

## **LANDSCAM: LANDSCAPE DYNAMICS SIMULATION CELLULAR AUTOMATA MODEL**

### **LANDSCAM: MODELO BASEADO EM AUTÔMATOS CELULARES PARA A SIMULAÇÃO DA DINÂMICA DA PAISAGEM**

Marcos Wellausen Dias de Freitas<sup>1</sup>, Pablo Muñoz<sup>2</sup>, João Roberto dos Santos<sup>3</sup>

<sup>1</sup> Universidade Federal do Rio Grande – FURG, Av. Itália, km 8, 96203-900 – Rio Grande, RS, Brasil, marcoswfreytas@gmail.com

<sup>2</sup> Instituto Tecnológico da Aeronáutica – ITA, Praça Marechal-do-Ar Eduardo Gomes, 50, 12228-900 – São José dos Campos, SP, Brasil, pablocus@gmail.com

<sup>3</sup> Instituto Nacional de Pesquisas Espaciais - INPE, Av. dos Astronautas, 1758, 12.227-010 - São José dos Campos, SP, Brasil, jroberto@dsr.inpe.br

#### **ABSTRACT**

This work presents a land use and cover change (LUCC) spatial dynamics model developed for the operationalization of the main procedures involved in the LUCC dynamics modelling such as: transition rates estimation and landscape dynamics processes mapping, generation of transition probabilities maps based on logistic regression and spatial allocation of the LUCC transitions with deterministic cellular automata and stochastic patches formation functions. The modular structure of the LanDSCAM model and its open source characteristic facilitate the modification and adaptation of this new LUCC and scenario generation model, which at this present study involved the following LUCC processes: forest degradation, grassland degradation, regeneration, intensification, extensification, silviculture expansion and urbanization.

*Keywords:* Spatial dynamics model, land use and cover change, cellular automata, landscape dynamics.

#### **RESUMO**

Este trabalho apresenta um modelo de dinâmica espacial de mudanças de uso e cobertura da terra (MUCT) desenvolvido para a implementação dos principais procedimentos envolvidos na modelagem dinâmica de MUCT, tais como: cálculo das demandas de transição e geração de mapas de transições, geração de mapas de probabilidade de transição, funções de alocação espacial dinâmica das mudanças com funções de autômato celular determinístico e de formação de manchas aleatórias, ferramentas para a calibração e validação dos modelos. A estrutura modular do LandSCAM e sua característica de código-aberto facilitam modificações e adaptações desse novo modelo de dinâmica de uso e cobertura e prognose de cenários, que no presente estudo de caso envolveram os seguintes processos de MUCT: degradação florestal, degradação de campos, regeneração, intensificação, extensificação, expansão da silvicultura e urbanização.

*Palavras-chave:* Modelo de dinâmica espacial, mudanças de uso e cobertura da terra, autômato celular.

#### **INTRODUÇÃO**

Neste trabalho, apresenta-se o modelo computacional de dinâmica espacial baseado em autômatos celulares, denominado LanDSCAM (*Landscape Dynamics Simulation Cellular Automata Model*), usando como exemplo e estudo de caso a modelagem de mudanças de uso e cobertura da terra (MUCT) no período 1986-2002 na região do Alto Uruguai (RS-SC) com base em dados Landsat-TM. O código deste modelo foi desenvolvido na linguagem Fortran 90 em ambiente Unix (MacOSX ou Linux) e é um código-aberto (*open source*) livre para uso e adaptação por outros usuários. A escolha do ambiente Unix se deve à falta de

modelos de MUCT que, em sua maioria, apresentam versões somente no sistema operacional Windows. Enquanto a escolha pelo código-aberto também se relaciona com a falta de modelos de MUCT com tais características, o que limita muito a interação do usuário e a possibilidade de maiores adaptações e desenvolvimentos sobre a base do modelo. Já a estrutura modular de programação do modelo LandSCAM e sua característica de código aberto facilitam a alteração do modelo de dinâmica espacial e a implementação de novas rotinas de modelagem aos usuários com conhecimento da linguagem Fortran 90.

