

[J. Gary Polhill, Dawn Parker, Daniel Brown and Volker Grimm \(2008\)](#)

## Using the ODD Protocol for Describing Three Agent-Based Social Simulation Models of Land-Use Change

*Journal of Artificial Societies and Social Simulation* vol. 11, no. 2 3  
<<http://jasss.soc.surrey.ac.uk/11/2/3.html>>

For information about citing this article, click [here](#)

Received: 03-Aug-2007 Accepted: 18-Dec-2007 Published: 31-Mar-2008



### Abstract

This article describes three agent-based social simulation models in the area of land-use change using a model documentation protocol, ODD, from the ecological literature. Our goal is to evaluate how well fitted it is to social simulations and how successful it might be in increasing communication between individual- and agent-based modellers. Such shared protocols can facilitate model review, comparison, and replication. We initially conclude that the framework holds promise as a standard communication mechanism, although some refinements may be needed.

### Keywords:

Model Communication, Model Documentation, ODD, LUCC

### Introduction

*Hercules himself must yield to ODDs*  
Henry VI part III, Act II Scene I

#### 1.1

As agent-based social simulation matures as a domain of research, there are various attempts to make sense of the plethora of models, software, simulations, results, frameworks and libraries arising from it. In the area of land use and cover change (LUCC), members of the community met in 2001 at Irvine, California, in a workshop to compare agent-based models under development at the time. The resulting report ([Parker, Berger, and Manson 2002](#)) was possibly one of the first efforts to document and compare models variously termed under the headings MAS/LUCC, ABM/LUCC and ABLUMs.

#### 1.2

Parker, Berger, and Manson ([2002](#)) conclude with a set of questions that are relevant for literature in the field, defining potential criteria for referees and journal editors (p. 88):

Is the author clear about the goals of the model? Are these goals appropriate? Has the model appropriately represented relevant spatial processes? Have standard techniques for verification and validation been used? Are the mechanisms of the model clearly communicated to the audience? Have the model mechanisms been appropriately verified? How does the model compare to other ongoing ABM/LUCC work?

#### 1.3

These questions are as relevant today as they were then. Unlike mathematical models, computer simulations have no standard language or protocol for communication ([Grimm and Railsback 2005](#)). A few mathematical equations and relevant derivations or conclusions from them can easily be included in a journal article, where they, and any implicit assumptions, are open to scrutiny by anyone fluent in mathematics. The equivalent in software is the source code of the simulation program, which may run into several hundred, if not thousands or tens of thousands of lines. These reams of material are hardly appropriate for inclusion in a journal article, and even if the source code is made available on the internet, the practicalities

involved make their detailed scrutiny by the academic community unlikely. Yet any one line of code could embed a critical, possibly implicit, assumption (or even a mistake!) that might affect the acceptability of the model.

#### 1.4

These issues are highlighted by those who report on reimplementation and replication work. Edmonds and Hales (2003) report that simply changing a '>=' in an if statement to a '>' in a model of Riolo, Cohen and Axelrod (2001) dramatically reduces the effect emergent in their results. Galan and Izquierdo (2005) also show that changing arbitrary assumptions in one of Axelrod's (1986) models and running the simulation for longer generates the opposite results to those reported. Many authors undertaking replication work report that a great deal of communication between the original developers and those building the replication was necessary (Axtell, Axelrod, Epstein and Cohen, 1995; Bigbee, Cioffi-Revilla and Luke 2005; Wilensky and Rand 2007), which further points to the inadequacy of model descriptions in the literature.

#### 1.5

Although not all of these issues can be addressed in the literature itself, standards for documenting models in journal articles would at least ensure that all the relevant ground has been covered (Richiardi, Leombruni, Saam and Sonnessa 2006). Together with links to further standardised supporting material elsewhere, it becomes possible to provide the kind of rigour and repeatability in simulation that a scientific endeavour ideally requires. To this end, Grimm et al. (2006) have proposed a standard protocol dubbed ODD (Overview, Design concepts, and Details) aimed at the description of individual-based models in ecology, with a view to its application in agent-based social simulation and other disciplines. Although limitations of the approach are acknowledged (p. 124), these limitations could potentially be remedied through refinement (rather than radical alteration) of the protocol as it becomes more widely applied. Interestingly, much of the material that ODD stipulates be provided addresses the questions cited above in Parker, Berger, and Manson (2002).

#### 1.6

The main focus of ODD is the description of individual- and agent-based models in the academic literature. It is designed to ensure that such descriptions are *readable* and *complete*. ODD is intended to facilitate readability through stipulating a structure for the description with a logical ordering. The structure also contains guidelines for the contents of the description, designed to ensure that all relevant information about a model is captured, at least for the purposes of understanding the application in the associated article, and for replication.

#### 1.7

Simulation models and programs have been around in the academic literature for a long time. It is not unreasonable to expect, therefore, that the issue of describing models in journal articles should have been resolved already. Various approaches for presenting the information in diagrammatic, tabular and text form have been suggested and used, including pseudocode, flowcharts, UML, systems dynamics diagrams and hypertext. Some, such as Agent UML (Bauer, Müller and Odell 2001) have been tailored for areas potentially relevant to the social simulation domain, whilst others are associated with specific development platforms; NetLogo (Wilensky 1999), for example, provides an information tab for developers to describe their models, and those development platforms using Java will be able to make use of Javadoc. Whilst these techniques may have their merits, none are specifically aimed at providing a complete, readable account of a model in a journal article. Further, the difficulties experienced by those doing replication suggest that, at least for this activity, they are not fulfilling the needs of the practitioners. It may also be the case that individual- and agent-based models have specific requirements for their description that justify a protocol tailored to them.

#### 1.8

ODD was developed in large part by the individual-based modelling community for description of models produced by its members. Individual-based models differ from agent-based models in that they generally model non-human entities interacting within an ecological system (Grimm and Railsback 2005), whereas agent-based models generally model human actors and decisions (Gilbert and Troitzsch 2005). The chief purpose of this paper is to assess the extent to which the ODD framework is sufficiently general to be of use to agent-based as well as individual-based modellers.

#### 1.9

There are many reasons to work towards better communication and collaboration between these two groups of researchers, especially as far as models of land use and natural resource management are concerned. First, more and more models are being developed that focus on the interactions between human decision makers and non-human organisms (for example Dreyfus-Leon and Kleiber 2001; Harper, Westervelt and Trame 2002; Mathevet, Bousquet, Le Page and Antona 2003). This means that individual- and agent-based modellers (I/ABM) have

a strong motivation to develop an efficient means of communication. In particular, if a single framework can be developed that comprehensively describes both the individual- and agent-based components of the model, model description and communication is significantly simplified.

**1.10** Second, a common framework for describing adaptive agents in general would be desirable because ecological and economic models often focus on similar processes: competition for scarce resources, and the evolution of successful strategies for resource acquisition. Many other concepts, such as fitness, increasing and diminishing returns, information gathering, expectation formation, and migration, are commonly investigated by both groups, suggesting the potential for sharing of model components. For example, individual-based models in ecology are increasingly taking into account adaptive decisions of animals and plants (Railsback and Harvey 2002; Goss-Custard et al. 2006; Grimm, 2007), and modelling adaptive behaviour is one of the central motivations for agent-based modelling (Berry, Kiel, and Elliot, 2002). ODD makes no distinction between individual- and agent-based models, and therefore it should be possible to use ODD to describe any bottom-up simulation model, including grid-based or cellular automaton models (Grimm et al. 2006). It remains, however, an open question whether the Design Concepts, a part of the ODD protocol formulated for organisms (e.g., fitness, adaptation, sensing), can also be applied to human agents and problems from the social sciences. However, given the overlap in the concepts studied by both fields, it seems reasonable to expect success in cross-disciplinary application of this section of the ODD protocol.

**1.11** Here, we consider the suitability of ODD to describing models that address LUCS questions through applying it to three models in the ABM/LUCS literature: SLUDGE (Simulated Land Use Dependent on eDGe Effect externalities) (Parker 1999; Parker and Meretsky, 2004; Parker 2005), SOME (SLUCE's Original Model for Experimentation) (Brown and Robinson 2006; Rand et al. 2003; Brown, Page, Riolo, Zellner and Rand 2005; Brown, Page, Riolo and Rand 2004), and FEARLUS (Framework for Evaluation and Assessment of Regional Land Use Scenarios) (Polhill, Gotts and Law 2001; Gotts, Polhill and Law 2003) with the ELMM (Endogenised Land Market Model) extension (Polhill, Parker and Gotts 2005; Polhill, Parker, and Gotts 2008). By so doing, we are also able to show some of the diversity with which the ODD protocol might be interpreted by authors of models in a similar application domain, even with the normative influence of its developer on the descriptions demonstrated. After an introduction to the ODD protocol in section 2, we provide descriptions of the three models in section 3, and discuss our experience in using ODD in section 4 before concluding.

## The ODD protocol

**2.1** The ODD protocol is described in detail in Grimm et al. (2006; see also <http://www.ufz.de/oesatools/odd/>). The following provides a brief overview, relating it to Parker, Berger, and Manson's (2002) questions cited above, and emphasising the key goals of ODD: readability and completeness. ODD stands for 'Overview, Design concepts and Details', which collectively comprise the three major categories of sections that ODD requires of a model description (Figure 1). Of these categories, the Overview category is broken down into three sections: 'Purpose', 'State Variables and Scales' and 'Process Overview and Scheduling'; and the Details category has sections: 'Initialization', 'Input' and 'Submodels'. The Design Concepts category is also a single section, providing a list of concepts that could be discussed.

<b>Overview</b>	<b>Purpose</b>
	<b>State Variables and Scales</b>
	<b>Process Overview and Scheduling</b>
<b>Design Concepts</b>	<b>Design Concepts</b>
<b>Details</b>	<b>Initialization</b>
	<b>Input</b>
	<b>Submodels</b>

**Figure 1.** The seven elements of the ODD protocol. The three categories on the left side are only for explaining the general structure of the protocol but are not used while describing a model. Rather, a model description following ODD has the seven sections listed on the right

## 2.2

The Overview category is intended to contain sufficient detail to create an outline of the model: the entities in the model (e.g. agents, collectives of agents, patches of the landscape), the processes, and the model's schedule. The 'Purpose' section is intended to explain what is to be done with the model, priming the reader's expectations of what is and what is not included in the ensuing model description, and addressing the questions about the goals of the model from Parker, Berger, and Manson ([2002](#)). If the model is designed to explain certain 'stylized facts' ([Kaldor 1961/1968](#)) or sets of observed 'patterns' ([Grimm et al. 2005](#)), the 'Purpose' section will help readers understanding why certain state variables and processes are included: to allow the same stylized facts or patterns to emerge in the model ('pattern-oriented modelling'; [Grimm and Railsback 2005](#)).

## 2.3

The 'State Variables and Scales' section outlines the structure of the model, specifying all types of entities, or 'things', that are in the model (e.g., types of agents, spatial units, environmental variables) and *all* their low-level state variables that constitute the state of these entities at a certain point in time. Low-level state variables cannot be deduced from other state variables. Age, for example, can be a low-level state variable of an agent, whereas average age is not low-level and therefore would not be listed as a state variable in ODD. Grimm et al. ([2006](#)) recommend the use of UML class diagrams, particularly if the number of state variables is large. The spatial (if applicable) and temporal scales addressed by the model are also covered in this section, with justifications. The 'Process Overview and Scheduling' section follows, listing all the processes that occur in the model and how they are scheduled: who is doing what and when and in which order, and when the state variables are updated. Clearly it would not be possible to discuss this adequately without first having covered what the entities in the model are in the 'State Variables and Scales' section. The Overview category should thus allow readers to re-implement the skeleton of the model, i.e. the entities, or objects therein, including their state variables, the schedule, and the headers of the functions or methods representing the processes. Collectively, the 'State Variables and Scales' and 'Process Overview and Scheduling' sections relate back to the questions of adequate spatial representation and communication of mechanisms in Parker, Berger, and Manson ([2002](#)).

## 2.4

The ODD protocol does not follow the object-oriented paradigm of programming: the state variables of the model's entities (objects) are presented separately from their processes (methods). The reason is that the developers of ODD considered model descriptions that clearly separate between the description of the state of the model system and its dynamics easier to read and understand. The state description sets the arena in which the equations and transition rules that can make the state variables change are applied. Nevertheless, it is easy to combine the structure of ODD and the object-oriented organisations of programs by using UML class diagrams. Moreover, the link between a model's entities and processes is provided in the description of the model's schedule, which specifies: who (i.e. which entity or entities) is doing what (i.e. performing which of its/their processes) at what time.

