

## Métricas de qualidade de modelos espaciais

Luiz Gustavo Diniz de Oliveira Veras<sup>1</sup>, Pedro Ribeiro de Andrade<sup>2</sup>, Gilberto Câmara<sup>2</sup>

<sup>1</sup>Programa de Mestrado em Computação Aplicada – CAP  
Instituto Nacional de Pesquisas Espaciais – INPE

<sup>2</sup>Centro de Ciências do Sistemas Terrestre – CST  
Instituto Nacional de Pesquisas Espaciais – INPE

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

**Abstract.** *In dynamic space modeling, geographic space is represented in simplified form in order to study the different spatial processes and their impacts on different scales. Because these models are an approximate representation of spatial dynamics, it is important to evaluate its outcome with real data. To this end, measures called Goodness-of-fit are used to estimate the degree of similarity between maps derived from simulation and reality. There are approaches to goodness-of-fit in one or several resolutions for comparison. The Euclidean Distance is a measure that uses a single spatial resolution, and represents the distance between two points or in the case of this work, a distance between two maps. The Multiple Resolution approach evaluates maps generated by spatial models in various resolutions. This paper presents a description of the methods Euclidean Distance and Multiple Resolutions and an analysis of the performance of these methods by applying them to a case study.*

**Resumo.** *Em modelagem dinâmica espacial, o espaço geográfico é representado de forma simplificada com o objetivo de estudar os diversos processos espaciais e seus impactos em diversas escalas. Devido a esses modelos serem uma representação aproximada da dinâmica espacial, é importante avaliar seu resultado considerando como parâmetros de comparação dados reais. Para esse fim, são utilizadas métricas de qualidade, usualmente chamadas de métricas de Goodness-of-fit, que são medidas que estimam o grau de similaridade entre mapas resultantes de simulação e de observações da realidade. Existem abordagens de métricas de qualidade com uma ou diversas resoluções de comparação. A Distância Euclidiana é uma medida que, que utiliza uma única resolução espacial, e representa a distância entre dois pontos, ou no caso deste trabalho, uma distância entre dois mapas. O Método de Múltiplas Resoluções apresenta uma abordagem que avalia mapas gerados por modelos espaciais em diversas resoluções. Este trabalho apresenta uma descrição dos métodos Distância Euclidiana e Múltiplas Resoluções e uma análise do desempenho desses métodos aplicando-os a um estudo de caso.*

**Palavras-chave:** *modelos espaciais, métricas de qualidade, goodness-of-fit.*

### 1. Introdução

Em modelagem espacial, o espaço geográfico é representado de forma simplificada com o objetivo de estudar os diversos processos espaciais e seus impactos em várias escalas. Os modelos podem ser estruturados por simples equações até interações baseadas em

um conjunto de premissas. Seus resultados geralmente são estimativas do futuro [Longley et al. 2007] que representam o impacto causado por processos do presente.

Devido a esses modelos serem uma representação aproximada da dinâmica espacial, é importante avaliar seu resultado considerando como parâmetros de comparação dados reais [Costanza 1988; Pontius et al. 2006]. Para esse fim são utilizadas métricas de qualidade, usualmente chamadas de métricas de *Goodness-of-fit*, que são medidas que estimam o grau de similaridade entre mapas resultantes de simulação e de observações da realidade.

Este trabalho apresenta uma descrição de duas métricas de qualidade de modelos dinâmicos espaciais, a Distância Euclidiana e a Múltiplas Resoluções [Costanza 1988], e uma análise do desempenho desses métodos aplicando-os a um estudo de caso. Os métodos foram implementados no framework de modelagem dinâmica espacial TerraME [Carneiro 2006], onde os mapas são representados por meio de espaços celulares. O artigo está organizado como segue. Na seção 2 é descrito o método da Distância Euclidiana. Na seção 3 o método de Múltiplas Resoluções é apresentado. Na seção 4 os resultados são discutidos.

## 2. Distância Euclidiana

A *Distância Euclidiana* é uma medida que representa a distância entre dois pontos, ou adaptando ao problema exposto, uma distância entre dois mapas. O cálculo é realizado de forma que os pixels de um espaço celular são comparados aos pixels correspondentes à um segundo espaço celular.

Para a aplicação da Distância Euclidiana, é necessário definir dois conjuntos que representam os espaços celulares:

$$Q = (q_1, q_2, q_3 \dots q_n)$$

$$P = (p_1, p_2, p_3 \dots p_n)$$

onde Q representa o espaço celular contendo os dados  $q_1, q_2, q_3 \dots q_n$  gerados por simulação e P representa o espaço celular onde os dados  $p_1, p_2, p_3, \dots p_n$  são reais. A partir destes conjuntos, pode-se determinar uma distância  $D(Q, P)$  como sendo:

$$D(Q, P) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 \dots (q_n - p_n)^2}$$

que pode ser resumida como:

$$D(Q, P) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

### 3. Múltiplas Resoluções

O método de múltiplas resoluções, introduzido por Costanza [1988], considera diversas resoluções de comparação, nas quais o cálculo sobre os mapas é realizado. A resolução inicial é definida pelo lado da menor unidade que compõe o espaço celular, ou seja, uma célula. A resolução final é a que corresponde à resolução nativa do espaço celular. A resolução inicial diminui em um lado de uma célula até atingir a resolução final, com o objetivo de captar o padrão de distribuição espacial entre dois mapas.