Este modelo se baseou na estrutura dos principais modelos de MUCT, DINAMICA – EGO (SOARES-FILHO *et al.*, 2002) e CLUE-S (VERBURG *et al.*, 2002), onde foram desenvolvidos módulos para as principais operações envolvidas na modelagem de MUCT num modelo de dinâmica espacial (cálculo das demandas de transição e geração de mapas de transições, geração de mapas de probabilidade de transição, funções de alocação espacial dinâmica das mudanças, ferramentas para a calibração e validação dos modelos). A principal diferença de concepção do LandSCAM em relação aos demais na estrutura de modelagem é: a) sua orientação para processos de MUCT, o que possibilita a agregação de diferentes tipos de transição de uso e cobertura da terra, b) o uso de regiões definidas pelo usuário para a alocação das transições e calibração dos modelos, c) o uso de uma função de alocação das transições dividida entre um autômato celular determinístico e uma função de geração de manchas de caráter estocástico.

## **METODOLOGIA E ESTRUTURA DE MODELAGEM**

A estrutura do modelo segue a da maioria dos modelos existentes de MUCT, sendo composta por cinco etapas, organizadas conforme módulos: a) cálculo de demandas de transição, b) geração de mapas de probabilidade de transição, c) alocação espacial das MUCT, d) calibração e validação dos resultados e e) geração de cenários de prognose de MUCT (Figura 1). A definição de taxas de transição é gerada de forma regionalizada, de acordo com regiões definidas pelo usuário que podem representar municípios, distritos, bacias hidrográficas e outras definições regionais. Este modelo foi implementado com o objetivo de modelagem de conjuntos de dados matriciais de grande extensão, com o foco na rapidez e economia de memória real alocada, permitindo a geração de modelos na resolução espacial original do Landsat para grandes áreas.

O modelo apresenta rotinas conjugadas em quatro módulos principais com a entrada dos dados matriciais em formato ASCII (*American Standard Code for Information Interchange*) ou em formato binário *raster* suportado pela biblioteca GDAL (*Geospatial Data Abstraction Library*), como GeoTIFF, img, asc e outros formatos suportados por tal biblioteca. O uso da biblioteca GDAL em ambiente de programação Fortran é possibilitado pela interface da biblioteca FortranGIS que realiza a conversão dos dados binários para as matrizes em memória durante a compilação do código em Fortran, sendo ambas bibliotecas de código-aberto. A seguir, pode-se ver em detalhes os métodos e técnicas implementadas nestas operações.

