



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INTRODUÇÃO AO GEOPROCESSAMENTO

ANÁLISE DE MOBILIDADE REGIONAL E DA SUA RELAÇÃO COM A
URBANIZAÇÃO DISPERSA A PARTIR DE MAPAS AUTO-ORGANIZÁVEIS EM
SÃO JOSÉ DOS CAMPOS, SÃO PAULO

Rayanna Barroso de Oliveira Alves

INPE
São José dos Campos
2021

RESUMO

A identificação de padrões e processos de movimentos pendulares nas cidades contemporâneas tem sido bastante estudadas e tem sido um grande desafio para a gestão urbana. Dessa forma, este trabalho permitiu a utilização de um banco de dados censitário e de entrevista de levantamentos a respeito de origem e destino da população residente no município de São José dos Campos em relação a municípios vizinhos, para fins de trabalho, estudo, comércio, serviços e lazer. Assim, foi desenvolvido um diagnóstico de caracterização da dinâmica intra-regional. Para tal, foram utilizadas redes neurais não supervisionadas para analisar conjuntamente aspectos socioeconômicos, natureza, tipologia e frequência (intensidade) dos deslocamentos, com fins à identificação de "*clusters*", ou seja, agrupamentos de ocorrência de deslocamentos. O aumento do poder de mobilidade conduz à nucleação de assentamentos urbanos, seja na forma de loteamentos isolados, tipologias condominiais diversificadas ou empreendimentos imobiliários de usos múltiplos, e à consequente fragmentação do tecido urbano, podendo causar impactos ambientais diversos.

1. INTRODUÇÃO

O crescimento contínuo das cidades brasileiras e o avanço desordenado da ocupação urbana exerce pressão sobre áreas que deveriam ser desenvolvidas adequadamente.

Segundo OJIMA (2007), o processo de crescimento urbano decorre de dois fatores. O primeiro é referente à população, já o segundo é relativo à expansão física das ocupações urbanas. O fator populacional desempenha e vem desempenhando grande peso na expansão da estrutura urbana.

No que diz respeito à expansão física das ocupações urbanas, o seu crescimento depende do padrão de ocupação territorial, sendo necessário identificar se esse crescimento ocorreu com maior ou menor custo social, refletindo assim, impactos sobre a configuração de reprodução social e sustentabilidade ambiental (UNFPA, 2007).

O crescimento urbano vem causando movimentos pendulares que são decorrentes de melhoria de condições de vida. Eles podem ser entre regiões – intra-regional - ou dentro de uma região – inter-regional, e estudos da origem e destino da população vêm permitindo analisar a direção dos movimentos da população e a sua relação com padrões de urbanização (QUEIROZ et al., 2015).

Ante ao exposto, é necessário utilizar ferramentas computacionais que contribuam para o desenvolvimento de análises comportamentais de cenários espaciais sob as condições do crescimento urbano. Assim, técnicas de inferência espacial são indispensáveis para a elaboração de sistemas geográficos de apoio a monitoramento e decisão (SILVA et al., 2002).

Os Sistemas de Informações Geográficas (SIGs) são constituídos de instrumentos significativamente importantes para o desenvolvimento de estudos urbanos, principalmente de áreas metropolitanas, que subsidiam o planejamento territorial. Atributos urbanos espaciais podem ser tratados algebricamente, analisados, combinados e visualizados através de mapas temáticos (LOMBARDO et.al, 2001)

Assim, o presente trabalho utiliza redes neurais, as quais segundo Câmara et. al., 2001 são consideradas “modelos empíricos baseados em dados e se utilizam de amostras disponíveis para desenvolver as inferências”, para a geração de mapas auto-organizáveis, a fim de verificar a mobilidade intra-regional da cidade de São José dos Campos, São Paulo, em relação a municípios vizinhos.

2. REVISÃO DE LITERATURA

2.1 URBANIZAÇÃO DISPERSA

Segundo Reis (2006), a urbanização dispersa se materializa no espaço intra-urbano e metropolitano ou regional no formato de novas áreas de expansão e renovação do tecido urbano, apresentando descontinuidades. Buchell et al. (1998) definem a dispersão como “intrusão de baixa densidade de residências e outros usos não residenciais na zona rural, bem como em áreas não desenvolvidas, constituindo-se em uma modalidade de consumo da terra de forma espalhada, imprevisível e, por vezes, segregadora”. Dessa forma, definições precisas dos padrões de dispersão no cenário urbano podem ser caracterizadas utilizando medidas espaciais.

O fenômeno da dispersão urbana reporta-se a duas escalas espaciais, ou seja, dois contextos distintos, porém, interrelacionados entre si. A primeira diz respeito às áreas metropolitanas, que se dispersam em crescentes núcleos ou polos, intervalados de vazios, com redução de densidades de ocupação, formando o que alguns autores chamam de nebulosa urbana (REIS, 2006).

A segunda escala é referente ao tecido urbano, onde se definem relações físicas e jurídicas entre o público e o privado do espaço urbano, sua produção material, bem como sua apropriação, uso e transformação. Nela, a dispersão se evidencia pelo esgarçamento crescente do tecido dos principais núcleos urbanos, em suas áreas periféricas, bem como pela formação de constelação ou nebulosas de núcleos urbanos e bairros isolados em meio ao campo, de diferentes dimensões, integrados em uma área metropolitana ou em um conjunto ou sistema de áreas metropolitanas (REIS, 2006).

A dispersão urbana é caracterizada na literatura internacional como uma forma distinta de crescimento urbano, que possui caráter difuso, polinucleado e ineficiente. Alguns estudos têm defendido os benefícios do estilo disperso de desenvolvimento urbano e argumentam, particularmente com relação aos moldes do modelo americano de dispersão, que os padrões de “suburbanização” são resultados das forças de mercado livre, escolha do consumidor e o reflexo do sistema democrático de controle da terra (CARLINER, 1999; EASTERBROOK, 1999).

Segundo Hodge (1992), há um crescente consenso em meio à comunidade científica de que o processo de globalização não somente influencia a estrutura econômica de grandes cidades, mas também a sua estrutura espacial de modo geral. De acordo com investigações nesta linha, estruturas urbanas emergentes apontam para uma transição do

modelo monocêntrico do pós-guerra para o que se denomina de “região urbana dispersa ou policêntrica” (CLARK et al., 1994).

2.2 MOBILIDADE INTRA-REGIONAL

Movimentos pendulares podem ser entre regiões – intra-regional - ou dentro de uma região – inter-regional. Em países subdesenvolvidos, a mobilidade interna se encontra muito bem caracterizada (FERREIRA, 1986).

Assim, estudos da origem e destino da população vêm permitindo analisar a direção dos movimentos da população e a sua relação com padrões de urbanização (QUEIROZ et al., 2015). Conforme Baeninger (2012), a temática envolvendo os movimentos pendulares entre os centros urbanos passaram a sofrer especificidades dadas as dinâmicas regionais.

A evidência de desigualdades sociais entre segmentos populacionais do espaço intra-urbano está presente no processo de desenvolvimento nacional. A configuração socioespacial que é resultante desse processo de estruturação da cidade é caracterizada pela formação de periferias. Essas desigualdades são atenuadas pela intensificação dos deslocamentos populacionais intra-urbanos, que são representados por fluxos de migração intrametropolitanos em direção à periferia regional. A intensificação da frequência desse deslocamento revela a existência de movimentos pendulares (CAIADO, 2005).

Os dados dos censos demográficos são uma forma de verificar as informações da dinâmica intra-regional. Os estudos relacionados a esses movimentos migratórios utilizam parâmetros de periodicidade, comparação e capacidade de identificação de diversos aspectos, como frequência e local de moradia e trabalho (AVILA, 2007).

De uma forma mais objetiva, a urbanização dispersa se materializa no espaço, intra-urbano, sob a forma de novas áreas de expansão e renovação do tecido urbano, apresentando descontinuidades, núcleos isolados, empreendimentos de grande porte ocupando vazios urbanos de áreas consolidadas, renovação urbana e reconstrução em áreas degradadas (REIS, 2006).

2.3 REDES NEURAIS E MAPAS AUTO-ORGANIZÁVEIS

2.3.1 CONCEITUAÇÃO

Modelos de inferência espacial são desenvolvidos por ferramentas operacionais para tomada de decisão através de modelos estatísticos. As aplicações das redes neurais para análise espacial são consideradas modelos propostos concebidos empiricamente, que permite autonomia em inúmeras áreas da análise espacial (SILVA, et al.,2002).

Assim, métodos de análise espacial estatística através das análises neurais e expilarias de maneira não supervisionada e semiautomática oferece boas perspectivas no âmbito da neurocomputação e SIG no suporte ao planejamento urbano (SILVA et al., 2004).

Redes neurais artificiais (RNA) podem ser basicamente definidas como uma estrutura computacional de processamento paralelo maciçamente distribuído, a qual consiste de unidades de processamento, também chamadas de neurônios ou nós, organizadas em uma sequência de camadas. Os neurônios são responsáveis pelo armazenamento do conhecimento adquirido pelo sistema, o qual deve ser disponibilizado para usos futuros (HAYKIN, 1999).

Segundo Fischer et al. (2000), as redes neurais são capazes de aprender e tomar decisões, baseando-se em informação incompleta, ruidosa ou nebulosa, e esta é precisamente a razão pela qual as mesmas podem ser facilmente adequadas para lidar com problemas espaciais.

