

Recovering Tobler's Views to Represent Full Toblerian Spaces in GisCiencas

Marcio Azeredo¹, Antônio M. V. Monteiro²

¹Diretoria de Serviço Geográfico (DSG)
Quartel General do Exército, Bloco "F", 70630 - 901, Brasília - DF - Brasil.

²Instituto Nacional de Pesquisas Espaciais (INPE)
Av. dos Astronautas 1758, 12227-001, São José dos Campos - SP - Brasil

{azeredo,miguel}@dpi.inpe.br

Abstract. *Many of the concepts, models, methods and techniques, employed in the current Geographical Information System (GIS), were driven by the use of a Euclidean metric inserted into a computer system of Cartesian representation. However, most spatial phenomena involving spatial relationships strongly anisotropic, spatially distributed and network connected, what requires possibilities for establishing new metrics of distance as an evolution needed for geographical analysis in GIS environment. Recovering Waldo Tobler's bidimensional regression technique from the 60's it's possible to represent spatial data in a non-euclidean distance metric space. This strategy is what we calling of Full Toblerian Space strategy (FTS-s) and hope that it provides new possibilities for explanations in the geography of the phenomena.*

Resumo. *Muitos dos conceitos, modelos, métodos e técnicas, empregados nos atuais Sistemas de Informação Geográfica (SIGs), foram impulsionados pela utilização de uma métrica euclidiana inserida em um sistema computacional cartesiano de representação. No entanto, grande parte dos fenômenos espaciais envolvem relacionamentos espaciais fortemente anisotrópicos, distribuídos no espaço e conectados por intermédio de redes, o que requer o uso de novas métricas de distância como uma necessidade de evolução na análise espacial dentro do ambiente de um SIG. Resgatando a técnica de regressão bidimensional proposta por Waldo Tobler na década de 60, é possível visualizar dados espaciais em um espaço que utiliza uma métrica de distância não-euclidiana. Essa estratégia é o que chamamos de Espaços Toblerianos Completos, e esperamos que forneça novas possibilidades de explicação da geografia dos fenômenos.*

1. Introdução

O termo GIScience surgiu recentemente como um conceito que abrange o estudo das tecnologias de tratamento, captação, interpretação, armazenamento, análise e intercâmbio de informações geográficas [UCGIS 1996]. Temas que se tornam cada vez mais importantes com o aumento da popularidade do uso dos Sistemas de Informação Geográfica (SIGs).

Nesses sistemas, muitos dos modelos, métodos e técnicas empregados, foram

impulsionados pela utilização do conceito de proximidade como ponto central na análise de fenômenos geográficos. Tais fenômenos, sejam eles naturais, ou sociais, apresentam uma relação entre si chamada de dependência espacial, que está diretamente relacionada à uma métrica de distância euclidiana.

Com o objetivo de simplificar a análise das interações espaciais, Waldo R. Tobler, geógrafo e matemático, assumiu que o valor de uma variável medido em uma dada localização, tende aos valores encontrados nas regiões mais próximas. Tal princípio ficou conhecido como a primeira lei da geografia, e diz que: "Todas as coisas são parecidas, mas as coisas mais próximas se parecem mais do que coisas mais distantes" [Tobler 1970]. Dois pontos estão mais próximos, ou são mais parecidos, quanto menor for a distância que os separa.

A utilização de uma métrica de distância euclidiana, parte do princípio de que os fenômenos espaciais apresentam um comportamento isotrópico, ou seja, que mantém suas propriedades constantes tomada qualquer direção dentro do espaço de representação. No entanto, o mundo não é sempre regular e previsível como o próprio Tobler concluiu [Tobler 2004].

Dessa forma, os fenômenos espaciais não consistentes com a lei de Tobler são regidos, na verdade, por relações de proximidade não euclidianas [Miller 2004]. Para esses casos, existem outras possibilidades que, combinadas à métrica euclidiana, podem representar com maior fidelidade a complexidade das interações espaciais.

Estabelecer novas métricas baseadas nas propriedades dos objetos espaciais, ou qualquer outro tipo de variável mensurável sobre o espaço, torna-se uma evolução necessária na análise geográfica dentro do ambiente de um GIS.

O presente estudo parte da premissa de que a regressão bidimensional proposta por Tobler favorece a análise dos dados espaciais utilizando uma métrica de distância não-euclidiana considerando para tal, as propriedades dos objetos envolvidos.

