

Agent placement for land change models in frontier regions

Sérgio Souza Costa^{1,2}
Ana Paula Aguiar²
Gilberto Câmara²
Tiago Carneiro²
Evaldinolia Moreira¹

¹ Instituto Federal de Educação, Ciência e Tecnologia – (IFMA)
São Luis – MA - Brasil
{skosta@gmail.com, evaldinolia@ifma.edu.br}

² National Institute for Space Research (INPE)
Av. dos Astronautas, 1758, 12227-001, São José dos Campos, SP, Brazil
{ana.aguiar@inpe.br, gilberto.camara@inpe.br }

³ Earth System Simulation Laboratory, Federal University of Ouro Preto (UFOP)
Campus Universitário, Morro do Cruzeiro, Ouro Preto, MG, 35900-000, Brazil.
{tiagogsc@gmail.com}

Abstract. Using agent-based models in large area studies is hard. Since researchers need local data and field knowledge to create empirical models, most ABM studies cover small areas. A great challenge is how the modellers place agents on the landscape (agent placement) and defines its attributes. This paper presents a method and a set of operators to support agent placement in models for frontier regions, such as Amazonia. These operations allow represent important process, like concentration of land ownership and land speculation.

Palavras-chave: ABM, Land Change, Models, Multi-agent

1. Introduction

Land use and cover changes result from complex social and biophysical causes. The implications of these changes in global warming have increased the interest in studying them. Differences in scientific discipline, tradition and research questions have resulted in different model approaches. In one hand, geographers and ecologists have focused on land change at the macroscale to identify social causes connected to the macroscale patterns. In another hand, researchers in the social sciences have a long tradition of study of the individual behaviour at the microlevel. The different level of analysis leads two main approaches in the literature for land change modelling, roughly split in pattern-based and agent-based models.

Pattern-based models do not represent human behaviour directly. Such models detach the quantity of change (how much change is expected) from location of change (where changes are likely to take place). In general, pattern-based models have three parts: the demand, potential and allocation modules. Demand for change is external. A global demand for food can increase the need for agricultural areas. Each place has a potential for transition between land use types; this potential depends on the relative importance of driving forces of change. Demand is spatially allocated based on each place's potential. Models that follow a pattern-based approach include CLUE (VELDKAMP, A. . L. F., 1996), CLUE-S (VERBURG, P. H. *et al.*, 2002), GEOMOD2 (PONTIUS, 2001) and DINAMICA (SOARES-FILHO, 2002). The similarities among pattern-oriented models enable the developing of meta-models, as LuccME, early discussed in (COSTA, S *et al.*, 2009).

Agent-based models (ABMs) originate from distributed artificial intelligence (DAI). In this approach, intelligence emerges from the interactions between several entities (or agents)

and the environment. One motivation for using ABMs is to describe complex patterns using simple rules. Early applications included the flocking behaviour of animals, human movement patterns, and the behaviour of economic and social agents. In the last decade, the land-change research community has explored ABMs with the aim of representing coupled socio-ecological systems (JANSSEN; OSTROM, 2006; PARKER, D C *et al.*, 2002; ROBINSON, DEREK T *et al.*, 2007; VERBURG, P., 2006). Agent-based models range from theoretical to empirical. Theoretical models use simple generalisable ideas, whereas empirical models require more complexity and case-specific data. Empirical models focus more than theoretical models on using data to simulate real-world problems (MANSON, S M *et al.*, 2012). Social scientists prefer using theoretical models to represent human behaviour (CONTE *et al.*, 2001; EPSTEIN; AXTELL, 1996; SAWYER, R.K., 2003). Land-change scientists prefer using empirical agent-based models to represent socio-ecological processes in a specific region and at a specific time (BERGER, THOMAS *et al.*, 2006; DEADMAN, PETER, 2005; HUIGEN, M. G. A., 2004; PARKER, D. *et al.*, 2008; VALBUENA *et al.*, 2009).

