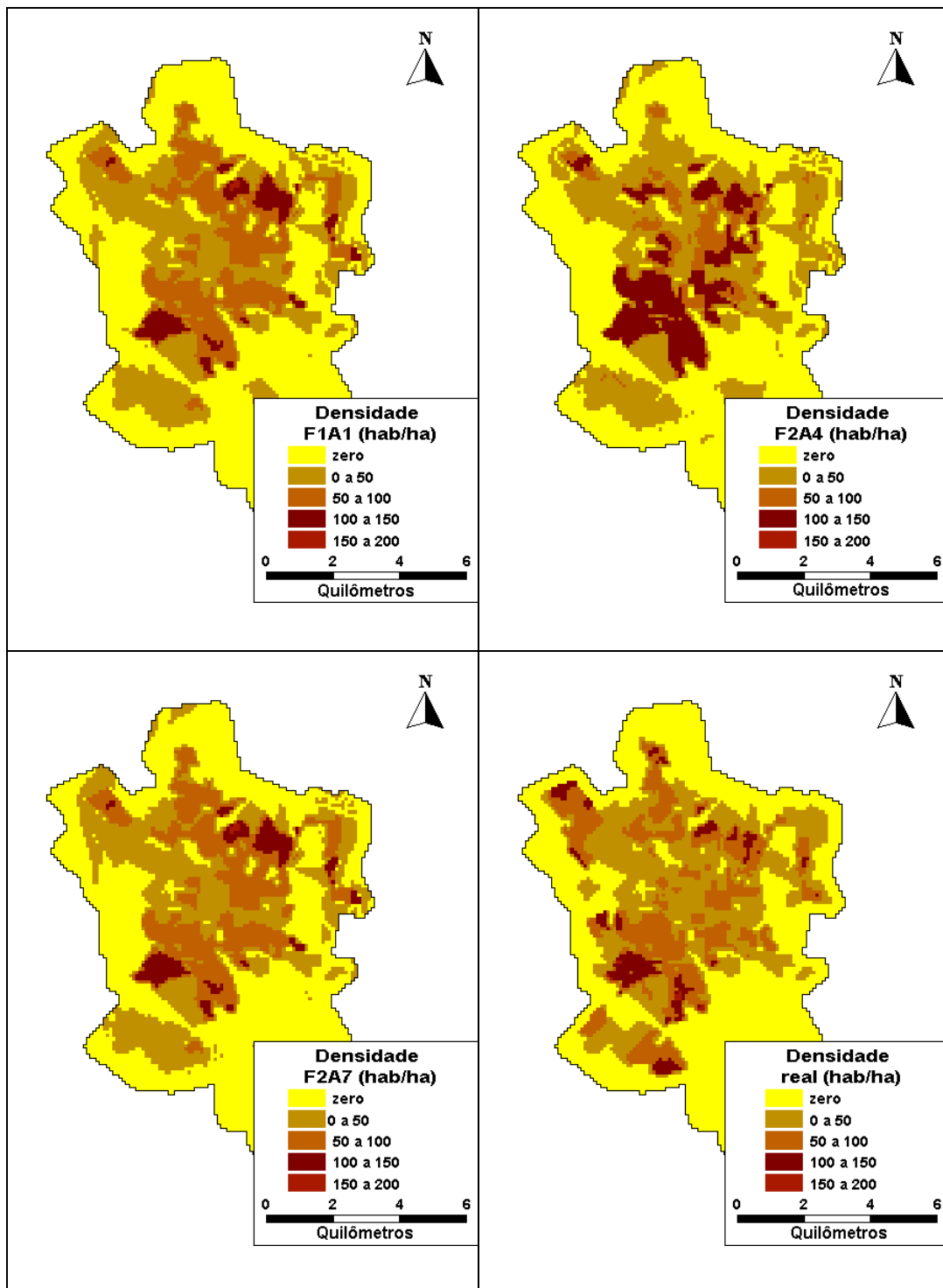


Figura 5.19 – Cenários de densidades populacionais gerados pelos modelos e reais para 2000.

5.2.3.1 Escolha do melhor modelo global de diferenciação

Os cenários gerados pelos modelos relativos aos anos de 1991 e 2000 foram comparados aos cenários reais através do coeficiente Fit. As Tabelas 5.16 e 5.17 apresentam os valores do coeficiente Fit obtido pelos modelos para

tamanhos de janela de 1 a 5, além do coeficiente médio para as fases de calibração e validação, respectivamente.

Tabela 5.16 – Valores do coeficiente Fit obtidos para o ano de 1991.

1991						
Modelo	Fit 1	Fit 2	Fit 3	Fit 4	Fit 5	Fit
F2A4	0,9150	0,9325	0,9429	0,9502	0,9558	0,9488
F2A7	0,9148	0,9329	0,9431	0,9505	0,9561	0,9491
F1A1	0,9151	0,9327	0,9430	0,9504	0,9561	0,9491

Tabela 5.17 – Valores do coeficiente Fit obtidos para o ano de 2000.

2000						
Modelo	Fit 1	Fit 2	Fit 3	Fit 4	Fit 5	Fit
F2A4	0,9049	0,9242	0,9344	0,9415	0,9470	0,9402
F2A7	0,9040	0,9246	0,9366	0,9454	0,9519	0,9432
F1A1	0,9043	0,9245	0,9363	0,9449	0,9512	0,9428

Os modelos apresentaram resultados muito semelhantes na diferenciação da dispersão, destacando-se o modelo F2A7, que apresentou melhor desempenho e que foi em seguida utilizado para obter a distribuição da população para o ano de 2010, apresentada na Figura 5.20.

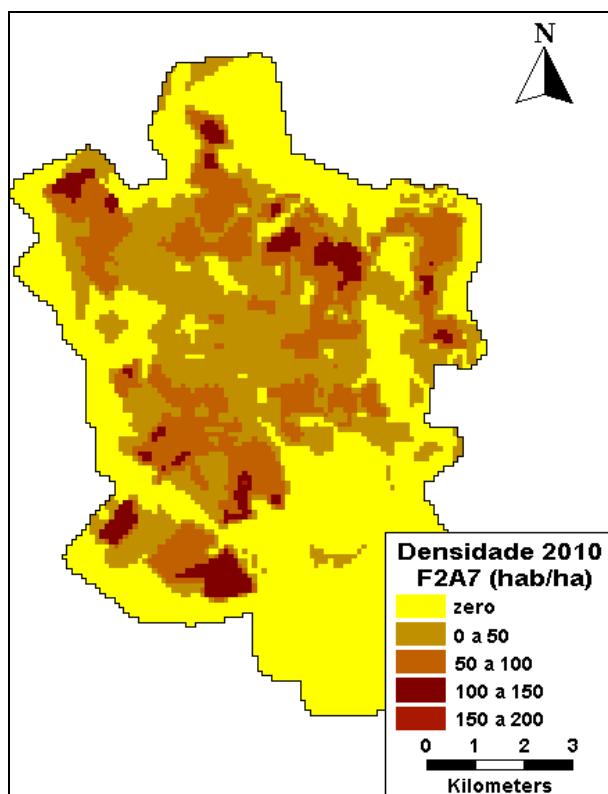


Figura 5.20 – Cenários de densidades populacionais gerados pelo modelo de *Cellular Automata* para 2010.