A esse conjunto de métodos baseados em métricas não-euclidianas, que estendem a capacidade de análise dos SIGs, é o que chamamos de Espaços Toblerianos Completos [Câmara et al. 2008], em homenagem a Tobler.

2. A métrica do espaço euclidiano

Um espaço métrico é composto por uma coleção de objetos e uma função de distância definida sobre eles [Navarro 1999]. No caso do espaço euclidiano, a métrica de distância é definida como sendo uma função $d: E \times E \rightarrow \mathbb{R}^n$ que associa a cada par de pontos $A\{a_1, a_2, a_3, \dots, a_n\}$ e $B\{b_1, b_2, b_3, \dots, b_n\} \in E$, um número real d , não negativo, chamado de distância do ponto A ao ponto B, tal que:

$$d(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2}$$

Usualmente, dados espaciais são representados em um plano euclidiano (duas dimensões $n=2$). Nesse plano, dois pontos estão mais próximos, quanto menor for o comprimento da linha reta que os conecta.

Quando construímos mentalmente as relações de proximidade, o fazemos

segundo uma métrica de distância euclidiana. Segundo Tobler [1993], grande parte da responsabilidade por essa forma de pensamento é das escolas do ensino fundamental, que lecionam a geometria de euclides, ao invés da geometria elíptica, mais indicada para o estudo da Terra.

Outras propriedades importantes do espaço euclidiano que merecem menção são:

- se $A = B$, então $d(A,B) = 0$ (espaço positivamente definido)
- se $A \neq B$, então $d(A,B) > 0$ (espaço positivamente definido)
- $d(A,B) = d(B,A)$ (espaço simétrico)
- para qualquer A, B e $C \in E$, $d(A,B) + d(B,C) \geq d(A,C)$
(espaço satisfaz a desigualdade triangular)

3. A anisotropia dos fenômenos geográficos

Grande parte dos estudos de fenômenos geográficos envolvem relações espaciais fortemente anisotrópicas [Câmara et al. 2008]. Ou seja, as propriedades dos elementos espaciais não são constantes, e variam no espaço de representação em função da direção em que ocorrem.

Como determinar, por exemplo, todos os pontos acessíveis partindo de um ponto central, dentro de um tempo limite?

Esse espaço anisotrópico resultante pode ser observado na figura 1. Nesse caso, as isolinhas de variam de cinco em cinco minutos.



Figura 1. Espaço anisotrópico com isolinhas de 5 minutos.
Fonte: [Tobler 1993]

O círculo construído permite observar que o deslocando-se no sentido norte é mais demorado do que no sentido sul.

Dependendo das variáveis consideradas na análise do problema, as isolinhas podem estar associadas à declividade do terreno, às condições de trafegabilidade, ao tipo de cobertura vegetal, ou qualquer outra propriedade que varie no espaço e que acabe por retardar o deslocamento em uma dada direção.

4. As redes de conexão e as relações espaciais

Outra característica importante das relações espaciais é a possibilidade de representação por intermédio de diferentes tipos de redes (ex. redes sociais, de comunicações e de transportes).

As características de construção dessas redes podem ser utilizadas para apontar as direções preferenciais de conexão (normalmente as que possuem um menor custo associado).

Como determinar em uma rede de transportes, por exemplo, qual a cidade mais próxima da cidade de São José dos Campos-SP, considerando as cidades de Resende-RJ, distante 180km, e São Paulo-SP, a 98km? A resposta parece simples (São Paulo) se usarmos o senso comum de que a proximidade é baseada em uma métrica de distância euclidiana.

E se o deslocamento para a cidade de São Paulo for realizado de trem, enquanto que para Resende, de avião? E se ambos os deslocamentos forem realizados de carro, sendo que uma obra na estrada para São Paulo acrescentou 1 hora ao tempo de viagem?

Nesse caso, foram consideradas as velocidades dos meios de transporte e uma métrica de distância não-euclidiana. Essa nova métrica em função do tempo de deslocamento entre as cidades indica que a cidade de Resende é a resposta. As cidades estão mais próximas, quanto menor for o tempo necessário para percorrer o espaço físico entre as mesmas.

5. As visões de Tobler - A regressão bidimensional

Tobler foi pioneiro ao trabalhar com métricas não-euclidianas no auxílio da representação dos fenômenos geográficos. Desde então, algumas de suas técnicas têm sido utilizadas em estudos de representação de mapas mentais [Cauvin 2002] [Friedman and Kohler 2003], de espaços urbanos [Cauvin 2001] [Ahmed and Miller 2007], de configurações cartográficas [Dias et al. 2000], de fluxos migratórios [Worboys et al. 1998] [Phan et al. 2005], entre outros.