In land-change studies, an agent-based model consists of autonomous entities named agents and an environment, which represents a real-world landscape (PARKER, DAWN C *et al.*, 2002). The agents are usually located on a cell or on a parcel of land (e.g., a farm) that is composed of a set of cells. ABMs represent the world in great detail, which makes them an important tool for explaining how cities, regions and the global system itself evolve and change (PARKER, DAWN C *et al.*, 2002). However, this representation requires much data to determine a parameterisation. Determining an empirical parameterisation is a great challenge for ABM, as noted by (BATTY; TORRENS, 2005; CROOKS, A. *et al.*, 2008; EVANS *et al.*, 2008; MANSON, STEVEN M, 2007; O'SULLIVAN *et al.*, 2012). Moreover, in the initialisation stage, some agents are typically already located in the research area. During the simulation, others agents arrive and search for a place to settle. Thus, the models require a method of agent placement for both the initialisation stage and during the simulation. In theoretical models, the agent placement method may be simple random placement (EPSTEIN; AXTELL, 1996). Empirical models require much data, such as census and cadastral data. In this section, we present the agent placement methods used in the example models.

Manson *et al.* (2012) argue that modellers rarely possess the individual-level data necessary to populate agents such as households. They have either a limited set of random samples (for example, household surveys or phone interviews) or, more often, spatially aggregated data (for example, census data or regional economic information). These data are collected for other purposes at different scales and by different government agencies. For example, the Brazilian agricultural census provides aggregate data regarding municipalities that are approximately 150,000 km² in size. Agent placement in models for frontier regions, such as Amazonia, is a significant challenge because land changes are frequent and immigration is high in these regions. For example, the numbers of farms in São Félix do Xingú increased from 1375 to 5890 in 10 years. The total area of farms increased from 2300 km² to 9600 km², an increased of approximately 7000 km². This process leads to the incorporation of new land into the productive system through deforestation and (both legal and illegal) land appropriation. In addition, these regions have great farm diversity. For example, in São Félix do Xingu in 2006, there were approximately 4000 farms of area less than 50 ha; the area of such farms was approximately 80,000 ha in total (IBGE, 2006). However, in the same period and region, a single farm could have an area greater than 60,000 ha. Thus, a location decision made by a single farmer can have a great impact on the resulting landscape because a single farm could occupy an area of thousands of hectares. These characteristics make our model different from models previously presented in the literature,

such as VALBUENA (2009) and LUCITA. For example in VALBUENA (2009), the research area is a small agricultural region (60,000 ha) and immigration is not modelled. LUCITA fixes the number of farms used in the simulation.

In this paper, we present a TerraME extension that aids the development of agent-based models for large frontier areas. TerraME is open-source modelling software that is coupled to geographical databases (CARNEIRO, T. G. S., 2006). The TerraME model is implemented in the Lua language (IERUSALIMSCHY, 2006) with extra functions and data types. These functions support multiscale models in cell spaces for both agent-based and automaton-based models (MOREIRA, E. *et al.*, 2009). In this paper, we propose a set of data types and functions written in the Lua language that extends the TerraME.

2. Agent placement in previous work

In this chapter we analyse three agent-based models. First, we present a survey of the models that describes the purpose, research area and scale of each model and how they place agents on the landscape (agent placement) and define its attributes.

LUCITA (Land-Use Change in the Amazon) is an important agent-based model for explaining spatial patterns in the Amazon (CABRERA, A R *et al.*, 2012; DEADMAN, PETER, 2005; ROBINSON, D.T., 2003). Implemented in RePast, the model provides an experimental laboratory to explore the effects of different parameters (for example, household composition and soil quality) on patterns of land-use change. The model uses a spatial resolution of 1 ha and the research area is 146,950 ha located in western Altamira, Pará, Brazil. One time step in the model represents 1 year, and the model run for 30 years, theoretically covering the time period 1970–2000. LUCITA initialises the landscape as virgin forest with no settled households and 3,916 available properties. The properties have an average size of approximately 100 ha and are arranged along the Trans-Amazonian Highway and a series of side roads that run vertical to the main highway for approximately 5 km. Each year, households arrive in the region and each occupy an available property. If several properties are available, a household selects the property that is closest to the main road that leads to Altamira. The household randomly chooses how many crops to plant and then randomly selects the crops. LUCITA does not simulate the creation, division or union of farms. Moreover, the land market is not directly modelled. When a household fails, it abandons the farm and the land becomes available for another household.