## 2.5

The 'Design Concepts' section does not describe the model itself, but the general concepts underlying its design. The purpose of this is to make sure that authors and readers respectively make and understand important design decisions consciously. The model description is thereby linked to existing common threads in the relevant literature, (addressing a similar question in Parker, Berger, and Manson ([2002](#))). Suggested material for this section includes:

- Emergence (a summary of emergent phenomena from the interaction of the agents),
- Adaptation (how the agents adapt their behaviour to their and their environments current state),
- Fitness/Objectives (a summary of the agents' goals),
- Prediction (how the agents predict the consequences of their decisions),
- Sensing (environmental variables perceived by the agents, which might include their own variables),
- Interactions,
- Stochasticity (if present, and the reasons for it),
- Collectives (whether the agents are grouped socially) and
- Observation (how data are gathered from the model).

## 2.6

Again, this discussion relies on the entities and processes described in the Overview section. In contrast to the other parts of the ODD protocol, the Design Concepts section provides some degrees of freedom: those concepts that simply do not apply to the model to be described can be left out, for example if the model does not include explicit predictions made by the individuals or agents. Further, depending on the model's design, the sequence in which

the concepts are discussed can be changed.

## 2.7

The Details category puts the flesh on the skeleton outlined in the Overview, and should enable complete re-implementation of the model. In some cases, the amount of information in this section could be too much to be included in a journal article, and would appear either in an appendix or a separate linked document. The 'Initialisation' section describes how the model is started, providing, if appropriate, references to any data used to give initial values to the state variables. The 'Input' section describes any other external inputs to the model (such as time-series data of environmental variables, e.g. temperature in ecological models or overall demand in economic models).

## 2.8

In first applications of ODD, 'Input' has sometimes been misinterpreted as parameters or initial values of state variables. The latter, however, are described in the 'Initialization' section and parameters are presented in the 'Submodels' section, where often it will be useful to include a table with all relevant parameters, their units, reference values, data sources, and ranges over which they were varied.

## 2.9

It may be necessary for the 'Input' section to reference an online archive where the data and even the original random number seed can be accessed to achieve full reproducibility. (FEARLUS-G is a prototype grid-based environment for providing this kind of functionality for one particular model; [Polhill, Pignotti, Gotts, Edwards and Preece 2007](#).)

## 2.10

The 'Submodels' section explains in detail all the processes outlined in the 'Process Overview and Scheduling' subsection of the Overview. This includes describing how parameter values were chosen, and the testing and calibration of the submodels. It partly addresses the issues of verification and validation in Parker, Berger, and Manson ([2002](#)), but only at the level of submodels. Model testing and validation at the level of the entire model are not part of ODD, because ODD is designed only to describe a model, not a particular analysis based on it.



## Three LUC models

### 3.1

In describing the models, we adopt the convention that entities in the model (as opposed to the real world) are given an initial capital letter. The original descriptions of the three models, which might be compared to the following ODD-based descriptions, are given in: Parker ([1999](#)) (SLUDGE), Brown et al. ([2004](#)) (SOME), and Polhill, Parker and Gotts ([2005](#)) (FEARLUS+ELMM). For each model, the description using the ODD protocol is followed by some brief reflections on how well ODD has been able to capture the essential information about the model.

### 3.2

All three models are spatially explicit models of land-related social processes. SLUDGE is presented first. It is a relatively simple model in comparison with the other two, and is thus a useful illustration of the application of ODD to an agent-based model. SOME and FEARLUS+ELMM are then presented as urban and rural models of land use change respectively. SOME is also rather more data-driven than FEARLUS+ELMM, and to a certain extent, the three models could be seen as covering the three classes of ABM described by Boero and Squazzoni ([2005](#)). FEARLUS (without the ELMM extension), Boero and Squazzoni suggest (paragraph 4.36), is a 'Typification', a class of more theoretical models designed to explore a range of phenomena and questions within a particular area of study. SLUDGE belongs more to the class of 'Theoretical Abstractions', models which are more metaphorical in their representations, aimed at building on theory. SOME, though arguably also a Typification, crosses more into the 'Case Based Models' class, since it has been fitted to specific case-studies at the micro and macro levels through application of survey data and validation of spatial patterns. Boero and Squazzoni do acknowledge their classes are a continuum rather than definitions of discrete, mutually-exclusive sets (paragraph 3.21).

### 3.3

Whilst it is useful to explore models along Boero and Squazzoni's continuum to gauge the coverage of ODD to observed classes of agent-based model, a second reason to apply ODD to these three models is that they have also been described side-by-side elsewhere in the context of a standardised meta-modelling framework for agent-based models of land use, the MR POTATOHEAD framework ([Parker, Brown, Polhill, Deadman and Manson 2008](#)). This offers the interested reader an opportunity to compare ODD with another approach that has been used to communicate models.

## SLUDGE

**Purpose**

**3.4**

SLUDGE (Simulated Land Use Dependent on eDGe Effect externalities) is a simple combined cellular automaton and agent-based model designed to explore the effects of positive and negative distance-dependent spatial externalities on economic and landscape pattern outcomes. Spatial externalities refer to land-use activities by one land owner that affect the payoffs to the land uses of surrounding neighbours. SLUDGE is an abstract model designed for theoretical exploration and hypothesis generation. Specifically, SLUDGE was designed to extend existing analytical microeconomic theory to examine relationships between externalities, market mechanisms, and the efficiency of free-market land use patterns.

**3.5**

Effectively, the SLUDGE model functions as a search mechanism for a static equilibrium — one in which the current land-use pattern and composition are such that no agent has an incentive to choose a different land use. Quantity and price behaviour in the model mirror that of the classic cobweb model in agricultural economics, with oversupply occurring in response to a high price signal, a fall in price due to oversupply, and undersupply occurring in the next time period. The magnitude of these over-and-undershoots diminishes as the model reaches equilibrium. In contrast to a-spatial models, productivity effects of landscape pattern influence this process.

**State Variables and Scales**

**3.6**

The fundamental low-level state variable in the SLUDGE model is the profit-maximizing Land Use choice of the Land Owner agent. Each Land Owner agent is identical.

---

<b>Variable name</b>	<b>Brief description</b>
<i>Land Use</i>	Main decision variable for the agent. Two types: zero (fixed price) or one (quantity-dependent price).
<i>Coordinates</i>	X and Y coordinates of the cell over which the land owner decides land use.

---

**3.7**

Land Owners are agents that occupy a single cell in a uniform cellular Landscape Model. The Landscape Model has parameters that affect the profitability of each Land Use, Landscape Dimension, and X-Y coordinates that link to Land Owner agents.

*Landscape model:*

---

<b>Variable name</b>	<b>Brief description</b>	<b>Parameter in SLUDGE GUI</b>
<i>Productivity</i>	Per-cell land-use specific output variable.	Productivity_0 Productivity_1
<i>Externalities</i>	Positive or negative Productivity change with each neighbouring Land Use (4 variables). These result in a gain or loss of production for the first Land Use when it shares a border with the second.	One_With_0 One_With_1 Zero_With_0 Zero_With_1
<i>Output Price 0</i>	Parametric output price for output of Land Use 0.	Output_Price_0
<i>Landscape dimension</i>	Height and width of the model Landscape, in cells	Board_Size
<i>Coordinates</i>	X and Y coordinates of the cell, which determine transportation costs to market locations.	

---

**3.8**

A simple Supply and Demand Model calculates an expected and realised price for Land Use 1, which depends inversely on the quantity of production of Land Use 1.

*Supply and Demand model:*

---

Variable name	Brief description	Parameter in SLUDGE GUI
<i>Demand function</i>	Downward-sloping function and scaling parameter that determines price of Output 1.	Demand_Parameter
<i>Expected Price 1</i>	Price anticipated by each agent based on current Landscape and Economic Conditions.	
<i>Output Price 1</i>	Realised price of Output 1 dependent on total output of Land use 1.	

### 3.9

Market locations for the products of both Land uses are specified by the user, and profits for each cell are reduced by Transport costs to market via Euclidean distance.

*Transport cost model:*

Variable name	Brief description	Parameter in SLUDGE GUI
<i>Coordinates</i>	X and Y coordinates of the cell representing the market location for output of each Land use.	Market_0_X; Market_0_Y Market_1_X; Market_1_Y
<i>Transport costs</i>	Land-use specific per-unit transport cost.	Transport_Costs_0 Transport_Cost_1

### 3.10

The size of the cell-based Landscape is specified by the user (*Board\_size*). Thus, there is no absolute concept of spatial scale in the model (either extent or resolution). A relatively coarse resolution can be implemented by using a small Landscape size and smaller Demand parameter; and a relatively finer resolution model can be implemented through a larger Landscape size and larger Demand parameter. There is also no absolute concept of temporal scale since, as stated above, the model is effectively a search mechanism for a static equilibrium. The number of time steps required to reach an equilibrium depends on the initial conditions of the model, but is usually less than 20.

#### Process Overview and Scheduling

### 3.11

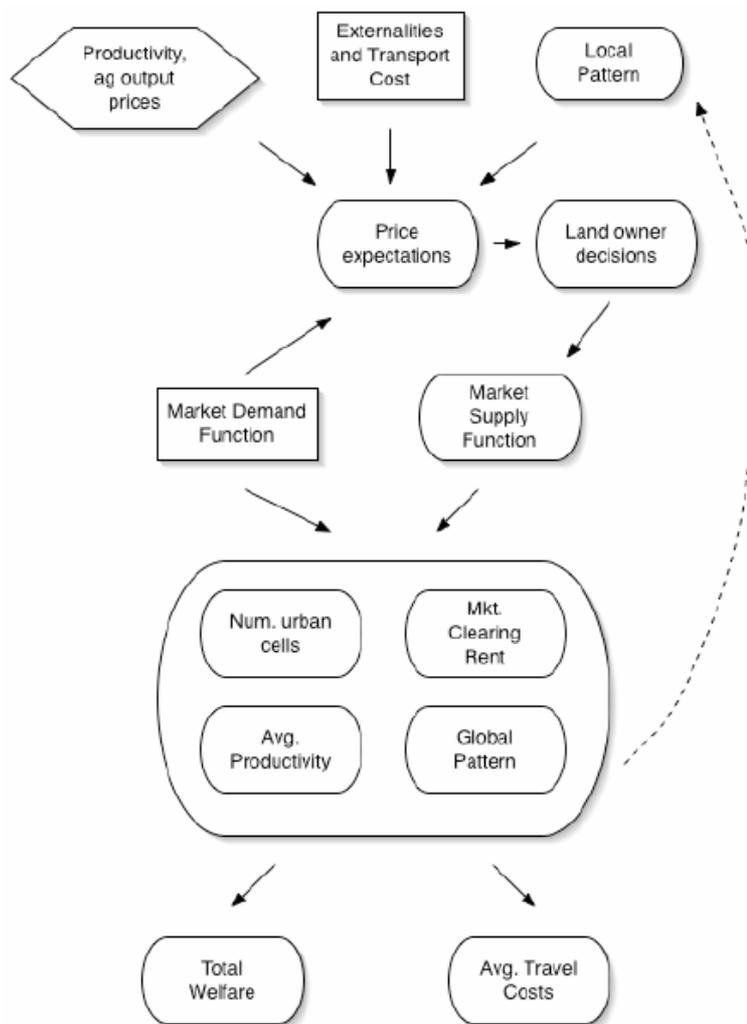
Basically, SLUDGE considers one main process, the profit-maximisation decision for each Land Owner agent for choosing Land Use zero or one. To make this decision, each agent first calculates an expected profit value for each of the two possible Land Uses, based on current Landscape and Market Conditions, calling the Demand and Supply model. To calculate an expected profit for Land Use one, the agent forms an Expected Price for that Land Use by estimating a market supply curve, created by the reservation prices (price at which each Land Owner would convert to Land Use one) for each agent, based on current Landscape pattern. The agent estimates an Expected Price for Land Use one by finding the intersection of this hypothetical supply curve with the parametrically set Demand curve.

### 3.12

Figure 2 illustrates the processes of SLUDGE at an aggregate level. Exogenous elements (initialisation parameters and Landscape Conditions set by the user) combine with the demand model to form price expectations for agents. Each agent then selects the highest-valued Land Use, and additional derived endogenous elements are calculated based on these choices. Resulting Landscape patterns then initialise the next time step.

### 3.13

SLUDGE uses a fixed event-sequencing mechanism, which allows every other cell to make a Land Use decision in each time step. This mechanism prevents oscillation, while introducing a minimal amount of additional path-dependency due to stochasticity. Agents are processed sequentially, but updating is synchronous, after each active agent (half the agents in each time step) has made a Land Use decision. When each agent is active, they simply choose the Land Use that has the highest expected profit, choosing Land Use one if profits are identical.



**Figure 2.** SLUDGE process overview (ovals are endogenous/emergent elements; lateral boxes are exogenous/initialisation elements)

## Design Concepts

### *Emergence*

**3.14** SLUDGE was designed to explore the relationship between two related emergent phenomena: landscape pattern and economic value. The landscape patterns and associated landscape productivity measures are emergent in the sense that they are the result of the decentralised decisions of autonomous land-manager agents.

