

Métrica de qualidade espaço-temporal de modelos dinâmicos espaciais

Luiz Gustavo Diniz de Oliveira Veras¹
Pedro Ribeiro de Andrade¹
Gilberto Câmara¹

¹ Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil
lgdo@dpi.inpe.com.br
{pedro.andrade, gilberto.camara}@inpe.com.br

Abstract. In the study and investigation of spatial phenomena, scientists need tools that can control and represent these phenomena, aiming analyze impacts caused by specific scenarios. Spatial models provide a representative form of this space, but its implementation without results evaluation isn't guarantee its capacity to generate accuracy results with the real world. To assist in this objective, methodologies called goodness-of-fit are used to compare the simulation's results with observed data. This paper presents a new goodness-of-fit metrics. It's based in the multiple resolution algorithm, introduced by Costanza (1989), inserting a temporal component into the methodology.

Palavras-chave: modelagem espacial, goodness-of-fit, métricas de qualidade, múltiplas resoluções.

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. Consequentemente, 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, dados utilizado na metodologia, resultados 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.

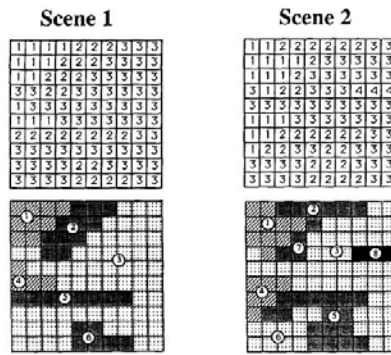


Figura 1. Exemplo de agregação de classes do método de múltiplas resoluções em duas cenas com 10x10 elementos e 4 categorias representativas. Fonte: (Costanza, 1989).

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. Nas metodologias de comparação, 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. Quando a janela de amostragem varrer todo o mapa, o ajuste geral é calculado pela equação da Figura 2.

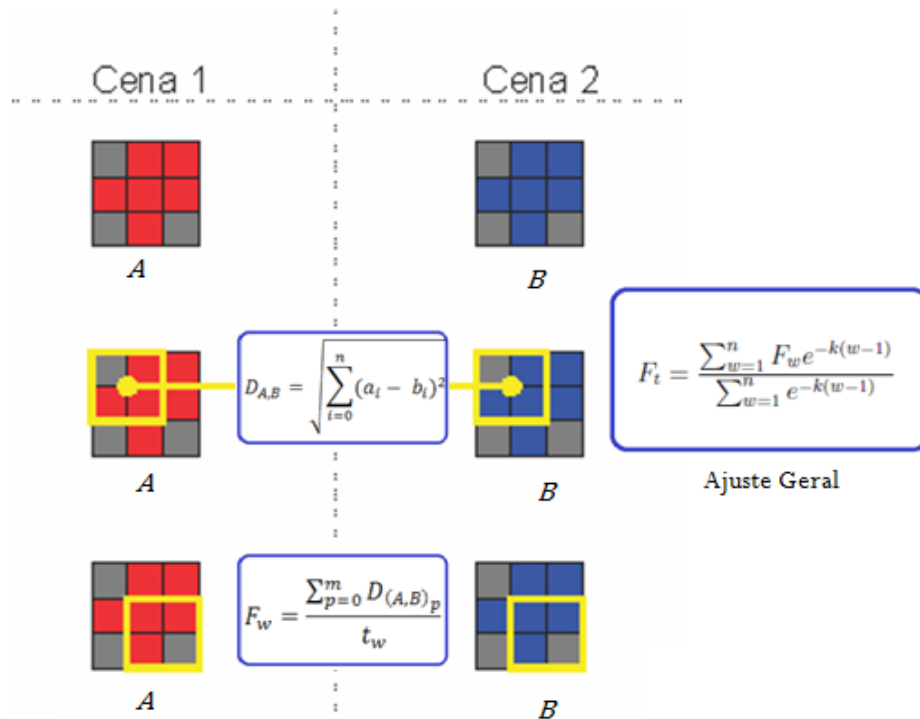


Figure 2. Agregação de células por meio de janelas de amostragem e equação do ajuste geral.

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, a fim 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 a linha vermelha, verde e azul os níveis temporais de comparação entre os mapas.