O cenário gerado pelo modelo de *Cellular Automata* estabelece um aumento da densificação da população nas regiões próximas à faixa de densidade

de 100 a 150 hab/ha, e os valores de densidade não ultrapassam a faixa de valores máxima observada nos dados referentes ao ano 2000. Nota-se que o padrão de ocupação vigente prioriza o desenvolvimento das regiões nordeste e sudoeste, a despeito do que se planeja no projeto de plano diretor da cidade, que pretende que a cidade se desenvolva a noroeste. Este fato enfatiza a necessidade de medidas legais para controlar a ocupação da cidade de São Carlos.

5.2.4 Estimativa da população de 4 a 6 anos

De posse do cenário de densidade populacional total determinado com o modelo de *Cellular Automata* iniciou-se a fase de estimativa da população de 4 a 6 anos, faixa etária que utiliza o serviço das EMEIs. Foi calculada a distância dos centróides dos setores censitários até o centro da cidade, assim como a proporção entre a população na faixa etária de 4 a 6 anos e a população total. Este procedimento justifica-se pelo fato de que as regiões da periferia são as que têm maior proporção de população jovem. A Figura 5.21 apresenta os pontos obtidos no cálculo e a reta ajustada obtida, que apresentou coeficiente de determinação $R^2 = 0,5607$.

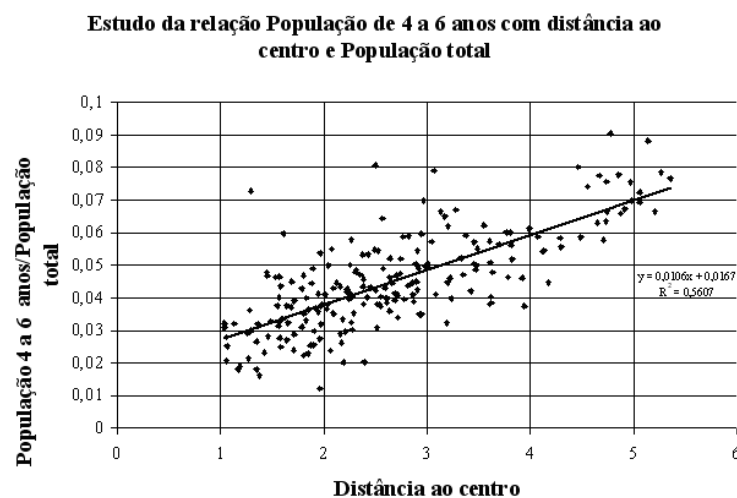


Figura 5.21 – Análise Espacial da População de 4 a 6 anos.

Estabelecida a função linear relacionando as duas variáveis calculadas, agregou-se a distribuição da população obtida pelo modelo de CA na fase anterior em setores quadrados de 500 metros de lado. Os setores que apresentaram população maior que 2000 habitantes (número máximo de habitantes por setor censitário do IBGE) foram divididos em dois. De posse da função e da população por setor foi realizado o cálculo da população e da

densidade populacional na faixa etária de 4 a 6 anos nos setores, apresentada na Figura 5.22.

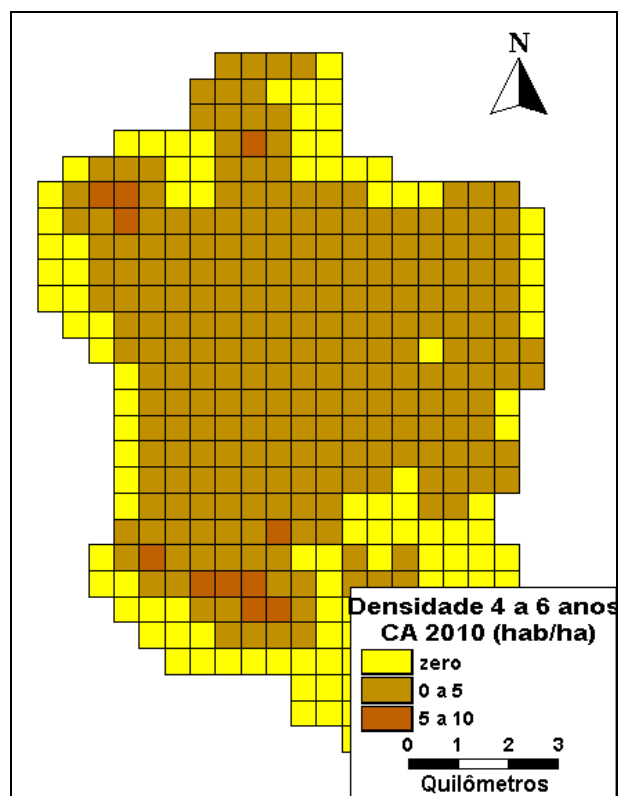


Figura 5.22 – População de 4 a 6 anos prevista pelo modelo de Cellular Automata e agregada em setores de 500 metros de lado.

5.3 Modelagem da demanda por EMEIs com modelos demográficos

Para fins comparativos foi feita também a estimativa da população na faixa etária de 4 a 6 anos utilizando um modelo demográfico de extrapolação de tendências.

A construção do modelo demográfico utilizou a estimativa de população para 2010 efetuada por Lima (2003), apresentada na Tabela 5.18 juntamente com dados de população nos anos considerados na pesquisa.

Tabela 5.18 – Dados e estimativas de população total para a cidade de São Carlos.

Ano	População
1980	109753
1991	147750
2000	183021
2010	219350

O passo inicial para a construção do modelo demográfico para previsão da distribuição espacial da população futura na faixa etária de EMEIs (4

a 6 anos) foi a construção de um modelo global, para estimar a distribuição espacial futura da população total da cidade. Com os dados de densidade populacional de São Carlos dos censos demográficos de 1980 e 1991 agregados por setores censitários (ajustados para eliminar grandes áreas vazias), obtidos da pesquisa de Lima, foi estimada a densidade populacional em 2000 em cada setor censitário. Após incluir os dados referentes ao ano de 2000, foi estimada a densidade populacional em 2010 em cada setor censitário. O cálculo, realizado em planilha eletrônica, utilizou uma curva de tendência (linear), observando um limite inferior igual a zero para não se estimar densidades negativas. A população total em 2000 prevista pelo modelo foi então comparada à população real em 2000 e a população total prevista para 2010 foi comparada à previsão de população urbana total da cidade, calculada por Lima. Nos dois casos efetuou-se um ajuste dos valores de densidade obtidos para que a população total prevista para a cidade se igualasse à obtida pelo modelo demográfico global.