### *Adaptation*

**3.15** The SLUDGE model does not model adaptation, although extensions of the model that include adaptation could be designed. Possible real-world adaptations to reduced productivity from surrounding land uses could include innovation into new land uses, development of mitigation strategies, or development of strategic behaviour related to initial location and/or extent of land use.

### *Fitness*

**3.16** At an individual agent level, fitness is measured by the profitability of the managed land cell, and agents hold a goal of choosing the land use that maximizes their fitness. The individual does not consider its current state when making future decisions, and fitness measures are fixed over the life of the agent. Global 'fitness' is measured through economic welfare statistics (producer surplus, consumer surplus, and total surplus), which depend on the degree of landscape fragmentation and market prices, as described above. One of the key roles of the SLUDGE model is to demonstrate that pursuit of fitness at the individual level will not necessarily maximize fitness at a global level.

### *Prediction*

**3.17** SLUDGE agents predict future outcomes in only one simple way. They form an Expected Price for urban land by anticipating the fitness (profit) maximising decisions of other agents and the price that would result if other agents behaved optimally. In making this calculation, they are assumed to know the locations and profit functions of all other agents. They do not, however, anticipate the effects that these relocation decisions would have on their own fitness/profits more than one time period in advance. In this sense they are more myopic than real-world agents are likely to be. They have no memory or learning mechanisms.

*Interaction*

**3.18** Agents interact indirectly at both a local and global level. At a local level, the land-use choice of one agent affects the profits of its four local neighbours. At a global level, the agent's choice affects the returns to urban land.

*Sensing*

**3.19** As mentioned in 'Prediction', agents are assumed to know the current landscape pattern and the profit-maximizing choices of other agents. Information is complete and certain.

*Stochasticity*

**3.20** SLUDGE deliberately avoids introducing stochasticity in order to minimize path dependence in the model, especially as it was designed to extend other deterministic models. The only potential source of stochasticity is the ability of the end user to generate an initial random landscape, based on a proportion of cells of each type.

*Observation*

**3.21** Observations include graphical display of land-use patterns, metrics measuring the fragmentation and Productivity characteristics of the aggregate landscape, and metrics that reflect economic outputs (outputs from each Land use, prices, transportation costs, and multiple economic fitness measures).

**Initialisation**

**3.22** The state of the model world is initialised by setting an initial configuration of land uses (for the Java model, either all land use 0 or a random distribution of 50% 0 and 50% 1).

**Input**

**3.23** Following initialisation, environmental conditions remain constant over space and time in the SLUDGE model. There is no imposed or modelled spatial or temporal heterogeneity, aside from the endogenous variations in landscape productivity, landscape pattern, and land-use choices that are the focus of the model.

**Submodels**

**3.24** The Landscape, Demand, and Supply submodels calculate information needed to estimate the profitability for each land use. The Landscape submodel calculates supply for each cell for each Land Use, based on the current neighbours of the cell. For Land Use 0, that productivity is multiplied by the fixed Output\_Price\_0 to calculate profits for Land Use 0 for that cell. These profits are given by:

$$\Pi_0 = \frac{\text{Output\_Price\_0} \times \left[ \text{Productivity\_0} + \left( \sum_{k=1}^4 (1 - LU_k) \times \text{Zero\_With\_0} \right) + \left( \sum_{k=1}^4 LU_k \times \text{Zero\_With\_1} \right) \right]}{-\text{Transport\_Cost\_0} \times \sqrt{(i - \text{Market\_0\_X})^2 + (j - \text{Market\_0\_Y})^2}} \quad (1)$$

**3.25** where  $LU = 1$  if cell  $k$  is in Land Use 1 and zero if in Land Use 0,  $k = 1$  to 4 represent the four neighbours of a cell at location  $(i, j)$ , and other parameters are as defined in section 3.1.2. The same logic is followed to calculate profits for Land Use 1; however, an expected price is estimated.

### 3.26

**Expected Price Estimation:** For each cell, the current profits from Land Use 0, the current potential productivity in Land Use 1, and the current transportation costs to Market 1 are recorded. This information is used to solve for the reservation price for converting to Land Use 1 for that cell—the price at which the profits from Land Use 1 are equal to those from Land Use 0. The following equation is used to solve for that reservation price:

$$\Pi_0 = \frac{\text{Reservation\_Price}_y \times \left[ \text{Productivity\_1} + \left( \sum_{k=1}^4 (1 - LU_k) \times \text{One\_With\_0} \right) + \left( \sum_{k=1}^4 LU_k \times \text{One\_With\_1} \right) \right]}{-\text{Transport\_Cost\_1} \times \sqrt{(i - \text{Market\_1\_X})^2 + (j - \text{Market\_1\_Y})^2}} \quad (2)$$

### 3.27

These reservation price/production quantity in Land Use 1 pairs are used to construct a market supply curve for Land Use 1 (a set of potential prices and the total production quantity expected to be supplied to market at those prices). An expected price for Land Use 1 is estimated by finding the price that corresponds to the intersection between the market supply curve and the parametric demand curve given by:

$$\text{Demand\_Price\_1} = \frac{\text{Demand\_Parameter} \times \text{Board\_Size}}{\text{Supply\_1}} \quad (3)$$

### 3.28

Once that single expected price is estimated, it is used by each agent to estimate a location-specific expected profit from Land Use 1:

$$\Pi_1 = \frac{\text{Expected\_Price\_1} \times \left[ \text{Productivity\_1} + \left( \sum_{k=1}^4 (1 - LU_k) \times \text{One\_With\_0} \right) + \left( \sum_{k=1}^4 LU_k \times \text{One\_With\_1} \right) \right]}{-\text{Transport\_Cost\_1} \times \sqrt{(i - \text{Market\_1\_X})^2 + (j - \text{Market\_1\_Y})^2}} \quad (4)$$

### 3.29

In each time step, each Land Owner then choose the most profitable land use, with Land Use 1 chosen if profits are equal.

### 3.30

A variety of outputs (economic and landscape pattern based) are calculated based on the simulated equilibrium landscape productivity, patterns, and economic values.

### 3.31

Additional documentation for the Repast/Java version of SLUDGE is available at <http://mason.gmu.edu/~dparker3/sludge/Sludge/doc/index.html> (created by Robert Najlis). Model code can be downloaded at <http://mason.gmu.edu/~dparker3/sludge/Sludge.zip>. A GUI executable is available at [http://mason.gmu.edu/~dparker3/sludge/sludge\\_gui/sludge\\_gui.zip](http://mason.gmu.edu/~dparker3/sludge/sludge_gui/sludge_gui.zip).

#### Reflections on using ODD (Dawn Parker)

### 3.32

Given that SLUDGE is the simplest of the three models we present, it would seem that its implementation in ODD should also be rather straightforward. However, that was not the case. Confusion arose by the SLUDGE designer (Parker) on many points. In particular, it was difficult to understand definitions of 'state variable' and 'entities,' and which of these are exogenous and endogenous. Consequently, arguments about whether the above description of SLUDGE's state variables and scales fully corresponds to what the creators of ODD had in mind could not be resolved.

### 3.33

I wanted to have a structured framework to communicate as well which exogenous elements are constant for all runs, and which can be modified or incremented by the user. I found it impossible to communicate the basic model mechanisms in the overview without also presenting model parameters. I struggled with a perceived mandate by ODD that model components be taken apart and presented separately, as I was then not sure as to how to coherently explain how they could be put back together. Given that SLUDGE has been implemented in two separate programming environments (Mathematica and Repast/Java), a choice also had to be made for which code base's parameter names should be used. Since an executable was available for the Repast/Java code base, those parameter names were used. Finally, I felt that it was difficult to avoid overlap between the Overview and Details sections of

ODD.

3.34

Some elements were quite straightforward to communicate, such as model scale, event sequencing mechanisms, design principles, and model initialisation. Completing the 'Details' section motivated me to provide more equations and mathematical details than were provided on other publications. Having seen students rely on the structured comparisons of ABM/LUCC models developed in [Parker et al. 2002](#), there is reason to believe that the new information laid out through ODD may help enhance communication of the model for users. However, only a replication effort will reveal whether the new detail is yet sufficient. Finally, I saw no limitation in ODD that stemmed from the fact that it was designed for non-human ecological systems. It seems there is sufficient commonality in models and research questions between the two fields that a shared model description protocol is an option.

**SOME**

**Purpose**

3.35

As part of the project called Spatial Land-use Change and Ecological Effects (SLUCE), a group at the University of Michigan created an initial model to represent urban growth processes, called SLUCE's Original Model for Experimentation (SOME). The purpose of the model was to support an exploration of the relationships between residential preferences, as observed through social surveys ([Fernandez, Brown, Marans and Nassauer 2005](#); [Marans 2003](#)) and represented in residential agents, and urban settlement patterns, as observed through remote sensing and parcel-based mapping. The model has been used to explore (1) the effects of agent heterogeneity on urban settlement patterns ([Brown and Robinson 2006](#)), (2) the degree to which agent preferences in the model can produce patterns that approximate the power-law distribution of settled patch sizes observed in real cities ([Rand et al. 2003](#)), (3) path dependence in the patterns of urban settlement and their implications for the evaluation and validation of spatial patterns ([Brown et al. 2005](#)), (4) the role of zoning in constraining the possible patterns of city formation ([Zellner et al. 2003](#)), and (5) the effects of greenbelts on a spreading urban pattern ([Brown et al. 2004](#)). Subsequent model development has led the group to expand the types of agents represented and the explicit effects of development on land cover. Given that the changes have led to a restructuring of the model, this expanded model has been given a new name (i.e., Dynamic Exurban Ecological Development, or DEED) and will be described in the literature as a different model. The description here focuses exclusively on the SOME model.

**State Variables and Scales**

3.36

SOME comprises a landscape consisting of Cells and two agent types: Residents and Service Centres, both of whose primary behaviours are to locate themselves on a grid-based landscape following a location decision-making process. For use in the model all state variables are scaled into the range [0,1].

---

*Landscape Cells*

---

<b>Variable name</b>	<b>Brief description</b>
<i>Coordinates</i>	X and Y coordinates of the Cell, that determine the Euclidean Distance to Service Centres.
<i>Aesthetic Quality</i>	A relative indicator of the visual attractiveness of a Cell to the Residents.

---



---

*Residents*

---

<b>Name</b>	<b>Brief description</b>
<i>Coordinates</i>	X and Y coordinates of the Cell where the Resident resides.
<i>Alpha Aesthetic Quality*</i>	Relative importance to resident of Aesthetic Quality.
<i>Alpha Distance to Service Centres*</i>	Relative importance to resident of Distance to Service Centres.

<i>Alpha Neighbourhood Density</i> *	Relative importance to Resident of Neighbourhood Density.
<i>Alpha Neighbourhood Similarity</i> *	Relative importance to Resident of Neighbourhood Similarity.
<i>Utility</i>	The overall level of satisfaction a Resident receives from a locational choice.

---

\*All alpha values are constrained to sum to 1.

---

### 3.37

Although the alpha values for each locational attribute are assigned once at the beginning of a run, they are different for each resident. For this reason, whether they should be classified as state variables or parameters is somewhat ambiguous. We list them as variables because the process description (below) relies fairly heavily on understanding how the characteristics of the Residents affects their choice of a Cell for location.

---

#### *Service Centres*

---

<b>Variable name</b>	<b>Brief description</b>
----------------------	--------------------------

<i>Coordinates</i>	X and Y coordinates of the Cell where the Service Centre is located.
--------------------	--

---

#### Process Overview and Scheduling

### 3.38

Several locational attributes are evaluated by the residents, including the state variable of a Cell, Aesthetic Quality, and three derived variables: Distance to Service Centres, resident Density, and Neighbourhood Similarity. All of these variables are updated and rescaled each time step or as needed when the underlying features change (described below). Each time step, processes are scheduled in the following way:

```

For 1 to the defined number of Residents to enter at each time step (specified by the user or file)
  Create a new Resident.
  For 1 to the number of Cells to test
    Randomly select an unoccupied Cell (without replacement)
    Calculate Neighbourhood Similarity, based on the difference between the Resident's
      importance values for each locational attribute ( $\alpha$ 's) and those of any
      residents already located in the eight neighbouring cells.
    Evaluate Utility at that Cell to that Resident: based on the combination of Cell
      attributes, weighted by the importance of those attributes to the Resident
      (represented by alpha).
    If it is the first Cell then
      Store the Cell and Utility as the best location.
    Else if it is not the first Cell evaluated by the Resident then
      If the current Cell's Utility > best Cell's Utility then
        Set the best Cell and Utility to those of the current Cell
      End if
    End if
  Next Test Cell
  Put Resident in the best Cell.
  Set Resident X,Y and Utility to those from the new Cell
  Calculate Neighbourhood Density for all Cells, as the number of occupied cells within the
    nine-cell neighbourhood around each Cell.
  If option is selected, update Aesthetic Quality near new Resident

  If the total number of Residents in the world divided by the specified number of Residents per
    Service Centre minus the number of existing Service Centres is >= 1 then
    Create and locate a new Service Centre (process described in details)
    Set Service Centre X,Y properties to those from the new Cell
    Calculate Distance to Service Centres for all Cells, based on Euclidean distance from
      each Cell to the nearest Service Centre
    If option is selected, update Aesthetic Quality near new Service Centre
  End IF
Next Resident

```

#### Design Concepts

##### *Emergence*

**3.39** SOME was designed to explore the processes that give rise to the spatial patterns of land settlement at the growing edge of a city, described in terms of their global and local properties, e.g., amount of development occurring at great distance from the city and clustering of developed patches. We have also evaluated distributions of utility achieved by residential agents. These patterns arise from the collective decisions of residential agents about where they wish to locate based on their preferences. While agents can have a preference for locating near service centres, and an initial service centre is seeded in the centre of the map in most applications, the patterns of service centres develop dynamically and, therefore, strong preference for nearness to service centres does not necessarily result in less sprawling development.