MameLuke uses the action-in-context method to construct a settlement model for the watershed in San Mariano, Philippines. The study area is 146,900 ha, and the model has a resolution of 1 ha. One time step in the model represents 1 year, and the model runs for 100 years, covering the time period 1900–2000. Implemented in RePast, the model aims to understand the relationship between population and deforestation (HUIGEN, M. G. A., 2004; HUIGEN, M. G. A. *et al.*, 2006). It models the dynamics between demography and ethnic identity in a spatially and temporally explicit model. The MameLuke model enables one to incorporate life histories from the field into a computerised model. The MameLuke initial conditions consist of approximately 1000 actors (approximately 200 households with an average of approximately five members each). The actor and household attributes are randomly assigned. The household spatial distribution is based on the location of ethnic groups. In each model step, a given number of households of a certain ethnicity arrive and find a place to settle. To select a location, they consider factors such as the proximity of co-members of their ethnic group, the slope of the land and the proximity of roads or rivers. When a household actor arrives, it searches for a location with a given slope that is within a Moore neighbourhood of the river or a road. The Moore neighbourhood size depends on ethnicity. If an agent does not find a satisfactory farm, it moves in a random direction and

repeats the process. Prior to the logging boom, households search for location along the river; afterward, they search for locations along a road. The PONs AlongRiver and AlongRoad represent, respectively, these behaviours, where the latter PON is activated after the construction of roads. In contrast to LUCITA, farms in MameLuke can be created and divided during the simulation. When the head of a household dies, the farm is divided amongst the children. If a household is childless, the farm becomes available for occupation by other households. However, similar to LUCITA, in MameLuke, the land market is not modelled directly.

Valbuena et al. (2009) present an ABM to analyse and explore regional land change that includes ideas such as agent typologies, farm trajectories and probabilistic decision-making processes. It simulates a region of area 600 km² in the Netherlands that is composed of 2700 small agricultural holdings. The model is implemented in Netlogo and runs for a period of 20 years, representing the time period 2005-2025. The model integrates data from different sources, including cadastral and census data. It describes three primary spatial processes in the region: farm cessation, farm expansion and the diversification of farming practices (for example, nature conservation, tourism and recreation practices). The model focuses on how external and internal factors, such as subsidies for landscape conservation, affect farmers' decisions. The Valbuena model starts in 2005 with 2741 agents and runs for 20 years. The final number of agents is 2302 because 16% of the agents stop farming during the course of the simulation. Cadastral data are used for spatially distributing the agents. The authors analyse census and survey data to develop an agent typology. In contrast to MameLuke and LUCITA, the land market is modelled directly. If there is a field or farm available, the closest agent can buy it. Similarly, the model selects which field an agent will sell depending on the field's distance to the owner. Nevertheless, during the simulation, no agent arrives and no farms are created, i.e., Valbuena does not model immigration.

3. Agent placement in models for frontier regions

The immigration in frontier regions leads to an intense land-market process, which includes buying, selling and illegal appropriation of land. The farm is the basic decision unit that refers to *Land-Market Decisions*. The *Farm* entity is a key entity for agent placement. It changes over time according to land-market decisions, in which farms are created, divided, and merged and farms' spatial relations are changed. Thus, we develop five farm operators:

- *Create operator*, which instantiates a new farm given an initial settlement cell and a farm area.
- *Border operator*, which identifies which cells represent the boundaries of the farm (a farm is represented as a set of cells).
- *Neighbourhood operator*, which identifies the spatial adjacency relations amongst the farms.
- *Split operator*, which divides a given region into several farms of specified areas.
- *Merge operator*, which joins two farms and updates the farm boundaries.

The *create operator* receives an initial cell and farm area. The choice of initial cell may depend on factors such as the proximity of rivers or roads (see Figure 1.a). The operator adds the available neighbouring cells until a given area is attained. The algorithm randomly selects neighbouring cells based on factors such as their proximity to roads. The shape of the farm depends on the farm's location, and it adjusts according to spatial limits, such as existing farms and large rivers, as illustrated in Figure 1.b.