As Tabelas 5.19 e 5.20 apresentam os resultados para as previsões realizadas para a população total para os anos 2000 e 2010, respectivamente, de forma resumida (são apresentados os resultados individuais de 10 setores e as somatórias para os 328 setores). As colunas D1980, D1991 e D2000 apresentam, respectivamente, os valores de densidade populacional nos anos de 1980, 1991 e 2000, para a divisão dos setores de 2000. A coluna D2010 apresenta os valores de densidade populacional para 2010 estimados pelo modelo demográfico e a P2010, os valores de população. A somatória dos valores obtidos de população para todos os setores é então comparada ao valor estimado de população urbana para 2010. O valor de população global estimada por Lima é 2,0 % maior que a estimada pelo modelo demográfico. Multiplicando-se os valores estimados de densidades populacionais pelo fator de ajuste 1,020 obtêm-se os valores de densidade e população ajustados, apresentados, respectivamente, nas colunas D2010aj e P2010aj. As Figuras 5.23 e 5.24 apresentam os cenários de densidades populacionais para 2000 e 2010 obtidos com o modelo demográfico.

Tabela 5.19 – População total prevista para 2000 pelo modelo demográfico.

ano de previsão: 2000		variável: população total					
setores		observados		estimados		ajustados	
ID	área (m ²)	D1980	D1991	D2000	P2000	D2000aj	P2000aj
3	60722	112,58	122,17	130,02	790	134,10	814
4	194714	10,34	19,63	27,23	530	28,08	547
9	150220	0,00	0,17	0,31	5	0,32	5
10	269835	0,00	6,98	12,69	343	13,09	353
11	301900	0,00	7,00	12,72	384	13,12	396
12	213789	0,00	6,76	12,29	263	12,68	271
13	826615	10,16	19,60	27,32	2259	28,18	2330
15	87029	0,40	109,77	199,26	1734	205,51	1789
19	41334	0,00	0,00	0,00	-	0,00	0
21	53727	124,24	169,65	206,80	1111	213,29	1146
...
...
TOTAIS (328 setores):					177454	183021	
valor total da população urbanizada					183021	D: densidade (hab.ha)	
fator de ajuste:					1,0314	P: população (hab)	

Tabela 5.20 – População total prevista para 2010 pelo modelo demográfico.

ano de previsão: 2010		variável: população total						
setores		observados		estimados		ajustados		
ID	área (m ²)	D1980	D1991	D2000	D2010	P2010	D2010aj	P2010aj
3	60722	112,58	122,17	108,53	111,22	675	113,48	689
4	194714	10,34	19,63	56,03	72,56	1413	74,04	1442
9	150220	0,00	0,17	88,07	112,86	1695	115,16	1730
10	269835	0,00	6,98	49,70	66,44	1793	67,79	1829
11	301900	0,00	7,00	47,73	63,92	1930	65,22	1969
12	213789	0,00	6,76	57,49	76,32	1632	77,88	1665
13	826615	10,16	19,60	16,25	21,72	1795	22,16	1832
15	87029	0,40	109,77	102,84	175,20	1525	178,77	1556
19	41334	0,00	0,00	142,26	182,20	753	185,91	768
21	53727	124,24	169,65	155,23	182,03	978	185,74	998
...
...
TOTAIS (328 setores):					214972	219350		
valor total da população urbanizada					219350	D: densidade (hab.ha)		
fator de ajuste:					1,020	P: população (hab)		

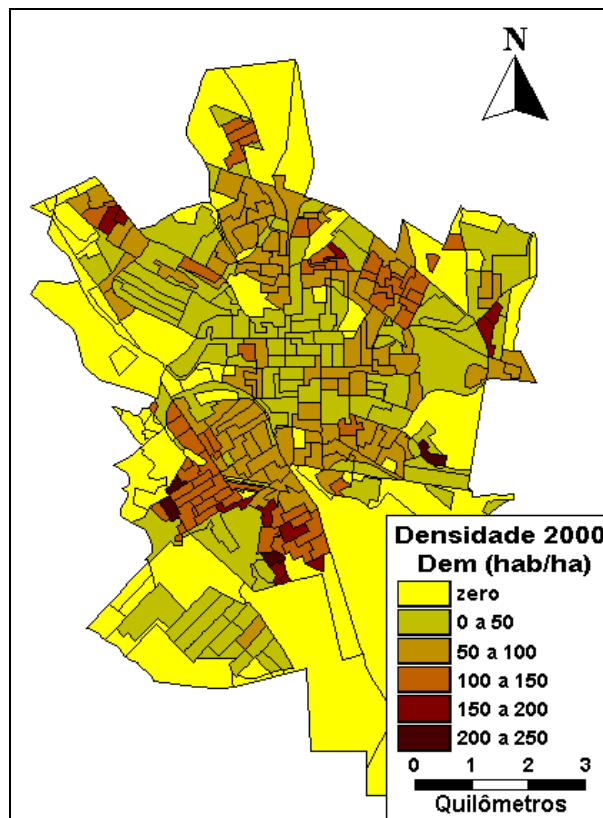


Figura 5.23 – Cenários de densidades populacionais gerados pelo modelo demográfico para 2000.

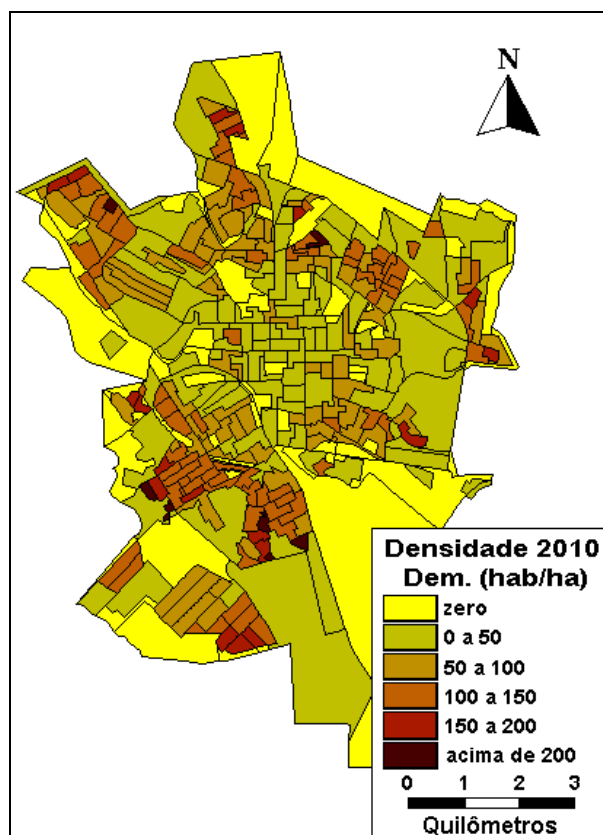


Figura 5.24 – Cenários de densidades populacionais gerados pelo modelo demográfico para 2010.

O cenário gerado pelo modelo demográfico determina, em algumas regiões, densidades demográficas numa faixa superior à encontrada nos valores de densidade populacional nos anos de 1980 a 2000. Como esperado, não apresenta expansão da mancha urbana.

Em seguida foram utilizados os modelos para projetar a demanda por EMEIs. A única diferença é que agora, ao invés de se trabalhar com a densidade global (população total), trabalha-se com a densidade de crianças (população de 4 a 6 anos), obtida por Lima (2003). O procedimento para estimativa da densidade populacional da faixa etária analisada é idêntico ao do modelo global. Os valores obtidos são então ajustados pelo fator de ajuste calculado para o modelo global. As Tabelas 5.21 e 5.22 apresentam os valores calculados e a população total de 4 a 6 anos previstas pelo modelo demográfico.