Embora RNA sejam comumente consideradas como pertencentes ao amplo ramo de inteligência artificial (AI), Fischer et al. (2000) postulam que implementações de redes neurais para análise e modelagem de dados geográficos deveriam ser preferencialmente atribuídas ao domínio de inteligência computacional (IC). Estes autores argumentam que as implementações de RNA sob a forma de aproximadores de funções e classificadores de padrões por anteropropagação conseguem apenas processar dados numéricos (baixo nível), ao passo que sistemas de IA constituem-se em sofisticadas plataformas de IC, dotadas de ferramentas especiais para incorporar conhecimento na forma de informações não-numéricas ou regras e restrições operacionais.

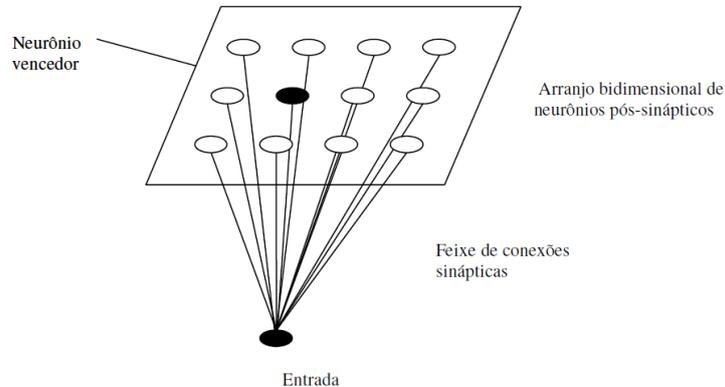
Mapas auto-organizáveis (ou “Self-Organizing Maps” – SOM) constituem uma classe especial de redes neurais (ou grades neurais), que se baseiam em aprendizagem competitiva. Os neurônios de saída da grade competem entre si para serem ativados ou disparados, com o resultado que apenas um neurônio de saída, ou um neurônio por grupo, está ligado em um instante de tempo. Um neurônio de saída que vence a competição é chamado de neurônio vencedor. Uma forma de induzir esta competição entre os neurônios

de saída é usar conexões laterais inibitórias, como, por exemplo caminhos de realimentação negativa (HAYKIN, 1999).

Segundo este mesmo autor, em um mapa auto-organizável, os neurônios estão colocados em nós de uma grade, que é normalmente uni ou bidimensional. Mapas de dimensionalidade mais alta são também possíveis, mas não são tão comuns. Os neurônios se tornam seletivamente sintonizados a vários padrões de entrada (estímulos) ou classes de padrões de entrada no decorrer de um processo de aprendizagem. As localizações dos neurônios assim sintonizados (neurônios vencedores) se tornam desordenadas entre si, de forma que um sistema de coordenadas significativo para diferentes características de entrada é criado sobre a grade (modelo de Kohonen).

Um mapa auto-organizável é, portanto, caracterizado pela formação de um mapa topográfico dos padrões de entrada, no qual as “localizações espaciais (coordenadas) dos neurônios na grade são indicativas das características estatísticas intrínsecas contidas nos padrões de entrada”, como apresentado na (Figura 1) (SANTOS et al., 2006).

Figura 1: Grades neurais no modelo de Kohonen.



Fonte: HAYKIN (1999).

Uma vez que a grade tenha sido apropriadamente inicializada, há três processos topológicos essenciais envolvidos na formação do mapa auto-organizável e serão matematicamente abordados na Seção 2.3.2:

1) Competição - Para cada padrão de entrada, os neurônios de grade calculam seus respectivos valores de uma função discriminante, a qual fornece a base para a competição entre os neurônios. O neurônio particular com o maior valor da função discriminante é declarado vencedor da competição.

2) *Cooperação* - O neurônio vencedor determina a localização espacial de uma vizinhança topológica de neurônios excitados, fornecendo assim a base para a cooperação entre os neurônios vizinhos.

3) *Adaptação Sináptica* - Este último mecanismo permite que os neurônios excitados aumentem seus valores individuais da função discriminante em relação ao padrão de entrada, através de ajustes adequados aplicados a seus pesos sinápticos. Os ajustes são feitos de modo que a resposta do neurônio vencedor à aplicação subsequente de um padrão de entrada similar seja melhorada” (HAYKIN, 1999).

Em outras palavras, o algoritmo SOM envolve três passos básicos – amostragem, casamento por similaridade e atualização – que são repetidos até que a formação do mapa de características esteja completa. O algoritmo pode ser resumido como segue:

1. *Inicialização*: Refere-se à escolha de valores aleatórios para os vetores de pesos iniciais $w_j(0)$. A única restrição é que os $w_j(0)$ sejam diferentes para $j = 1, 2, \dots, l$, em que l é o número de neurônios na grade. Pode ser desejável manter a magnitude dos pesos pequena. Um outro modo de inicializar o algoritmo é selecionar os vetores de peso $\{w_j(0)\}_{j=1}^l$ a partir do conjunto disponível dos vetores de entrada $\{x_i\}_{i=1}^N$ de uma maneira aleatória.

2. *Amostragem*: Retirada de uma amostra x do espaço de entrada com uma certa probabilidade; o vetor x representa o padrão de ativação que é aplicado à grade. A dimensão do vetor x é igual a m .

3. *Casamento por Similaridade*: Encontro do neurônio com o melhor casamento (vencedor) $i(x)$ no passo de tempo n usando o critério da mínima distância euclidiana:

$$i(x) = \arg \min_j \|x(n) - w_j\|, j=1,2,\dots,l . \quad (1)$$

4. *Atualização*: Ajuste os vetores de peso sináptico de todos os neurônios usando a fórmula de atualização

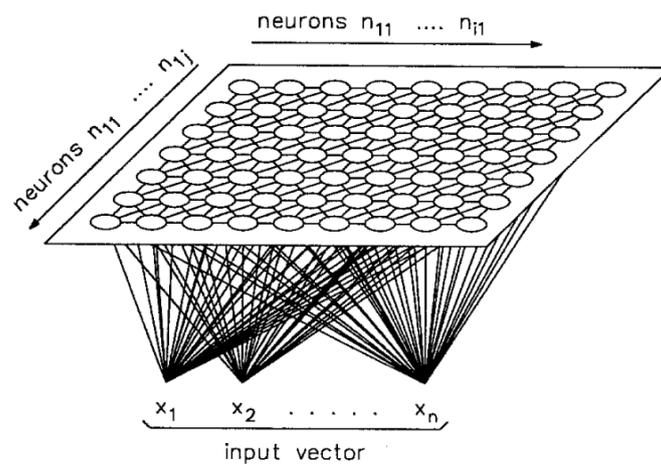
$$w_j(n+1) = w_j(n) + \eta(n) h_{j,i(x)}(n) (x(n) - w_j(n)), \quad (2)$$

em que $\eta(n)$ é o parâmetro da taxa de aprendizagem e $h_{j,i(x)}(n)$ é a função de vizinhança centrada em torno do neurônio vencedor $i(x)$; ambos $h(n)$ e $h_{j,i(x)}(n)$ são variados dinamicamente durante a aprendizagem para obter melhores resultados.

5. *Continuação*: Continua-se com o passo 2, até que não sejam observadas modificações significativas no mapa de características.

Em suma, o mapa auto-organizável proposto por Kohonen é uma rede neural engenhosa construída em torno de uma grade uni ou bidimensional de neurônios para capturar as características importantes contidas em um espaço de entrada (dados) de interesse. Dessa forma, ele fornece uma representação estrutural dos dados de entrada pelos vetores de peso dos neurônios como protótipos, como apresentado no Figura 2 (LOHNINGER, 1992).

Figura 2: Mapa de recurso auto-organizável bidimensional.



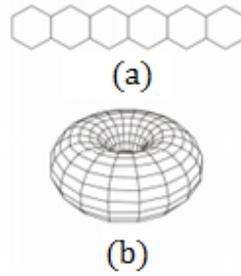
Fonte: LOHNINGER (1992).

O algoritmo SOM é inspirado na neurobiologia, incorporando todos os mecanismos que são básicos para a auto-organização: competição, cooperação e auto-amplificação, como abordado anteriormente. Ele pode servir como modelo genérico para descrever a emergência dos fenômenos de ordenação coletiva em sistemas complexos após iniciar a partir de uma desordem total (HAYKIN, 1999).

2.3.2 TOPOLOGIA

Na rede SOM, os dados são redimensionados segundo a topologia e dimensões de rede. A topologia permite preservar a estrutura dos dados de entrada e padrões próximos no conjunto de dados de amostra que estão associados a determinados neurônios da grade. Existem diferentes topologias para a arquitetura de mapas auto-organizáveis, sendo a mais comum a topologia bidimensional, a qual é subdividida em unidimensional (a) e toroidal (b), como apresentado na Figura 3.

Figura 3: a) SOM unidimensional. b) SOM toroidal.



Fonte: SILVA et al. (2011).

Há três processos topológicos (três fases) essenciais envolvidos na formação do mapa auto-organizável como explicados na seção 3.2.1. Na primeira fase (competitiva), cada x_k é apresentado a todos os neurônios em busca do melhor Math Unit (BMU), usando a distância euclidiana como uma referência de medição do recurso de distância (SILVA et al., 2011).

