Introduction to agent-based and cellular automaton models of land-use change

Dawn Cassandra Parker
Assistant Professor
Dept. of Computational Social Science
Center for Social Complexity
George Mason University
With contributions also from many others!

4 Feb. 2009
Universita' Ca' Foscari di Venezia
First Changes Winter School on *System Analysis and Integrated Modelling in Climate Change Research*

Outline

- Social and ecological consequences of land-use change
- Modeling concepts and definitions
- Questions to answer when building a land-use model
- Three modeling techniques: CA, statistical, and agent-based
- More on agent-based!

Social consequences of land-use change

- **Economic**: property values, commuting costs, level of business and industrial activity, mix of outputs, efficiency of land use, future income potential from land
- **Quality of life**: environmental amenities recreational opportunities, social interactions, accessibility to work, school, shopping, etc., health
- **Political**: changes in relative political influence of regions, changes in makeup of stakeholder groups, trigger of political action for stakeholder groups

Environmental/Ecological consequences of land-use change

- Changes in biodiversity
- Changes in habitat structure and species abundance
- Changes in ecosystem function/services
- Land degradation/restoration
- Changes in hydrologic networks
- Transport and fate of pollutants
- Global climate impacts

Modeling Definitions and Classifications

"Model" (from Briassoulis, 4.1)

- "the formal representation of some theory of a system of interest (Wilson 1974, 4)"
- An abstract, symbolic representation of a real-world system, potentially based on broader theory or concept
- "an idealized and structured representation of the real" (Johnston et al. 1994, 385, 622)
- "an experimental design based on a theory" (Harris 1966, 258; see also, Romanos 1977, 135)

"Theory" and "model" are not equivalent terms.

Theoretical vs. empirical models

- Distinction often not clarified in reviews of LUCC models, *especially* in the economics literature
- Theoretical models are based on abstract or algebraic representations and are not designed to apply to real-world cases
- Empirical models are structured and parameterized to apply to, or reflect a particular time and place

Types of models

- Analytical: based on a formal and complete mathematical (algebraic) representation of a system.
- Simulation: based on a mathematical representation of the behavior of individual entities in the system and their interrelationships.
- Statistical: estimated from real-world data assuming sampling from an uncertain distribution (stochasticity).

These models are generally complementary.

Scope of LUCC models (from Parker, Berger, Manson 2002a and b)

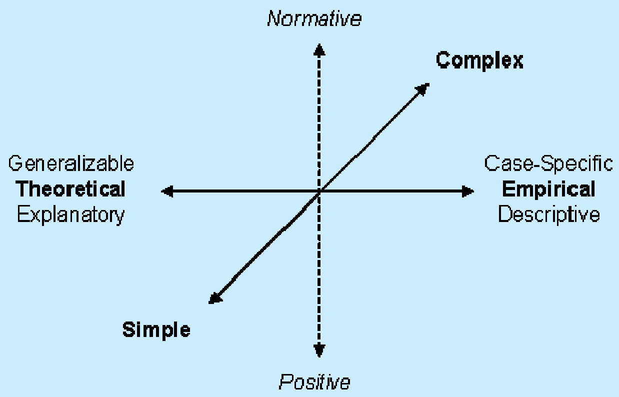


Figure 1. Continua for Categorizing Agent-Based Models

Uses of LUCC models (Briassoulis 1999, Parker, Berger, Mason 2002a and b)

- Explanatory (less so for empirical models)
- Exploratory
- Decision support/policy analysis
- Impact or scenario analysis
- Tell you what *won't* happen (Berger)
- Prediction/forecasting (MANY caveats here!)
- Prescriptive (normative: tell you what *should* happen)

Questions to answer when building a LUCC model

1. What is the model for?
2. What is the ideal spatial and/or temporal representation (structure and resolution) for the model?
3. What data are available for model construction?
4. What real-world spatial, temporal, and behavioral processes do you strive to represent?
5. What is the minimum level of complexity for the model?
6. What modeling methodologies are most appropriate?

Additional practical questions for model implementation

7. What software will be used, and how will software be integrated?
8. What techniques will be used for calibration, verification, and validation?
9. How will model input data, code, and results be handled?