Tabela 5.21 – População de 4 a 6 anos prevista para 2000 pelo modelo demográfico.

ano de previsão: 2000		variável: população 4-6 anos (EMEIs)					
setores	observados	estimados				ajustados	
ID	área (m ²)	D1980em	D1991em	D2000em	P2000em	D2000em_aj	P2000em_aj
3	60722	8,40	5,54	3,20	19	3,30	20
4	194714	0,87	1,33	1,70	33	1,76	34
9	150220	0,00	0,01	0,03	0	0,03	0
10	269835	0,00	0,60	1,09	29	1,12	30
11	301900	0,00	0,60	1,09	33	1,13	34
12	213789	0,00	0,58	1,06	23	1,09	23
13	826615	0,86	1,33	1,71	142	1,77	146
15	87029	0,01	9,41	17,10	149	17,64	153
19	41334	0,00	0,00	0,00	0	0,00	0
21	53727	10,73	8,88	7,37	40	7,60	41
...
...
TOTAIS (328 setores):					9031	9314	
valor total da população urbanizada					D: densidade (hab.ha)		
fator de ajuste:					1,0314	P: população (hab)	

Tabela 5.22 – População de 4 a 6 anos prevista para 2010 pelo modelo demográfico.

ano de previsão: 2010		variável: população 4-6 anos (EMEI)						
setores		observados			estimados		Ajustados	
ID	área (m ²)	D1980em	D1991em	D2000em	D2010em	P2010em	D2010aj_em	P2010aj_em
3	60722	8,40	5,54	4,45	2,19	13	2,24	14
4	194714	0,87	1,33	3,65	4,61	90	4,70	92
9	150220	0,00	0,01	6,86	8,79	132	8,97	135
10	269835	0,00	0,60	3,85	5,18	140	5,28	142
11	301900	0,00	0,60	3,61	4,86	147	4,96	150
12	213789	0,00	0,58	4,02	5,38	115	5,49	117
13	826615	0,86	1,33	0,75	0,91	75	0,93	77
15	87029	0,01	9,41	6,89	12,57	109	12,83	112
19	41334	0,00	0,00	5,81	7,44	31	7,59	31
21	53727	10,73	8,88	7,63	6,02	32	6,15	33
...
...
TOTAIS (328 setores):						9565	9759	
fator de ajuste:						1,020	D: densidade (hab.ha)	
valor total ajustado						9759	P: população (hab)	

A Figura 5.25 apresenta o cenário de densidade da população de 4 a 6 anos, obtido pelo modelo demográfico.

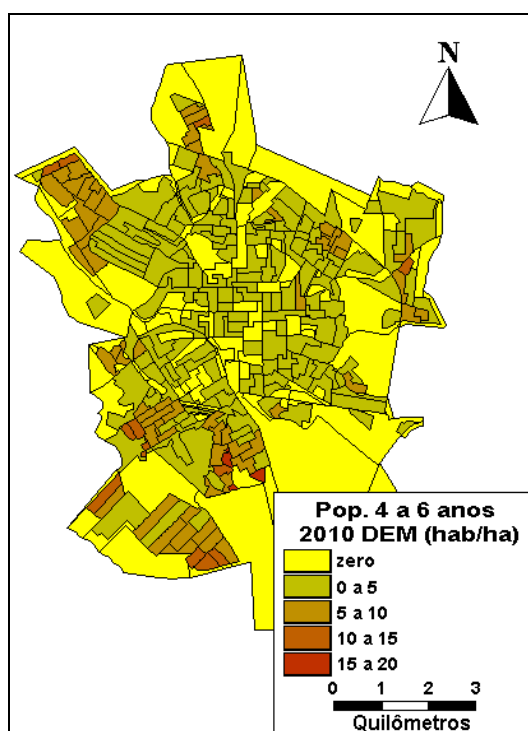


Figura 5.25 – População de 4 a 6 anos prevista pelo modelo demográfico.

Observando-se o resultado global das previsões de população de 4 a 6 anos, apresentado na Tabela 5.23, nota-se que para o ano 2000 ambos os modelos subestimam a população de 4 a 6 anos, apresentando valores semelhantes. Em ambas as previsões nota-se decréscimo na demanda total.

Conseqüentemente a proporção de pessoas na faixa etária de 4 a 6 anos em relação à população total (apresentada na Tabela 5.23) decresce com o tempo, uma situação coerente com as tendências demográficas de envelhecimento da população brasileira, embora diferente do que foi observado particularmente na cidade de São Carlos no ano 2000.

Tabela 5.23 – Proporção de pessoas de 4 a 6 anos .

ano	População de 4 a 6 anos	Cenários	
		População total	Proporção
1980	8570	109753	0,0781
1991	10769	147750	0,0729
2000	13238	183021	0,0723
2000 (CA)	9824	183021	0,0537
2000 (DEM)	9314	183021	0,0509
2010 (CA)	13255	219350	0,0604
2010 (DEM)	9759	219350	0,0436

Estimada a distribuição espacial da demanda pelo modelo, ajustou-se a produção em cada setor para que o fluxo total fosse igual à capacidade total das creches. Para isto os valores de população dos setores, previstos por cada modelo foram multiplicados pela taxa de atendimento.

5.4 Distribuição de viagens

Com a demanda agregada nos setores censitários ajustados e nos setores do modelo de *Cellular Automata* iniciou-se a fase de distribuição de viagens para as EMEIs.

Inicialmente foi realizada a exportação dos centróides dos setores. Realizou-se a seleção, na camada de pontos associada à malha viária, dos nós mais próximo aos centróides e às EMEIs. Após criar arquivo de rede com base no arquivo geográfico de linhas representando a malha viária, gerou-se uma matriz de menores distâncias da primeira seleção de pontos (centróides) à segunda (EMEIs).

Os dados da matriz foram então exportados para a planilha que reproduz a rede neural do modelo de distribuição de viagens de Akamine e Silva (2004).

- Para calcular o total de viagens geradas nos pontos de origem multiplicou-se a população de 4 a 6 anos obtida por cada modelo por uma porcentagem de atendimento tal que o

total de viagens produzidas em todos os setores fosse igual à capacidade total das escolas.

- Para o total de viagens atraídas nos pontos de destino foi considerada a capacidade das escolas;
- A distância entre os pontos de origem-destino é a exportada do TransCAD.

De posse destes dados determinou-se o fluxo entre todas as combinações entre os setores e escolas assim como o número de alunos que frequentarão cada escola segundo o modelo de *Cellular Automata* e demográfico. Como o número total de viagens produzidas deveria ser igual ao número total de vagas, considerou-se que apenas uma parcela da demanda total pelo serviço das EMEIs, constituída por toda a população de 4 a 6 anos no setor, seria atendida. Esta parcela é representada pela porcentagem de atendimento, obtida através da divisão do número de vagas oferecido pelo número de vagas necessário para atender a todas as crianças na faixa etária de 4 a 6 anos na cidade. Para obter os fluxos reais associou-se o endereço real de cada aluno ao setor (criado) a que pertencia e calculou-se o número de alunos que viajavam de cada setor para cada EMEI. Os fluxos obtidos pelos modelos foram comparados a estes dados de fluxo reais.