Na segunda fase (cooperativa), uma relação entre o BMU e os outros neurônios é definida para uma atualização suavizada de cada w_j . Finalmente, na última fase (adaptativa), todos os covetores w_j serão atualizados usando algum tipo de regra de adaptação, em geral sequencial ou em lote, conforme as Equações 3 e 4 para a regra de atualização em lote usada neste trabalho (VESANTO, 1999)

$$s_i(t) = \sum_j^{n_{V_i}} x_j \quad (3)$$

$$w_i(t+1) = \frac{\sum_j^m h_{ji}(t) s_j(t)}{\sum_j^m n_{V_j} h_{ji}(t)} \quad (4)$$

em que s_i representa uma soma de padrão de entrada para a i -ésima região de Voronoi, V_i ; e n_{V_i} é o número de amostras para o conjunto de dados do i -ésimo neurônio de Voronoi.

Na terceira fase (adaptação sináptica), é necessário realizar a adaptação dos pesos em relação à entrada, de forma iterativa. A Equação 5 corresponde à equação utilizada para realizar a adaptação.

$$w_{ji}(t+1) = w_{ji}(t) + \alpha(t) h_{ij}(t) [x_{ik}(t) - w_{ij}(t)] \quad (5)$$

em que $\alpha(t)$, $\alpha(0) < 1$, representa a função de taxa de aprendizagem. As funções de taxa de aprendizagem $\alpha(t)$, $h_{ij}(t)$ e $\delta(t)$ integram a categoria de funções monotonicamente decrescentes (SILVA, 2011).

O SOM cria um espaço de saída discreto, onde relações topológicas dentro das vizinhanças do espaço de entrada são preservadas. A seleção dos parâmetros de entrada é fundamental para preservar a topologia, e para a ordenação dessa topologia existem dois métodos de análise. O primeiro corresponde à análise através da definição de vizinhança – fase ordenação topológica por pesos; e a segunda ocorre de acordo com a distribuição de entrada – fase de convergência (HAYKLIN, 2001).

A fase de ordenação, segundo Hayklin (2001), realiza interseções com a vizinhança em que o número de iterações (T) decresce com o raio (d), em que N_0 corresponde ao valor inicial definido por:

$$d(t) = 1 + (N_0 - 1)\left(1 - \frac{t}{T}\right) \quad (3)$$

Na equação supracitada, é possível verificar que o tamanho da vizinhança decai linearmente de um valor inicial até chegar a 1, uma vez que $t \leq T$. De maneira similar, a taxa de aprendizagem $\eta(t)$ também decai linearmente de um valor inicial (η_0) até um final (η_T) para permitir que a rede se autoajuste, de acordo com a Equação 4.

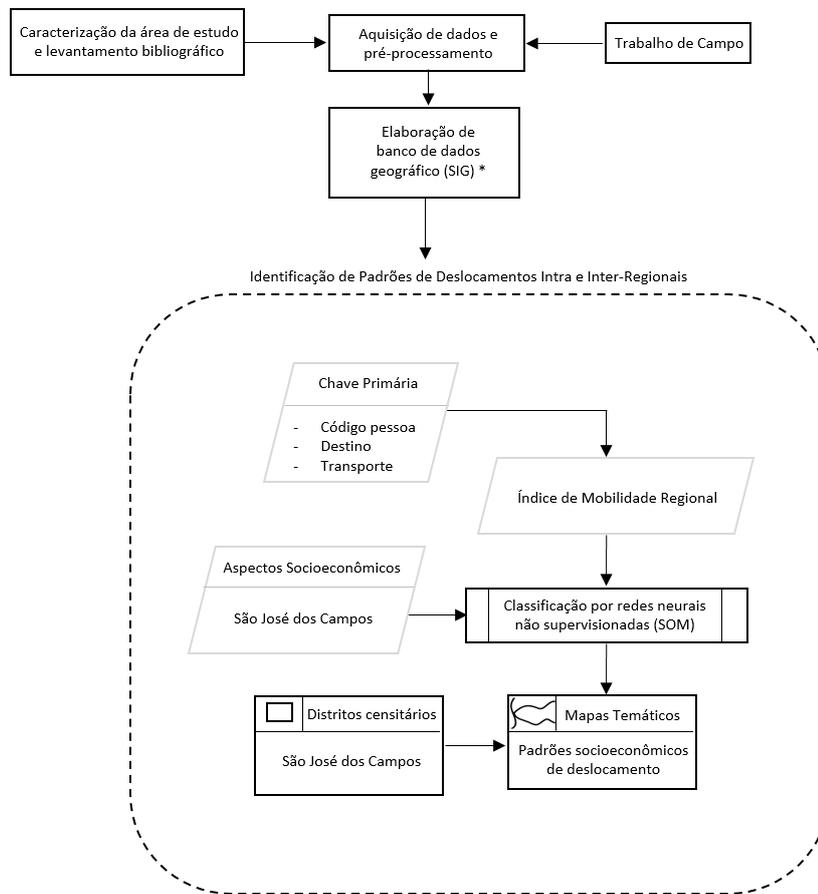
$$\eta(t) = \eta_T + (\eta_0 - \eta_T)\left(1 - \frac{t}{T}\right) \quad (4)$$

A fase de convergência corresponde à escolha de números de neurônios de forma experimental. O número de saídas afeta a precisão do mapeamento, aumentando o número de neurônios, sua resolução e, conseqüentemente, o tempo de treinamento. Em sua maioria, são geradas a partir de uma rede para representar uma grade de pontos (HAYKLIN, 2001).

3. METODOLOGIA

O OMT-G da metodologia desenvolvida para o trabalho está apresentado na Figura 4.

Figura 4: Fluxograma OMT-G.



Fonte: Autor.

3.1 AMOSTRAGEM ESTRATIFICADA POR SETORES CENSITÁRIOS DE SÃO JOSÉ DOS CAMPOS

Todos os dados sobre mobilidade intra-regional relativos ao município de São José dos Campos foram obtidos por pesquisa de campo em formato de questionários. Foram coletados um total de 4.260 questionários no período de agosto de 2005 até abril de 2006, com uma média de duas entrevistas por quadra de tamanho pequeno ou médio, e de três a seis no caso de quadras de grandes extensões (SANTOS et al., 2006) e repassados para o desenvolvimento do presente trabalho.

Optou-se pela aplicação de questionários origem-destino (O-D) junto à população de São José dos Campos. Como um dos interesses da pesquisa era cruzar as informações de mobilidade regional com dados socioeconômicos obtidos do censo de 2000, toda a

coleta de dados foi feita por amostragem estratificada com base na delimitação dos setores censitários definidos para a cidade (área urbanizada do município) em 2000.

Os questionários O-D regionais foram estruturados em três seções:

I) Dados Pessoais – contendo informações de cunho pessoal, bem como dados de escolaridade e nível de ocupação;

II) Dados sobre Percurso de Vida – contendo informações sobre local de nascimento, migração intra e inter-urbana e respectivas razões;

III) Dados sobre Deslocamentos Regionais – compreendendo informações sobre deslocamentos em função da finalidade (estudo, trabalho, comércio, serviços, lazer e segunda residência ou residência de veraneio), da frequência (diária ou pendular, até 3 vezes/semana, semanal em dia útil ou em fim-de-semana, quinzenal, mensal, bimestral etc.) e da modalidade de transporte (veículo particular, ônibus interurbano ou de turismo, van, veículo de empresa, ônibus fretado, aeronave etc.).

Na condução dos questionários, o procedimento era selecionar de forma aleatória, em cada quadra, uma pessoa residente (ou que trabalhasse) com padrão socioeconômico semelhante à média da quadra. Dentro de um mesmo setor, procurava-se variar a gama de sexo, idade e nível de ocupação dos entrevistados, visando-se evitar enviesamento dos dados (SANTOS et al., 2006).

3.2 MOBILIDADE REGIONAL

3.2.1 GERAÇÃO DO ÍNDICE DE MOBILIDADE REGIONAL

O Índice de Mobilidade Regional (IMR), em uma tentativa de simplificação conceitual, e conseqüentemente matemática, se utiliza apenas de dois parâmetros: finalidade e frequência dos deslocamentos. Era preciso que o índice expressasse o padrão subliminar do agente do deslocamento, reconhecendo que algumas finalidades implicam maior sofisticação que outras, e sobretudo, que a maior frequência do deslocamento era indicativa de um maior poder de mobilidade (SANTOS et al., 2006).

A análise de padrões de mobilidade regional foi conduzida em associação com padrões socioeconômicos através do uso de redes neurais não-supervisionadas (SOM). Para tanto, foi necessário que toda a informação levantada sobre o perfil dos deslocamentos fosse condensada em único número. Neste sentido, foi concebido o “Índice de Mobilidade Regional – IMR”, calculado de forma agregada para cada setor

cenitário, uma vez que os próprios dados socioeconômicos utilizados na pesquisa são os produzidos pelo IBGE por setor censitário.

O IMR foi desenvolvido, obedecendo a uma tabela de frequência e finalidades de deslocamento.

Tabela 1: Finalidade do deslocamento.

<i>FINALIDADE DO DESLOCAMENTO</i>	<i>VALOR CORRESPONDENTE</i>
<i>Estudo</i>	<i>4</i>
<i>Trabalho</i>	<i>4</i>
<i>Comércio</i>	<i>5</i>
<i>Serviços</i>	<i>4,5</i>
<i>Lazer/Recreação</i>	<i>3</i>
<i>Veraneio (Segunda Residência)</i>	<i>2</i>
<i>Ausência de Deslocamento</i>	<i>0</i>

Fonte: SANTOS et al. (2006).