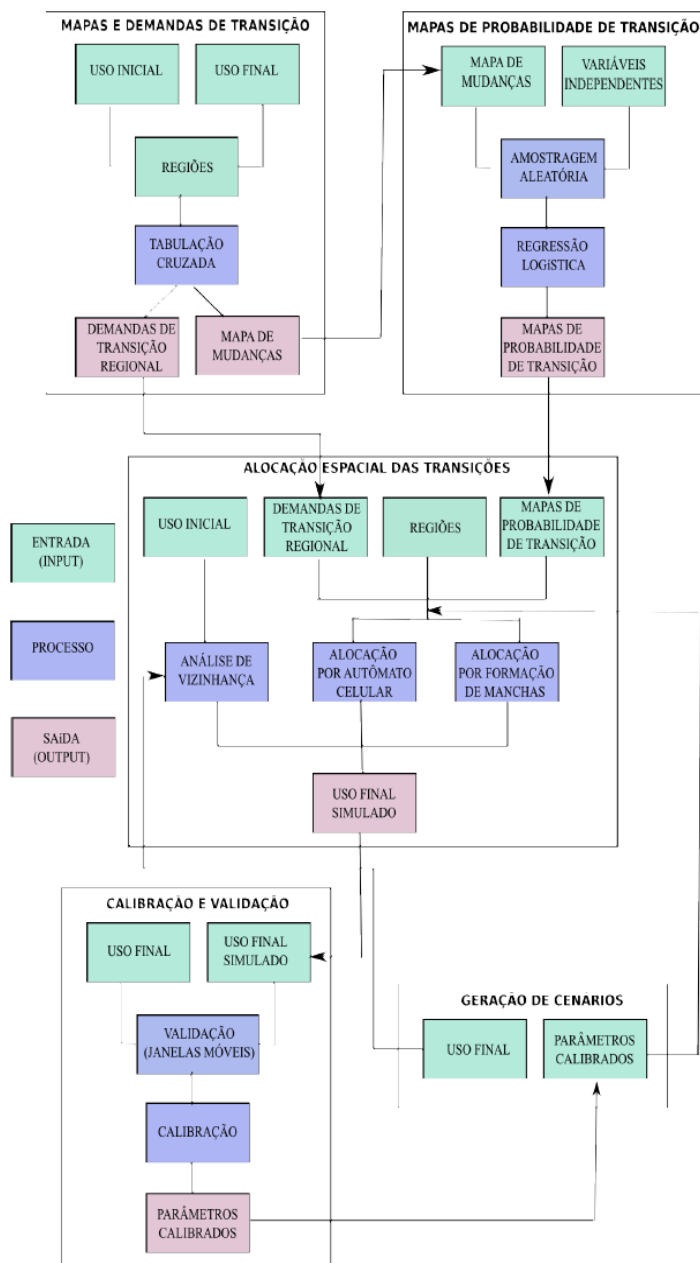


Figure 1. Flowchart of the LanDSCAM modelling structure.

Figura 1. Fluxograma representando a estrutura de modelagem do LanDSCAM.

**a) Cálculo de demandas de transição e mapas de mudanças da paisagem**

O primeiro módulo do modelo é composto por duas funções para a preparação da entrada no modelo das demandas de transição e dos mapas de mudança s de uso e cobertura da terra entre a data inicial e a final de cada intervalo de tempo. Tais funções apresentam característica complementar com o uso do resultado da primeira etapa de cálculo das demandas de transição para a definição da codificação das mudança s.

O cálculo das demandas de transição é gerado por uma operação de tabulação cruzada dos mapas de uso e cobertura da terra inicial e final analisados no passo de tempo utilizado para a calibração do modelo. O resultado desta operação é uma matriz de probabilidades de transição que permite a definição das taxas de transição baseadas nas mudança s ocorrentes entre as datas utilizadas para a calibração. Tais taxas de transição são calculadas com base em cadeias de Markov (MARKOV, 2006). A função de definição das

cadeias de Markov apresenta como entrada as regiões e os mapas de uso e cobertura da terra inicial e final. Nesta, a definição das Cadeias de Markov a partir de tabulação cruzada é gerada em um passo único entre a data inicial e final, que pode ser dividida em múltiplos passos de tempo no módulo de alocação espacial em si, por cada região.

A função de geração de mapas de mudança s deve ser realizada com base no resultado da função anterior gerado com uma única região total representando a área de estudo, para não se correr o risco de deixar transições sem classificação. O modelador pode definir a codificação das transições existentes com um valor único para cada uma ou ainda com a possibilidade de agrupamento de transições em processos de MUCT, como foi realizado no nosso estudo de caso.

A Figura 2 apresenta a matriz de processos gerada neste procedimento de classificação das transições ocorridas no intervalo de 1986-2002, tendo como área de estudo, a região do Alto Uruguai no Brasil na divisa entre os Estados de Santa Catarina e Rio Grande do Sul. Esta matriz de transição é utilizada no módulo de alocação espacial das transições para identificar os mapas de probabilidade de transição referentes a cada transição e identificados por números diferentes. No caso específico desse estudo de operacionalização do modelo, com 7 classes, seriam possíveis 42 transições (i.e., o número de classes ao quadrado menos o número de classes). Como ocorreram apenas 18 transições, classificadas em 7 processos, estas foram numericamente classificadas de 1 a 7 (o número de processos) e as transições não existentes foram identificadas pelo valor 0.