Some modeling techniques used for fine-scale LUCC modeling:

- Statistical and econometric models
- Cellular automata
- Agent-based models

Cellular Automaton Models

- CA models are dynamic simulation models, where cell transitions are based on the state of the current cell and the states of neighboring cells.
- “Neighbors” can be very broadly defined, and may include multi-scale influences
- Transitions may also depend on cell history (Markov models)
- Cellular structures are generally grids, but can be any cellular structure, in principle

Example: DUEM model (Batty, Xie, and Sun)

Purpose of the model:

- Demonstrate how urban sprawl can occur without population growth

Model Mechanisms

- Model inputs real-world raster layers to define initial land uses
- Possible land-use classes include housing, commercial, industrial, vacant, and roads
- Transition rules are influenced by spatial influences and temporal constraints

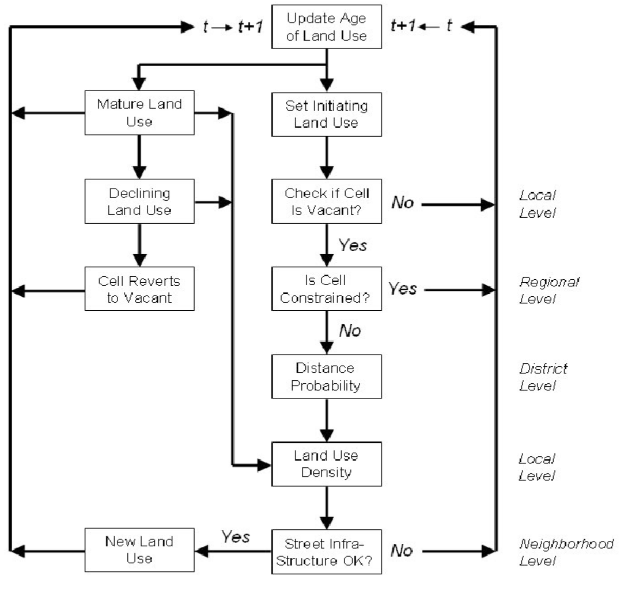


*Figure 1: The Sequence of Operations in the **DUEM** Model*

Spatial Complexity

- The urban system is represented by three nested scales: Neighborhood, Field, and Region
- Transitions are influenced by:
  - the other land uses in the local neighborhood
  - the density of land uses in the district
  - constraints on development in the region
- The size of the neighborhood is determined by the user

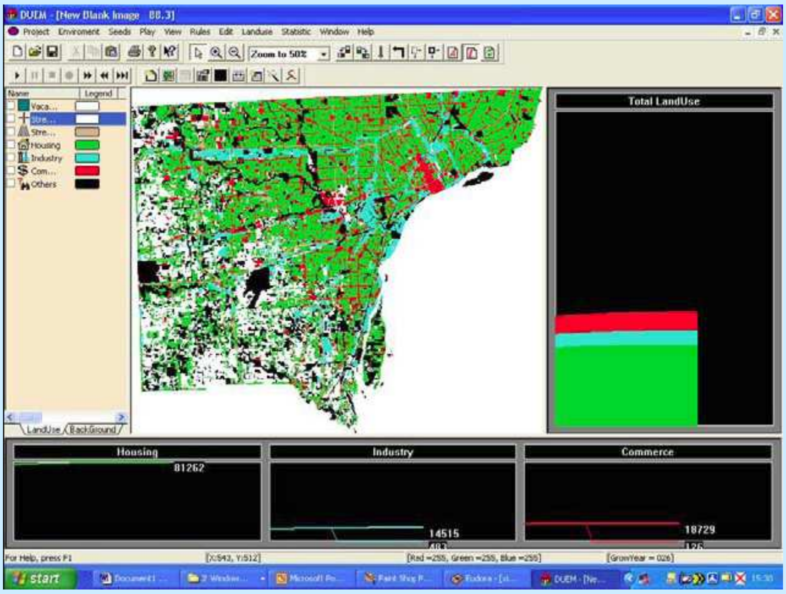
Temporal Complexity

- Land-use generation sequences are restricted (for example, streets generate only streets)
- Land uses have life cycles, and can revert to vacant land
- Only new land uses generate other new land uses

Behavioral Complexity

- No explicit behavioral complexity in this model, as there are no explicit decision-making agents
- Effects of agent decisions are implicitly represented through transition probabilities

DUEM interface and output (Batty et al, 2005)



Some strengths and weaknesses of cellular automaton models

- Strengths:
  - Models are very strong at representing local spatial processes
  - Models tend to do well at replicating real-world spatial patterns, especially fractal structures
- Weaknesses:
  - Models may place too much emphasis on local interactions
  - Models are not strong at representing behavior, especially when agents are mobile
  - Often, models require projections of rates and quantities of change to run

Other CA projects

- Many innovative projects at RIKS (CA, integrated models, and V&V
- SLEUTH model applications around the world
- Work by CASA alumni (including Benenson and Torrens, O' Sullivan, etc.