Tabela 2: Frequência do deslocamento.

<i>FREQÜÊNCIA DO DESLOCAMENTO</i>	<i>VALOR CORRESPONDENTE</i>
<i>Diária ou Pendular</i>	<i>100</i>
<i>Até 3 vezes por semana</i>	<i>80</i>
<i>Semanal em Dia Útil</i>	<i>70</i>
<i>Semanal em Fim-de-Semana</i>	<i>50</i>
<i>Quinzenal</i>	<i>35</i>
<i>Mensal</i>	<i>25</i>
<i>Bimestral</i>	<i>15</i>
<i>Outros</i>	<i>10</i>
<i>Ausência de Deslocamento</i>	<i>1</i>

Fonte: SANTOS et al. (2006).

Sendo assim, o índice confere maior peso a deslocamentos para fins de comércio, porque a pessoa que se desloca a partir de São José dos Campos para compras via de regra o faz para fins de comércio especializado, comumente em São Paulo. As demais finalidades, contudo, recebem um peso apenas ligeiramente inferior, porque não expressam menor poder de mobilidade. Por outro lado, a frequência dos deslocamentos é de fato um fator preponderante na qualificação do maior ou menor poder de mobilidade regional e, por este motivo, os valores numéricos que foram associados aos tipos de

frequência situam-se em um intervalo maior, [1, 100], e deste modo, procuram diferenciar sobremaneira os valores do IMR.

3.2.2 IDENTIFICAÇÃO DE *CLUSTERS* DO IMR EM ASSOCIAÇÃO COM DADOS SOCIOECONÔMICOS

Os *clusters* do IMR e dos dados socioeconômicos foram identificados a partir da seleção de 11 variáveis obtidas a partir do censo do IBGE 2000 para São José dos Campos, como apresentado na Tabela 3.

Tabela 3: Variáveis socioeconômicas simples e agregadas para a análise de *clusters*.

COMBINAÇÃO DE VARIÁVEIS ISOLADAS	VARIÁVEL AGREGADA CORRESPONDENTE
<i>(Rendimento mensal de mulheres chefes de família) / (Nº de mulheres chefes de família*)</i>	<i>Renda Média Mensal de Mulheres Chefes de Família</i>
<i>(Rendimento mensal de homens chefes de família) / (Nº de homens chefes de família*)</i>	<i>Renda Média Mensal de Homens Chefes de Família</i>
<i>(Rendimento mensal de chefes de família) / (Nº de chefes de família*)</i>	<i>Renda Média Mensal de Moradores no Setor</i>
<i>(Nº de chefes de família com rendimento até 0,5 SM**) / (Nº de chefes de família*)</i>	<i>Percentual de Chefes de Família com Renda Média Mensal até 0,5 SM</i>
<i>(Nº de chefes de família com rendimento de 0,5 a 1 SM**) / (Nº de chefes de família*)</i>	<i>Percentual de Chefes de Família com Renda Média Mensal de 0,5 a 1 SM</i>
<i>(Nº de chefes de família com rendimento de 1 a 2 SM**) / (Nº de chefes de família*)</i>	<i>Percentual de Chefes de Família com Renda Média Mensal de 1 a 2 SM</i>
<i>(Nº de chefes de família com rendimento > 20 SM**) / (Nº de chefes de família*)</i>	<i>Percentual de Chefes de Família com Renda Média Mensal > 20 SM</i>
<i>(Nº de chefes de família sem rendimento) / (Nº de chefes de família*)</i>	<i>Percentual de Chefes de Família sem Rendimento</i>
<i>(Nº de chefes de família sem instrução ou com menos de 1 ano de estudo) / (Nº de chefes de família*)</i>	<i>Percentual de Chefes de Família sem Instrução ou com Menos de Um Ano de Estudo</i>
<i>(Nº de chefes de família com 17 anos de estudo ou mais) / (Nº de chefes de família*)</i>	<i>Percentual de Chefes de Família com Dezesete Anos de Estudo ou Mais</i>

* com e sem rendimento.

** SM = salário mínimo.

Fonte: SANTOS et al. (2006).

Essas variáveis, juntamente com o IMR, foram normalizadas através da subtração de cada índice da média geral (obtida em relação aos 729 setores censitários) e da subsequente divisão pelo respectivo desvio padrão, como apresentado na Figura 5.

Figura 5: Variáveis normalizadas.

11												
1.042058	1.09115	0.918509	-0.43078	-0.17804	-0.72158	-0.58798	0.915119	-0.9836	0.326349		-0.50648882	354990405000001
0.235162	0.158407	0.106081	-0.43078	-0.1151	-0.67229	-0.47869	-0.03186	-0.74347	-0.19174		-0.70157614	354990405000002
-0.29147	-0.21068	-0.29164	-0.43078	0.345162	0.350655	1.065373	-0.20494	-0.43707	-0.4361		-0.23181069	354990405000003
-0.65346	-0.58258	-0.61565	3.384569	2.585046	0.712039	0.756309	-0.43307	1.173211	-0.28676		2.31727044	354990405000005
-0.09804	-0.12135	-0.18681	-0.43078	0.481889	-0.79006	0.990731	-0.14885	-0.25716	-0.52113		-0.68007300	354990405000006
0.079288	0.018684	-0.05492	-0.43078	0.245155	-0.31145	0.1659	0.190926	-0.42461	-0.01612		-0.22336696	354990405000007
0.012702	-0.19123	-0.18962	-0.43078	0.587448	-0.60487	-0.04287	-0.06081	-1.12315	0.40843		-0.94875705	354990405000008
0.386996	0.27191	0.170521	-0.43078	-0.4034	-0.7147	0.521127	0.177599	-0.58051	-0.05037		-1.03953221	354990405000009
0.178363	0.252601	0.151437	-0.43078	-0.60305	-0.22876	-0.52597	0.216138	0.251143	-0.38866		1.59189741	354990405000010
0.506035	0.227096	0.236481	-0.43078	-0.65576	-1.00078	-0.48892	0.06697	-0.98267	-0.03363		-0.52210888	354990405000011
0.123256	-0.10696	-0.11107	-0.43078	0.538619	-0.37183	-0.47299	-0.18467	-0.2886	-0.34013		-1.01516245	354990405000012
-0.0455	-0.37473	-0.34439	-0.43078	-0.5227	0.629645	0.023664	-0.40978	-0.60919	-0.07525		-0.39938118	354990405000013
-0.29132	-0.37608	-0.38964	0.729444	0.86504	0.178645	-0.49889	-0.36149	-0.507	-0.41423		-0.74993756	354990405000014
-0.37582	-0.43605	-0.43123	-0.43078	0.391299	0.443668	0.618391	-0.34889	0.260397	-0.4288		-1.01516245	354990405000015
-0.4494	-0.55478	-0.54077	-0.43078	1.011271	2.25732	-0.09251	-0.47757	-0.06953	-0.40688		-1.01516245	354990405000016
-0.75953	-0.78296	-0.77839	-0.43078	1.947307	0.803341	2.651784	-0.51989	2.802117	-0.52113		-0.97326316	354990405000017
-0.49576	-0.62863	-0.62322	0.786388	0.411584	1.400443	1.105892	-0.51989	1.333152	-0.40898		-1.01863623	354990405000018
-0.59134	-0.47482	-0.49187	-0.43078	0.839516	0.581782	1.215078	-0.39951	1.311893	-0.52113		-0.46688984	354990405000019
-0.47909	-0.44819	-0.51064	2.857587	3.359857	1.18368	-0.30482	-0.40765	2.136682	-0.31913		-0.94862822	354990405000020
-0.2476	0.021623	-0.02596	1.397775	-0.17248	0.332948	-0.59188	-0.05182	-0.24918	0.068589		-0.67473242	354990405000021
-0.22446	-0.3836	-0.38898	-0.43078	1.25608	0.255531	-0.37241	-0.41571	0.605874	-0.52113		0.85245621	354990405000022
0.023141	-0.01655	-0.05196	-0.43078	0.832168	-0.70729	1.077381	0.177775	0.026258	0.697628		-0.68598467	354990405000023
-0.3964	-0.45705	-0.47053	-0.43078	0.366624	1.086785	-0.08058	-0.41882	-0.17961	-0.52113		-0.44686511	354990405000024
-0.30062	-0.351	-0.36561	0.248667	0.73101	-0.0584	0.115389	-0.35756	0.753167	-0.14549		0.24592820	354990405000025
2.004221	1.658675	1.704801	-0.43078	-1.09039	-1.5705	-1.11229	2.418662	-1.12315	1.720929		-0.49735167	354990405000027
1.451636	1.816116	1.747623	-0.43078	-1.07826	-1.5705	-0.95939	2.489641	-1.12315	1.140815		-0.74993272	354990405000028
2.043496	2.312402	2.183802	-0.43078	-0.32454	-0.83632	-0.95347	2.272459	-0.68867	1.36355		0.86304774	354990405000029
1.004845	0.780828	0.708093	-0.43078	-0.5625	-0.84009	-0.336	0.652092	-1.12315	0.650753		0.86304774	354990405000030
0.390378	0.288574	0.255565	0.25572	-0.38098	-0.53554	-0.90231	0.229891	-0.53983	0.048183		-0.67504447	354990405000035
1.401356	1.811272	1.669768	-0.43078	0.099548	-0.80884	-0.56283	2.232141	-0.55971	1.238576		-0.89137642	354990405000036
0.47775	0.103477	0.138736	0.203081	-0.56474	-0.56939	-0.88332	0.020962	-0.92118	0.121344		1.79266881	354990405000037
0.011269	-0.09812	-0.12144	-0.43078	0.22694	-0.15453	-0.49889	-0.2031	-0.26053	-0.41423		-0.44309133	354990405000038
-0.27032	-0.22189	-0.25955	-0.43078	0.433827	0.078212	-0.744	-0.30216	0.028722	0.066693		-0.69489457	354990405000039
0.25363	0.280585	0.205558	-0.43078	-0.21151	-0.33645	-0.88741	0.189035	-0.43848	0.407837		0.84151530	354990405000041

Fonte: SANTOS et al. (2006).