Para cada resolução é definida uma janela de amostragem com as mesmas dimensões. A função dessa janela consiste em realizar o cálculo de comparação entre os mapas. Iniciando da célula localizada na extremidade superior esquerda do espaço celular, essa janela varre todo esse espaço, com deslocamento de um lado de uma célula. Ao final da varredura, o coeficiente de diferença naquela resolução é definido por:

$$F_w = \frac{\sum_{i=1}^{t_w} J_w}{t_w}$$

onde  $w$  é o tamanho de um lado da janela de amostragem,  $F_w$  é o coeficiente da janela com tamanho  $w$ ,  $J_w$  é a equação utilizada na janela de tamanho  $w$ ,  $t_w$  é a quantidade de janelas de amostragem para a janela de tamanho  $w$ .

### 4. Resultados

Como suporte para aplicação dos métodos foram utilizados dados do PRODES dos anos de 2000 à 2005 e de 2000 à 2009 [PRODES 2009]. A partir destes as variáveis *area\_defor2005* e *area\_defor2009* foram geradas. A visualização dos dados é apresentada na Figura 1.

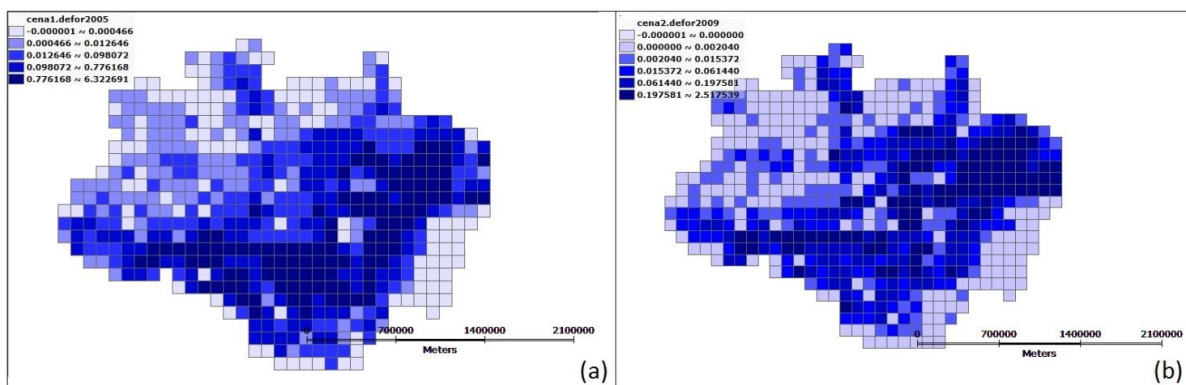
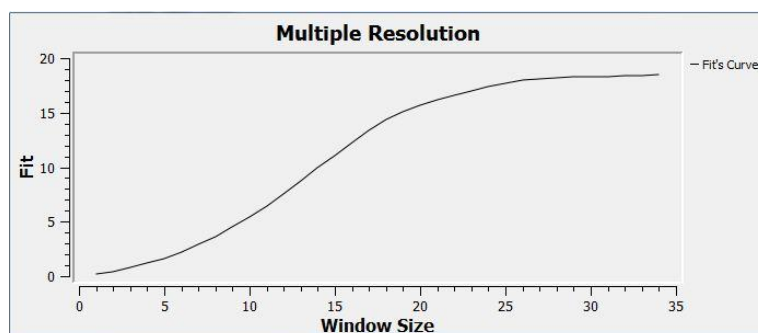


Figura 1: Visualização das variáveis *area\_defor2005* (a) e *área\_defor2009* (b).

O método da Distância Euclidiana apresentou um coeficiente de diferença no valor de 18,4442. Isso representa, de forma sintetizada, a distância numérica entre as variáveis utilizadas no procedimento, ou seja, o grau da diferença entre os dois mapas.

O método de múltiplas resoluções resulta em um gráfico que apresenta as janelas de

amostragem com seus respectivos coeficientes de diferença, como mostra a Figura 2. As janelas de tamanhos menores apresentam menor coeficiente de diferença, como apresentado pelas 10 primeiras janelas. Após esse tamanho a diferença entre elas cresce abruptamente. Isso ocorre devido à diminuição da resolução de comparação, tornando o padrão espacial aparente, permitindo maior agregação de valores dentro de uma janela de amostragem.



**Figura 2: Gráfico das janelas de amostragem e seus respectivos coeficientes de diferença**

## Referências

Blilie, C. The Promise and Limits of Computer Modeling. [S.l.]: World Scientific, 2007.

Costanza, R. (1989) Model goodness of fit: a multiple resolution procedure. Ecological Modelling.

Carneiro, T. G. de S. (2006) Nested-CA: A Foundation For Multiscale Modelling of Land Use and Land Cover Change. Tese (Doutorado) - Instituto Nacional de Pesquisas Espaciais – INPE.

Longley, Paul A., Goodchild, Michael F., Maguire, David J., and David W. Rhind. 2005. Geographic Information Systems and Science, 2nd ed., John Wiley and Sons, Toronto.

Pontius, R.; Cornell, J.; Hall, C. (2001) Modeling the spatial pattern of land-use change with geomod2: application and validation for costa rica. Agriculture, Ecosystems & Environment, v. 85.

PRODES/INPE. (2004/09) Projeto de desmatamento do instituto nacional de pesquisas espaciais. Banco de Dados. Sit: <http://www.obt.inpe.br/prodes/index.html> Acessado em: 25 de Julho de 2011.