*Adaptation*

**3.40** The SOME model includes adaptation only in the settlement process by choosing the best location for settlement in terms of Utility. Extensions of the model that include more adaptation could be designed. For example, Residents could adjust their preferences based on experience (i.e., where they settle initially) or the preferences of their neighbours (through a social learning process). Such adjustments could influence their calculations of Utility or, if relocation were permitted in the model, subsequent location decisions.

*Fitness*

**3.41** Each Resident is assigned preferences when they are created and uses them to weigh alternative locations. In the current version of the model these preferences do not change during the course of the model run. The assignment is based on a random draw from a pre-defined distribution, intended to represent the distribution of preferences within the population.

*Prediction*

**3.42** In SOME, the Residents calculate Utility based on the current state of the landscape. The predictions the Residents make at the time they make their location decision, therefore, assume no change in the future. This is a form of tacit (and very naïve) prediction.

*Interaction*

**3.43** Agents interact directly and indirectly. Indirect interactions involve the land-use choices of agents affecting the landscape characteristics that subsequent agents evaluate and the calculated Utility of their neighbours. Residents interact directly when an incoming agent factors the preference characteristics of neighbouring Residents into its determination of the Utility it will get from a particular Cell. We assume that actual residents get clues about the preferences of their potential neighbours, when seeking a place to live, by looking at the aesthetic characteristics of their home and landscape.

*Sensing*

**3.44** Agents are assumed to know perfectly the current landscape Aesthetic Quality, Distance to Service Centres, Neighbourhood Similarity, and Neighbourhood Density of all Cells they sample. However, their ability to sample Cells is restricted by a parameter that controls the incompleteness of information available about the housing market. They maximize the Utility from among the choices available to them.

*Stochasticity*

**3.45** SOME uses stochasticity to represent (a) the preferences of the Residents (defined by predefined distributions), (b) patterns of Aesthetic Quality (in some applications), and (c) the location of the Service Centres after the first one (guided by the location of a Resident).

*Collectives*

**3.46** One type of collective exists within SOME. The collective, which affects only agent characteristics, involves, in some experiments, the identification of categories of Residents that have their separate distributions from which preferences are drawn.

*Observation*

3.47

Observations include graphical display of land-use patterns, metrics measuring the sprawl and fragmentation of the settlement patterns, and metrics that describe distribution of achieved Utility levels across the Resident population.

#### Initialisation

3.48

The environment of the model is initialised by creating the geographic grid on which the model runs, based on parameters describing its x and y dimensions, establishing the initial pattern of aesthetic quality, locating an initial Service Centre (usually in the centre of the grid, but can also be read from GIS files), and calculating distance of every Cell to the initial Service Centre. No further initialisation is required for the Service Centres. The pattern of aesthetic quality can be initialised using a random assignment of values based on a specified distribution. Spatial autocorrelation can be introduced through smoothing. Additionally, an option exists to read the an aesthetic quality map from a file output from a GIS.

3.49

For the Residents, initialisation involves specification of the distribution of alpha and beta state variables. One of our key interests was to evaluate the effects of alternative specifications of agent diversity on the resulting settlement patterns. In the study of Brown and Robinson (2006), we specified the alpha values as uniform (i.e. no variation among residents), drawn from a normal distribution with a specified mean and standard deviation, uniform within each of seven different clusters of agents (but different between the clusters), and drawn from a normal distribution with mean and standard deviation specified separately for each of the seven clusters. In that study, we used survey data on locational preferences of residents within Southeastern Michigan to characterize the distributional parameters that were used to initialise the agents. In addition to alpha values, the residents use a beta value for calculation of utility, which specifies their ideal value on each location attribute they are evaluating. In each case, the alpha values were normalized so that they summed to one for each Resident. The beta values are typically specified once and uniformly for all agents, but the capability exists to allow these to vary as well.

3.50

Additional parameters controlling the dynamics of the model are specified at the beginning of a run, describing the rate at which Residents enter (i.e., how many per time step), the form of the utility function they use, how many Cells each Resident samples before evaluating utility and choosing one, the rate at which Service Centres are created, and whether or not Residents and Service Centres alter the aesthetic quality of a location.

#### Input

3.51

In most applications of the model, further inputs are not required once the model is initialised. However, for one application of the model (Brown et al. 2005), we sought to compare the patterns of development from the model with those observed in Washtenaw County, Michigan. Using data from the US Census, we identified periods with different rates of Resident households entering the model and different rates of Service Centre creation. The model was run with these rates as input to establish the relevant parameter values for the different time periods.

#### Submodels

3.52

The key processes in the model are associated with the location choices of Residents and Service Centres. The Residents calculate their utility for each of a small number of sampled Cells, and select the Cell with the highest utility, using either an additive or multiplicative combination of locational attribute values. In the additive case, the utility function takes the form:

$$u_{r(x,y)} = \sum_{i=1}^m \alpha_{ir} \left( 1 - \left| \beta_i - \gamma_{i(x,y)} \right| \right) \quad (5)$$

where  $u_{r(x,y)}$  is the utility of location  $(x,y)$  for resident  $r$ ;  $\alpha_{ir}$  is the weight the resident  $r$  places on factor  $i$ ;  $\beta_i$  is the preferred value on component  $i$  and is usually assumed constant for all residents, e.g., all residents desire the most aesthetic quality and shortest distance to service centers;  $\gamma_{i(x,y)}$  is the value of component  $i$  at location  $(x,y)$ , and  $m$  is the number of location attributes evaluated. The form of the multiplicative function is:

$$u_{r(x,y)} = \prod_{i=1}^m \left( 1 - \left| \beta_i - \gamma_{i(x,y)} \right| \right)^{\alpha_{ir}} \quad (6)$$

Choice of whether to use the additive or multiplicative utility function is driven by theoretical and experimental considerations ([Brown and Robinson 2006](#)).

### 3.53

Service Centres are located in a much simpler manner, to reflect a general dependence on where residents are locating and to introduce a degree of stochasticity into their locations. The stochasticity is introduced through the decision to locate near the last resident, which means the service centre will locate near a resident with, in most cases, a stochastically determined set of locational preferences and a stochastically determined set of locations to choose from. When a new Service Centre is created it is located by the following process:

```

Select a random adjacent Cell next to the last Resident
Do until a Cell is selected for the Service Centre
    To get a new Cell spiral outwards from the last Resident Location, while
        checking for edge effects
    If the Cell is not occupied then
        Select the Cell.
    End if
End Do

```

#### Reflections on using ODD (Dan Brown)

### 3.54

The process of developing an ODD-based description of the SOME model was useful to its developers. In fact, many, if not most, of the details of the model as prescribed by ODD have been published along with SOME-based analyses and experiments. The key difference is that ODD provides a structure that allows for commonality and comparison with other models, like SLUDGE and FEARLUS. So, the application of ODD to a land-use-oriented social simulation like SOME offers significant promise.

### 3.55

Probably the biggest sources of confusion in the process, like that involving SLUDGE and FEARLUS, was in the definitions of state variables and parameters. Clearly, as stated previously, these need to be described before the process overview can be understandable to the reader. It was not always clear that describing state variables was sufficient to provide the needed level of clarity. As with FEARLUS, the aspects of the SOME model aimed at Typification necessitate experimentation with key parameter values. So, some parameters should be defined before the process overview, although ODD does not call for parameters to be defined in the first section. Furthermore, there was some disagreement about whether the alpha values of residents are state variables or parameters, the general rule for state variables being that they are dynamic (which the alpha values are not) and that they vary from agent to agent (which the alpha values do). In this case, it seems that decision about whether to define something as a state variable, especially when it is ambiguous, might best rest on whether it needs to be defined before the process overview can be clearly described.

#### FEARLUS+ELMM

##### Purpose

### 3.56

FEARLUS was conceived out of the recognition that traditional spatial modelling techniques assuming (fiscally) economically rational behaviour could not always adequately account for observed spatial patterns. Land use change is a complex process requiring integrated social, economic and biophysical models.

### 3.57

The FEARLUS software was developed with a deliberately modular architecture, to facilitate a broad range of land use change-related studies. FEARLUS is thus regarded by its developers as a modelling framework that can be configured to produce a particular model of interest, rather than a specific model in itself. Earlier work with FEARLUS involved comparing the relative success of innovating and imitating farmer land use decision-making strategies in environments with different rates of spatial and temporal change in factors influencing the economic returns from selected land uses. More recently, FEARLUS has been coupled with a water model ([Koo, Dunn and Ferrier 2004](#)) to explore the potential influence of catchment-scale programmes of measures and social approval on diffuse pollution from land use ([Davies et al. 2006](#)).

### 3.58

For the purposes of this paper, we focus on describing FEARLUS for re-evaluating the relative success of innovating and imitating decision strategies for farmers when using a more realistic land market model (Polhill, Parker and Gotts 2005; Polhill, Parker, and Gotts 2007). The specific question addressed by this work is whether more realistic land markets than those used earlier give a greater advantage to innovators than was suggested in the earlier work. The model will be referred to as FEARLUS+ELMM to distinguish it from FEARLUS as a framework. We focus on describing FEARLUS+ELMM using ODD as though it is to be used for some of the experiments done in Polhill, Parker and Gotts (2007). (Specifically, the 'type 1' and 'type 4' experiments with ELMM.)

#### State Variables and Scales

3.59

The FEARLUS+ELMM model consists of the following core entities, which are described in a series of tables. One time step in the model is a single Year. Simulations typically run for 300 years.

---

##### *Land Parcel*: Field scale

---

<b>Variable name</b>	<b>Brief description</b>
<i>Environment</i>	The Environment this Land Parcel appears in (see below).
<i>Biophysical Characteristics</i>	An abstract representation of the local properties of the Land Parcel, which could conceptually be understood as such things as soil type, aspect, altitude or slope, though these are not specifically represented.
<i>Land Use</i>	An abstract representation of the land use this field has been applied to.
<i>Yield</i>	The Yield of the most recently applied Land Use.
<i>Owner</i>	The Land Manager responsible for making decisions for the Parcel, and harvesting the Yield from it.

---



---

##### *Environment*: Catchment or regional scale

---

<b>Variable name</b>	<b>Brief description</b>
<i>Spatial Topology</i>	Determines the neighbourhood of each Land Parcel (toroidal/bounded; von Neumann/Moore).
<i>External Conditions</i>	An abstract representation of spatially homogeneous conditions that change over time and affect Economic Returns to Land Managers. These could be thought of as including such things as exchange rates, regional climate, and demand, though such things are not specifically modelled.
<i>External Conditions Flip Probability Array</i>	Probabilities determining the rate of change of the External Conditions from one Year to the next.
<i>Biophysical Characteristics Clumped?</i>	Whether or not to make the Biophysical Characteristics of neighbouring Land Parcels more similar to each other using a clumping algorithm.
<i>Break-Even Threshold</i>	The amount of Yield a Land Manager has to make from a Land Parcel to break even.
<i>Land Uses</i>	The set of Land Uses that Land Managers may apply to Land Parcels.

---



---

##### *Land Manager*: Farm scale

---

<b>Variable name</b>	<b>Brief description</b>
<i>Land Parcel List</i>	List of Land Parcels owned by the Land Manager.