Statistical/Regression Models

- These models find a set of best-fit model coefficients that express a statistical relationship between a *dependent variable* (often land use or cover) and a series of *independent variables* (representing drivers of LUCC)
- Models produce a transition probability, conditional on states of independent variables
- Models are only dynamic when some set of rules is used to generate transitioned landscapes using those estimated probabilities

Example: CLUE-S model (Verburg et al.)

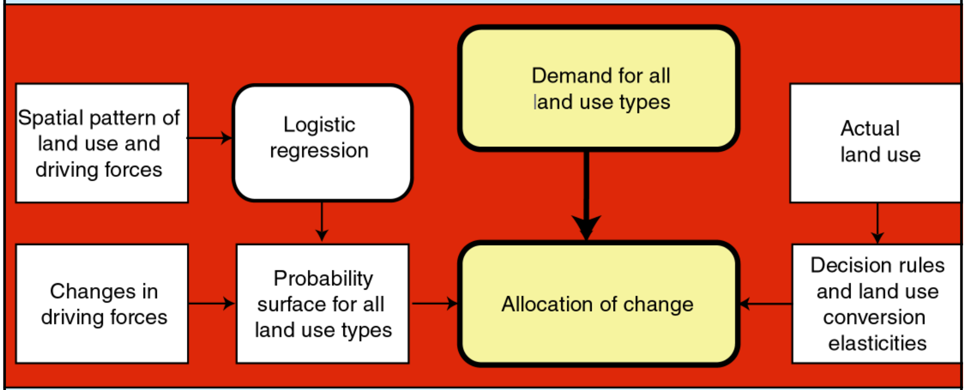
Purpose of the model:

- Projections of land-use change under status quo
- Scenario analysis
- Hypothetical impacts of new protected area
- Identify possible hot-spots of land-use change

Model structure

- Non-spatial model determines aggregate demand for land
- Spatial statistical model determines transition probabilities for particular land uses
- User-determined conversion rules limit possible transitions, in order to correctly reflect temporal dynamics
- Dynamic allocation protocol allocates change based on estimated transition probabilities and conversion elasticities

Model structure



Spatial Complexity

Interaction through accessibility

➔ The suitability of a location is (partly) determined by its access to facilities and/or other land uses

Direct interaction

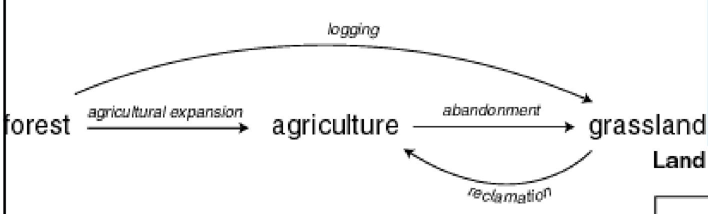
➔ Influence of neighboring land uses:

- Centripetal forces: economies of scale, labor markets etc.
- Centrifugal forces: congestion, environmental pollution etc.

Temporal Complexity

- Not all land use conversions are reversible:
  - Urban area and residential area
  - Deforestation of primary rain forest
- Other conversions are very costly:
  - Fruit tree plantations
- Some locations are converted after a short time period:
  - Abandonment after shifting cultivation (nutrient depletion)

Land use change sequence



Land use conversion matrix

present land use ↓ / future land use →	Forest	Agriculture	Grassland
Forest	+	+	+
Agriculture	-	+	+
Grassland	-	+	+

\+ conversion possible
\- conversion not possible

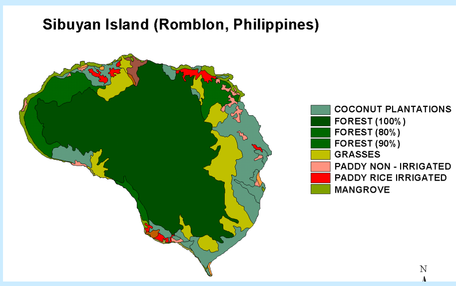
Example of the translation of a land use change sequence into a land use conversion matrix