3.2.3 MANIPULAÇÃO DOS NEURÔNIOS A PARTIR DO SOMPAK

No trabalho de Santos (2006), o *software* SOMPAK foi alimentado com as 11 variáveis em formato texto. Inicialmente, foi solicitado ao programa que produzisse um resultado preliminar, agrupando os mais de 700 setores censitários de São José dos Campos em 25 *clusters* ou neurônios, de acordo com afinidades em termos de padrões socioeconômicos e de mobilidade regional.

Esses neurônios foram reagrupados visualmente, de modo a se obter apenas seis grandes grupos ou *clusters* de setores censitários. O reagrupamento obedeceu a critérios de proximidade (ou similaridade) entre os neurônios, correspondente a tons de cinza mais claros. Após esse reagrupamento, por meio da identificação dos neurônios por códigos de coluna-linha e de rotinas no programa Excel, os seis grupos finais foram mapeados no aplicativo TerraView7 sobre o mapa de setores censitários de São José enquanto “classes de padrões socioeconômicos e de mobilidade regional” (SANTOS et al., 2006).

Neste trabalho, foi utilizado o *software* R para melhor visualização e validação da escolha dos valores de *clusters* para a análise.

3.2.4 MANIPULAÇÃO DOS NEURÔNIOS A PARTIR DO R STUDIO

Para iniciar o desenvolvimento do SOM no *software* R Studio, é necessário instalar o *software* R. Foi utilizado o RStudio Desktop 1.4.1717. Primeiramente, foram instalados os pacotes necessários para o desenvolvimento do SOM, sendo eles: o pacote *kohonen* para o desenvolvimento da metodologia utilizada, pacote para a manipulação dos dados, pacote estatístico de distância, pacote de algoritmo hierárquico, pacote para plotar os vetores de pesos do SOM e pacote de paleta de cores.

Importou-se a planilha de dados do IMR das 11 variáveis normalizadas com a verificação do delimitador *double*, e se iniciou a execução do *somgrid* com duas matrizes (*x* e *y*) (Figura 6) de valor 5 cada uma, para criação de 25 neurônios, definindo também a forma da matriz como “retangular” e análise de vizinhança *bubble*.

Figura 6: Dados de entrada para elaboração do SOM no R Studio.

x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12
1.042058151	1.091149882	0.918508935	-0.4307786	-0.17803503	-0.72157812	-0.587976935	0.915119272	-0.983597349	0.326349329	0.21124152	3.5499e+14
0.235162401	0.158406777	0.106081409	-0.4307786	-0.11510448	-0.67229124	-0.478687521	-0.031856342	-0.743465768	-0.191738477	0.15674793	3.5499e+14
-0.291471604	-0.210682033	-0.291642079	-0.4307786	0.34516158	0.35065461	1.065372711	-0.204937388	-0.437069216	-0.436104908	0.28796715	3.5499e+14
-0.653455538	-0.582582667	-0.615647564	3.3845692	2.58504591	0.71203931	0.756308928	-0.433074794	1.173210737	-0.286757383	1.000000000	3.5499e+14
-0.098037021	-0.121347715	-0.186807710	-0.4307786	0.48188854	-0.79006058	0.990731111	-0.148850267	-0.257157553	-0.521134886	0.16275439	3.5499e+14
0.079288430	0.018683508	-0.054923146	-0.4307786	0.24515463	-0.31144724	0.165899563	0.190926130	-0.424614706	-0.016122587	0.29032573	3.5499e+14
0.012702260	-0.191233724	-0.189616659	-0.4307786	0.58744834	-0.60487105	-0.042865288	-0.060807381	-1.123150257	0.408430125	0.08770308	3.5499e+14
0.386995933	0.271910058	0.170521384	-0.4307786	-0.40340035	-0.71469605	0.521126816	0.177599366	-0.580511052	-0.050368056	0.06234692	3.5499e+14
0.178363295	0.252601016	0.151436956	-0.4307786	-0.60305211	-0.22876474	-0.525965230	0.216138078	0.251142512	-0.388660645	0.79738212	3.5499e+14
0.506035420	0.227096011	0.236481143	-0.4307786	-0.65576470	-1.00077980	-0.488921755	0.066970097	-0.982666996	-0.033629680	0.20687838	3.5499e+14
0.123255757	-0.106958476	-0.111072657	-0.4307786	0.53861909	-0.37182889	-0.472987762	-0.184668280	-0.288596232	-0.340130478	0.06915411	3.5499e+14
-0.045497201	-0.374731679	-0.344387368	-0.4307786	-0.52269755	0.62964513	0.023663576	-0.409783408	-0.609187107	-0.075245978	0.24115981	3.5499e+14
-0.291315892	-0.376082957	-0.389643745	0.7294441	0.86503977	0.17864541	-0.498886666	-0.361491857	-0.506995604	-0.414225849	0.14323917	3.5499e+14
-0.375820712	-0.436046847	-0.431234025	-0.4307786	0.39129903	0.44366777	0.618391300	-0.348892153	0.260397009	-0.428804354	0.06915411	3.5499e+14
-0.449401180	-0.554776252	-0.540770396	-0.4307786	1.01127097	2.25731988	-0.092505115	-0.477566636	-0.069525801	-0.406875853	0.06915411	3.5499e+14
-0.759527583	-0.782964186	-0.778385612	-0.4307786	1.94730705	0.80334097	2.651783910	-0.519888144	2.802117324	-0.521134886	0.08085781	3.5499e+14
-0.495760642	-0.628634039	-0.623223641	0.7863876	0.41158441	1.40044262	1.105892041	-0.519888144	1.333152156	-0.408978780	0.06818378	3.5499e+14
-0.591336769	-0.474823172	-0.491873915	-0.4307786	0.83951578	0.58178228	1.215078127	-0.399506966	1.311892930	-0.521134886	0.22230267	3.5499e+14
-0.479088090	-0.448186868	-0.510639412	2.8575875	3.35985715	1.18368048	-0.304818637	-0.407654311	2.136682314	-0.319129966	0.08773907	3.5499e+14
-0.247596953	0.021622879	-0.025956036	1.3977752	-0.17248498	0.33294809	-0.591877201	-0.051816168	-0.249176054	0.068589153	0.16424617	3.5499e+14
-0.224461772	-0.383604633	-0.388982588	-0.4307786	1.25608041	0.25553102	-0.372408942	-0.415712124	0.605874492	-0.521134886	0.59083458	3.5499e+14
0.023140938	-0.016549428	-0.051962582	-0.4307786	0.83216797	-0.70728647	1.077381167	0.177775499	0.026258241	0.697628128	0.16110308	3.5499e+14
-0.396396953	-0.457045429	-0.470531183	-0.4307786	0.36662353	1.08678493	-0.080583474	-0.418821856	-0.179605968	-0.521134886	0.22786823	3.5499e+14
-0.300617995	-0.351003732	-0.365612255	0.2486669	0.73100968	-0.05839660	0.115388706	-0.357559073	0.753167268	-0.145488751	0.42141358	3.5499e+14

Fonte: Autor.

Os dados de entrada foram transformados em matriz (*.asmatrix*) por causa do tipo de dado que é aceito no pacote *kohonen*. Com a finalidade de manter estaticamente a entrada dos dados no programa, ou seja, de forma não aleatória, utilizou-se a função *set.seed(x)*.

Para realizar o *plot* principal para visualizar os padrões dos neurônios, que mostrasse a quantidade de amostras associadas a cada neurônio, utilizou-se o *som_cluster* em formato de linha.

Para a entrada do pacote com a função de parametrização, o qual é uma extensão de mapas auto-organizáveis para múltiplas camadas de dados, que possuem diferentes tipos de variáveis (com mesmo número de objetos), foi calculada a distância para se determinar as unidades vencedoras, utilizando-se o número de vezes em que o conjunto de dados completo seria apresentado à rede, como apresentado na Figura 7.

Figura 7: Código para iniciar treinamento do SOM.

```
som_cluster <- supersom(  
  data = input_data,  
  grid = somgrid,  
  rlen = 1000,  
  keep.data = TRUE  
)
```

Fonte: Autor.

Utilizou-se o *codeRendering* para realizar a renderização dos neurônios e plotar os subgrupos separados por agrupamento padrão (25). Dessa forma, foi permitida a visualização dos conjuntos de neurônio interligados.

Para visualização do *plot* dos vetores de peso e das configurações internas de cada neurônio, foi realizada a operação *.plot.zoo*.

A representação diagramática dos neurônios foi realizada através do diagrama que representa uma árvore: o dendrograma, através da operação *.asdendrogram*.