<i>Subpopulation</i>	The Subpopulation this Land Manager belongs to, which determines the settings of its parameters.
<i>Land Market</i>	A pointer to the Land Market to which bids for Land Parcels should be sent.
<i>Account</i>	The accumulated wealth of the Land Manager.
<i>Aspiration Threshold</i>	The amount of Yield the Land Manager hopes to obtain from each Land Parcel.
<i>Imitative Strategy</i>	The algorithm to use for choosing a Land Use by imitating neighbouring Land Uses if the Aspiration Threshold is not achieved.
<i>Memory size</i>	How many Years in the past the Land Manager may look for data in the Imitative Strategy algorithm.
<i>Innovative Strategy</i>	The algorithm to use for choosing a Land Use without imitating neighbouring Land Uses if the Aspiration Threshold is not achieved.
<i>Imitation Probability</i>	The probability of choosing the Imitative Strategy if the Aspiration Threshold is not achieved.
<i>Land Offer Threshold</i>	The amount that must be in the Account before the Land Manager will bid for Land Parcels.
<i>Bidding Strategy</i>	The algorithm to use to generate a Price to offer for Land Parcels available for sale.
<i>Selection Strategy</i>	The algorithm to use to decide which bids for Land Parcels to actually make.

*Subpopulation:* Land Manager collective

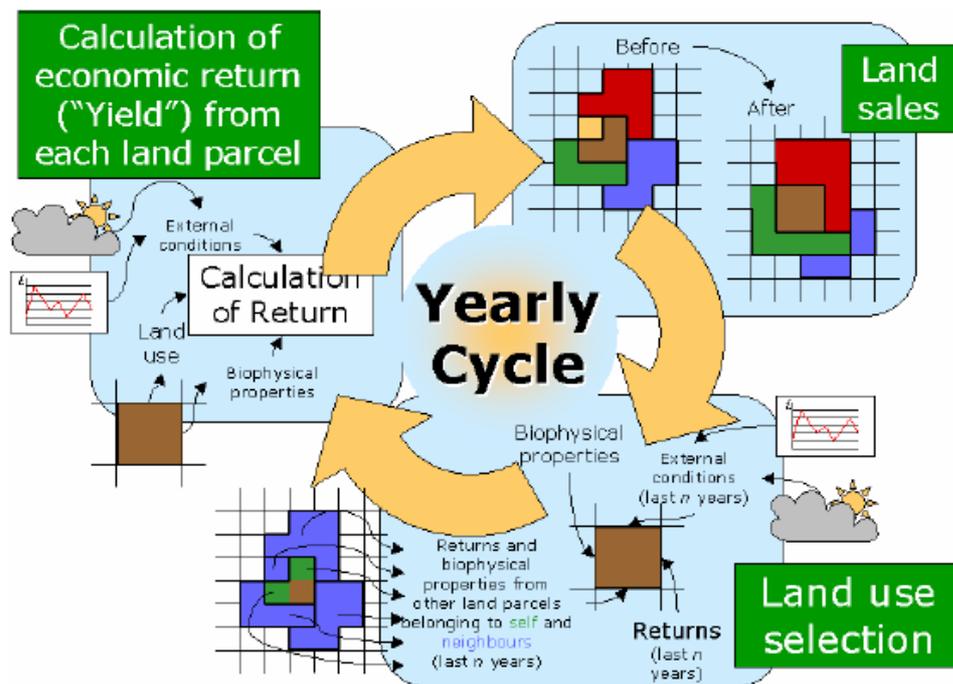
<b>Variable name</b>	<b>Brief description</b>
<i>Land Manager List</i>	List of Land Managers belonging to this Subpopulation.
<i>Incomer Offer Price Distribution</i>	Determines the distribution from which new Land Managers of this Subpopulation will create an offer price for Land Parcels — e.g. Normal(Mean, Variance).
<i>Imitative Probability Distribution</i>	Distribution from which the Imitative Probabilities of new Land Managers of this Subpopulation are taken.
<i>Aspiration Threshold Distribution</i>	Distribution from which the Aspiration Thresholds of new Land Managers of this Subpopulation are taken.
<i>Land Offer Threshold Distribution</i>	Distribution from which the Land Offer Thresholds of new Land Managers of this Subpopulation are taken.
<i>Bidding Strategy Configuration</i>	Configuration string for the Bidding Strategy of new Land Managers of this Subpopulation. This depends on the particular Bidding Strategy algorithm used. For example, for a Wealth Multiple Bidding Strategy, this contains the distribution of the fraction of the Account that member Land Managers will use to generate a bid from.

*Land Market:* Responsible for organising the exchange of Land Parcels among Land Managers.

<b>Variable name</b>	<b>Brief description</b>
<i>Auction Type</i>	First price sealed bid or Vickrey auction.
<i>Parcels for Sale List</i>	A list of the Land Parcels available for sale in the current Year.
<i>Bid Lst</i>	A list of bids made by Land Managers for members of the Land Parcels for Sale List.

### 3.60

A high-level diagram showing an overview of the yearly schedule in FEARLUS is given in Figure 3. More detail on each section in the diagram is provided afterwards. Where the schedule involves calls to Land Managers, the order in which the Managers are called is arbitrary but non-random (the model is so written [1] that Land Managers' decisions do not interfere with each other).



**Figure 3.** The Yearly cycle in FEARLUS. Starting at the bottom right, Land Managers choose Land Uses for their Land Parcels, the Economic Return is then calculated (top left), then Land Parcel exchanged (top right)

#### *Land Use Selection*

### 3.61

Land Managers do the following for each Land Parcel they own:

- If the Yield of the Land Parcel meets or exceeds the Aspiration Threshold of the Land Manager, then the Land Use is unchanged.
- If not, then if a random number in the range [0, 1] is less than the Imitative Probability, the Land Manager uses their Imitative Strategy to determine the new Land Use of the Land Parcel.
- Otherwise, the Land Manager uses their Innovative Strategy to determine a new Land Use of the Land Parcel

Examples of Imitative Strategies are:

- Choose a Land Use weighted by the number of times it occurs in the Land Manager's neighbourhood.
- Choose the Land Use having the highest mean Yield in the neighbourhood.

The Innovative Strategy is usually to choose uniformly randomly among the full set of Land Uses contained in the array in the Environment object.

#### *Calculation of Return*

### 3.62

The External Conditions is determined for this year. The Yield for each Land Parcel is then calculated, based how good a match the chosen Land Use is with the local Biophysical Characteristics and the global External Conditions. Land Managers then increment their Account for each Land Parcel they own by the Yield minus the Break Even Threshold.

#### *Land Sales*

### 3.63

The Land Sales mechanism is discussed in detail in Polhill et al. (2005). Essentially, Land Managers must put all their Land Parcels up for sale if their Account is below zero at the end of the Year. These Land Parcels are then bought by their neighbours, or by an in-migrating Land Manager. Neighbouring Land Managers use a Bidding Strategy to decide a Bid for each

Land Parcel for sale. Examples would be to bid a certain proportion of the Account, to bid a proportion of the estimated Profit from the Land Parcel, or to bid a fixed price. A Selection Strategy is then used to decide which of the Bids they will actually make. An example would be to bid for the highest-priced Land Parcels first, as these are the Parcels of most value to the bidding Land Manager. A Subpopulation will also be selected at random to generate an in-migrant Bid for a Land Parcel for sale.

### 3.64

When the Land Allocator has received all the Bids, an auction is used to determine which of the Bids wins. The auction can be a first-price sealed-bid auction (ignoring the complication that it is not rational to bid your valuation in such an auction), or a Vickrey auction, in which the highest bidder wins, but the price paid is that of the second highest bid.

#### Design Concepts

##### *Emergence*

### 3.65

The emergent effect for the purposes described in this paper is the performance of Land Managers employing different Land Use Selection Algorithms, i.e. pure imitation or including both imitation and innovating components (Imitation Probability = 1 or < 1, respectively). Currently, performance of Subpopulations is measured by comparing the numbers of Land Parcels owned by member Land Managers. Performance emerges from the interaction of the Land Managers with each other and the landscape. It can be influenced, but not entirely imposed, by the "External Conditions Flip Probability Array" and "Biophysical Characteristics Clumped?" parameters.

##### *Adaptation*

### 3.66

Land Managers in FEARLUS adapt by choosing new Land Uses to respond to the changing External Conditions. Adaptive strategies are not directly aimed at making more money or acquiring more Land (the measures of fitness), but are instead heuristic. Thus, Random Experimentation could be seen as an effort to try anything new, whilst imitation could be seen as hoping that neighbours are making good choices of Land Use. Two forms of imitation have been tried in FEARLUS+ELMM: Simple Imitation, which chooses the modal Land Use in the neighbourhood, and Best Mean Imitation, which chooses the Land Use with the highest average Yield in the neighbourhood.

### 3.67

Land Managers also adapt by buying Land Parcels. More Land means that one poorly performing Land Parcel can be compensated for by another that performs well, and allows Land Managers to spread risk by diversifying the Land Uses employed.

##### *Fitness*

### 3.68

Fitness of individual Land Managers is determined by their Account. If the Account drops below zero, Land Managers are bankrupt, and no longer participate in the simulation. The fitness of a Subpopulation of Land Managers is determined by the number of Land Parcels they collectively own. The Account and the number of Land Parcels owned are related by the Land Market. Land Managers with more money are able to buy more Land.

##### *Prediction*

### 3.69

Prediction is implicit in heuristics used by some types of Land Managers. Simple Imitators assume that the most popular Land Use must be the best performing, whilst Best Mean Imitators use the average Yield as a proxy for estimated profit. Imitators are also implicitly assuming that successful Land Uses in the previous Memory Size Years are likely to be successful in the coming Year.

### 3.70

When the Yield from a Land Parcel meets the Land Manager's Aspiration Threshold, the Land Use is retained on the Parcel from one Year into the next. Both innovators and imitators may do this, and it essentially entails the same assumption as above, but for the previous Year only.

### 3.71

Another implicit prediction is related to investment in a new Land Parcel. When Land is exchanged using a full Land Market, Land Managers use discounting to create a bid based on an estimated profit from the Parcel for sale that is a function of the Land Manager's risk perception (a parameter of the Discounting Bidding Strategy), the profit the Manager has

made on other Parcels and the average Yield over the past  $n$  Years of the Parcel for sale (where  $n$  is a parameter in the Bidding Strategy Configuration).

#### *Interaction*

**3.72** Imitating Land Managers interact with their neighbours to find out about the performance of Land Uses in their area, i.e. from all Land Parcels owned by Land Managers owning a neighbouring Parcel. This set of Parcels is the 'social neighbourhood' of a Land Manager, and though the Parcels are queried directly, it is intended to represent Land Managers exchanging information.

**3.73** All Land Managers also have mediated interaction in the Land Market, where they make competing bids for Land Parcels.

#### *Sensing*

**3.74** Land Managers are assumed to have access to Yield, Land Use, External Conditions and Biophysical Characteristics data for all Land Parcels they make decisions about (including whether or not to buy them), and in the case of imitating Land Managers, for all neighbouring Land Parcels. Land Managers also know the state of their Account. Whether this information is used depends on the strategy employed. Simple Imitators just use Land Use information, while Best Mean Imitators use Yield information as well. In the Land Market, Land Managers use Yield and Account information when computing bids for Parcels using discounting.

#### *Stochasticity*

**3.75** Stochasticity is used to simulate spatial variability in Land Parcels' Biophysical Characteristics, and temporal variability in External Conditions. Land Uses are also initially randomly assembled. Clearly, stochasticity is involved in Random Experimentation for innovating Land Managers, and imitating Land Managers use stochasticity to select among Land Uses with equal maximum ordering, or according to the weight they are given, depending on the method. In the Land Market, stochasticity is used to select a new owner for a Land Parcel with equal maximum bids from different Land Managers. If a simulation involves more than one Subpopulation, then the Subpopulation from which an in-migrant bid is generated for a Land Parcel for sale is selected at random.

#### *Collectives*

**3.76** Subpopulations are collectives of Land Managers, represented explicitly with state variables used to create new Land Managers.

#### *Observation*

**3.77** Various observations are available in FEARLUS+ELMM, from an omniscient perspective. However, the key observation as far as the work described herein is concerned is the number of Land Parcels owned by members of each Subpopulation at termination of the simulation.

#### **Initialisation**

**3.78** In FEARLUS+ELMM, initialisation is a stochastic process, driven by a specified seed for a pseudo-random number generator. Initial values are thus chosen arbitrarily, with several runs used to determine the typical behaviour of the model using a statistical test.

**3.79** At the start of the simulation, the Environment and Land Market objects are created. The Environment then creates the Land Parcels, using the selected Spatial Topology (Toroidal-Moore) to determine their neighbours, and assigning them random initial Biophysical Characteristics. If the Biophysical Characteristics are to be spatially correlated, a clumping algorithm is performed. Next, the Land Uses are created. The initial External Conditions are then determined. After this, the initial population of Managers is created: For each Land Parcel, a random Subpopulation is selected (with equal probability), and that Subpopulation builds a member Land Manager to own the Parcel using random samples from the various distribution parameters to set the initial parameters of the Land Manager. The initial Account of each Land Manager is zero. The Land Managers then make a random initial selection of Land Use. This is then used to calculate an initial Yield for each Land Parcel.