### b) Geração de mapas de probabilidade de transição

Como em outros modelos de MUCT, a análise espacial utilizada para a geração dos mapas de probabilidades de uso foi baseada na regressão logística binária multivariada (HOSMER e LEMESHOW, 2000) que é usada para indicar a probabilidade de uma célula assumir uma transição determinada, dado um conjunto de variáveis independentes contínuas ou categóricas. A metodologia adotada na geração dos mapas de probabilidade é composta por três etapas: a) a amostragem espacial aleatória das variáveis dependentes e independentes realizada por um módulo do modelo, b) a análise de regressão logística propriamente dita realizada no ambiente de programação R (R DEVELOPMENT CORE TEAM, 2008) e c) o cálculo dos mapas de probabilidade de transição por outro módulo do modelo.

		USO FINAL								
		FOM	CAMPO-NAT	SILV	AGRIC-FAM	AGRIC-GP	URB	ÁGUA		
I N I C I A L	FOM	0	0	4	1	1	7	0	1	DEGRADAÇÃO FOM
	CAMPO-NAT	0	0	4	2	2	7	0	2	DEGRADAÇÃO CAMPO
	SILV	0	0	0	5	6	7	0	3	REGENERAÇÃO
	AGRIC-FAM	3	0	4	0	6	7	0	4	EXPANSÃO DA SILVICULTURA
	AGRIC-GP	0	0	4	5	0	7	0	5	EXTENSIFICAÇÃO
	URB	0	0	0	0	0	0	0	6	INTENSIFICAÇÃO
	ÁGUA	0	0	0	0	0	0	0	7	URBANIZAÇÃO

Figure 2. Transition matrix used for classifying the land use transitions occurred between the initial and final land use maps at the period 1986-2002.

Figura 2. Matriz de transição utilizada para a classificação das transições ocorridas entre os mapas iniciais e finais no período 1986-2002. Abreviaturas: FOM é floresta, CAMPO-NAT é campo nativo, SILV é silvicultura, AGRIC-FAM é agricultura familiar, AGRIC-GP é agricultura de grande porte, URB é água, DEGRADAÇÃO FOM é degradação florestal.

A primeira fase da regressão logística é a amostragem das variáveis dependentes e independentes que foi implementada no modelo com inspiração na ferramenta de amostragem aleatória do modelo CLUE-S. O módulo de amostragem aleatória se divide em duas etapas: a) a definição dos índices (i e j) de cada pixel amostrado e b) a geração das tabelas com os dados das variáveis independentes.

Na primeira etapa da amostragem aleatória, a entrada é a matriz de inteiros representando as variáveis independentes onde, neste caso, foram usados os processos de MUCT, mas podem ser todas as transições ou ainda classes de uso e cobertura da terra da data inicial. O algoritmo implementado faz de forma automática a categorização binária (0 e 1) para cada processo ou valor da matriz de entrada. Os parâmetros definidos pelo usuário são a porcentagem de cada classe a ser amostrada, a proporção entre o número de amostras de 0 e 1 e a distância entre cada pixel amostrado.

A segunda etapa da amostragem aleatória é composta por duas entradas: a saída do procedimento anterior (duas listas de índices i e j, os com valores 1 e os com valor 0) e as matrizes referentes às variáveis independentes (separadas ou num cubo tridimensional). Esta etapa faz a “coleta” do valor de cada variável independente de acordo com a posição definida pela lista de índices para cada processo de forma individual, gerando a saída que é uma tabela com as variáveis dependentes e independentes a ser utilizada na análise de regressão logística no ambiente R.

A análise de regressão logística foi realizada no ambiente de programação de código-aberto voltado à estatística R (R DEVELOPMENT CORE TEAM, 2008) com o uso dois pacotes com funções para construção, seleção de variáveis e avaliação de performance: o pacote base do R e o pacote Design (HARREL JR., 2012).

