

Métricas de qualidade de modelos espaciais

Luiz Gustavo Veras¹, Pedro Ribeiro de Andrade², Gilberto Câmara²

¹Programa de Mestrado em Computação Aplicada - INPE
São José dos Campos – SP – Brasil

²Centro de Ciências do Sistemas Terrestre – CST/INPE
José dos Campos – SP – Brasil

lgdo@dpi.inpe.br, {pedro, gilberto}@inpe.br

Abstract. *In the study and investigation of spatial phenomena, scientists need tools capable of representing and control such phenomena in order to analyze the impacts of a specific scenario. Spatial models provide a representative form of this space, but its implementation alone does not guarantee its ability to generate results close to reality. To this purpose, was introduced goodness-of-fit techniques applied to the results obtained by these models, comparing the generated maps with actual maps, resulting coefficients representing the degree of similarity between model and reality. Based on the methodology of multiple resolutions presented by Costanza (1989), two methodologies were developed, a multitemporal approach of multiple resolutions, and one based on the agregation of neighborhoods, called aggregation method neighborhoods.*

Resumo. *No estudo e investigação de fenômenos espaciais, cientistas necessitam de ferramentas capazes de representar e controlar tais fenômenos com o objetivo de analisar os impactos causados por um cenário específico. Modelos espaciais fornecem uma forma representativa deste espaço, mas sua implementação por si só não fornece informações sobre a sua capacidade de gerar resultados próximos da realidade. Para este fim, foram introduzidas métricas de qualidade aplicadas aos resultados obtidos por estes modelos, comparando os mapas gerados com mapas reais, resultando coeficientes que representam o grau de similaridade entre o modelo e a realidade. Com base na metodologia de múltiplas resoluções apresentada por Costanza(1989), foram desenvolvidas duas metodologias, uma abordagem multitemporal das múltiplas resoluções, e uma com base na formação de vizinhanças, denominado método de agregação de vizinhanças.*

Palavras-chave: *modelos dinâmicos espaciais, métricas de qualidade, modelagem computacional.*

1. Introdução

Modelos espaciais são representações simplificadas do sistema terrestre, o que envolve a representação de processos, agentes e o ambiente nos quais eles atuam. Mas o espaço geográfico se encontra em constante mudança, e os elementos citados não são suficientes para entender as relações espaciais e sua complexidade, a qual se estende também a uma componente temporal. Nos casos em que o tempo também é considerado, denominamos estas representações de modelos dinâmicos espaciais (Carneiro, 2006).

No processo de modelagem, o comportamento dos componentes espaciais são

identificados e representados em ambiente computacional, com o intuito de reproduzir com máxima aproximação os fenômenos espaciais observados em meio real. Devido a essa "aproximação", é inerente aos modelos espaciais um grau de erro ou de dessemelhança em relação ao sistema representando. Conseqüentemente, de forma complementar, estes possuirão um grau de semelhança com o mundo real.

Um dos desafios atuais em modelagem dinâmica espacial é aferir, ou pelo menos verificar, os resultados obtidos. Muitos cientistas ignoram a verificação dos dados gerados pela simulação dos modelos espaciais, bem como as etapas de calibração e validação, levando a ignorar erros e induzindo à falsa impressão que seus modelos são confiáveis o bastante para o objeto proposto. A verificação de dados é procedimento de suma importância, no que diz respeito a quantificar a qualidade do dado gerado em relação à realidade, e a crescente abundância de dados em diversos campos de estudo reforça a demanda por ferramentas de avaliação de desempenho de modelos dinâmicos espaciais (Pontius, 2004; Costanza, 1989).

Atualmente, uma grande quantidade de dados espaciais, em diversas áreas de aplicação (uso e cobertura da terra, ecologia, economia, etc), se encontra disponível em formato digital. Essas informações são de grande importância para que procedimentos de calibração e validação de modelos possam atuar sobre mais de uma resolução temporal. Com o crescente aumento de modelos dinâmicos espaciais na última década, é possível aplicar as metodologias de goodness-of-fit para verificação dos resultados desses modelos perante os dados reais dos anos os quais eles estimaram. Portanto, este trabalho foi motivado pela tentativa de fortalecer a discussão do tema, expandindo a metodologia apresentada por Costanza (1989), a múltipla resolução, adicionando uma componente temporal e introduzindo uma nova abordagem de comparação de modelos baseada em vizinhança.