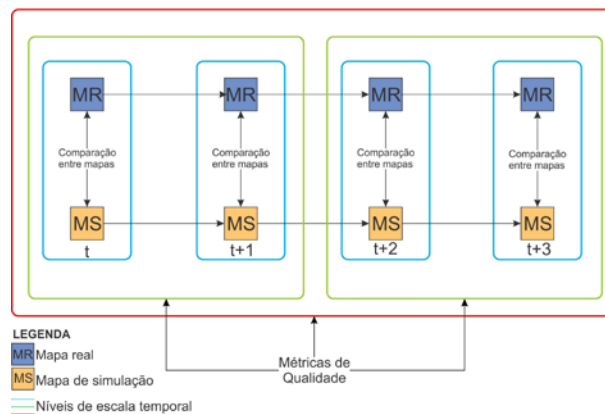


Figura 3. Modelo esquemático do método de múltiplas resoluções espaço-temporal.

Para cada resolução temporal, o método de múltiplas resoluções gera uma métrica. Essas métricas individuais em cada resolução de tempo representam somente a componente espacial da comparação. Após as métricas serem geradas para cada tempo do intervalo de comparação, uma nova métrica é gerada, resumizando o desempenho do modelo em um único coeficiente referente ao período de tempo considerado.

5. Dados utilizados na metodologia

Para a execução do método que calcula a métrica espaço-temporal, são necessário dados de observação (mapa de referência) e dados gerados por simulação (mapa de comparação). Os dados utilizados como mapas de referência são provenientes do programa Prodes(2012). Os mapas utilizados como mapas de comparação são oriundos de resultados de dois trabalhos: **The future of brazilian Amazon** publicado por Pontius (2009), e o trabalho de tese de doutorado **Modelagem de mudança do uso da terra na Amazônia: explorando a heterogeneidade intra-regional** defendida por Aguiar (2006).

O Prodes (2009) é um programa de monitoramento do desmatamento na Amazônia, utilizando imageamento remoto e ferramentas de geoprocessamento. Segue uma metodologia que se baseia na análise de desmatamento de anos anteriores para estimar ocorrência onde não é possível detecção devido à resolução do sensor imageador dos satélites a ocorrência de nuvens. É considerado um dos melhores programas de detecção de desmatamento no mundo. Devido ao longo período de execução do Prodes (desde 1989), foi possível obter dados temporais suficientes para analisar a nova métrica.

Laurance (2001) estudou o desmatamento na Amazônia enfatizando como principal fator deste fenômeno as estradas e obras de desenvolvimento econômico-social como usinas e hidrelétricas. Foram estimadas taxas de desmatamento anuais por meio de análise do desmatamento de anos anteriores. Com essa taxa o desmatamento da Amazônia foi estimado até 2020.

Aguiar (2006) enfatiza que as conexões a mercados é que possuem maior participação no desmatamento. É uma extensão do arcabouço de modelagem dinâmica CLUE para explorar como fatores alternativos, públicas e condições de mercado que influenciam no processo de ocupação da Amazônia. Utilizando um quantidade de demanda anual, níveis de desmatamento são estimados também até 2020.

O intervalo das comparações foi estabelecido entre 2001 e 2008. O início do período representa o dado mais antigo gerado por Laurance (2001), e o fim o dado mais recente do Prodes até a realização dos experimentos deste trabalho. A Figura 4 apresenta uma esquematização das comparações realizadas. Com o objetivo de obter o desempenho de um modelo nulo, o dado do Prodes do ano de 2001 foi utilizado como tal.

Os mapas dos respectivos modelos são representados por meio de espaço celular, com células na resolução de 25x25 km². O desmatamento é representado uma porcentagem de área em cada uma dessas células. Para implementação da metodologia foi utilizado o framework TerraME, que utiliza a linguagem de script Lua. O banco de dados utilizado foi montado sobre a interface de dados espaciais TerraLib.