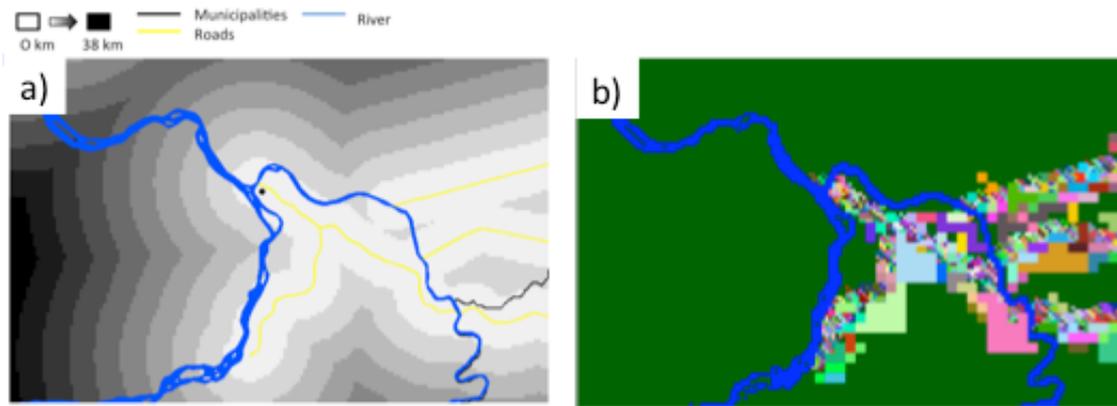


Figure 1. Distance to roads in 1985 (a). Hypothetical land tenure, where the distance to roads is used as a location factor (b)

In Figure 1.b we randomly located 8,000 farms, where 80% had areas of 25-50 ha, 18% had areas of 50-500 ha, and 2% had areas of 500-1000 ha. The algorithm yielded some emergent trends, such as small farms being placed closer to roads than large farms. This pattern is common in the region, as discussed in (ESCADA *et al.*, 2005). This pattern depends on the proportion of small and large farms. Because the number of small farms is greater than that of large farms, the small farms occupy the regions closest the roads, whereas large farms are located farther away.

After the farm creation, the initial settlement cell is marked as the location of the agent's house and for each cell, the algorithm calculates the distance of the cell from the house. This distance is important, for example, because when changing a set of cells from forested to deforested, the farmer minimises the distance to the house. Similar to (CABRERA, ARTHUR RAYMOND, 2009), we use the Chebyshev distance heuristic to produce square or nearly square land allocations. In contrast, a Euclidean distance heuristic would produce more circular allocations, as demonstrated in Figure 2.

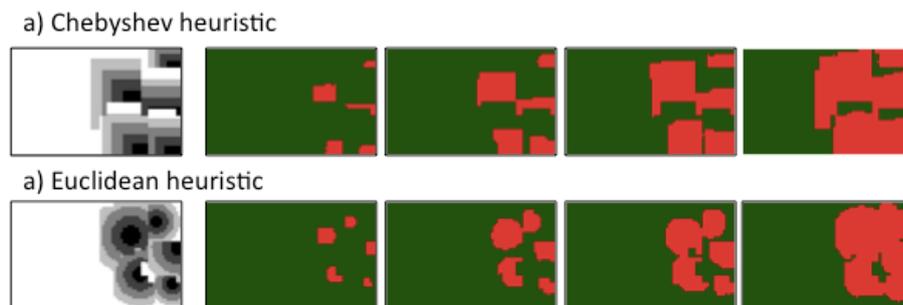


Figure 2 Different patterns yielded by different distance heuristics

Given the theoretical land tenure shown in Figure 1.b and considering that each agent deforests 10% of its farms, the model produces the deforestation pattern illustrated in Figure 3.

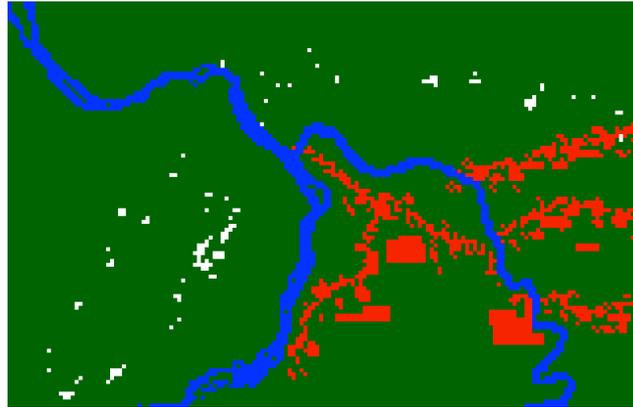


Figure 3. Deforestation pattern.

For each new farm, we perform the *border operator*. Because each farm is a set of cells, we must know their boundaries to identify the spatial relations amongst them. We use an attribute to represent the farm boundaries; an attribute value of 1 indicates that there is a border, as illustrated in Figure 4.

0	0	0	0	0	0	0	0
0	1	1	1	1	1	0	0
0	1	0	0	0	1	0	0
0	1	0	0	0	1	1	0
0	1	1	1	0	0	1	0
0	0	0	1	1	1	1	0

Figure 4. The farm boundary representation

The *neighbour* operator uses the border attributes to identify the spatial relations amongst the farms. If the border of two farms touches, then the farms are neighbours. For example, in Figure 5, farm B is a neighbour of farms A and C.