Uma de suas técnicas em particular, a regressão bidimensional, foi desenvolvida na década de 70 tendo como base o trabalho "On Growth and Form" de D'Arcy Thompson. Seu uso permite avaliar o grau de semelhança entre duas, ou mais, configurações planares de um mesmo conjunto de pontos, cujas coordenadas são conhecidas em cada representação [Friedman 2003].

Inicialmente, o objetivo da regressão bidimensional era resolver o problema de comparação entre mapas de uma mesma região, construídos em épocas diferentes. Porém, acabou por ser utilizada em vários outros tipos de comparação de formas.

Segundo Tobler [1994], a regressão bidimensional é uma extensão da regressão unidimensional para o caso onde as duas variáveis consideradas, dependente e independente, possuem duas componentes de representação, ou seja, duas coordenadas.

O objetivo é relacionar um espaço de referência representado pela variável independente $Z\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, com um espaço imagem representado pela variável dependente $W\{(u_1, v_1), (u_2, v_2), \dots, (u_n, v_n)\}$.

Para cada par de Z é associado um par de W segundo uma função de transformação do tipo $W' = f(Z)$ (figura 2) de forma a permitir que o mapeamento seja o mais próximo de $Z \rightarrow W$ [Tobler 1994].

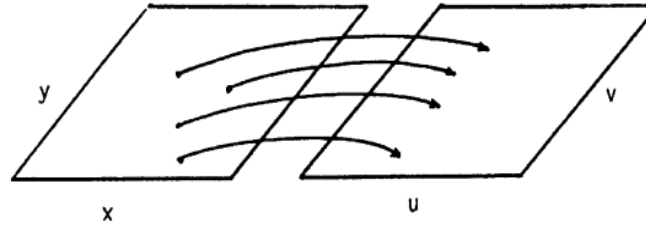


Figura 2. Transformação $Z \rightarrow W$. Fonte: [Tobler 1994]

Uma verificação da "qualidade" do $W'(u', v')$ estimado pode ser realizada utilizando-se o coeficiente de correlação bidimensional de Tobler (r^2). Quanto maior for o coeficiente, mais semelhantes são as configurações planares.

$$r^2 = 1 - \sqrt{\frac{\sum (u' - x)^2 + (v' - y)^2}{\sum (x - \bar{x})^2 + (y - \bar{y})^2}}$$

A técnica da Regressão Bidimensional pode ser dividida em duas fases principais: um ajustamento das observações de W para Z (W') e uma interpolação de forma a generalizar os resultados para todo o espaço (W'') [Cauvin 2005].

Durante a transformação entre os sistemas Z e W , as transformações geométricas sofridas pelas configurações de pontos podem ser modeladas utilizando-se parâmetros de rotação, translação, cisalhamento e escala. Tais parâmetros podem ocorrer simultaneamente, ou de forma independente.

A transformação mais abrangente é a transformação afim geral, onde são considerados seis parâmetros (dois de translação, dois de escala, um de rotação e um de cisalhamento).

Após algum cálculo algébrico, os parâmetros da transformação afim podem ser reduzidos às constantes $\omega_1, \omega_2, \beta_1, \beta_2, \beta_3$ e β_4 na equação abaixo.

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} + \begin{pmatrix} \beta_1 & -\beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

A substituição dos pontos x e y na equação permite a montagem de um sistema de equações lineares cuja solução fornece os valores dos parâmetros da transformação.

Determinados os parâmetros, o conjunto de dados é ajustado pelo método dos

mínimos quadrados de forma a minimizar a soma dos quadrados dos resíduos, ou seja, a diferença entre as coordenadas $Z(x,y)$ e $W'(u',v')$.

O próximo passo após o ajustamento das observações é a interpolação dos resultados obtidos para todo o espaço de representação.

Um resumo dessas etapas da regressão bidimensional pode ser observado na figura 3.