As Figuras 5.26, 5.27 e 5.28 apresentam a comparação dos fluxos de 2000 que foram obtidos pelos modelos com os fluxos reais e a comparação entre os fluxos obtidos pelo modelo de *Cellular Automata* e Demográfico. Nota-se que o desempenho dos modelos, que seriam perfeitos se os pontos estivessem alinhadas em uma reta com 45° a partir da origem dos eixos, é bastante semelhante na previsão do fluxo em 2000. Esta tendência se mantém quando comparados os fluxos no ano de 2010. Apesar de os modelos apresentarem valores de demanda muito diferentes, quando a produção é ajustada pela porcentagem de atendimento, os resultados de distribuição de viagens é muito parecido. É importante notar, entretanto, que no caso do modelo de *Cellular Automata*, como a demanda prevista é maior e a oferta é a mesma, a porcentagem de atendimento é muito inferior à do modelo demográfico, ou seja, no cenário de *Cellular Automata* a demanda não atendida pelas EMEIs é muito maior, como pode ser observado na Tabela 5.24.

Na prática, no entanto, esta condição dificilmente seria observada, pois a oferta muito provavelmente seria ampliada. Independente do local onde

ocorreria este aumento na oferta, que poderia ser dar tanto nas escolas já existentes como em novas escolas, o impacto sobre os transportes seria melhor caracterizado no caso dos modelos de CA, cujas estimativas de demanda são significativamente maiores e com a possibilidade de ocupação de novas áreas urbanas.

Tabela 5.24 – Proporção de pessoas de 4 a 6 anos.

ano	Cenários		
	População de 4 a 6 anos	Demanda considerada	Porcentagem de atendimento
2000	13238	5879	44,4%
2000 (CA)	9824	5879	59,8%
2000 (DEM)	9314	5879	63,1%
2010 (CA)	13255	5879	44,4%
2010 (Dem.)	9759	5879	60,2%

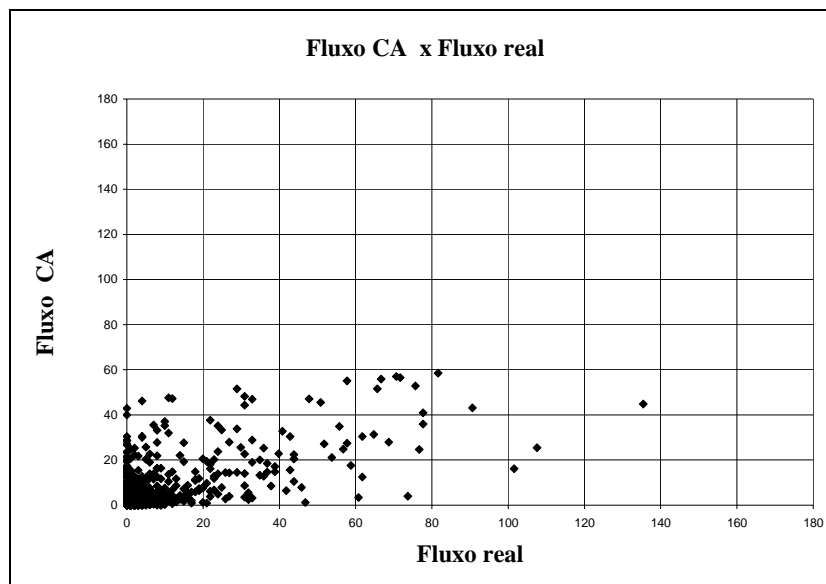


Figura 5.26 – Comparação entre fluxos do modelo de *Cellular Automata* e real em 2000.

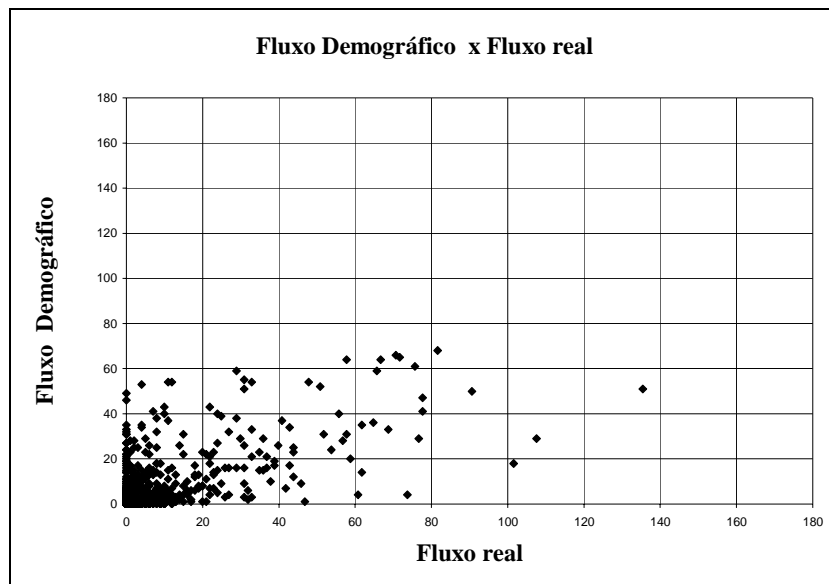


Figura 5.27 – Comparação entre fluxos do modelo demográfico e real em 2000.

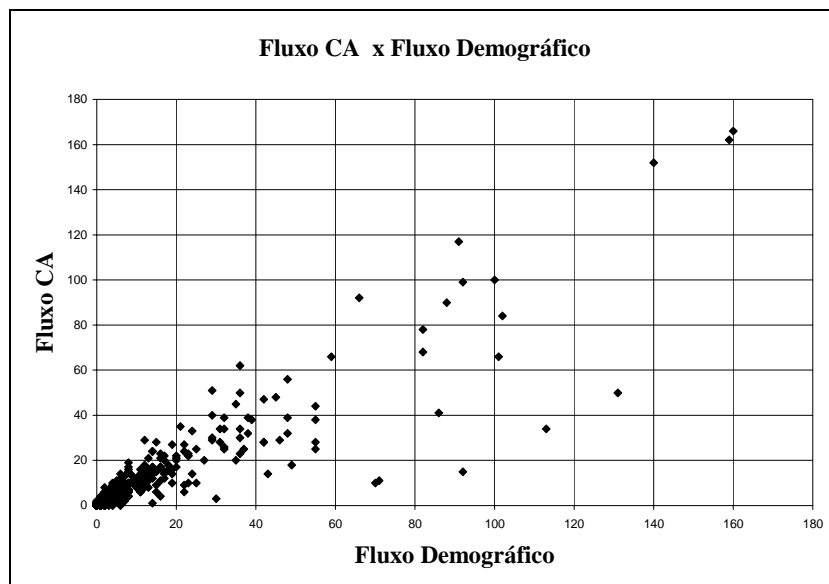


Figura 5.28 – Comparação entre fluxos do modelo demográfico e real em 2010

Nota-se nos cenários de distribuição gerados para o ano de 2000 pelos dois modelos que os fluxos estimados restringem-se ao valor máximo de 60 enquanto os fluxos reais atingem valores próximos a 140. Portanto, há superestimativa dos fluxos menores enquanto que os fluxos maiores são subestimados. A análise do gráfico de comparação entre os fluxos estimados para o ano 2010 pelo modelo de *Cellular Automata* e demográfico revela a continuidade da semelhança entre as previsões de fluxos, o que evidentemente não seria o caso na condição provável de ampliação da oferta, onde os modelos de CA forneceriam previsões aparentemente muito mais realistas.