O trabalho está organizado da seguinte forma: trabalhos relacionados, componentes das métricas de qualidade, múltiplas resoluções ao longo do tempo, comparação por agregação de vizinhanças e conclusão.

2. Trabalhos relacionados

Os métodos apresentados neste trabalho se baseiam na metodologia de múltiplas resoluções introduzida por Costanza (1989). O objetivo principal do método é agregar os elementos dos mapas por níveis de resolução, desde sua resolução nativa (resolução natural do mapa) até a menor resolução possível (agregação total de todos os elementos do mapa em um só). Com esta abordagem Costanza (1989) conseguiu comparar dois mapas por proporção de suas classes, bem como pelo padrão espacial formado pela disposição das mesmas. A Figura 1 apresenta a identificação desses padrões em um mapa com resolução de 10x10 elementos e 4 categorias representativas.

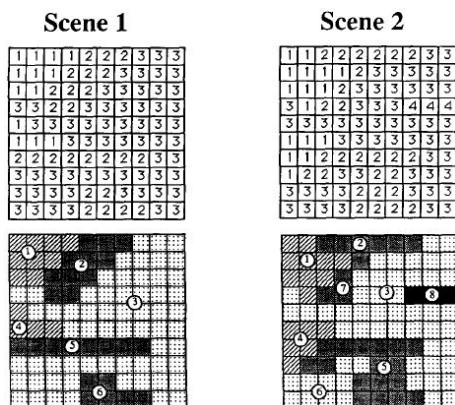


Figure 1. Exemplo de agregação de classes do método de múltiplas resoluções em duas cenas com 10x10 elementos e 4 categorias representativas. Fonte: (Costanza, 1989)

3. Componentes das métricas de qualidade

As métricas de qualidade são aplicadas a mapas representados por espaços celulares, que podem ser definidos como sendo um conjunto de células com relações de vizinhanças (Andrade et al, 2009). As unidades que formam estes espaços, as células, armazenam atributos e comportamentos que podem ser alterados ao longo do tempo. Na metodologia de comparação, proposta por Costanza (1989), as células são agregadas por geometrias limitadoras, denominadas janelas de amostragem. A cada passo de tempo da execução dos métodos, as janelas aumentam de acordo com um critério de agregação de células. A Figura 2 mostra como as janelas de amostragem agregam um conjunto de células para comparação.

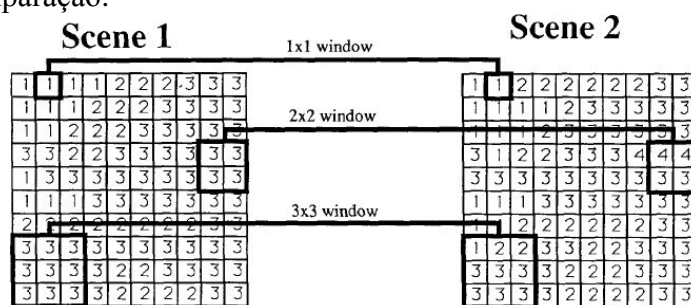


Figure 2. Agregação de células por meio de janelas de amostragem. Fonte: (Costanza, 1989)

4. Múltiplas resoluções ao longo do tempo

O método de múltiplas resoluções permite expressar por meio de coeficientes a similaridade entre componentes espaciais de dois mapas. Mas esta abordagem aplica as métricas de qualidade a dois mapas em apenas um instante no tempo, e portanto não considera o grau de correspondência temporal. A proposta deste método é estender o método de múltiplas resoluções, estimando o ajuste para vários tempos, afim de detectar erros de alocação não somente em escala espacial, mas também temporal. A Figura 3 apresenta um modelo esquemático do método das múltiplas resoluções espaço-temporais. A sigla MR identifica os mapas reais, a sigla MS identifica os mapas resultantes da simulação dos modelos dinâmicos espaciais, e as linhas vermelha, verde e

azul os níveis temporais de comparação entre os mapas.