### c) Módulo de cálculo de mapas de probabilidade de transição

O módulo de cálculo de mapas de probabilidade de transição tem como entrada os dados matriciais com as variáveis independentes e uma tabela em formato ASCII com uma lista com os coeficientes do intercepto e das variáveis independentes. Este é um módulo individual, onde cada modelo é calculado separadamente. A partir destes dados de entrada são produzidos mapas de probabilidades de transição de acordo com o modelo de regressão logística:

$$P_{i,j} = \frac{e^{\beta_0 + \beta_1 X_{1,i,j} + \dots + \beta_n X_{n,i,j}}}{1 + e^{\beta_0 + \beta_1 X_{1,i,j} + \dots + \beta_n X_{n,i,j}}}$$

onde  $P_{i,j}$  é a probabilidade de transição em cada pixel,  $\beta_0$  é o intercepto do modelo,  $\beta_{1,n}$  é o coeficiente de cada variável independente e  $X_{n,i,j}$  é o valor de cada variável independente em cada pixel.

A saída deste módulo corresponde aos mapas de probabilidade de transição que são utilizados como entrada no módulo de alocação espacial. Neste trabalho, os mapas corresponderam à probabilidade de transição de um tema para outro, em acordo com os 7 processos de MUCT identificados na área de estudo (Figura 2): degradação florestal, degradação de campos, regeneração, intensificação, extensificação, expansão da silvicultura e urbanização.

### d) Alocação espacial das mudanças de uso e cobertura da terra

A alocação espacial das transições foi baseada na abordagem de autômatos celulares através de um processo de ordenação e de escolha dos pixels candidatos a transição de acordo com o número de vizinhos e com a probabilidade de transição que busca preencher as demandas de transição nos passos de tempo do modelo. Neste modelo, o total das demandas de cada transição é dividido entre dois processos de alocação: a) um processo de escolha ordenada de candidatos baseado em regras determinísticas de autômatos celulares e b) um processo de formação de manchas por crescimento de regiões a partir de sementes (*seeds*) estocásticas.

A entrada do módulo de alocação espacial é composta pelo mapa de uso inicial, o mapa de regiões, os mapas de probabilidades de transição, a matriz de processos e a tabela de demandas de transição por região. Os parâmetros definidos pelo modelador são o número de vizinhos (de 1 até 8), o número de passos de tempo e a proporção do total da demanda de cada transição a ser satisfeita pelos dois procedimentos de alocação (de 0 a 1). Estes parâmetros são tratados de forma dinâmica no modelo, podendo ser flexibilizados durante a execução do módulo para a garantia de satisfação das demandas de transição ao final do período modelado.

O procedimento inicial na alocação espacial das transições é a geração de demandas de transição em múltiplos passos de tempo conforme o número destes definido pelo modelador. Após isto, ocorre a divisão proporcional da matriz de demanda total de transição de acordo com a porcentagem da função de alocação determinística por autômato celular definida para cada transição de acordo da seguinte forma:

$$Dt_r = \left( \frac{Pa_{tr} \times (\text{Autômato Celular})}{100} + \frac{Pb_{tr} \times (\text{Formação de Manchas})}{100} \right) / nt \quad (2)$$

onde  $Dt_r$  é a demanda total da transição  $t$  na região  $r$  num passo de tempo único,  $Pa$  é a proporção para a função determinística de autômato celular,  $Pb$  é a proporção para a função aleatória de formação de manchas, onde  $Pb = 100 - Pa$  e  $nt$  é o número de passos de tempo.