1	1	1	1	1	1	1	1	0	0	0	0	A
1	0	0	1	1	0	0	1	0	0	0	0	B	———
1	0	0	1	1	0	0	1	0	0	0	0	C	- · -
1	0	0	1	1	0	0	1	0	0	0	0		
1	1	1	1	1	1	0	0	0	0	0	0		
0	0	0	0	1	0	0	1	0	0	0	0		
0	0	0	0	1	1	1	1	0	0	0	0		
1	1	1	1	1	1	1	1	1	0	0	0		
1	0	0	0	0	0	0	0	1	0	0	0		
1	0	0	0	0	0	0	0	1	0	0	0		
1	1	1	1	1	1	1	1	1	0	0	0		

Figure 5 Representation of the spatial relations amongst farms

A land-market decision may consider the spatial relations amongst farms because a farmer may prefer to buy neighbouring farms rather than distant farms. When a farmer buys a neighbouring farm, the farms can be joined, which leads to concentration of land ownership. Concentration of land ownership is common in Amazonia. For example, between 1996 and

2006, Tucumã, PA, had significant concentration of land ownership. In 1996, farms with individual areas greater than 1000 ha accounted for 8% of the total area of farms. In 2006, they accounted for greater than 60% of the total (IBGE, 2006). In the same period, the number of farms decreased from 2518 to 1039, which suggests that the owners of several small farms sold their farms to the owners of large farms. The *merge* operator enables the model to represent this process, in which the owners of large farms buy neighbouring farms and thereby increase their area. Because a farm is a set of cells, the merge operation is simple. The operator adds the source farm cells to the target farm. Next, it updates the cell attributes and farm boundary using the *border* operator. Another important process in Amazonia is land speculation, in which an agent purchases a large region and later divides it into small farms. The algorithm used to represent this process is similar to the *create* operation, where the primary difference is that the speculation operator creates farms from existent farms.

4. Final remarks

This paper shows how to use a set of farm operators to support agent placement in models for frontier regions. These operators are essential for the simulation and initialisation when cadastral data are not available. When these data are available, we can use other approaches, such as that presented in (ASSIS, 2012), for initial agent placement. However, agent-based models for frontier regions must deal with intense land market processes. Thus, these operators are still necessary during the simulation. In a submitted paper, (COSTA, SÉRGIO *et al.*, 2013), the authors use these operators to support an agent-based model for a large frontier region of 60,000 km² in southeast of Pará.

References

- ASSIS, T. O. **Relações entre agentes e o espaço em modelos de mudança de uso e cobertura da terra baseados em agente**. [S.l.]: Instituto Nacional de Pesquisas Espaciais, 2012.
- BERGER, THOMAS; SCHREINEMACHERS, P.; ARNOLD, T. Mathematical programming-based multi-agent systems to simulate sustainable resource use in agriculture and forestry. **Internal document, University of Hohenheim, Stuttgart, Germany**, 2006.
- CABRERA, A R; DEADMAN, P; MORAN, E; BRONDÍZIO, E. S.; VANWEY, L. K. Exploring demographic and lot effects in an ABM/LUCC of agriculture in the Brazilian Amazon. **Agent-Based Models of Geographical Systems**, p. 663-676, 2012.
- CABRERA, ARTHUR RAYMOND. Land Use Change and Economic Opportunity in Amazonia : An Agent-based Model by. 2009.
- CARNEIRO, T. G. S. NESTED-CA: A FOUNDATION FOR MULTISCALE MODELLING OF LAND USE AND LAND COVER CHANGE. 2006.
- CONTE, R.; EDMONDS, B.; MOSS, S.; SAWYER, R K. Sociology and social theory in agent based social simulation: A symposium. **Computational & Mathematical Organization Theory**, v. 7, n. 3, p. 183-205, 2001.
- COSTA, S; AGUIAR, A. P.; CÂMARA, G.; MOREIRA, E. G. **Common Concepts to Development of the Top-Down Models of Land Changes**. XIV SBSR Simpósio Brasileiro de Sensoriamento Remoto. **Anais...** Natal, RN: INPE. , 2009
- COSTA, SÉRGIO; DUTRA AGUIAR, A. P.; CÂMARA, GILBERTO; *et al.* Using institutional arrangements and hybrid automata for regional scale agent-based modelling of land change (submitted). 2013.
- DEADMAN, PETER. Household Decision Making and Patterns of Land Use Change in LUCITA: An Agent Based Simulation of the Altamira Region, Brazil. **Simulation**, p. 120-126, 2005.
- EPSTEIN, J. M.; AXTELL, R. **Growing Artificial Societies: Social Science from the Bottom Up**. [S.l.]: The MIT Press, 1996. v. 64p. 208