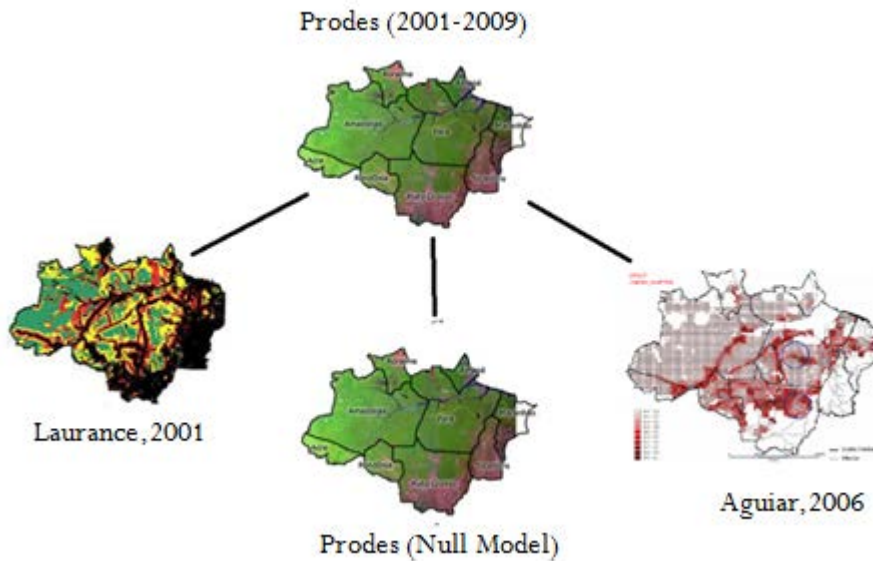


Figura 4. Esquema das comparações realizadas entre os mapas de comparação e o de referência.

6. Resultados

A quantidade apresentada por cada resultado de simulação e do mapa de referência é apresentado na Figura 5. A curva do Prodes (na cor azul) é a meta de qualidade (em termos quantitativos) que os modelos devem alcançar. A curva de Aguiar (na cor vermelha) apresenta variações entre cada ano, aproximando-se dos resultados (em termos quantitativos) do Prodes somente nos últimos anos de comparação. A de Laurance (na cor laranja) mantém uma constante queda na geração do desmatamento. Os pontos com maior proximidade com a curva do Prodes são nos anos de detecção de menor desmatamento (2001, 2003, 2004 e 2005).

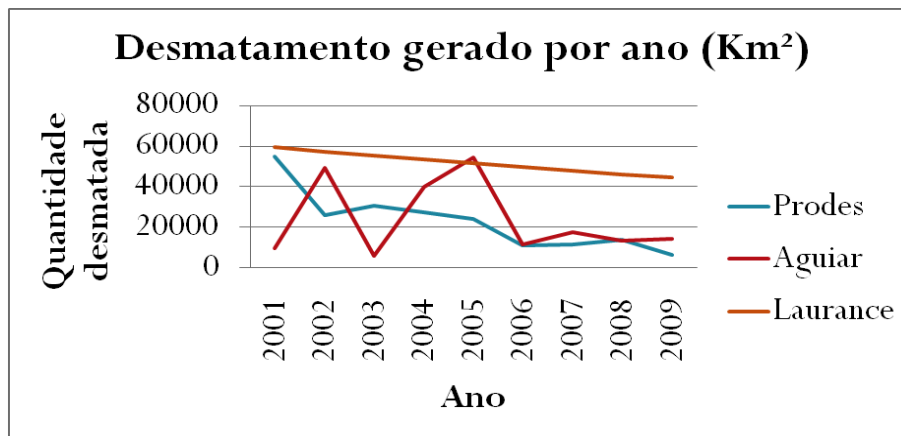


Figura 5. Quantidade de desmatamento encontrado em cada um dos mapas utilizados na comparação.

A Figura 6 apresenta as curvas de comparação do modelo de Laurance (2001), Aguiar (2006) e do Prodes nulo. Observou-se alta porcentagem de acerto dos modelos em geral. As curvas iniciais possuem menor índice de acerto, devido às janelas de amostragem concentrar uma menor quantidade de unidades de mapa por comparação. Apesar dos valores dos ajustes dos

modelos serem próximos entre si, por meio do ajuste geral é possível verificar de forma mais evidenciada o desempenho entre eles.