### **Alocação espacial baseada em autômatos celulares:**

O procedimento de alocação espacial baseada em autômatos celulares buscou a satisfação das demandas de transição definidas pela proporção das demandas totais de cada transição definida a este procedimento de alocação. Este procedimento foi implementado por região a partir de regras determinísticas de vizinhança e a ordenação das probabilidades de transição, sendo realizado em separado para cada transição. O algoritmo inicia com uma filtragem de candidatos a partir de uma análise de vizinhança baseada no tipo de vizinhança Moore (8 vizinhos numa janela de 3 x 3 pixels) que é utilizado na maioria dos modelos de MUCT baseados em autômatos celulares. Este procedimento foi gerado com base na resolução espacial de 30 metros, conferindo um nível maior de detalhamento das MUCT em relação aos modelos baseados em dados generalizados. A análise de vizinhança realiza a definição do número de vizinhos da mesma classe do pixel central da janela e do número de vizinhos de cada uma das classes. A partir desta análise de vizinhança, inicia-se o processo de filtragem de candidatos a transição com a geração de uma lista dos índices dos pixels candidatos. Este processo é definido essencialmente de acordo com o parâmetro de número de vizinhos definido pelo modelador na regra de autômato celular. A regra de autômato celular implementada na filtragem inicial dos pixels candidatos a transição pode ser definida a partir da seguinte cláusula condicional do tipo se-então (*if-then*):

$$\text{If } nl \geq v \text{ then } m_{ijr} = l_{ijr}$$

onde  $nl$  é o número de células da classe  $l$  encontrada na janela de 3 x 3 pixels,  $m_{ijr}$  é o pixel central da janela com classe  $m$  na região  $r$ ,  $l_{ijr}$  é um pixel presente na lista de candidatos à transição da classe  $m$  para a classe  $l$  e  $v$  é o parâmetro de número de vizinhos definido pelo modelador.

Por exemplo, no caso do parâmetro  $v$  ser definido como igual a 4, apenas os pixels da classe inicial com mais de 4 pixel vizinhos da classe final serão considerados para a geração de uma transição para a classe final de interesse. Assim, para a definição dos candidatos a uma transição, são filtrados os pixels da classe inicial de transição que apresentam um número de vizinhos maior do que o mínimo da classe final de transição. Desta maneira são escolhidos os elementos nas bordas de uma classe inicial que apresentam um número acima do mínimo de vizinhos de outras classes, gerando um processo de expansão das manchas anteriores da classe final de transição.

Após esta filtragem inicial, os pixels candidatos são ordenados em ordem decrescente. Isso, de acordo com os mapas de probabilidade de transição para o processo referente à cada transição modelada (identificados pela matriz de processos da Figura 2). E, por fim, os pixels candidatos a transição são escolhidos até a satisfação das demandas requeridas para cada transição em cada região. A satisfação das demandas pode não ser completamente satisfeita devido à oferta de candidatos em relação às demandas requeridas. Desta forma os parâmetros de número de vizinhos e de passos de tempo do modelos foram flexibilizados para a satisfação mais próxima das demandas. O procedimento de alocação espacial das transições adotado de filtragem de pixels candidatos foi dirigido à eficiência computacional do código, pois há uma vantagem em termos de rapidez ao uso de dados escalares em comparação aos dados matriciais nos procedimentos de iteração (*loops*) na linguagem Fortran 90, a qual foi concebida para operações com dados escalares.

### **Alocação espacial com formação de manchas:**

O procedimento de alocação espacial com formação de manchas foi baseado numa abordagem estocástica onde estas são geradas através de uma técnica de crescimento de regiões a partir de sementes escolhidas de forma aleatória. Os parâmetros definidos pelo usuário são a média e o desvio-padrão da área de cada tipo de transição que devem ser calculados com o uso de métricas de paisagem do *software* Fragstats (MACGARIGAL e MARKS, 1995) ou em ambiente de SIG. O resultado deste procedimento deve ser representado, para a entrada no modelo, através de tabelas com a média e o desvio-padrão das áreas de cada transição, dispostas como a matriz de processos, i.e., a cada transição corresponde sua média e desvio-padrão de área da mancha a ser gerada em metros quadrados.