Foi realizada a separação dos neurônios por cores, a fim de se visualizar os grupos separados por agrupamento hierárquico, através da função *codeRendering* associada ao tipo de retorno (linha), árvore e paleta de cor; e utilizado o modo de separação de subgrupos através da função *kohonen*.

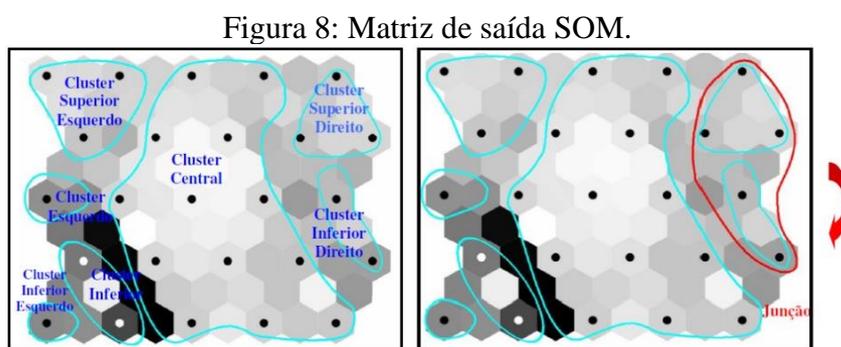
Através da função *cbind*, foi realizada a visualização da saída como um *tibble* com o ID do neurônio associado à alocação do respectivo neurônio.

Criou-se também um *set* dos neurônios associados a cada *cluster*, através do *data.frame*. Para exploração por cluster individual, foi desenvolvido o código por grupo de *cluster*, o qual retorna os neurônios inseridos nele.

4. RESULTADOS E DISCUSSÕES

4.1 ANÁLISE DO IMR ATRAVÉS DO ALGORÍTMO SOM (SOMPAK)

O processamento dos dados socioeconômicos e de mobilidade regional através do algoritmo SOM no ambiente SOMPAK produziu como saída uma matriz com 25 neurônios. A reclassificação destes neurônios em seis grupos não foi imediata. A princípio, delinearam-se sete grupos, e foi verificado que haveria três possíveis junções entre o *cluster* central, o superior direito e o inferior direito (Figura 8).

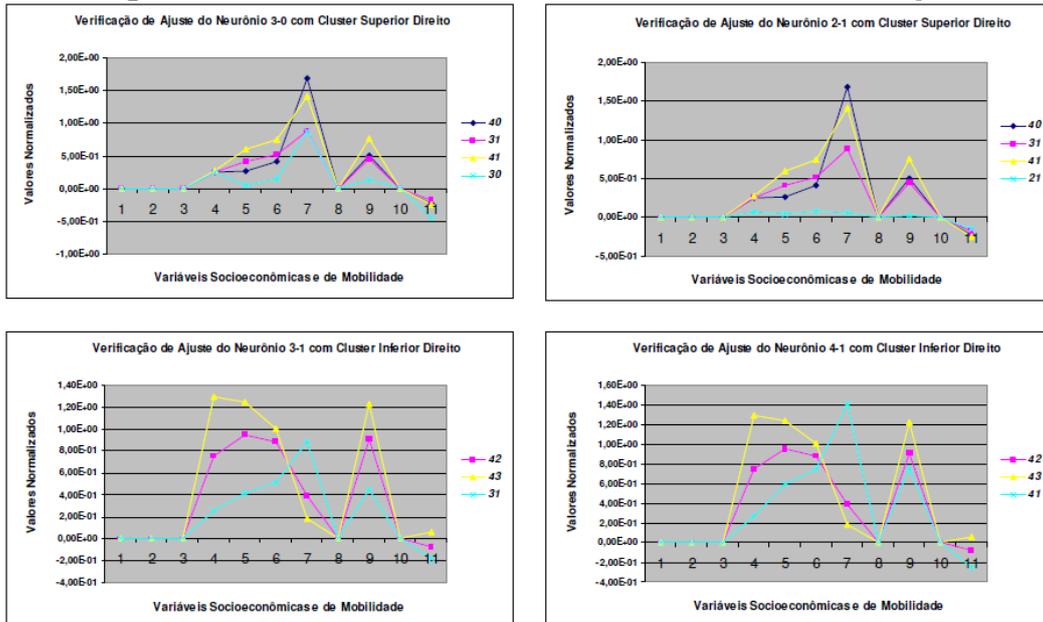


Fonte: SANTOS et al. (2006).

Uma das possibilidades (*cluster* central e *cluster* inferior direito) foi descartada inicialmente, visto que os hexágonos de fronteira entre esses dois agrupamentos possuíam tons de cinza demasiado escuros. Assim sendo, para se escolher entre (i) a junção do *cluster* central com o superior direito, ou (ii) a junção do *cluster* superior direito com o inferior direito, recorreu-se a uma análise de gráficos, nos quais se plotaram o comportamento da média de cada neurônio do *cluster* considerado em relação à média de cada um dos neurônios de fronteira do *cluster* a ser agregado.

A Figura 9 apresenta quatro desses gráficos e indica que a junção mais conflituosa seria entre o *cluster* superior direito e o neurônio 2-1 (coluna-linha), visto que a curva referente à média deste neurônio, em cor azul claro, para as onze variáveis utilizadas, apresenta um comportamento por demais anômalo em relação às demais curvas. Em vista disso, a solução foi a junção entre o *cluster* superior direito e o inferior direito.

Figura 9: Gráficos de análises de *clusters* com neurônios fronteiros.

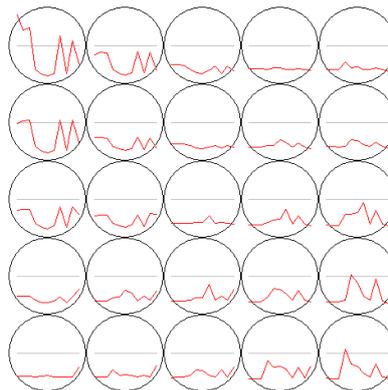


Fonte: SANTOS et al. (2006).

4.2 ANÁLISE DO IMR ATRAVÉS DO ALGORÍTMO SOM (R STUDIO)

A padronização das 11 variáveis normalizadas permitiu a visualização padrão dos neurônios de acordo com os dados do IMR como apresentado na Figura 10.

Figura 10: Visualização dos neurônios no R Studio.

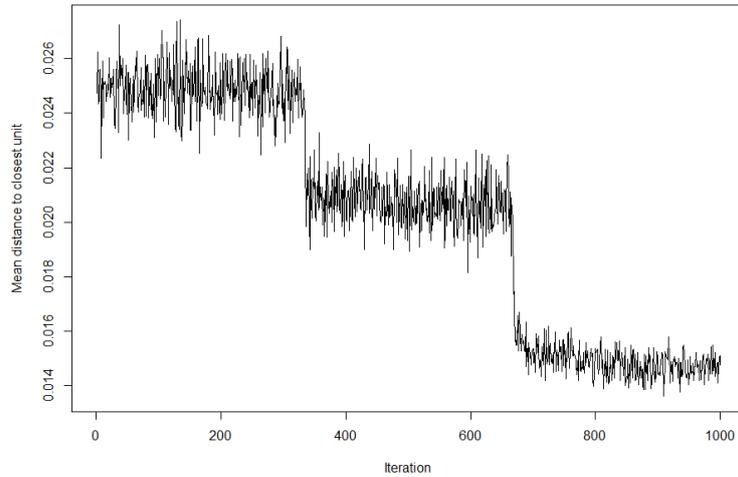


Fonte: Autor.

É possível visualizar que os dados de entrada foram agrupados por similaridade, ou seja, agrupamento dos setores censitários que possuem características semelhantes. Pelo fato de o SOM ser um algoritmo de redução de dimensionalidade, os parâmetros são representados por grupos semelhantes vizinhos.

Para verificar a convergência dos neurônios, a Figura 11 apresentou a interação entre os neurônios.

Figura 11: Distância média para as unidades próximas.

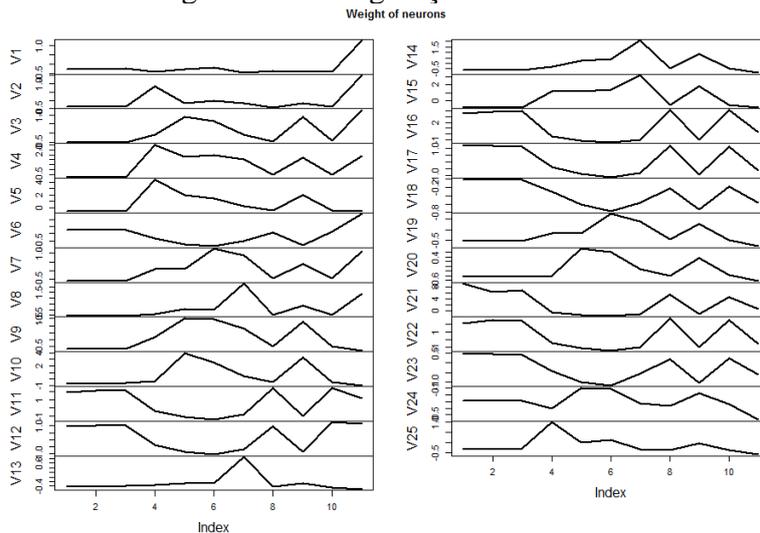


Fonte: Autor.