**3.80**

Certain kinds of experiment (in particular the 'type 4' experiments) use paired runs, which endeavour to replicate as much as possible of the initial conditions in two similar runs with a control parameter changed. (For example, to address the question of whether innovating Land Managers are more successful against imitators if everyone uses a Discounting rather than a Fixed Price Bidding Strategy when buying Land Parcels. Here, one run of each pair would involve all Land managers using a Discounting strategy, and the other a Fixed Price Bidding Strategy, with everything else kept as similar as possible.) In such cases, the first of the pair saves the initial conditions to a file, which are then loaded by the second of the pair, rather than generating them randomly. The information saved to the file consists of the Biophysical Characteristics of each Land Parcel, the Land Uses, and the External Conditions time series.

**Input**

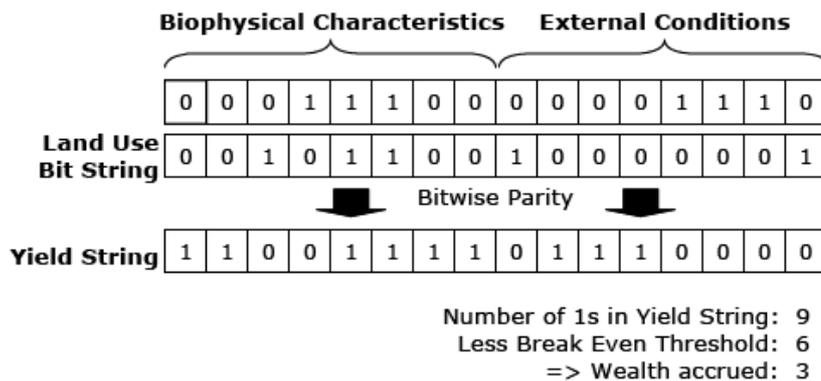
**3.81** The time series of the External Conditions is the only exogenous input to FEARLUS+ELMM. This may be computed stochastically using the Flip Probability Array, or for paired runs (see above), the time series can be loaded from a file.

**Submodels**

**3.82** The three main processes in FEARLUS+ELMM are Land Use Selection, Calculation of Return and Land Sales. To understand how they work it is also necessary to understand the abstract land use choice outcome submodel.

*Land Use Choice Outcome Submodel*

**3.83** This submodel in FEARLUS+ELMM is used to derive the outcome of the Land Use choices for each Land Manager. For the purposes of the investigations conducted, all that matters is that there are a finite number of different Land Use choices, the outcome of which may vary in space and in time. For this reason, a bitstring-based representation is used, in which the Biophysical Characteristics, External Conditions and Land Uses are represented using these data structures. The length of the Land Use bitstring is equal to the sum of the lengths of the Biophysical Characteristics and External Conditions bitstrings; these two lengths being parameters of the model. Conceptually, a Land Use is a better match for the time and place in which it is applied the more bits in its bitstring have the same value as the corresponding bits in the Biophysical Characteristics and External Conditions bitstrings (Figure 4). The number of Land Uses available is a parameter, and is customarily less than  $2^{BU}$ , where  $BU$  is the number of bits in the Land Use bitstring, meaning that for most Land Use decisions, it will not be possible to choose a perfectly matching Land Use for the expected situation in the forthcoming Year.



**Figure 4.** Illustration of bitstring-based model of Land Use choice outcome.

**3.84** During initialisation, the Land Parcels are each assigned random Biophysical Characteristics bitstrings with an equal probability of 1 or 0 for each bit. If the Biophysical Characteristics are to be clumped, then a clumping algorithm is used that, for each bit in turn, swaps bits between Land Parcels if doing so will increase the similarity to their neighbours' Biophysical Characteristics until no such pair of Land Parcels can be found. The Land Uses' and initial External Conditions' bitstrings are also created randomly with an equal probability of 1 or 0 for each bit.

*Land Use Selection*

**3.85**

The following is repeated by each Land Manager for each Land Parcel they own:

- a. Determination of strategy to use:
  - If the Yield of the Land Parcel meets or exceeds the Aspiration Threshold of the Land Manager, then the Land Use is unchanged.
  - If not, then if a random number in the range [0, 1] is less than the Imitative Probability, the Land Manager uses their Imitative Strategy to determine the new Land Use of the Land Parcel.
  - Otherwise, the Land Manager uses their Innovative Strategy to determine a new Land Use of the Land Parcel
- b. Application of strategy
  - If the Land Use is to be unchanged, then new Land Use = previous Land Use.
  - If the Innovative Strategy is to be used, then apply the Innovative Strategy of the Land Manager. Typically a Random Innovative Strategy is used: choose a new Land Use uniformly randomly among the full set of Land Uses. Note that this means there is a  $1/n$  chance that the Land Use will not change, where  $n$  is the number of Land Uses.
  - If the Imitative Strategy is to be used, then apply the Imitative Strategy of the Land Manager:
    - If the Imitative Strategy is Simple Imitation, then for each Land Use, initialise a score to zero. Next, loop through each Land Parcel owned by the Land Manager or by the owner of any neighbouring Parcel, and add 1 to the score for the Land Use applied on that Parcel. Finally, select the Land Use at random in proportion to its score by looping through the Land Uses in order and setting the weighted score of the Land Use to the sum of its score and that of the previous Land Use in the order, then choosing a uniform random number in the range 0–total score and selecting the first Land Use with weighted score greater than the random number.
    - If the Imitative Strategy is Best Mean Imitation, then for each Land Use, initialise a score and a counter to zero. Next, loop through each Land Parcel owned by the Land Manager or by the owner of any neighbouring Parcel, and add the Yield of the Parcel to the score for its Land Use, and increment the counter of the Land Use. Finally, select the Land Use to use by choosing uniformly randomly among those Land Uses with counter > 0 having equal maximum average Yield (score divided by counter).

Note that in the Imitative Strategy the neighbourhood is a 'social neighbourhood': All Land Parcels belonging to a neighbour are examined even if these Land Parcels do not physically neighbour those of the Land Manager making the decision.

#### *Calculation of Return*

- a. Determination of External Conditions
  - If the External Conditions bitstring is not being loaded from a file, or if the end of the file has been reached, then for each bit in the External Conditions bitstring, if a uniform random number between 0 and 1 is less than the Flip Probability of that bit, then set the bit to its complement, else leave the bit unchanged.
  - Else load the next External Conditions bitstring from the file.
- b. Computation of Yield
  - The Yield for each Land Parcel is the number of bits having the same value in the bitstring of the Land Use applied and a bitstring comprised of a concatenation of the Biophysical Characteristics and External Conditions bitstrings.
- c. Computation of Economic Return and update of Land Managers' Accounts
  - Each Land Manager, for each Land Parcel they own, increments their Account by the Yield of the Land Parcel minus the Break Even Threshold.

#### *Land Sales*

- a. Determination of bankruptcy
  - A Land Manager is bankrupt if their Account is less than zero.
- b. Notification of Land Parcels for sale
  - All bankrupt Land Managers notify the Land Market that all their Parcels are available for sale.
  - The Land Market notifies all non-bankrupt Land Managers owning a Parcel neighbouring those for sale of the availability of these Parcels.
- c. Creation and selection of bids for Land Parcels
  - For each Land Parcel for sale, select a Subpopulation at random with equal probability to generate a Bid for the Parcel using a random sample from the Incomer Offer Price Distribution.
  - For each Land Manager notified of a Parcel for sale, use the Bidding Strategy to generate a potential Bid for the Parcel:
    - If the Bidding Strategy is a Fixed Price, then use the Fixed Price assigned from the Subpopulation's Bidding Strategy Configuration.

- If the Bidding Strategy is a Discounting Bidding strategy, then the Price is equal to the Estimated Profit divided by the Rate Multiple assigned from the Subpopulation's Bidding Strategy Configuration. The Estimated Profit is taken from the following formula (where  $EP$  is the Estimated Profit,  $w$  is a weighting parameter assigned from the Bidding Strategy Configuration,  $M$  is a 'memory' parameter also assigned from the Bidding Strategy Configuration,  $A_{mi}$  is the Account of Manager  $m$  in Year  $i$ ,  $y$  is the current Year, and  $Y_p$  is the most recent Yield of Parcel  $p$ ):

$$EP = \frac{w}{M} \sum_{i=y-M+1, y} A_{mi} + (1-w)Y_p \quad (7)$$

- For each Land Manager notified of a Parcel for sale, use the Selection Strategy to decide which of the potential Bids will actually be made. Typically a 'Buy Dearest' Selection Strategy is used:
  - Set a Total Bid to zero.
  - Sort the potential Bids in descending order of Price into a Bid List.
  - Increment the Total Bid by the Price of the first item on the Bid List.
  - While the Land Manager's Account minus the Total Bid is greater than the Manager's Land Offer Threshold, remove the first item on the Bid List, and notify the Land Market that this Bid is made, then increment the Total Bid by the Price of the new first item on the Bid List and repeat.
- d. Determination of new owner for each Parcel for sale
  - For each Land Parcel for sale, sort the bids made for the Parcel in descending order of Price into a list.
  - The first item on the list is the winning Bid.
  - If the auction is a Vickrey auction, and there are two or more Bids made for the Parcel, then the winning Price is the Price of the second Bid on the list. Else, the winning Price is the Price of the winning Bid.
  - Subtract the winning Price from the Account of the winning Land Manager. (If the winning Land Manager is an incomer, the result of this operation is to set their Account to zero.)
  - Assign the Land Parcel to the winning Land Manager.

### 3.86

Further details can be found in the [FEARLUS 0-6-8+ELMM0-2 user guide](#).

Reflections on using ODD (Gary Polhill)

### 3.87

As a model belonging to the Typification class of Boero and Squazzoni (2005), FEARLUS+ELMM is not straightforward to describe using ODD. Models belonging to the Typification class may have large numbers of optional parameter settings and algorithmic components that are difficult to describe concisely using ODD as it stands. ODD is much more focused on one specific model and associated parameter settings, but explorations of questions using models such as FEARLUS+ELMM may involve several algorithmic variants within a theme under investigation each studied with a subset of parameter space, and we are not so accustomed to thinking specifically about just one model and parameter setting. The work in Polhill, Parker and Gotts (2008) is a case in point, as it explores the use of several different land exchange protocols. Nonetheless, describing the model using ODD was a useful exercise, as it forced us to think more carefully about how to communicate it to readers, through providing a skeleton of the information that should be provided. With some minor adjustments, ODD could be used to clearly outline a baseline and alternatives used in explorations with a model of this kind.



## Discussion

### 4.1

We applied the ODD protocol for describing individual- or agent-based models to three agent-based models of land-use change. The purpose of this exercise was to test whether ODD, which was formulated by ecologists who usually do not include human agents in their models, is also useful for agent-based models of the social sciences. A final answer to this question is not possible here, because it requires adoption by a broader community and will be evidenced in such things as:

- Use of ODD to describe models in future publications by the authors of this article. This will demonstrate that we at least have seen merit in using ODD enough to apply it in earnest ourselves (examples in ecology include [Piou et al. 2007](#) and [Wang and Grimm 2007](#)).
- A desire to see ODD being used in other publications that is sufficient to provoke

comments in peer reviews when it is absent (which is already the case in the ecological literature).

- A successful replication of a social simulation model described using ODD without other communication with the authors.

#### 4.2

Be that as it may, this exercise has demonstrated that agent-based models can be described using ODD, as shown by the three examples. Whilst the effectiveness of this description can really only be judged impartially by the reader, it does at least attest to the feasibility of applying ODD to agent-based models: for these examples, there were no issues that arose suggesting a critical deficiency with ODD.

#### 4.3

This is not to say that the process was entirely smooth. The three example models were first independently described according to ODD by those who were authors or co-authors of these models. The resulting descriptions were quite different and inconsistent, in particular the Overview parts. This problem has also been reported by Grimm et al. (2006) for the 19 model descriptions they used to test ODD for ecological models. We needed to revise our first descriptions and had to rely on the feedback and input of one of the developers of ODD (Grimm). Thus, ODD does not necessarily lead to consistent model descriptions per se, but requires active interactions with other users of ODD, or imitation of existing uses of ODD, which do not yet exist in social sciences. This current limitation of ODD was anticipated by Grimm et al. (2006, p. 124):

Still, ..., differences in the style of the presentation are likely to remain. We have to accept this at the current stage, because the protocol has to compromise between being general enough to include all kinds of individual- or agent-based models and being specific enough to fulfil its purpose.

#### 4.4

We also found differences in understanding as to what should be included in the Overview category. In ecology, the concept of 'low-level state variables' was rather easily understood and applied, but in social sciences this seems to be more difficult. There were also issues over the definitions of other terms used in ODD: What, exactly, should be considered as an entity? Can we always clearly distinguish between state variables and parameters? Should exogenous as well as endogenous variables be included? It should be noted that our above description of ODD already includes answers to some of these questions, but they were not answered in the original publication describing ODD (Grimm et al. 2006) used as a basis for the descriptions herein. Obviously, as anticipated by the developers of ODD, the description and explanation of ODD needs to be refined to take into account different terminologies and mindsets in different disciplines. Providing more precise definitions of these terms is important in assisting users of ODD to put the right information in the right place. While this will be the subject of future work, achieving the general applicability for which ODD is intended will inevitably mean there is some ambiguity.