Behavioral Complexity

- Location and community specific information such as population density, literacy and income enter the statistical model
- There is no explicit decision-making function
- There are no explicit agent-agent-interactions

Applications

Sibuyan island, Philippines



Some strengths and weaknesses of spatial statistical/regression models

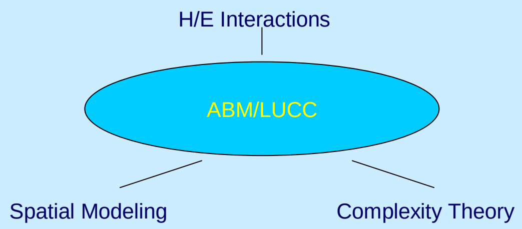
- Strengths:
  - Models provide information on key drivers of change
  - Spatial and temporal lags can be incorporated
  - Data can be entered at multiple scales
- Weaknesses:
  - Models themselves don't produce projections of spatial change
  - Arbitrary transition rules may lead to different change projections for the same data
  - Simulations of change require projections of rates and quantities of change
  - Models may have little out-of-sample power

Other statistical models …

- Many in the economics literature, especially hedonic property value models
- See for instance the work from U. Maryland, USA and extensions (Bockstael, Irwin, Geoghegan, etc.)
- Many developing country models using RS data (see my course bibliography for examples)

What are ABM models?

- ABM or MAS models are simulation models
- Generally implemented as computer code
- ABM's do not have a set of equilibrium conditions imposed on the model; generally, you do not "solve" or "estimate" the model
- ABMs can both complement and substitute for other modeling techniques



Spatial agent-based models of H-E consist of:

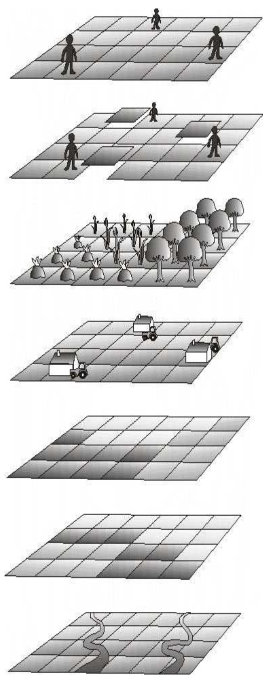
- An electronic representation of a landscape
- An "agent-based" simulation of decision-makers whose choices alter the landscape (usually a computer program)

(Parker et. al 2003; Parker, Berger, Manson 2002)

Cellular/Spatial Landscape Model

- May or may not be based on real-world maps via geographic information systems layers
- May contain a variety of geographic and socioeconomic features such as:
  - Slope, elevation, vegetative cover, soil types, zoning restrictions
  - Road and rail networks, information on social networks (who knows who)
  - Models of "spatial diffusion," such as how air pollution spreads and disperses across a region

Figure 7, LUCC Report #6



Agent-Based Model

- Autonomous decision-making agents
- Interaction environment
- Interdependencies among agents, their environment, or both
- Rules governing sequencing of actions and information flows

What is an “agent”?

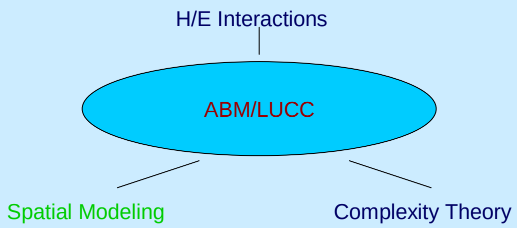
- Goal-oriented entity
- Model of cognition that links goals and behavior:
- Capable of autonomous action
- Capable of responding to changes in its environment

Agent-Based Model of Decision Making

- Each individual decision maker is represented through a set of rules that link information about his/her environment to a decision
- Decisions often depend on the agent's physical environment (the landscape)
- Decisions may also depend on what other agents do as well -- can lead to "path dependence"