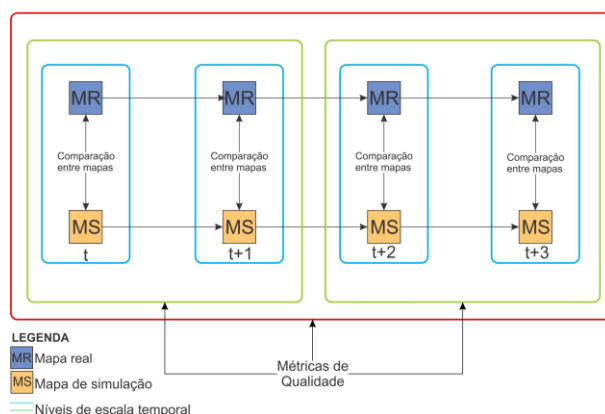


Figure 3. Modelo esquemático do método de múltiplas resoluções espaço-temporal.

5. Comparação por agregação de vizinhanças (CAV)

O método de comparação por agregação de vizinhanças (CAV) utiliza o relacionamento de vizinhança entre as células como agregador (contraste) do padrão espacial. Esta metodologia é diferente dos métodos de comparação padrão, que utilizam janelas para encapsular uma área de comparação entre mapas. Enquanto a grande parte utiliza como critério o aumento das janelas de amostragem o acréscimo de seus lados em uma unidade de célula, a agregação por vizinhanças forma as áreas de comparação agregando as células vizinhas.

Similar ao método de múltiplas resoluções, um novo nível acarreta em uma maior agregação de células, onde a agregação é gerada pelo acumulo dos vizinhos da célula de um dado nível. É fácil notar o padrão espacial formado pela CAV em relação ao método de múltiplas resoluções, principalmente quando a relação de vizinhança não segue um padrão de adjacência entre vizinhos, como é apresentada na Figura 4.

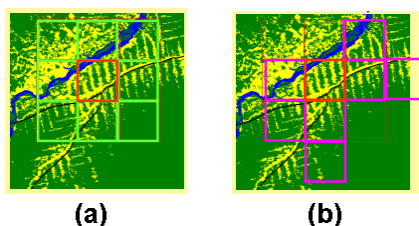


Figure 4. Vizinhança por adjacência (a) e vizinhança com células não adjacentes (b). Fonte: (Aguiar, Câmara, Cartaxo, 2003).

5.1 Descrição do método

A agregação parte de um nível inicial, definido pelas vizinhanças nativas do modelo, que são utilizadas para gerar os novos níveis de células. Uma nova célula neste método é formada pela agregação de vizinhanças conectadas por meio de suas componentes (células da vizinhança). A cada nível, é verificado se as componentes da vizinhança corrente participam de outra vizinhança. Se sim, a vizinhança é incorporada formando uma vizinhança. O método tende a agrupar todas as vizinhanças até que não existam mais conexões entre elas.

Podemos definir um conjunto V para representar todas as vizinhanças contidas em um mapa, onde $V = \{V_1, V_2, V_3, \dots, V_n\}$. Cada subconjunto V_n pode ser definido como $V_n = \{v_1, v_2, v_3, \dots, v_n\}$, onde cada elemento v_n é uma célula contida na vizinhança. A tendência é que à medida que as vizinhanças são agregadas, o conjunto V tende a diminuir, e os subconjuntos V_n que agregam a si outros subconjuntos aumentem em tamanho.

Como meio de evitar casos onde uma célula que já tenha sido agregada naquele nível seja agregada novamente (já que uma célula pode ser vizinha de várias outras), algumas restrições na agregação são atribuídas ao algoritmo CAV:

- uma célula identificará um agregado de células. Ela servirá como o identificador de uma célula, e servirá como meio de acesso àquele agregado;
- existem dois conceitos atribuídas à metodologia: o de célula agregada, e o de célula visitada;
- uma célula é agregada quando ela já foi incorporada em uma vizinhança, portanto não pode mais desempenhar a função de identificador de uma célula de um nível maior;
- uma célula é visitada quando em uma iteração, ela já tenha sido utilizada, de acordo com alguns critérios, para a formação de um novo nível de vizinhanças. As células são classificadas como visitadas se forem identificadoras de uma vizinhança ou são constatadas como componentes de uma.

A Figura 5 mostra o resultado da aplicação do método CAV em passos de tempos discretos em um espaço celular em que cada célula contém uma vizinhança de Moore.