#### 4.5

Clearly, for ODD to be successful, it needs to be easy to do, and not to require communication with its creator each time it is used. This, however, is a general issue pertaining to establishing a community of practice with ODD (something that will happen in time if ODD is more widely accepted), and is not specifically a problem with its application to agent- rather than individual-based models. Currently, an update not of ODD itself but of its description is being developed, with an accompanying set of examples, and will be made available on the Internet (<http://www.ufz.de/oesatools/odd/>; see also <http://www.openabm.org/>).

#### 4.6

Though we found the Design Concepts section relatively easy to complete, it is not without controversy. Certain terms, such as 'fitness', that are quite acceptable in the ecological modelling community, have political connotations in the social sciences that some may find uncomfortable. The Design Concepts section is the least structured of those in ODD. Some of the matters it addresses seem trivial for certain applications, others, such as emergence, much more fundamental. However, the point of the section is, at least in part, to raise common threads occurring in the literature (Railsback 2001; Grimm and Railsback 2005). By its very nature, the Design Concepts section will have to be fluid in structure, and for some applications, certain elements will inevitably seem trivial. It is nonetheless worth documenting them to facilitate comparison along various dimensions. We kept the original set of elements in the Design Concepts sections given in the examples, and they do seem successful in bringing out key differences among the models.

#### 4.7

As such, Design Concepts could be seen as acting rather like an FAQ. The set of topics covered for social simulation in the Design Concepts section may, if ODD is more widely adopted, be different from those for ecological models (though dialogue will be encouraged by retaining as many linked terms and common questions as possible). The questionnaire of

Richiardi et al. (2006) may serve as a starting point for the contribution of social simulation to this melting pot, and for LUCC models, the MR POTATOHEAD meta-modelling framework (Parker, Brown, Polhill, Deadman and Manson 2008) provides further dimensions of comparison.

#### 4.8

Richiardi et al.'s (2006) questionnaire merits more detailed consideration, since it, much like ODD, is aimed at increasing methodological rigour and replicability in I/ABMs, but is an initiative from the social rather than ecological simulation community. While the questionnaire does not follow the same logical structure as ODD, there are some overlaps in the information it requires. The first three questions on the questionnaire consist of a checklist of material some of which is covered (in rather more detail) in the Overview category of ODD and some in the Design Concepts category, though much of the analysis question (3) belongs more in the Details category. Question 4 of the questionnaire focuses on availability of software and relevant documentation, and question 5 covers more detailed information that ODD also covers in the Details category. As suggested above, the questionnaire may provide material that a revision to ODD could draw on, particularly if such a revision were to bear in mind issues specific to social rather than ecological simulation. However, as it stands, the questionnaire does not provide, like ODD, a structure for a complete and readable account of a model. Though this is not the intention of the questionnaire, we would recommend that the working group discussed in Richiardi et al.'s paper (section 6) bear ODD in mind as part of the voluntary methodological standard for agent-based simulations it is to propose in step 3 of the outlined strategy.

#### 4.9

Two issues do pose questions for ODD's general application in agent-based modelling, which did not arise in the exercises here due to similarities in the authors' approach to putting their various models together. First, the models described here are all procedural, rather than declarative models.<sup>[2]</sup> There is an established body of work using declarative modelling in agent-based social simulation (in particular, that involving the use of Moss, Gaylard, Wallis and Edmonds' (1998) Strictly Declarative Modelling Language) that perhaps does not exist in the individual-based modelling community. With declarative modelling, it may not be meaningful to talk about state-variables, as variables in declarative programs have rather different properties to those in procedural programs. Further, in declarative models, the schedule (or order of events) 'emerges' from the question posed to the inference engine (Moss, pers. comm.), and is largely beyond the control of the developers. While the question of whether ODD can be effectively used to describe declarative models remains open until a practitioner using that approach tries it, clearly such descriptions will have differences from those of models in procedural languages that may affect the comparability of such models, and the ability to replicate declarative models procedurally and vice versa.

#### 4.10

In light of this, it is somewhat ironic that ODD does not follow the object-oriented paradigm, where the instance variables of model entities are described together with their processes. The reason for this approach was that the processes of an agent often would refer to other entities, which might not yet have been described, so that forward references would be necessary, reducing readability.

#### 4.11

It could be countered that object-oriented (OO) frameworks are more appropriate for description of I/ABMs. OO frameworks can be used to express hierarchical taxonomies, which can more effectively facilitate comparison between models, or expression of a series of models as special instances of a general class of models. The use of taxonomies and definitions of classes of model to describe agent-based models is already under way. Hare and Deadman (2004) have made the first steps towards creating a taxonomy of land-use change models, and Cioffi-Revilla and Gotts (2003) have defined the TRAP<sup>2</sup> class of models.

#### 4.12

On the other hand, ODD is not specifically designed for model comparison (though it could help with this) but is focused on communication, and should ideally not be related to the underlying software design. Indeed, perhaps the issue with declarative models suggests ODD is already too closely tied to a particular modelling technology. Tying ODD more closely to OO at this early stage would limit its generality, and confuse software design with model description too much.

#### 4.13

Nevertheless, OO is an important concept, and a compromise could be, as has been suggested by Grimm et al. (2006), that for those models that have been implemented using OO languages, UML class diagrams could be used to communicate both the model entities' state variables and their processes in the same diagram. A UML class diagram used in Polhill, Parker and Gotts (2007) to communicate the structure of FEARLUS+ELMM (without using ODD) is shown in Figure 5. One of the weaknesses of this approach in comparison with the

ODD description given above is that the purpose of the variables is not explained in detail. In ODD itself, the link between entities and processes is communicated in the schedule. Here, a general format still has to emerge, but a description of the schedule should not only include *what* is done at *what* time, but also *who* is doing it, and in which order. A good example is given in Pitt, Box and Knowlton (2003), who use the schedule of their Swarm implementation to describe who is doing what at what time.

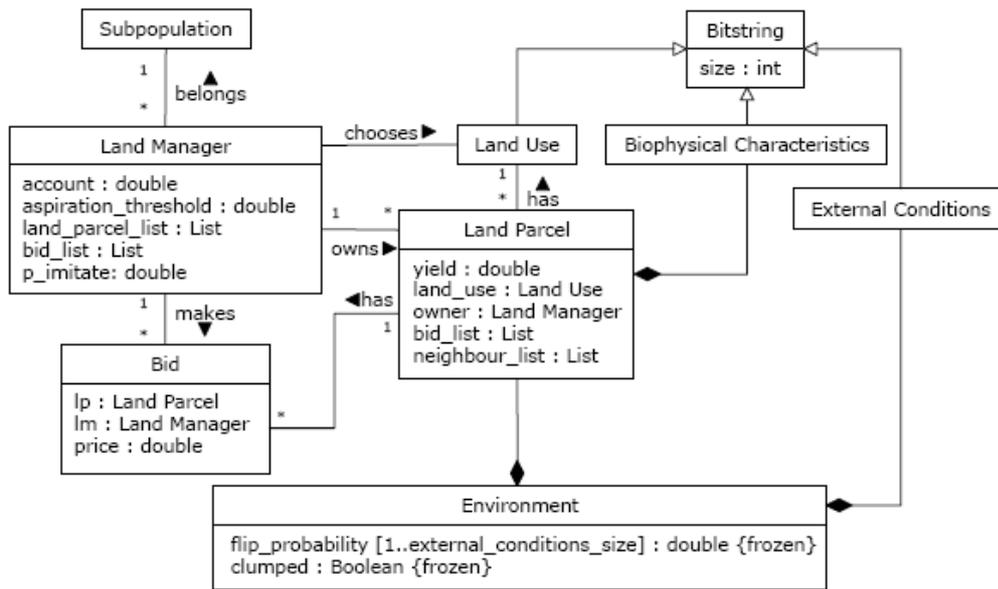


Figure 5. UML class diagram depicting part of the structure of FEARLUS+ELMM. (Copyright 2008 IGI Global. Reprint by permission of the publisher.)

4.14 The second issue pertains to the approach to designing the models taken in different fields. Ecologists and social scientists may differ in the way they view modelling activities. For example, the companion modelling approach of Barreteau et al. (2003) is subjective, and, by design, highly context-dependent, in contrast to the kinds of models produced by ecologists. The purpose of the model is to facilitate a dialogue among stakeholder factions, rather than to establish any authoritative or predictive narrative of a scenario. Here, though replication and comprehension of the model itself may seem less important, ODD may still serve in documenting for posterity the outcome of that dialogue, providing a complete and readable account of the world-view(s) proposed by that community. Again, however, until such models are described using ODD, this possibility cannot be fully evaluated.

4.15 A more trivial point that might be raised about ODD is the case where a model has so many state variables and processes that merely describing them in the kind of shallow detail required in the Overview would not be feasible in the context of a journal. This is an interesting question, but it is not specific to ODD. The same issue could equally arise for describing the model completely using text, UML diagrams, pseudocode or any other technique. Whilst the issue ultimately raises the question of the limitations of traditional publication streams for communicating simulation work, a model so vast that it cannot feasibly be described to allow others to fully understand it would surely be unlikely to pass the peer-review process.

4.16 Related to this, however, is the question of what happens when someone adapts a pre-existing model already described with ODD. Surely then one is only interested in the differences between the original model and the adapted model, which would be obscured by a full ODD should it be insisted this be produced for the adapted model? Clearly there is potential demand for a ΔODD, a structure to note changes to an existing ODD. Such a ΔODD, a possible subject of future work, may also be useful for noting incremental adaptations to a model and integration of submodels; both the kinds of activity that could potentially result in a model too large for description in one journal article. ΔODD might also help solving the problem that with complex models it can be difficult to say what "the" model is, because alternative submodels of the same processes, or a large number of model simplifications were tested, as is the case for FEARLUS+ELMM.

## Conclusion

5.1 The goals of ODD are simple: a complete and readable account of a model in a journal article.

As such, it should not be confused with various other approaches to describing software, simulations and models that do not have these specific goals. In particular, it would be a mistake to view ODD as akin to a meta-model (such as the MR POTATOHEAD framework created for agent-based models of land-use change), or language used for meta-modelling (such as OWL (Antoniou and van Harmelen 2004)), which has been used to create ontologies describing agent-based models and agent-based modelling (e.g. Christley, Xiang and Madey 2004). ODD is, and should be, nothing more than a document structure; a partial response to a number of normative calls for greater openness and integration in agent-based modelling (e.g. Alessa, Laituri and Barton 2006; Schweik, Evans and Grove 2005).

## 5.2

We conclude that ODD should further be considered for use with social simulation models. However, more and full applications of ODD are needed to better understand its specific potentials and limitations when dealing with models that include human agents. As more and more examples of ODD appear in the literature, it will be easier for others to follow the protocol without having the advice of its developers, and easier to see where changes are needed to clarify requirements, facilitate the model description process and fit the description needs of different models. We could confirm the status of ODD described in Grimm et al. (2006) as being only the first step on the way to establishing a general protocol for describing individual- and agent-based models, though a nonetheless promising step. The success of ODD depends on it being used, scrutinized, and developed. For ODD, a forum for this is provided at via <http://www.ufz.de/oesatools/odd/> and at <http://www.openabm.org>.



## Acknowledgements

This work was funded in part by the Scottish Government Rural and Environment Research and Analysis Directorate, the US National Science Foundation (grant #BCS-0119804), and the European Commission (Nest project PATRES; Contract No 043268). We thank Marc Artzrouni and the participants of M2M2007, particularly Nils Ferrand, Luis Izquierdo, Ken Kahn and Juliette Rouchier, and three anonymous reviewers for useful comments and feedback on earlier drafts of this paper.



## Notes

<sup>1</sup> Specifically, this is achieved using duplicates of instance variables changed by Land Managers' decisions, which are then updated in the next schedule step.

<sup>2</sup> For those unfamiliar with this distinction, procedural programming languages consist of a series of statements that (loops, conditionals and function or method calls aside) are executed in the order written. Declarative programming languages consist of a series of statements that could be considered to be logically axiomatic, the order of which need not matter (though this depends on the particular language being used). The program is then run by asking a question of an inference engine, which attempts to derive an answer to the question by executing an appropriate selection of the statements in an order determined by the principles on which it is built.



## References

ALESSA L N, Laituri M and Barton M (2006). An "all hands" call to the social science community: Establishing a community framework for complexity modeling using agent based models and cyberinfrastructure. *Journal of Artificial Societies and Social Simulation* 9 (4) 6. <http://jasss.soc.surrey.ac.uk/9/4/6.html>.