6 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivos modelar a dinâmica populacional intra-urbana de uma cidade de médio porte brasileira (São Carlos) através da combinação de modelos com *Cellular Automata* (CA) e Avaliação Multicritério (AMC), como alternativa para antecipar necessidades de infra-estruturas pontuais e modelar a geração e distribuição de viagens para uma infra-estrutura pontual específica, as Escolas Municipais de Educação Infantil (EMEI).

Na revisão bibliográfica que se conduziu a princípio, tentando identificar procedimentos que auxiliassem na elaboração do modelo aqui proposto, foram encontrados diversos trabalhos que colaboraram para a definição dos critérios utilizados na construção do modelo. Os trabalhos de Wu e Webster (1998), Sasaki *et al.* (2001) e Antoni (2002a) merecem destaque na contribuição ao método utilizado na modelagem.

Na aplicação prática da pesquisa analisou-se a cidade de São Carlos, localizada no estado de São Paulo, Brasil. Inicialmente foi modelada a dinâmica populacional intra-urbana na cidade de São Carlos como um todo com um modelo de *Cellular Automata* e um modelo demográfico de extrapolação de tendências. O modelo urbano global de *Cellular Automata* conseguiu captar a dinâmica da ocupação urbana de São Carlos, que é diferente da meta definida pelo plano diretor da cidade, o que reforça a necessidade de medidas legais para controlar a ocupação urbana. O modelo apresenta bons resultados tanto na localização das áreas que seriam incorporadas à mancha urbana da cidade quanto na determinação da densidade populacional na mancha urbana resultante.

O modelo global de *Cellular Automata* apresentou um grande avanço em sua precisão em relação ao modelo criado por Lima (2003). No modelo de Lima (2003) as redes neurais ficavam responsáveis tanto pela localização quanto pela diferenciação da dispersão, que era realizada em termos absolutos, prevendo diretamente os valores de densidade populacional. No estudo

anterior verificou-se que acrescentar muitas células à periferia da mancha urbana da cidade dificultava o treinamento das redes neurais. Isto era uma restrição importante, pois o que mais diferencia os modelos de *Cellular Automata* dos outros modelos urbanos é justamente sua capacidade de expansão da mancha urbana. Delegar à rede neural apenas a diferenciação da dispersão trouxe resultados muito mais precisos.

Quanto às limitações do modelo aqui construído ressalta-se que o mesmo não é capaz de prever diretamente a densidade populacional, necessitando de uma previsão global de população, ao contrário do modelo demográfico, que é capaz de fazer uma estimativa da população total, podendo posteriormente ser corrigido por uma previsão de população. Além disso, existe a necessidade de obtenção de dados com especialistas, o que limita consideravelmente a criação os cenários de localização da dispersão.

Outro aspecto importante deste estudo é a introdução da independência da quantidade de células acrescentadas à periferia durante a avaliação do modelo. Numa comparação entre cenários, estas células poderiam levar à falsa impressão de que o modelo é melhor do que realmente é, pois não haveria diferença entre os cenários reais e modelados em toda a área periférica, que representa uma área significativa dos mesmos. Esta característica permite maior liberdade ao acrescentar células à periferia em estudos futuros.

No caso do modelo demográfico notou-se que, devido à restrição da mancha urbana à área previamente ocupada, surgem valores acima da faixa de densidade populacional observada na série histórica e nos dados reais, o que representa uma superestimativa da demanda em algumas regiões da cidade.

Durante a caracterização da demanda por EMEIs foram criados cenários de distribuição de viagens em 2000 e 2010 para as EMEIs existentes no ano de 2000. Observou-se que os modelos apresentaram comportamentos distintos na caracterização da demanda, que foi expressivamente maior no modelo de *Cellular Automata*. No entanto, quando se modela o fluxo para a demanda que as EMEIs efetivamente podem atender, os resultados obtidos são muito semelhantes, residindo a diferença principal na porcentagem de atendimento, muito menor no modelo de *Cellular Automata*.

Uma recomendação para trabalhos posteriores seria a inclusão das EMEIs que foram abertas pela prefeitura após o ano de 2000 no estudo da distribuição de viagens dos modelos de *Cellular Automata*.

Quanto à localização da dispersão, recomenda-se o estudo de maneiras alternativas de localização das áreas que serão incorporadas à mancha urbana, utilizando algumas das técnicas sugeridas na revisão da literatura, sendo particularmente úteis os estudos de Yeh e Li (2001a) e Liu e Phinn (2003).

Na determinação da densidade populacional da mancha urbana prevista pode-se utilizar o método desenvolvido por Ohgai *et al.* (2001), um dos poucos trabalhos que realiza esta fase da modelagem urbana.

Sugere-se ainda utilizar o modelo global para o planejamento estratégico de outras infra-estruturas pontuais cuja demanda é toda a população, tais como unidades do corpo de bombeiro e postos de saúde ou modelar outras faixas etárias específicas, obtendo, por exemplo a demanda por creches ou escolas municipais de ensino básico. Uma outra aplicação com grande interesse para os transportes seria a modelagem de padrões de crescimento e mudança, ao longo do tempo, de agrupamentos de Pólos Geradores de Viagens. Outra sugestão seria criar redes neurais específicas para a modelagem da faixa etária de interesse, eliminando assim a necessidade de estabelecer, a posteriori, uma relação matemática entre a população total e a da faixa etária modelada.