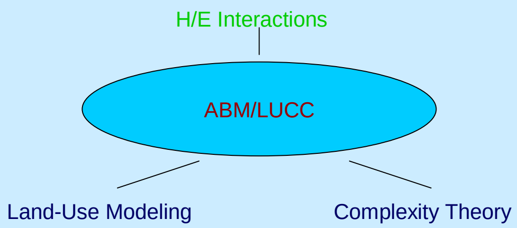
Potential Decision Models

- Simple heuristics
  - Classifier/rule based (LUCITA)
  - Imitative behavior (FEARLUS)
- Boundedly rational profit or utility maximization or risk minimization
  - Analytical implementations for theoretical models (SLUDGE, SOME)
  - Genetic algorithms (SYPRIA)
  - Mathematical programming (Berger, Balmann)



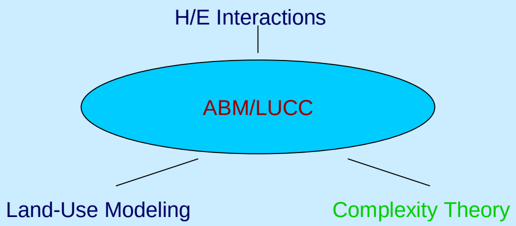
Advantages of ABM/LUCC for spatial modeling:

- Your model can have a realistic (and appropriate) geographic representation
- Potential links with geographic information systems for data input and output visualization
- Modeling of structures that are nested in time and space (cross-scale)



Modeling human/environment interactions:

- Socioeconomic and biophysical models can be linked spatially
- Simulation approach allows for feedbacks between dynamic social and environmental processes
- Applications include crop yields, hydrology, forest growth, pest species modeling, endangered species populations



Human/environmental landscapes are *complex:*

- Characterized by:
  - Interdependencies (one agent's action depend on what another has done previously)
  - Heterogeneity (diverse variation in the same type of object: object-oriented programmers think "subclass")
  - nested hierarchies (overlapping structures in time and space)

Key sources of agent heterogeneity:

- Pecuniary and non-pecuniary motivations: profits, preservation of family farm, environmental ethic
- Experience and knowledge
- Financial, physical, and human capital
- Access to credit
- Expectation formation mechanisms
- Decision strategies

Types of Interactions

Agent-Agent

- Information transfer
- Technology diffusion
- Land markets
- Local labor exchange
- Community-based resource management

Agent-Environment

- Hydrology (ground and surface)
- Erosion
- Deforestation
- Transport of pollutants
- Species migration
- Soil fertility

Sequencing and information transfer

- Pre-determined:
  - Synchronous
  - Asynchronous
- Event-driven

Note: Any event sequencing mechanism might introduce path-dependency into model outcomes

From abstraction to the real world …

Environment \ Agents	Designed	Analyzed
Designed	Existence proof; Discovery of new relationships	Laboratory experiments; Role-playing games
Analyzed	Explanation	Explanation; Projection; Scenario analysis

Berger/Parker (2002).

How might ABM be used to study H-E interactions?

- To link socioeconomic drivers of resource use to their biophysical impacts
- To explore the effects of feedbacks between humans and their environment
- To examine whether current systems of resource use are sustainable
- To design policies to encourage more sustainable resource use

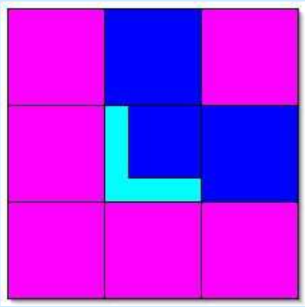
Example: SLUDGE (Simulated land use dependent on edge-effect externalities) (Parker et al.)

Purpose of the model--Investigate:

- Can spatial externalities lead to economically inefficient levels of landscape fragmentation?
- How do initial land-use patterns influence future land-use fragmentation?
- Are spatial externalities sufficient to produce "urban sprawl"?
- How to interactions between spatial externalities and transportation costs influence sprawl?

What is a "spatial externality"?

Costs or benefit to a neighbor of a surrounding land use, which are not taken into account when the generating neighbor makes a decision about land use



Model Mechanisms

- Model operates as a hybrid CA/ABM model, with a single agent in each cell/polygon
- Agents form expectations regarding land-use profitability, based on neighboring land uses and current land use pattern and composition
- Agents choose the highest-valued land use in each time period
- A landscape evolves where no agent can do better by changing land uses
- The model reports land rents, economic welfare measures, and measures of landscape fragmentation

Spatial Complexity

- Neighborhood effects: payoffs to a given land use depend on the actions of multiple neighbors
- Spatial pattern effects: Landscape productivity depends on the spatial arrangement of land uses, as well as amount of land in each use
- Transportation costs affect payoffs to each land use