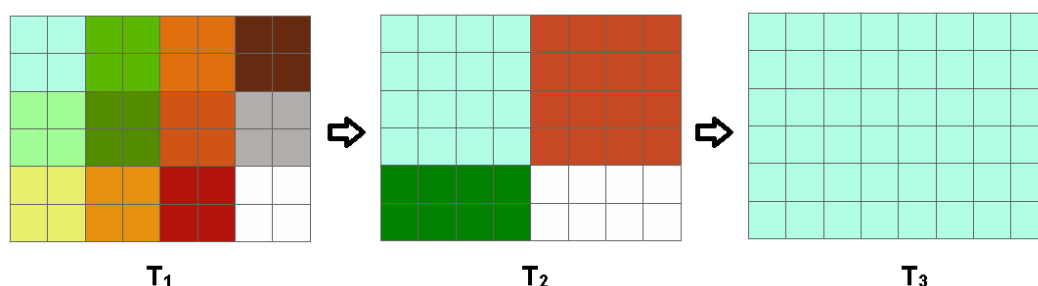


Figure 5. Padrão de agregação de células considerando a vizinhança de Moore em passos discretos de tempo.

6. Conclusão

Foram apresentadas duas novas abordagens para construção de métricas de qualidade de modelos dinâmicos espaciais. A primeira é uma extensão do método de múltiplas resoluções introduzida por Costanza (1989). A segunda introduz uma nova abordagem de agregações de padrões espaciais, considerando a vizinhança das células para formar novos níveis de comparação. Com isto deseja-se fortalecer a discussão sobre avaliação de desempenho de modelos espaciais, tema imprescindível para gerar confiabilidade nos resultados gerados pelos modelos, que auxiliam tanto nas tomadas de decisão e nos estudos da dinâmica espacial.

Testes adicionais estão sendo realizados utilizando dados de modelo e cobertura da terra, com foco no desmatamento na área da Amazônia Legal, devido à relevância do tema para a conservação da floresta Amazônica (Aguiar et al., 2007; Laurance et al., 2001).

Referências

- Andrade, P. R., A. Monteiro, M. V., Câmara, G., Sandri, S.. Games on Cellular Spaces: How Mobility Affects Equilibrium. *Journal of Artificial Societies and Social Simulation (JASSS)*. 12(1) p. 5. Jan/2009.
- Aguiar, A. P. D. ; Câmara, G.; Cartaxo, R.. Modeling Spatial Relations by Generalized Proximity Matrices. In: *Brazilian Symposium on Geoinformatics, 2003*, Campos do Jordão. Anais do Geoinfo 2003. São José dos Campos: INPE, 2003.
- Aguiar, A. P. D.; Câmara, G.; Escada, M. I. S. Spatial statistical analysis of land-use determinants in the brazilian amazonia: Exploring intra-regional heterogeneity. *ecological modelling*, v. 209, p. 169 - 188, 2007.
- Carneiro, T. G. de S. Nested-CA: A Foundation For Multiscale Modelling of Land Use and Land Cover Change. Tese (Doutorado). Instituto Nacional de Pesquisas Espaciais - INPE, 2006.
- Costanza, R. Model goodness of fit: a multiple resolution procedure. *Ecological Modelling*, v. 47, p. 199 - 215, 1989.
- Lambin, E. F.; Geist, H.; Rindfuss, R. R. Land-use and landcover change: local processes and global impacts. In: . [S.l.]: Springer, Berlim, 2006. cap. Introduction: Local Processes with Global Impacts, p. 18.
- Laurance, W. F.; Cochrane, M. A.; Bergen, S.; M.Fearnside, P.; Delamônica, P.; Barber, C.; D'angelo, S.; Fernandes, T. The future of the brazilian amazon. *Science*, v. 291, p. 438 - 439, 2001.
- Pimenta, P. F. Uma proposta de um framework para modelos computacionais de uso e cobertura da terra. Tese (Mestrado). Universidade Federal do Pará, 2010.
- Pontius, R. G.; Huffaker, D.; Denman, K. Useful techniques of validation for spatially explicit land-change models. *Ecological Modelling*, v. 179, p. 445, 461, 2004.