ANTONIOU G and van Harmelen F (2004). Web Ontology Language: OWL. In Staab S and Studer R (Eds.) *Handbook on Ontologies*. Berlin: Springer. pp. 67–92.

AXELROD R M (1986). An evolutionary approach to norms. *American Political Science Review*. 80 (4), 1095–1111.

AXTELL R, Axelrod R, Epstein J, and Cohen M D (1995). Aligning simulation models: A case study and results. *Computational and Mathematical Organization Theory* 1 (1), 123–141.

BARRETEAU O and others (2003). Our companion modelling approach. *Journal of Artificial Societies and Social Simulation* 6 (2) 1. <http://jasss.soc.surrey.ac.uk/6/2/1.html>

BAUER B, Müller J and Odell J (2001). Agent UML: A formalism for specifying multiagent interaction. In Ciancarini P and Wooldridge M (Eds.) *Agent-Oriented Software Engineering: First International Workshop AOSE 2000, Limerick, Ireland, June 10, 2000. Lecture Notes in*

Computer Science 1957. Berlin: Springer-Verlag. pp. 91-103.

BERRY B, Kiel L, and Elliot E (2002). Adaptive agents, intelligence, and emergent human organization: Capturing complexity through agent-based modeling. *Proceedings of the National Academy of Sciences* 99 (Supplement 3), 7178-88.

BIGBEE A, Cioffi-Revilla C and Luke S (2005). Replication of Sugarscape using MASON. In Troitzsch K G (Ed.) *Representing Social Reality: Pre-proceedings of the Third Conference of the European Social Simulation Association, Koblenz, September 5-9, 2005*. Koblenz: Verlag Dietmar Fölbach. pp. 6-15.

BOERO R and Squazzoni F (2005). Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science. *Journal of Artificial Societies and Social Simulation* 8 (4) 6. <http://jasss.soc.surrey.ac.uk/8/4/6.html>

BROWN D G, Page S E, Riolo R and Rand W (2004). Agent based and analytical modeling to evaluate the effectiveness of greenbelts. *Environmental Modelling and Software* 19, 1097-1109.

BROWN D G, Page S E, Riolo R, Zellner M and Rand W (2005). Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographic Information Systems* 19, 153-174.

BROWND G and Robinson D T (2006). Effects of heterogeneity in preferences on an agent-based model of urban sprawl. *Ecology and Society* 11 (1): 46, <http://www.ecologyandsociety.org/vol11/iss1/art46/>.

CHRISTLEY S, Xiang X and Madey G (2004). Ontology for agent-based modeling and simulation. In Macal C M, Sallach D and North M J (Eds.) *Proceedings of the Agent 2004 Conference on Social Dynamics: Interaction, Reflexivity and Emergence, co-sponsored by Argonne National Laboratory and the University of Chicago, October 7-9*. <http://www.agent2005.anl.gov/Agent2004.pdf>.

CIOFFI-REVILLA C and Gotts N (2003). Comparative analysis of agent-based simulations: Geosim and FEARLUS models. *Journal of Artificial Societies and Social Simulation* 6 (4) 10. <http://jasss.soc.surrey.ac.uk/6/4/10.html>.

DAVIES B, Koo B, Hunsberger C, Rothman D, Polhill G, Blackstock K, Izquierdo L, Gotts N, Ferrier R and Dunn S (2006). Story and simulation approaches to appraise possible programmes of measures for a sub basin in Scotland. *Third Harmoni-CA Forum and Conference, 5-7 April 2006, Osnabrück, Germany*. [http://www.harmoni-ca.info/Conferences/Past\\_Meetings/3rd\\_Harmoni-CA\\_Forum\\_and\\_Conference/abstract%20Blackstock.pdf](http://www.harmoni-ca.info/Conferences/Past_Meetings/3rd_Harmoni-CA_Forum_and_Conference/abstract%20Blackstock.pdf).

DREYFUS-LEON M and Kleiber P (2001). A spatial individual behaviour-based model approach of the yellowfin tuna fishery in the eastern Pacific Ocean. *Ecological Modelling* 146 (1-3), 47-56.

EDMONDS B and Hales D (2003). Replication, replication and replication: Some hard lessons from model alignment. *Journal of Artificial Societies and Social Simulation* 6 (4) 11, <http://jasss.soc.surrey.ac.uk/6/4/11.html>.

FERNANDEZ L, Brown D G, Marans R and Nassauer J (2005). Characterizing location preferences in an exurban population: Implications for agent based modeling. *Environment and Planning B*, 32 (6), 799-820.

GALAN J M and Izquierdo L R (2005). Appearances can be deceiving: Lessons learned re-implementing Axelrod's 'Evolutionary Approach to Norms'. *Journal of Artificial Societies and Social Simulation* 8 (3) 2, <http://jasss.soc.surrey.ac.uk/8/3/2.html>.

GILBERT N and Troitzsch K G (2005). *Simulation for the social scientist*. Open University Press.

GOSS-CUSTARD J D, Burton N H K, Clark N A, Ferns P N, McGroarty S, Reading C J, Rehfish M M, Stillman R A, Townend I, West A D and Worrall D H (2006). Test of a behavior-based individual-based model: response of shorebird mortality to habitat loss. *Ecological Applications*, 16, 2215-2222.

GOTTS N M, Polhill J G and Law A N R (2003). Aspiration levels in a land use simulation. *Cybernetics and Systems* 34, 663-683.

GRIMM V (2007). Ecological models: Individual-based models. In Jørgensen S-E (Ed.) *Encyclopedia of Ecology*. Elsevier (in press).

GRIMM V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, Goss-Custard J, Grand T, Heinz S K, Huse G, Huth A, Jepsen J U, Jørgensen C, Mooij W M, Müller B, Pe'er G, Piou C, Railsback S

F, Robbins A M, Robbins M M, Rossmannith E, Rüger N, Strand E, Souissi S, Stillman R A, Vabø R, Visser U and DeAngelis D L (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198 (1–2), 115–126.

GRIMM V and Railsback S F (2005). *Individual-based Modeling and Ecology*. Princeton University Press, Princeton, NJ.

GRIMM V, Revilla E, Berger U, Jeltsch F, Mooij W M, Railsback S F, Thulke H–H, Weiner J, Wiegand T and DeAngelis D L (2005). Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science* 310, 987–991.

HARE M and Deadman P J (2004). Further towards a taxonomy of agent based simulation models in environmental management. *Mathematics and Computers in Simulation* 64 (1), 25–40.

HARPER, S J, Westervelt J D and Trame A–M (2002). Management application of an agent-based model: control of cowbirds at the landscape scale. In Gimblett H R (Ed.) *Integrating Geographic Information Systems and Agent-Based Modeling Techniques for Understanding Social and Ecological Processes*. Oxford University Press.

KALDOR N (1961/1968). Capital Accumulation and Economic Growth. In Lutz, F A and Hague, D C (Ed.): *The Theory of Capital*. Reprint. London: Macmillan, pp. 177–222.

KOO B K, Dunn S M and Ferrier R C (2004). A spatially-distributed conceptual model for reactive transport of phosphorus from diffuse sources: an object-orientated approach. In Pahl–Wostl C, Schmidt S and Jakeman T (Eds.) *IEMSs 2004 International Congress: Complexity and Integrated Resources Management, International Environmental Modelling and Software Society, Osnabrück, Germany, 14–17 June 2004*.

MARANS R W (2003). Understanding environmental quality through quality of life studies: The 2001 DAS and its use of subjective and objective indicators. *Landscape and Urban Planning* 65 (1–2), 75–85.

MATHEVET R, Bousquet F, Le Page C and Antona M (2003). Agent-based simulations of interactions between duck population, farming decisions and leasing of hunting rights in the Camargue (Southern France). *Ecological Modelling* 165: 107–126.

MOSS S, Gaylard H, Wallis S and Edmonds B (1998). SDML: A multi-agent language for organizational modelling. *Computational & Mathematical Organization Theory* 4 (1), 43–69.

PARKER D C (1999). Landscape outcomes in a model of edge-effect externalities: A computational economics approach. *SFI Publication 99–07–051*. Santa Fe, NM: Santa Fe Institute. <http://www.santafe.edu/research/publications/workingpapers/99-07-051.pdf>.

PARKER D C (2005). Agent-based modelling to explore linkages between preferences for open space, fragmentation at the urban-rural fringe, and economic welfare. *Paper presented at The Role of Open Space and Green Amenities in the Residential Move from Cities, December 14–16 2005, Dijon, France*.

PARKER D C, Berger T and Manson S M (2002). Agent-Based Models of Land-Use and Land Cover Change: Report and review of an international workshop, October 4–7, 2001, Irvine, California, USA. *LUCC Report Series No. 6*. Anthropological Center for Training and Research on Global Environment Change, Indiana University, USA: LUCC Focus 1 Office.

PARKER D C, Brown D G, Polhill J G, Deadman P J and Manson S M (2008). Illustrating a new 'conceptual design pattern' for agent-based models and land use via five case studies: The MR POTATOHEAD framework. In Lopez Paredes A and Hernandez Iglesias C (Eds.) *Agent-Based Modelling in Natural Resource Management Valladolid, Spain: Universidad de Valladolid*, pp. 23–51.

PARKER D C and Meretsky V (2004). Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. *Agriculture, Ecosystems and Environment* 101, 233–250.

PIOU P, Berger U, Hildenbrandt H, Grimm V, Diele K and D'Lima C (2007). Simulating cryptic movements of a mangrove crab: recovery phenomena after small scale fishery. *Ecological Modelling* 205, 110–122.

PITT W C, Box P W and Knowlton F F (2003). An individual-based model of canid populations: modelling territoriality and social structure. *Ecological Modelling*, 166, 109–121.

POLHILL J G, Gotts N M and Law A N R (2001). Imitative versus nonimitative strategies in a land use simulation. *Cybernetics and Systems* 32 (1–2), 285–307.

POLHILL J G, Parker D C and Gotts N M (2005). Introducing land markets to and agent based

models of land use change. In Troitzsch K G (Ed.) *Representing Social Reality: Pre-proceedings of the Third Conference of the European Social Simulation Association*, Koblenz, September 5–9, 2005. Koblenz: Verlag Dietmar Fölbach. pp.150–157.

POLHILL J G, Parker D C and Gotts N M (2008). Effects of land markets on competition between innovators and imitators in land use: results from FEARLUS–ELMM. In: Hernandez, C., Troitzsch, K. and Edmonds, B. (Eds.), *Social Simulation Technologies: Advances and New Discoveries*. New York: IGI Global. pp. 81–97.

POLHILL J G, Pignotti E, Gotts N M, Edwards P and Preece A (2007). A semantic grid service for experimentation with an agent–based model of land use change. *Journal of Artificial Societies and Social Simulation* 10 (2) 2, <http://jasss.soc.surrey.ac.uk/10/2/2.html>.

RAILSBACK S F (2001). Concepts from complex adaptive systems as a framework for individual–based modelling. *Ecological Modelling*, 139, 47–62.

RAILSBACK S F and Harvey, B C (2002). Analysis of habitat selection rules using an individual–based model. *Ecology*, 83, 1817–1830.

RAND W, Brown D G, Page S E, Riolo R, Fernandez L E and Zellner M (2003). Statistical validation of spatial patterns in agent–based models. *Proceedings of Agent Based Simulation 4*, Montpellier, France, 2003.

RICHIARDI M, R Leombruni R, Saam N and Sonnessa M (2006). A Common Protocol for Agent–Based Social Simulation. *Journal of Artificial Societies and Social Simulation* 9 (1) 15. <http://jasss.soc.surrey.ac.uk/9/1/15.html>.

RIOLO R L, Cohen M D and Axelrod R (2001). Evolution of cooperation without reciprocity. *Nature* 411, 441–443.

SCHWEIK C, Evans T and Grove J M (2005). Open source and open content: A framework for global collaboration in social–ecological research. *Ecology and Society* 10 (1): 33. <http://www.ecologyandsociety.org/vol10/iss1/art33/>.

WANG M and Grimm V (2007). Home range dynamics and population regulation: an individual–based model of the common shrew. *Ecological Modelling* 205, 397–409.

WILENSKY U (1999). Netlogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer–Based Modeling, Northwestern University, Evanston, IL.

WILENSKY U and RAND W (2007) Making models match: replicating an agent–based model. *Journal of Artificial Societies and Social Simulation* 10 (4) 2. <http://jasss.soc.surrey.ac.uk/10/4/2.html>

ZELLNER M L, Riolo R, Rand W, Page S E, Brown D G and Fernandez L E (2003). The interaction between zoning regulations and residential preferences as a driver of urban form. 2003 UTEP Distinguished Faculty and Student Symposium, Urban and Regional Planning Program, University of Michigan. April, 2003.

---

[Return to Contents of this issue](#)

© Copyright *Journal of Artificial Societies and Social Simulation*, [2008]