Estes dados são utilizados em uma função de geração de manchas onde o tamanho de cada mancha a ser gerada e o número de sementes são definidos com base numa distribuição gaussiana com a média e o desvio-padrão de cada transição em particular. Esta função funciona como um gerador aleatório de áreas de manchas com base numa distribuição gaussiana definida a partir da média e do desvio-padrão das áreas de manchas inseridos pelo usuário. Assim, as sementes da classe inicial são selecionadas de forma aleatória e os pixels vizinhos da mesma classe vão sendo escolhidos até cumprirem o tamanho da mancha sorteada na função de distribuição gaussiana. Esta função de crescimento de regiões foi baseada num padrão radial de expansão das manchas que assumem um formato circular.

### **e) Calibração e validação dos modelos**

A técnica de validação utilizada foi o procedimento de validação multi-resolução baseado em janelas móveis [Costanza 1989] que foi implementado para a avaliação da qualidade de modelos de simulação ecológicos e é amplamente utilizado em modelos de MUCT que utilizam dados categóricos (SOARES-FILHO *et al.*, 2002; VERBURG *et al.* 2002, CASTELLA e VERBURG, 2007).

### **Estudo de caso**

O resultado do modelo de mudança de uso e cobertura da terra no intervalo 1986-2002 na área do Alto Uruguai (Figura 3) apresentou resultados superiores aos do modelo nulo com um ajuste médio total, mensurado pelo método de janelas móveis (janelas de  $1 \times 1$  a  $15 \times 15$ ), de 0.771 em relação ao ajuste de 0.703 do modelo nulo. A comparação das áreas das classes no resultado do modelo em relação ao uso real em 2002 apresentou erros inferiores a 2.8%, com exceção da classe urbano que apresentou um erro de 9.57%.