7 REFERÊNCIAS BIBLIOGRÁFICAS

- Akamine, A. e Silva, A.N.R. (2004) An evaluation of neural spatial interection models based on a practical application. **Design and Decision Support Systems in Urban Planning**. In: Van Leeuwen, J.P. & Timmermans H.J.P. (eds) Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, p. 19-32.
- Almeida, C. M.; Monteiro; A. M. V.; Câmara, G.; Soares-Filho, B. S.; Cerqueira, G. C.; Pennachin, C. L. e Batty, M. (2002a). Modelos espaço-temporais em planejamento urbano: desafios para a nova geração de SIG. In: **GIS Brasil 2002, Mostra do talento Científico, 2**, Curitiba, 2002. Anais eletrônicos.
- Almeida, C. M.; Monteiro; A. M. V.; Câmara, G.; Soares-Filho, B. S.; Cerqueira, G. C.; Pennachin, C. L. e Batty, M. (2002b). Empiricism and Stochastics in Cellular Automaton Modeling of Urban Land Use Dynamics. Paper 42 **Centre for Advanced Spatial Analysis (CASA)**, University College Londres, Inglaterra, http://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. (2003). Modelos celulares de dinâmicas espaço-temporais urbanas. In Câmara, G.; Monteiro, A. M. V. (2003) Introdução à modelagem dinâmica espacial. **Simpósio Brasileiro de Sensoriamento Remoto**, capítulo 3, Belo Horizonte.
- Antoni, J. P. (2002a) Urban Sprawl Modelling: Combining Models to make decisions. **Design and Decision Support Systems in Urban Planning**. In Timmermans, H. (Ed.) Architecture and Urban Planning, Ellecom, Holanda.
- Antoni, J. P. (2002b) Urban Sprawl Modelling: A methodological approach. **Cybergeo 2002**, <http://www.cybergeo.presse.fr/ectqg12/antoni/antoni.htm>.
- Batty, M.; Xie, Y. e Sun, Z. (1999) Modeling urban dynamics through GIS-based cellular automata **Computers, Environments and Urban Systems**, v. 23 p. 205-233.
- Bonham-Carter, G. F. (1994) **Geographic Information System for geoscientists: Modelling with GIS**. Ontario: Pergamon. 414 p.

- Bossard, E. G. (1999) Envisioning Neighborhood Quality of Life Using Conditions in the Neighborhood Access to and from Conditions in the Surrounding Region. *In: Rizzi, P. (ed.), Computers in Urban Planning and Urban Management on the Edge of the Millenium*. FrancoAngeli, Veneza, Itália.
- Carver, S. J. (1991) Integrating Multi-criteria Evaluation with Geographical Information Systems. **International Journal of Geographical Information Systems**, v. 5, n. 3, p. 321-339.
- Cecchini, A. e Rizzi, P. (2001) The reasons why Cellular Automata are a useful tool in the working-kit for the millennium urban planner in governing the territory. Anais do **VII International Conference on Computers in Urban Planning and Urban Management, CUPUM**, Honolulu, EUA. Anais eletrônicos.
- Constanza, R. (1989) Model goodness of fit: a multiple resolution procedure. **Ecological Modelling**, v.47, p.199 – 215.
- Costa, M. S. (2003) *Mobilidade urbana sustentável: um estudo comparativo e as bases de um sistema de gestão para Brasil e Portugal*. 184 p. Dissertação (Mestrado) – **Escola de Engenharia de São Carlos**, Universidade de São Paulo, São Carlos. 2003.
- Dewdney, A. K. (1989) *The Turing Omnibus*, **Computer Science Press**. Rockville, MD, EUA.
- Diappi, L.; Bolchi, P e Buscema, M. (2004) Improved understanding of urban sprawl using Neural Networks. *In: Van Leeuwen, J. P.; Timmermans, H. J. P. Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning*, Sint-Michielsgestel, Holanda, , p. 33-49.
- Eastman, J. R. (1997) *IDRISI for Windows: User's Guide. Version 2.0*. **Clark University - Graduate School of Geography**, Worcester, MA, EUA.
- Eastman, J. R. e Jiang, H. (1996) Fuzzy Measures in Multi-Criteria Evaluation. *In 2nd International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Studies*, May 21-23, Fort Collins, EUA, p. 527-534.
- Eastman, J. R.; Jiang, H. e Toledano, J. (1998) Multi-criteria and Multi-objective Decision Making for Land Allocation Using GIS. *In: Beinat, E. e Nijkamp, P. (eds.) Multicriteria Analysis for Land-Use Management*. **Kluwer Academic Publishers**, Dordrecht, Holanda, p. 227-251.
- Haque, S. M. e Okumura, M. (2003) Simulation of Land Use changes for Strategic Urban Management with a GIS based Statistical Model. Anais do **VIII International Conference on Computers in Urban Planning and Urban Management, CUPUM**, Sendai, Japão. Anais eletrônicos.
- IBGE (2000) Tabela de Indicadores Sociais Municipais – 2000, <http://www.ibge.gov.br>.
- IBGE (2004) IBGE cidades@, <http://www.ibge.gov.br/cidadesat/>.

- Lima, R. S. (2003) *Bases para uma metodologia de apoio à decisão para serviços de educação e saúde sob a ótica dos transportes*. 200 p. Tese (Doutorado) – **Escola de Engenharia de São Carlos**, Universidade de São Paulo, São Carlos.
- Liu, Y e Phinn, S. R. (2003) Modelling urban development with cellular automata incorporating fuzzy-set approaches. **Computers, Environments and Urban Systems**, v. 27 p. 637-658.
- Malczewski, J. (1999) GIS and Multicriteria Decision Analysis. **John Wiley & Sons**, New York, NY, EUA.
- Mendes, J. F. G. (2000) Decision Strategy Spectrum for the Evaluation of Quality of Life in Cities. In: **Proceedings of the International Conference on Quality of Life in Cities - ICQOLC 2000**, Singapura.
- Novaes, A. G. (1982) Modelos em planejamento urbano, regional e de transportes. **Editores Blücher**, São Paulo, Brasil, 290 p.
- Novak, H. (1998) As cidades do Brasil do futuro. **Revista de Administração Municipal**, Rio de Janeiro, 186 p.
- Ohgai, A.; Igarashi, M; Ma, S. R. e Ikaruga, S. (2001) A Cellular Automata Model to Investigate the Influence of land Use Policies on Future Distributions of Population and Employees. Anais do **VII International Conference on Computers in Urban Planning and Urban Management, CUPUM**, Honolulu, Havaí - EUA. Anais eletrônicos.
- O'Sullivan, D. e Torrens, P. M. (2001). Cellular Automata in urban simulation: where do we go from here? **Environment and Planning B: Planning and Design**, v. 28, p. 163-168.
- Press, W. H.; Flannery, B. P.; Teukolsky, S. A. E e Vetterlint, W. T. (1986) Numerical Recipes: The Art of Scientific Computing: **Cambridge University Press**, Cambridge, 818 p.
- Ramos, R. A. R. (2000) *Localização Industrial. Um Modelo Espacial para o Noroeste de Portugal*. 299 p. Tese (Doutorado) - **Universidade do Minho**, Braga, Portugal.
- Ramos, R. A. R. e Silva, A. N. R. (2003) A data-driven approach for the definition of metropolitan regions. **Anais do VIII International Conference on Computers in Urban Planning and Urban Management, CUPUM**, Sendai, Japão. Anais eletrônicos.
- Ramos, R. A. R.; Silva, A. N. R. (2002) Oportunidades e desafios de técnicas emergentes para o planejamento urbano: o caso dos modelos de Cellular Automata. In: **Encontro de Utilizadores de Informação Geográfica, Comunicações**. Associação de Utilizadores de Informação Geográfica, Lisboa, Portugal.
- Rosenthal, R. E. (1985) Concepts, Theory and Techniques: Principles of Multiobjective Optimisation. **Decision Sciences**, v. 16, n. 2, p. 133-152.