Calculando-se o valor de 1.000 iterações para o *rlen*, o mesmo correspondeu ao erro da distância, ou seja, quanto maior o valor, os dados são continuamente realocados, a fim de se obter mais similaridade e homogeneidade internamente aos neurônios.

No que diz respeito ao peso e às configurações internas de cada neurônio na visualização de forma gráfica, observou-se que o neurônio 21 está separado dos demais, como apresentado na Figura 12.

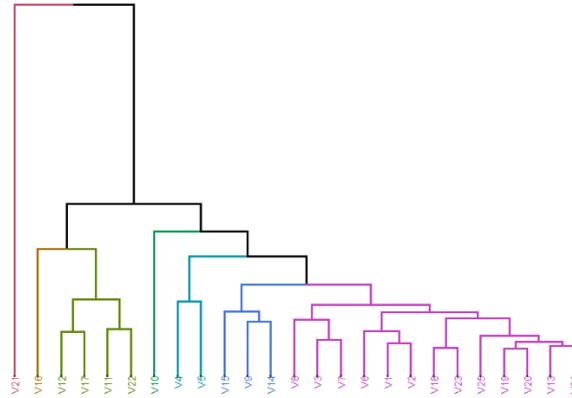
Figura 12: Configurações neuronais.



Fonte: Autor.

Para melhor visualização, a árvore de neurônios elaborada através do dendrograma encontra-se disposta na Figura 13.

Figura 13: Dendrograma para 7 clusters.

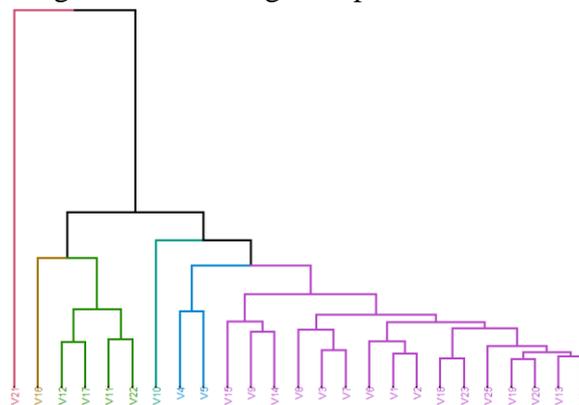


Fonte: Autor.

Dessa forma, é possível verificar que, apesar de estarem próximos, o neurônio 21 possui alguma variável que o separa dos demais neurônios, ou seja, possui alguma característica que não permite que esteja nas árvores 12, 17, 11 e 22, pois possui uma representação semelhante em seu peso e formato como observado na Figura 12.

Como observado por Santos (2006), esse neurônio apresenta um comportamento diverso, sendo necessário o desenvolvimento e a análise com 6 subgrupos. Assim, a Figura 14 apresenta o dendrograma dos subgrupos supracitados.

Figura 14: Dendrograma para 6 clusters.

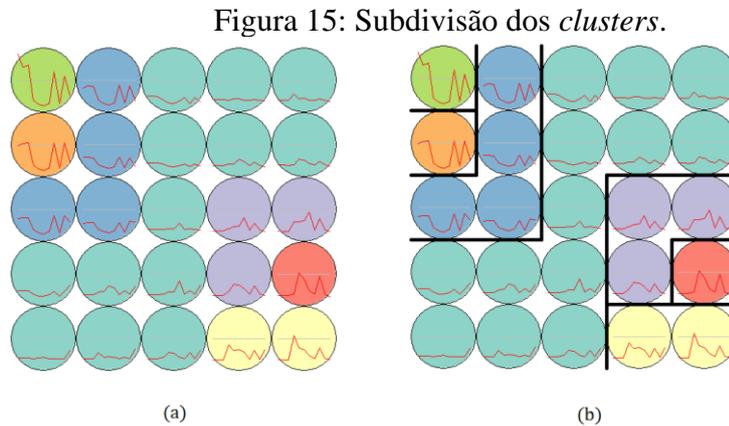


Fonte: Autor.

Com isso, é possível verificar que o neurônio 21 possui os parâmetros corretos, permanecendo na mesma posição e sem integrar-se a outro subgrupo. Observou-se também que os neurônios 9, 14 e 15 passaram a fazer parte do subgrupo de coloração

roxa – primeiro subgrupo do lado direito. Dessa forma, optou-se por continuar as análises com um total de 7 subgrupos.

A subdivisão dos *clusters* possibilitou a visualização padrão dos neurônios Figura 15 (a) e em função dos limites de Kohonen (b).



Fonte: Autor.

A criação dos neurônios em seus subgrupos como mostrado anteriormente permite a obtenção da saída dos neurônios e seus respectivos *clusters*. Na Figura 16, observa-se que os neurônios 1, 2 e 3 estão associados ao *cluster* 1, da mesma forma que o 4 e 5 estão associados ao *cluster* 2, e, assim, sucessivamente.

Figura 16: *Clusters* com neurônios associados.

	cluster	neuron		cluster	neuron
v1	1	1	v15	3	15
v2	1	2	v16	5	16
v3	1	3	v17	6	17
v4	2	4	v18	6	18
v5	2	5	v19	1	19
v6	1	6	v20	1	20
v7	1	7	v21	7	21
v8	1	8	v22	5	22
v9	3	9	v23	6	23
v10	4	10	v24	1	24
v11	5	11	v25	1	25
v12	6	12			
v13	1	13			
v14	3	14			

Fonte: Autor.

A saída final e principal da realização do SOM corresponde a uma tabela com as variáveis do ID do setor censitário e ID do neurônio, como apresentado na Figura 17.

Figura 17: Saída da Rede SOM.

id_data	id_setorcensitario	id_neuron
1	3,55E+18	17
2	3,55E+18	18
3	3,55E+18	13
4	3,55E+18	4
5	3,55E+18	13
6	3,55E+18	24
7	3,55E+18	18
8	3,55E+18	18
9	3,55E+18	1
10	3,55E+18	23
11	3,55E+18	24
12	3,55E+18	24
13	3,55E+18	25
14	3,55E+18	19
15	3,55E+18	20
16	3,55E+18	15

Fonte: Autor.

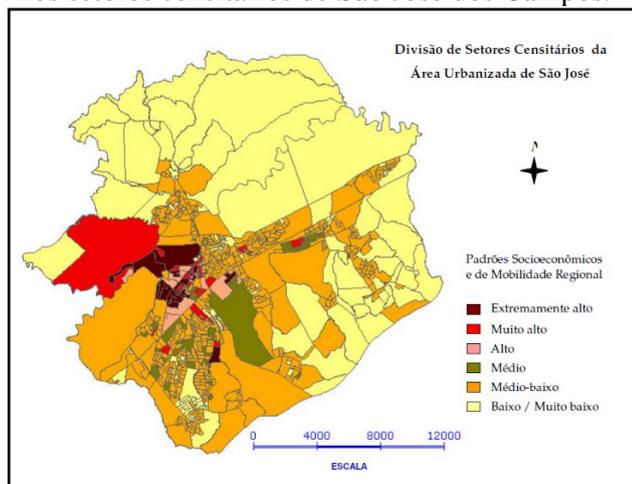
Essa saída final resultou no agrupamento dos 730 setores censitários de São José dos Campos em 25 neurônios, de acordo com afinidades em termos de padrões socioeconômicos e de mobilidade regional.

Assim, os parâmetros foram inseridos na tabela de atributos desses setores censitários para geração e espacialização dos mesmos.

4.1 ANÁLISE DA ESPACIALIZAÇÃO DOS SETORES CENSITÁRIOS

A espacialização destes seis grupos supracitados no mapa de setores censitários de São José dos Campos, feita através do aplicativo TerraView, está apresentada na Figura 18.

Figura 18: Espacialização das classes de padrão socioeconômico e de mobilidade regional nos setores censitários de São José dos Campos.



Fonte: Adaptado de Santos (2006), Almeida et al. (2007).

A cor amarelo-clara, referente ao padrão muito baixo ou baixo, corresponde aos bairros periféricos e zonas rurais do município, incluindo bairros como o Campo dos Alemães, a Chácara Araújo, Sol Nascente, entre outros, bem como as favelas do Banhado e da margem direita da Dutra, sentido São Paulo – Rio.

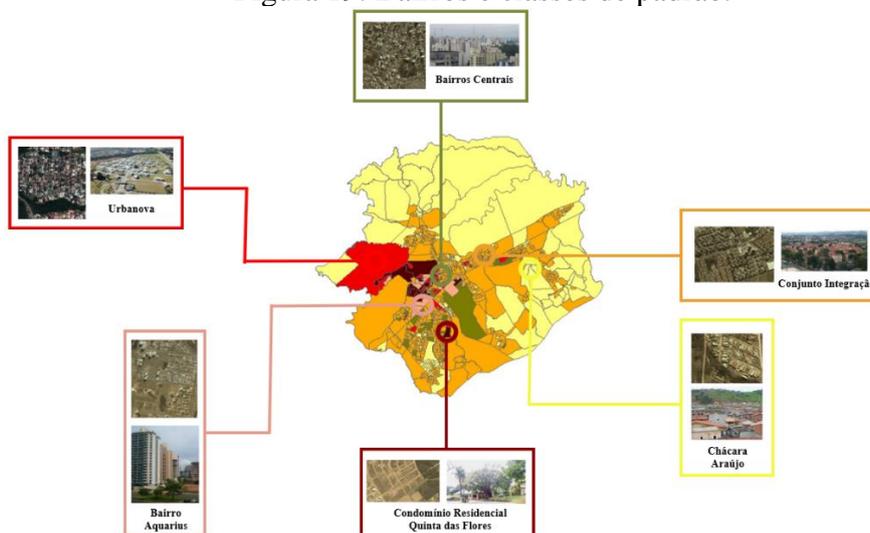
As áreas relativas ao padrão médio-baixo, em tom laranja, se referem aos bairros também periféricos de classe média-baixa, tais como Santana e Alto da Ponte ao norte; Vila Industrial, Jardim Diamante e Novo Horizonte à leste; Jardim Portugal, Jardim Madeira e Jardim Estoril ao sul; e Jardim das Indústrias e Jd. Por do Sol a oeste.