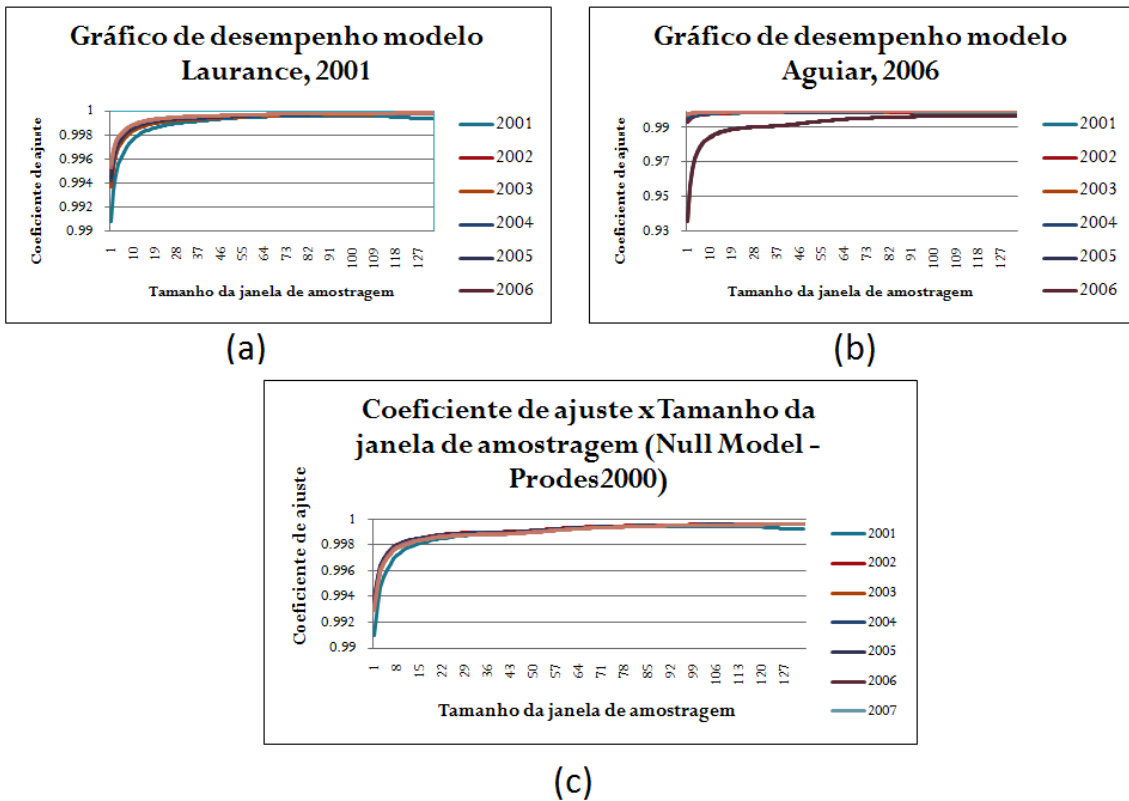


Figura 6. (a) Gráfico de desempenho do modelo de Laurance (2001), (b) do modelo de Aguiar et al.(2006) e (c) do modelo Nulo.

A Figura 7 apresenta os ajustes gerais de cada um dos testes. O desempenho dos modelos ao longo dos anos é bastante semelhante. Na curva de Aguiar (de cor azul), ocorre uma queda brusca nos anos de 2005 e 2006 comparadas ao desempenho que os modelos em geral apresentaram. A ocorrência desta queda é provavelmente devido ao mecanismo de alocação de demanda do modelo em questão.

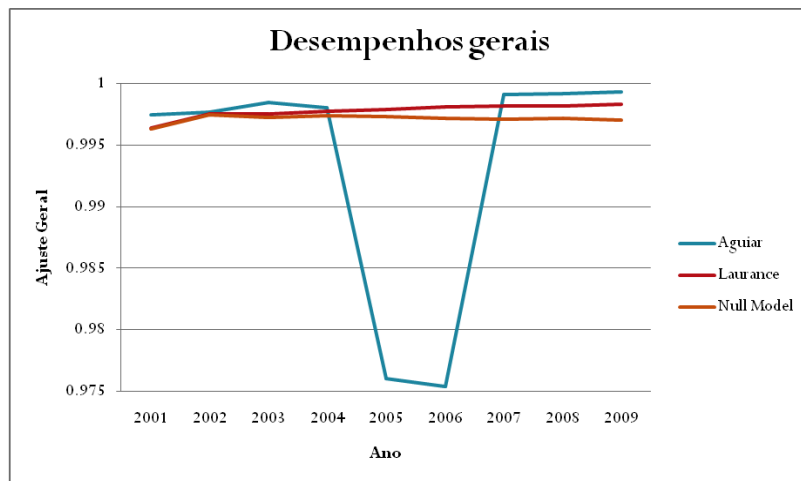


Figura 7. Gráfico de desempenho geral dos modelos.

7. Conclusão

Foi apresentada uma nova abordagem de métricas de qualidade de modelos dinâmicos espaciais. Ela é uma extensão do método de múltiplas resoluções introduzido por Costanza (1989). 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). Os próximos passos são desenvolver novas métricas e verificar o alto grau de acerto por parte dos modelos em relação ao mapa de referência, se isso é devido à natureza constante dos dados utilizados ou se realmente é uma verificação válida do desempenho das simulações.

8. 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. **Modelagem de mudança do uso da terra na amazônia: explorando a heterogeneidade intra-regional**. Tese de doutorado, INPE, 2006.
- 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, Berlin, 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.
- Programa de Cálculo do Desflorestamento da Amazônia (PRODES). **PRODES – dados 2001-2009**. 2009. Disponível em < <http://www.dpi.inpe.br/prodesdigital/prodes.php>>. Acesso em 05/04/2010.