ESCADA, M. I. S.; VIEIRA, I. C. G.; KAMPEL, S. A. *et al.* Processos de ocupação nas novas fronteiras da Amazônia: o interflúvio do Xingu/Iriri. **Estudos Avançados**, v. 19, n. 54, p. 9–23, 2005.

HUIGEN, M. G. A. First principles of the MameLuke multi-actor modelling framework for land use change, illustrated with a Philippine case study. **Journal of environmental management**, v. 72, n. 1-2, p. 5-21, ago 2004.

IBGE. **Censo Agropecuario 2006 (Agricultural Census 2006), Instituto Brasileiro de Geographia e Estatística (Brazilian Census Bureau).**

IERUSALIMSCHY, R. **Programming in Lua, Second Edition.** [S.l.]: Lua.org, 2006.

JANSSEN, M. A.; OSTROM, E. Empirically Based , Agent-based models. **Search**, v. 11, n. 2, p. 1-12, 2006.

MANSON, S M; SUN, S.; BONSAI, D. Agent-based modeling and complexity. **Agent-based models of geographical systems.** [S.l.]: Springer, 2012. p. 125-139.

MOREIRA, E.; COSTA, SÉRGIO; AGUIAR, A. P.; CÂMARA, G.; CARNEIRO, T. Dynamical coupling of multiscale land change models. **Landscape Ecology**, v. 24, n. 9, p. 1183-1194, set 2009.

PARKER, D C; BERGER, T; MANSON, S M; OTHERS. Agent-based models of land-use and land-cover change. **LUCC Report Series**, v. 6, n. 6, 2002.

PARKER, D.; ENTWISLE, B.; RINDFUSS, R. *et al.* Case studies, cross-site comparisons, and the challenge of generalization: comparing agent-based models of land-use change in frontier regions. **Journal of Land Use Science**, v. 3, n. 1, p. 41-72, 2008.

PARKER, DAWN C; BERGER, THOMAS; MANSON, STEVEN M; *et al.* Meeting the Challenge of Complexity Proceedings of a Special Workshop on Land-Use / Land-Cover Change. n. 6, 2002.

PONTIUS, R. Modeling the spatial pattern of land-use change with GEOMOD2: application and validation for Costa Rica. **Agriculture, Ecosystems & Environment**, v. 85, n. 1-3, p. 191-203, jun 2001.

ROBINSON, D.T. Modelling Farmer Household Decision-making and Its Effects on Land use/cover Change in the Altamira Region, Pará, Brazil. **master's thesis, Department of Geography, Faculty of Environmental Studies, University of Waterloo, Waterloo, Ontario, 2003.**

ROBINSON, DEREK T; BROWN, D. G.; PARKER, DAWN C; *et al.* Comparison of empirical methods for building agent-based models in land use science. **Journal of Land Use Science**, v. 2, n. 1, p. 31-55, mar 2007.

SAWYER, R.K; Artificial Societies: Multiagent Systems and the Micro-Macro Link in Sociological Theory. **Sociological Methods & Research**, v. 31, n. 3, p. 325-363, 2003.

SOARES-FILHO, B. 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, set 2002.

VALBUENA, D.; VERBURG, P. H.; BREGT, A. K.; LIGTENBERG, A. An agent-based approach to model land-use change at a regional scale. **Landscape Ecology**, 2009.

VELDKAMP, A. . L. F. CLUE: a conceptual model to study the Conversion of Land Use and its Effects. **Ecological Modelling**, v. 85, n. 2-3, p. 253-270, mar 1996.

VERBURG, P. Simulating feedbacks in land use and land cover change models. **Landscape Ecology**, v. 21, n. 8, p. 1171-1183, 2006.

VERBURG, P. H.; SOEPBOER, W.; VELDKAMP, A; *et al.* Modeling the spatial dynamics of regional land use: the CLUE-S model. **Environmental management**, v. 30, n. 3, p. 391-405, set 2002.