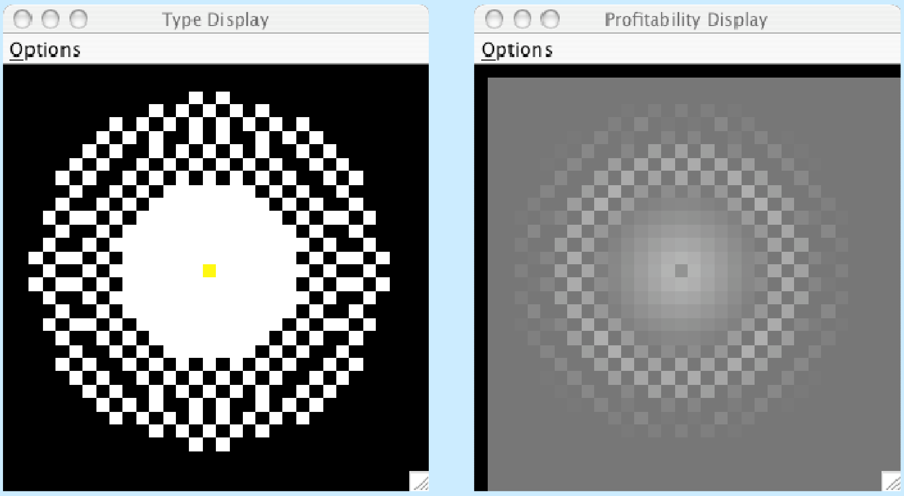
Temporal Complexity

- Agents form expectations about the future productivity of the “urban” land use
- Little other temporal complexity:
  - No constraints on land-use transitions
  - No land-use life cycles
- However, these factors can be included in agent models

Behavioral Complexity

- Prices, landscape pattern, and actions of neighbors influence agent decisions
- Agents have fairly mathematically sophisticated decision rules
- However:
  - Agents are homogeneous
  - Agents are not forward-looking
  - Agents interact with other agents indirectly
- Next generation model later today!

Landscape pattern effects of open-space amenities



- Result is fragmentation, leapfrogging

Some strengths and weaknesses of agent-based models

- Strengths:
  - Models can incorporate important sources of spatial, temporal, and behavioral complexity
  - Very strong format for integrated models (human-environment interactions
  - Potentially strong for cross-scale feedbacks
  - Can link human actions to landscape pattern
- Weaknesses:
  - Can be difficult to map and communicate model mechanisms and outcomes
  - Error propagation is potentially very high
  - Can be very data hungry

Identified Challenges (Introductory readings, Framing LUCC meetings)

- Understanding complex models
- Modeling cross-scale feedbacks and dynamics, land markets, and institutions
- Modeling and verifying agent decision-making
- Verification and validation of landscape outcomes (if possible)
- Communication of model mechanisms and results (esp. to policy makers)

More challenges and areas of progress

- Using survey data to develop/parameterize agent models
- Identifying complementarities between statistical and ABM models
- Combining ABM and lab experiments to understand agent behavior
- Integrating ABM and GIS
- Building a community modeling platform
- Participatory modeling

ABM, knowledge synthesis, and discovery

- Can contribute to knowledge synthesis by integrating interdisciplinary sub-models
- Can contribute to knowledge discovery
  - During model building (esp. stakeholder participation and companion modeling)
  - Through experiments that link human and biophysical drivers of LUCC and outcomes
- Proposed as a framework to extend and synthesize LUCC theory

The MR POTATOHEAD framework (Parker et al. 2008 a,b)

- Expressed as a conceptual, object-oriented classification of elements that are essential for an ABM/LUCC model, with alternative instances
- Seven ABM/LUCC models described via MR POTATOHEAD as special cases of the general meta-model
- Model instantiated in Web Ontology Language with the goal of model comparison and meta-model creation

MR POTATOHEAD: longer-term goals

- Create a simple code base that nest the 7 models (perhaps more) as special cases
  - Allows for formal comparison of model structures and results
  - Facilitates comparative experiments to explore alternative parameter spaces
  - Provides code base for new model creation
- Build a graphical modeling front-end
  - Educational use
  - Reduces barriers to entry to the field
- Start over with a new model built around land use/management, land cover, and landscape function?