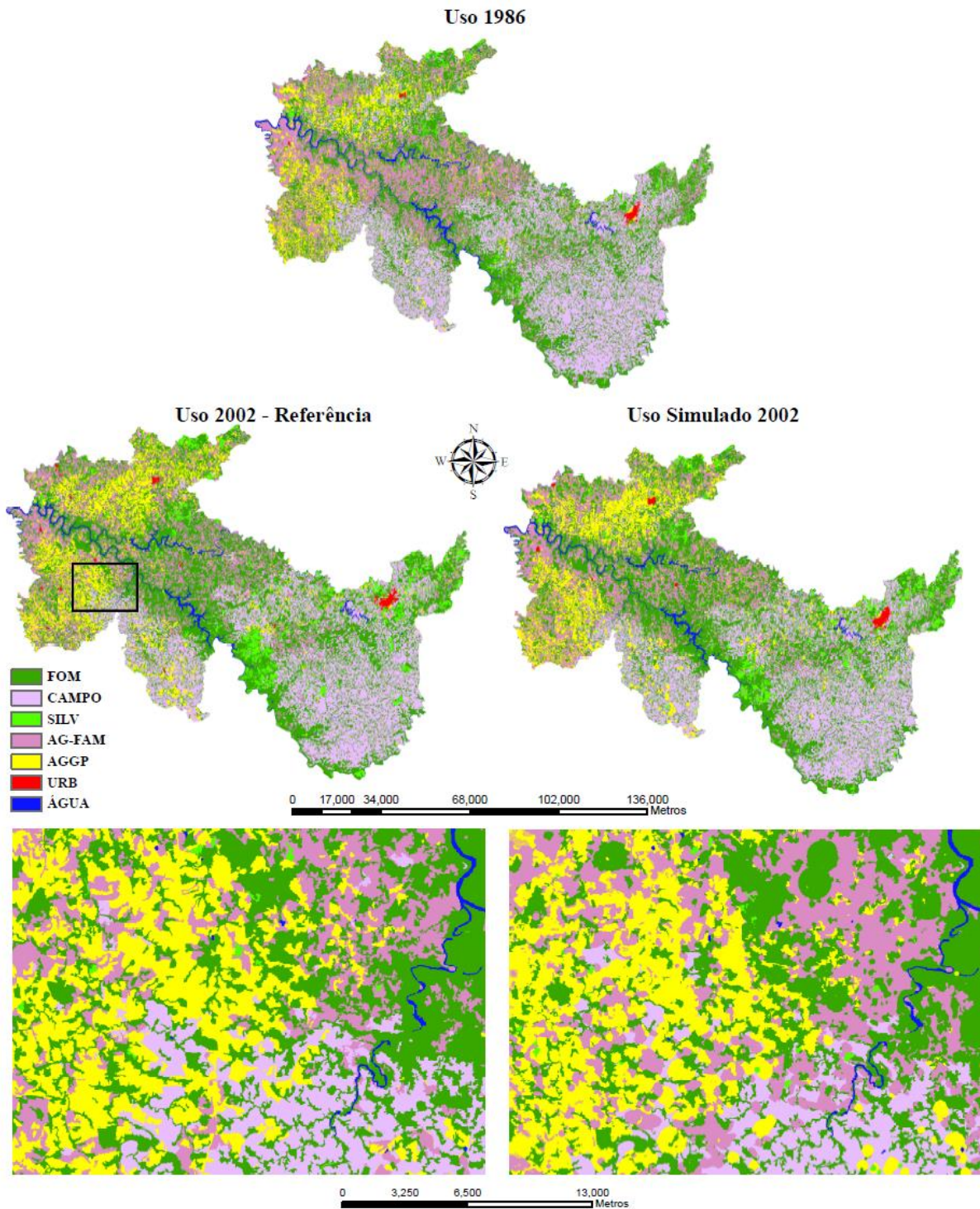


Figure 3. Results from the LUC model in the interval 1986-2002.

Figura 3. Resultado do modelo de MUCT no intervalo 1986-2002. Classes de floresta (FOM), campo nativo (CAMPO), silvicultura (SILV), agricultura familiar (AG-FAM), agricultura de grande porte (AGGP), urbano (URB) e água.

Em relação ao particionamento entre as funções de alocação de autômato celular determinística e de formação de manchas, a maior parte das transições foi alocada com a primeira função. Já a função de formação de manchas foi mais utilizada nas transições relacionadas com o processo de expansão da silvicultura e também na transição de campo nativo para agricultura de grande porte, transições estas que apresentam uma maior proporção de novas manchas.



## CONSIDERAÇÕES GERAIS

O modelo LandSCAM apresentado neste trabalho se caracteriza como uma ferramenta para a modelagem das mudanças de uso e cobertura da terra, a qual é um tema fundamental em diversos campos relacionados com o planejamento territorial dos espaços rurais e urbanos, em especial para atividades florestais, agrícolas, agroindustriais e outras que se preocupem com a adoção de práticas que aliem a produção e o desenvolvimento econômico com a conservação dos recursos naturais e qualidade ambiental.

## REFERÊNCIAS

- CASTELLA, J.; VERBURG, P.H. Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. *Ecological Modelling*, v. 202, p. 410-420, 2007.
- COSTANZA, R. Model goodness of fit: a multiple resolution procedure. *Ecological Modelling*, v. 47, p. 199-215, 1989.
- HARREL JR., F.E. *Design package*. 2012.
- HOSMER, D.W.; LEMESHOW, S. *Applied logistic regression*. John Wiley Sons, New York. 391 p, 2000.
- MACGARIGAL, K.; MARKS, B.J. *Fragstats*. US Department of Agriculture, Portland, OR. 122 p, 1995.
- MARKOV, A.A. An example of statistical investigation of the text Eugene Onegin concerning the connection of samples in chains. *Science in Context*, v. 19, p. 591-600, 2006.
- R DEVELOPMENT CORE TEAM (2008). R: a language and environment for statistical computing.
- SOARES-FILHO, B.S.; CERQUEIRA, G.; PENNACHIN, C.L. DINAMICA: a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling*, v. 154, n. 3, p. 217-235, 2002.
- VERBURG, P.H.; SOEPBOER, W.; VELDKAMP, A.; LIMPIADA, R.; ESPALDON, V.; MASTURAS, S. S. A. Modeling the spatial dynamics of regional land use: the CLUE-S model. *Environmental Management*, v. 30, n. 3, p. 391-405, 2002.