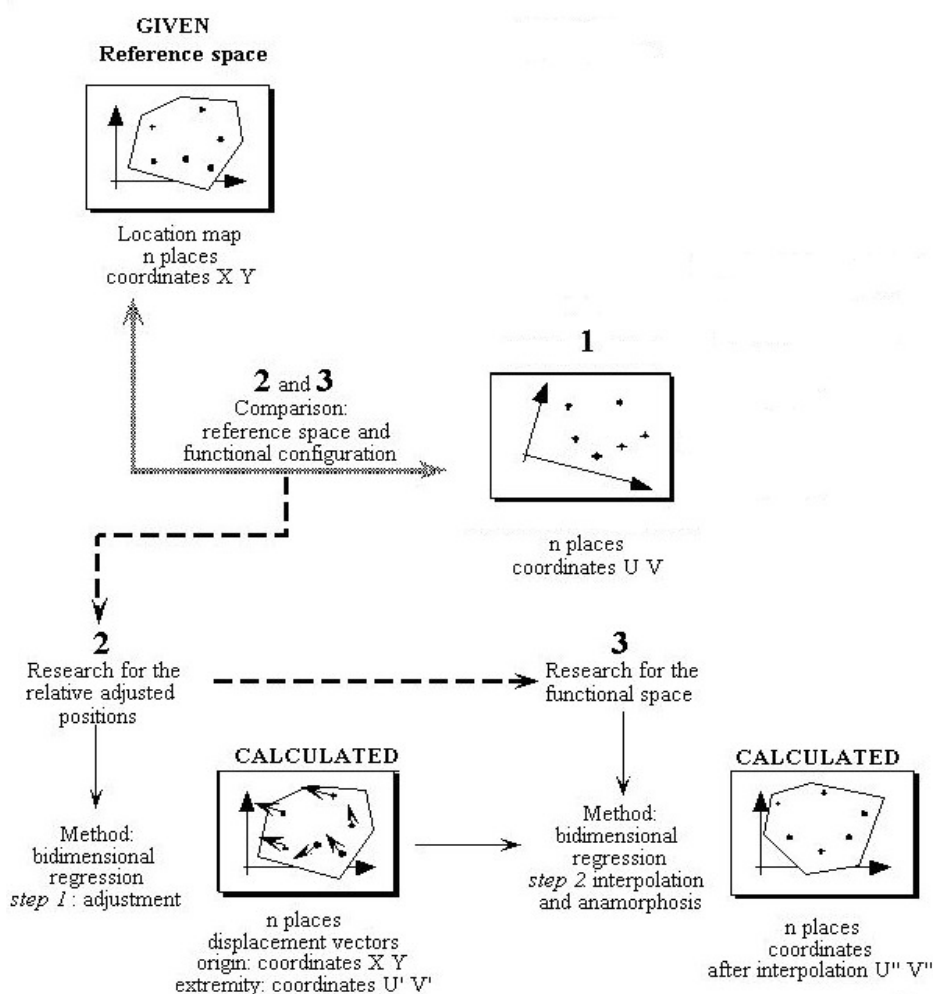


Figura 3. Metodologia da Regressão Bidimensional. Adaptado de: [Cauvin 2005]

Um exemplo prático é encontrado em Dias et al. [2000]. Nesse trabalho, a regressão bidimensional é utilizada para comparação de uma carta da região de Lisboa de 1757, com a Carta Militar de Portugal, usada como atual referência.

Após a primeira fase de ajustamento, as diferenças entre os valores das coordenadas podem ser observados na forma de vetores de distorção (figura 4.a) A seguir, os vetores são interpolados para todo o espaço de representação (figura 4.b). Por último, o espaço é distorcido segundo os vetores calculados na fase de interpolação (figura 4.c).