O padrão alto, em cor salmão, abrange setores localizados em bairros como Aquarius, Jardim Esplanada, Jardim Nova América, bem como setores da ala residencial de oficiais do CTA. A cor vermelha, referente ao padrão muito alto, corresponde a setores como o Condomínio Floresta, os Condomínios Urbanova I, III a VII, Condomínio Residencial Planalto, além de setores localizados nos bairros Floradas de São José, Jardim Residencial (vizinho ao Condomínio Quinta das Flores), Anchieta-Esplanada, Jardim Oswaldo Cruz, entre outros.

Por fim, o padrão extremamente alto, indicado com a cor marrom escuro, refere-se a setores como os Condomínios Urbanova II, Quinta das Flores, Chácara dos Eucaliptos, Esplanada do Sol, Bosque Imperial, Altos do Esplanada, Parque Residencial Aquarius, o bairro Serimbura, partes do Jd. Esplanada II, do Jd. Apolo I e II e da Vila Adyana, além da ala residencial de brigadeiros do CTA.

A Figura 19, apresenta Espacialização das classes de padrão socioeconômico e de mobilidade regional nos setores censitários de São José dos Campos.

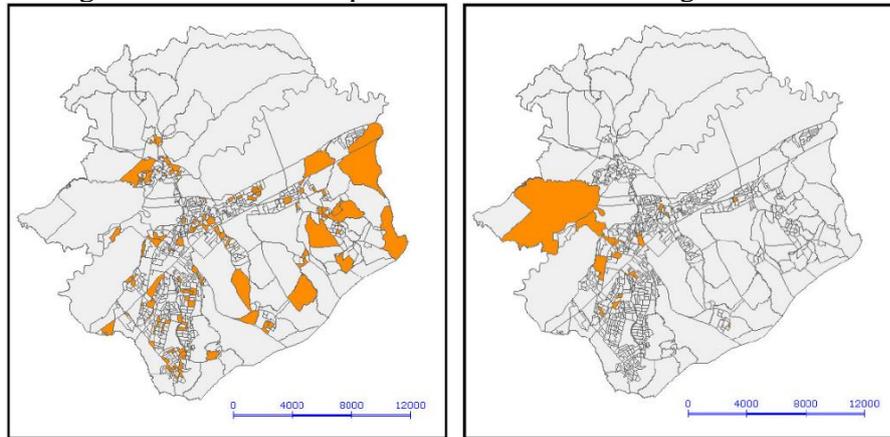
Figura 19: Bairros e classes de padrão.



Fonte: Adaptado de Santos (2006), Almeida et al. (2007).

Foram conduzidas consultas espaciais no TerraView, a fim de se verificar os setores censitários geradores de deslocamentos regionais específicos, indicativos de padrões socioeconômicos (Figura 20).

Figura 20: Consultas espaciais de deslocamento regional.



Fonte: Santos (2006).

Na consulta espacial referente aos deslocamentos para o sul de Minas, grande parte dos setores geradores (em cor laranja) correspondem quase que exclusivamente aos setores com padrão baixo/muito baixo e médio-baixo. Por sua vez, os deslocamentos a São Paulo referem-se majoritariamente aos setores com padrão médio, alto, muito alto e extremamente alto.

4. CONSIDERAÇÕES FINAIS

Este trabalho objetivou descrever e analisar quantitativamente os fenômenos correlatos de mobilidade intra-regional e urbanização dispersa, baseando-se em formulações e postulados teóricos acerca destes fenômenos encontrados na literatura científica no âmbito de urbanismo.

A metodologia se utilizou de ferramentas técnicas e metodológicas aportadas pelo Sensoriamento Remoto e pela Geoinformação.

A análise espacial da mobilidade regional juntamente com indicadores de renda e educacionais foi realizada através do método de mapas auto-organizáveis, que se constitui em uma modalidade de redes neurais não supervisionadas, revelando padrões coerentes de mobilidade e socioeconômicos da cidade de São José dos Campos.

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, C. M.; ALVES, C. D.; PEREIRA, M. N.; MIGLIORANZA, E.; TOMAS, L. R. Morphological Metrics and Unsupervised Neural Networks to Analyse Urban Sprawl and Intercity Commuting. In: XIII Simpósio Brasileiro de Sensoriamento Remoto, 2007, Florianópolis, SC. **Anais do XIII Simpósio Brasileiro de Sensoriamento Remoto**. São José dos Campos, SP: Instituto Nacional de Pesquisas Espaciais - INPE, 2007. v. 1. p. 5059-5066.

AVILA, C.F.D. O Brasil diante da dinâmica migratória intra-regional vigente na América Latina e Caribe: tendências, perspectivas e oportunidades em uma nova era. **Revista Brasileira de Política Internacional**. 2007. v. 50, n. 2. p. 118-128.

BEANINGER, R. Rotatividade migratória: um novo olhar para as migrações internas no Brasil. **Revista Interdisciplinar de Mobilidade Humana**, 2012. v. 20 n. 39. p. 77-100.

CAIADO, M. C. S. Deslocamentos intra-urbanos e estruturação socioespacial na metrópole paulista. **São Paulo em Perspectiva**. 2005. v. 19, n.4. p.64-77.

CARLINER, M. S. “Retracting suburbia: smart growth and the future of housing”, **Housing Policy Debate**. 1999. v. 10, n. 3, p. 549-553

CLARK, W. A. V. Kuijpers-Linde, M. Commuting in restructuring urban regions, **Urban Studies** (Special Issue: Globalisation, World Cities and the Randstad). 1994. v. 31, n. 3, p. 465-484.

EASTERBROOK, G. “Retracting suburbia: smart growth and the future of housing”, **Housing Policy Debate**, 1999. v. 10, n. 3, p. 541-547.

FERREIRA, A. H. B. Migrações internas e subdesenvolvimento: uma discussão. **Revista de Economia Política**. 1986. São Paulo, p. 98-124.

FISCHER, M. M. ABRAHART, R. J. Neurocomputing – tools for geographers. In: Openshaw, S.; Abrahart, R. J. ed. **GeoComputation**. New York: Taylor & Francis. 2000. p. 187- 217.

HAYKIN, S. S. **Neurais networks: a comprehensive foundation**. Upper Saddle River, NJ: Prentice Hall. 1999. 842p.

HAYKLIN, S.S. Mapas auto-organizáveis. (Ed.) **Redes Neurais Princípios e Práticas**. Bookman. 2001. cap. 9. p. 483-519.

HODGE, D. C. Urban congestion: reshaping urban life, **Urban Geography**. 1992. v. 13, p. 577-588.

LOHNINGER, H., STANEL. F. Comparing the performance of neural networks to well-established methods of multivariate data analysis: the classification of mass spectral data. **Fresenius J Anal Chem**. 1992. v. 344. p. 186-189.

LOMBARDO, M. A., MACHADO, R. P. P. Aplicação das técnicas de sensoriamento remoto e sistemas de informações geográficas nos estudos urbanos. **Revista do Departamento de Geografia**. 2011. n. 10, p. 5-11

OJIMA, R. Dimensões da urbanização dispersa e proposta metodológica para estudos comparativos: uma abordagem socioespacial em aglomerações urbanas brasileiras. **Revista Brasileira de Estudos de População**. 2007. v. 24, n. 2. p. 277-300.

REIS, N. G. **Notas sobre urbanização dispersa e novas formas de tecido urbano**. São Paulo: Via das Artes, 2006. 201p.

SANTOS, M.A.S. Neurocomputação e SIG no suporte ao planejamento urbano. In: II Simpósio Regional de Geoprocessamento e Sensoriamento Remoto, 2004. Anais eletrônicos... Aracajú, 2004.

SANTOS, L.F., Redes neurais não-supervisionadas e fractais para a análise da mobilidade regional e da urbanização dispersa em São José dos Campos. PIBIC/CNPq. São José dos Campos: INPE. 2006.

SILVA, M.A.S., MONTEIRO, A.M.V., MEDEIROS, J.S. Geografia computacional em inferências geográficas com redes neurais artificiais. URLib MTC-M SID. São José dos Campos: INPE. 2002.

SILVA, M.A.S. SIQUEIRA, E.R. TEIXEIRA, O.A. , MANOS, M.G.L. MONTEIRO, A.M.V. Using self-organizing maps of rural territorial typology. (Ed.) **Computational methods for agricultural research: advances and applications**. Information Science reference. 2011. cap. 7. p. 107-126.

UNFPA – United Nations Population Fund. State of world population 2007: unleashing the potencial of urban growth. New York: UNFPA, 2007.

VESANTO, J. SOM based data visualization methods. *Intelligent Data Analysis*, n. 3, p. 111–126. 1999.