- Roy, G. G. e Snickars, F (1996) Citylife: A study of cellular automata in urban dynamics In Fisher, M.; Scholten, H. J. e Unwin , D **Spatial Analytical Perspectives on GIS.**, Taylor e Francis, Londres p. 213-218.
- Saaty, T. L. E e Vargas, L. G. (1991) Prediction, Projection and Forecasting. **Kluwer Academic Publishers**, Boston, MA - USA.
- Saaty, T. L. (1977) A Scaling Method for Priorities in Hierarchical Structures. **Journal of Mathematical Psychology**, v. 15, n. 3, p. 234-281.
- Saaty, T. L. (1980) The Analytical Hierarchy Process: Planning, Priority Setting, Resource Allocation. **McGraw-Hill**, New York, NY, USA.
- Saaty, T. L. (1987) Concepts, Theory, and Techniques: Rank Generation, Preservation, and Reversal in the Analytic Hierarchy Decision Process. **Decision Sciences**, v. 18, n. 2, p. 157-177.
- Sasaki, K; Tamura, T. e Saitoh, K. (2001). Application of Neural Cellular Automata to Land Use-Population Structure- In case of Metro Manila. **IX World Conference on Transport Research, WCTR**, Seul, Coréia.
- Silva, A. N. R.; Ramos, R. A. R.; Souza, L. C. L.; Rodrigues, D. S. e Mendes, J. F. G. (2004) 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, SP**, Edição dos Autores. 227 p.
- Taylor, C. E. (1992). "Fleshing Out" Artificial Life II. In **Artificial Life**, Santa Fe, New Mexico: p. 25-38.
- Torrens, P. M. (2000). How cellular models of urban systems work. WP-28, **Centre for Advanced Spatial Analysis (CASA)**, University College Londres - Inglaterra, http://www.casa.ucl.ac.uk/how_ca_work.pdf
- Viher, B.; Dobnikar, A.; Zazula, D. (1998). Cellular Automata and Follicle Recognition Problem and Possibilities of Using Cellular Automata for Image Recognition Purposes. **International Journal of Medical Informatics**, v. 49, p. 231-241.
- Voogd, H. (1983) Multicriteria Evaluation for Urban and Regional Planning. **Pion**, Londres, Reino Unido.
- von Winterfeldt, D. e Edwards, W. (1986) Decision Analysis and Behavioural Research. **Cambridge University Press**, Cambridge, Reino Unido.
- Wegener, M. (1994) Operational urban models. **Journal of the American Planning Association**, v. 60, n. 1; p. 17-27.
- Wu, F. e Webster, C. J. (1998) Simulation of land development through the integration of cellular automata and multicriteria evaluation. **Environment and Planning B: Planning and Design**, v. 25, p. 103-126.

- Yager, R. R. (1988) On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decision Making. **IEEE Transactions on Systems, Man and Cybernetics**, v. 8, n. 1, p. 183-190.
- Yeh, A. G. e Li, X. (2001a) A constrained CA model for the simulation and planning of sustainable urban forms by using GIS. **Environment and Planning B: Planning and Design**, v. 28, p. 733-753.
- Yeh, A. G. e Li, X. (2001b) Neural Network CA simulation Urban Forms. Anais do **VII International Conference on Computers in Urban Planning and Urban Management, CUPUM**, Havai, EUA. Anais eletrônicos.
- Yeh, A. G.; Li, X., (2002) A cellular automata model to simulate development density for urban planning. **Environment and Planning B: Planning and Design**, v. 29, p. 431-450.
- Zadeh, L. A. (1965) Fuzzy Sets. **Information and Control**, v. 8, p. 338-353.

ANEXO 1 - Matrizes de comparação par a par

Matriz 1 - Matriz de comparação par a par do avaliador 1.

	dc	mv	dwl	dviap	docup80	ddem
dc	1	1/3	1/3	1/5	1	1
mv	3	1	3	1	1	3
dwl	3	1/3	1	1/9	1/3	1
dviap	5	1	9	1	3	5
docup80	1	1	3	1/3	1	5
ddem	1	1/3	1	1/5	1/5	1

Matriz 2 - Matriz de comparação par a par do avaliador 2.

	Dc	mv	dwl	dviap	docup80	ddem
dc	1	5	3	3	5	1/3
mv	1/5	1	1/5	1/3	1/3	1/7
dwl	1/3	5	1	3	5	1/3
dviap	1/3	3	1/3	1	3	1/5
docup80	1/5	3	1/5	1/3	1	1/5
ddem	3	7	3	5	5	1

Matriz 3 - Matriz de comparação par a par do avaliador 3.

	Dc	mv	dwl	dviap	docup80	ddem
dc	1	7	9	5	5	5
mv	1/7	1	1/3	1/3	1/3	1/3
dwl	1/9	3	1	1/5	1/5	1/5
dviap	1/5	3	5	1	1	1
docup80	1/5	3	5	1	1	1/3
ddem	1/5	3	5	1	3	1

Matriz 4 - Matriz de comparação par a par do avaliador 4.

	Dc	mv	dwl	dviap	docup80	ddem
dc	1	5	9	7	7	5
mv	1/5	1	7	5	5	3
dwl	1/9	1/7	1	1/3	1/5	1/3
dviap	1/7	1/5	3	1	1	1
docup80	1/7	1/5	5	1	1	1/3
ddem	1/5	1/3	3	1	3	1

Matriz 5 - Matriz de comparação par a par do avaliador 5.

	Dc	mv	dwl	dviap	docup80	ddem
dc	1	1	1/3	1/5	1/7	1/7
mv	1	1	1/3	1/5	1/7	1/7
dwl	3	3	1	1/3	1/5	1/5
dviap	5	5	3	1	1/3	1/3
docup80	7	7	5	3	1	1
ddem	7	7	5	3	1	1

Matriz 6 - Matriz de comparação par a par do avaliador 6.

	Dc	mv	dwl	dviap	docup80	ddem
dc	1	3	3	3	3	1/3
mv	1/3	1	3	3	3	1/3
dwl	1/3	1/3	1	1/3	1/3	1/3
dviap	1/3	1/3	3	1	3	1/3
docup80	1/3	1/3	3	1/3	1	1/3
ddem	3	3	3	3	3	1

Matriz 7 - Matriz de comparação par a par do avaliador 7.

	Dc	mv	dwl	dviap	docup80	ddem
dc	1	1/5	7	3	1/3	3
mv	5	1	9	7	3	7
dwl	1/7	1/9	1	1/3	1/7	1/3
dviap	1/3	1/7	3	1	1/5	1/3
docup80	3	1/3	7	5	1	5
ddem	1/3	1/7	3	3	1/5	1

Matriz 8 - Matriz de comparação par a par do avaliador 8.

	Dc	mv	dwl	dviap	docup80	ddem
dc	1	3	5	1/5	1/3	5
mv	1/3	1	3	1/7	1/5	5
dwl	1/5	1/3	1	1/9	1/7	3
dviap	5	7	9	1	3	9
docup80	3	5	7	1/3	1	7
ddem	1/5	1/5	1/3	1/9	1/7	1