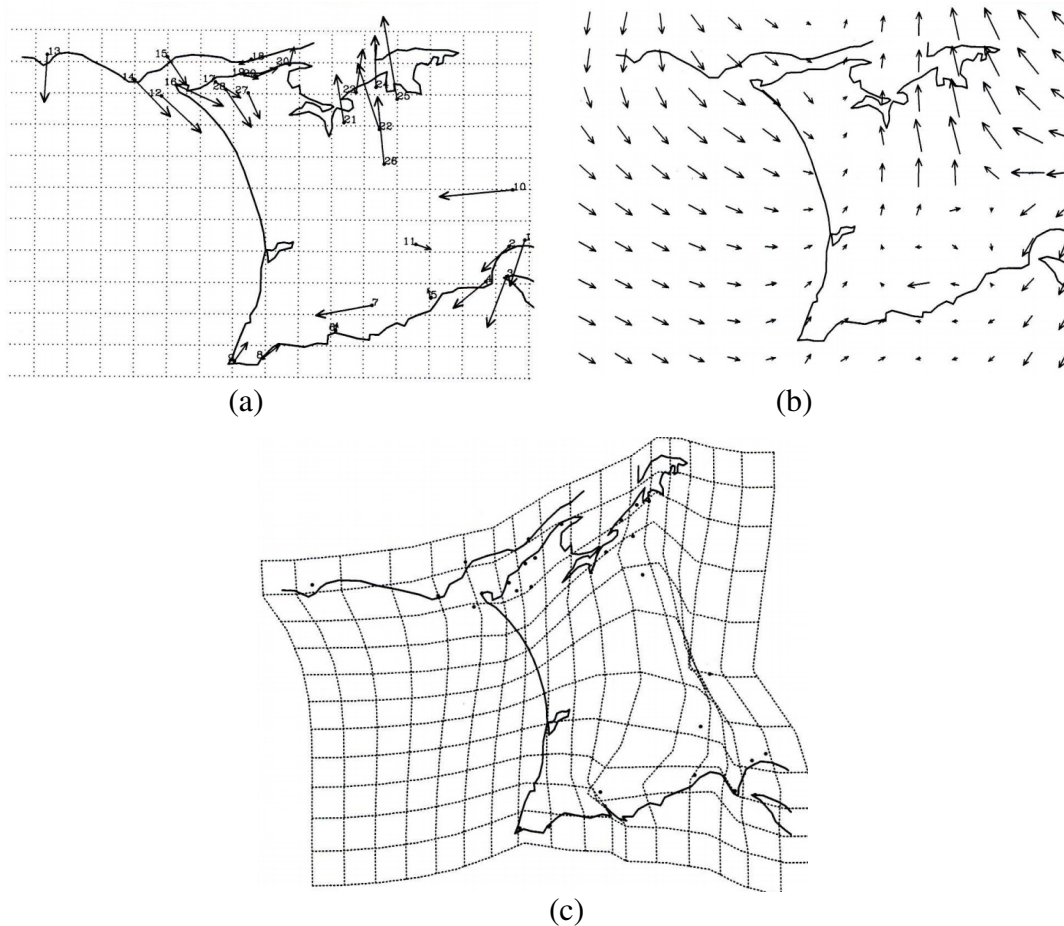


Figura 4. Comparação entre configurações cartográficas (a) Vetores de deformação. (b) Interpolação para todo o espaço. (c) Mapa resultante. Fonte: [Dias 2000]

Analisando as transformações geométricas obtidas, alguns questionamentos podem ser feitos. Quais os fatores que poderiam ter influenciado tais distorções? Elas configuram um padrão sobre os mapas da mesma região, ou mesma época?

Outro exemplo interessante do uso da regressão bidimensional é encontrado em Ahmed and Miller [2007], onde o espaço da cidade americana de Salt Lake é representado em função do tempo de acesso entre os nós de uma rede de transporte .

Aqui fica a observação de um acréscimo à metodologia apresentada na figura 3, originalmente proposta por Cauvin [2005]. Como a rede só fornece os valores atribuídos às interações espaciais (custo associado às arestas), é necessária a utilização de um método de análise multidimensional de proximidade (ex. escalonamento multidimensional - MDS) para obtenção das coordenadas u e v .

Como resultado é criado um espaço em função do tempo de acesso entre os nós da rede (figura 5).

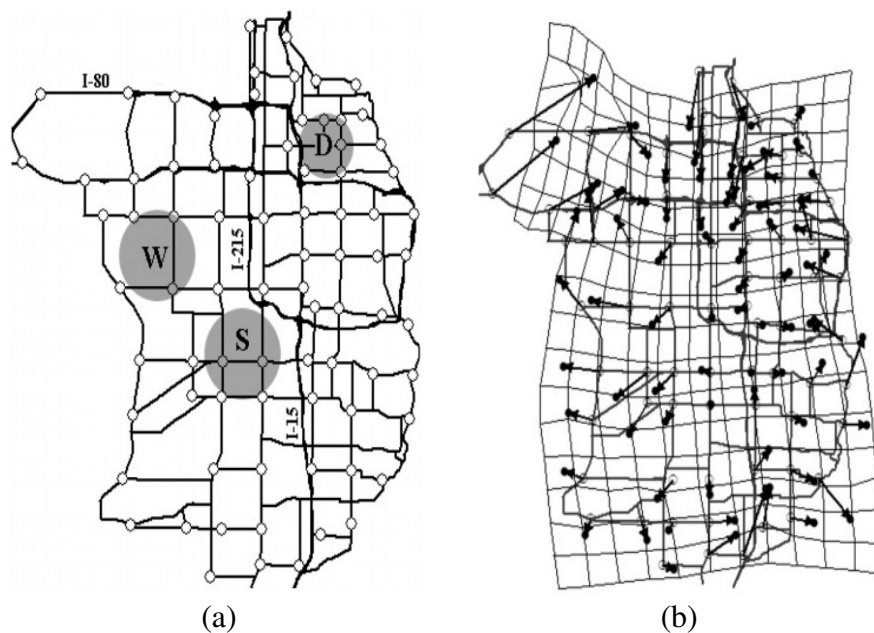


Figura 5. (a) Rede de transportes. (b) Vetores de deformação e espaço resultante em 2001. Fonte: [Ahmed and Miller 2007]

As transformações geométricas geradas pela rede podem fornecer informações importantes para o planejamento de redes de infraestrutura de transporte.

6. Conclusão

Não há mais sentido em limitar a capacidade de análise dos fenômenos espaciais ao espaço euclidiano. As novas gerações de SIGs devem ter seu foco no desenvolvimento de métodos capazes de capturar, interpretar, armazenar e analisar dados espaciais baseados em métricas não euclidianas.

Nesse sentido, a regressão bidimensional proposta por Tobler tem se mostrado uma técnica interessante, pois expande o conceito trivial de proximidade ao representar de interações espaciais utilizando métricas baseadas nas propriedades dos elementos envolvidos.

O grande desafio é redirecionar a atenção dos mapas que apresentam respostas, para os mapas que fomentam a busca pelas mesmas (MacEachren and Monmonier 1992).

Métricas que fornecem novos meios de visualizar e analisar o espaço geográfico conduzem a novas formas de questionar a realidade.

7. Referências Bibliográficas

- Ahmed, N.; Miller, H. J. (2007). Time-space transformations of geographic space for exploring, analysing and visualizing transportation systems. *Journal of Transport Geography*, p. 2–17.
- Câmara, G.; Aguiar, A. P.; Carneiro, T.; Costa, S. S.; Neto, P.; Monteiro, A. M.; Amaral, S. (2008). Non-toblerian geographical spaces.

- Cauvin, C. (2002). Cognitive and Cartographic representations: Towards A Comprehensive Approach. *Cybergeo*, Artigo 206.
- Cauvin, C. (2005) A Systemic Approach to Transport Accessibility. A Methodology Developed in Strasbourg: 1982-2002. *Cybergeo*. Artigo 311.
- Dias, M. H.; Cauvin, C.; Alegria, M. F. (2000). Comparação de configurações cartográficas através da regressão bidimensional. *Finisterra*, v. 35, n. 69, p. 95–107.
- Friedman, A.; Kohler, B. (2003). Bidimensional regression: Assessing the configural similarity and accuracy of cognitive maps and other two-dimensional data sets. *Psychological Methods*, v. 8, n. 4, p. 468–491.
- Golledge, R. G.; Hubert L. J. (1982) Some comments on non-euclidean mental maps. *Environment and Planning*, v. 14, p. 107–118.
- Miller, H. J. (2004) Tobler's First Law and Spatial Analysis. *Annals of the Association of American Geographers*, v.94, N° 2, p. 284-289.
- MacEachren, A. M. and Monmonier, M. (1992). Introduction (to special edition on Geographic Visualization). *Cartographic Information Systems*, 19, p. 197.
- Navarro, G. (1999). Searching in Metric Spaces by Spatial Approximation. *The VLDB Journal*, v.11 : p. 141-148.
- Phan, D.; Xiao, L.; Hanrahan, P; Winograd, T. (2005). Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization, p. 29.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, v.46: p. 234-240.
- Tobler, W. R. (1987). Experiments in Migration Mapping by Computer. *American Cartographer*, 1987.
- Tobler, W. R. (1993). Three Presentations on Geographical Analysis and Modeling. Technical Report 93-1.
- Tobler, W. R. (1994). Bidimensional regression. *Geographical Analysis*, v. 26, n. 3, p. 187–212.
- UCGIS (1996). Research Priorities for Geographic Information Science. *Cartography and Geographic Information Systems*, v.23, n.3.
- Worboys, M.; Mason, K.; Lingham, J. (1998). Computational techniques for non-Euclidean planar spatial data applied to migrant flows. Accepted for publication in *Innovations in Geographical Information Systems*: 4, Carver, C. (ed), Taylor and Francis.