Statistical Validation of Spatial Patterns in Agent-Based Models

William Rand
Daniel G. Brown
Scott E. Page
Rick Riolo
Luis E. Fernandez
Moira Zellner

Project Sluce
Modeling Group
http://www.cscs.umich.edu/
sluce@umich.edu

Center for the Study of Complex Systems

University of Michigan

NSF
Overview of Presentation

- Project Overview and Urban Modelling
- Validation of ABMs
- SOME
  - Agents
  - Environment
  - Interaction
- Valid Results :)
- Conclusions and Other Work
Project Sluce
*(Spatial Land Use Change and Ecological Effects)*

- Project Sluce focuses on land-use dynamics at the Urban-Rural Fringe
  - Greater Detroit Metropolitan Area
- Uses an ABM in synthesis with other modeling techniques
- Results will be compared with historical data and theoretical models
Modelling of Urban Development

- Understanding the Processes that result in Urban and Exurban Patterns
- Suburban Sprawl = Negative Ecosystem Impacts
  - Habitat Destruction
  - Migration Corridor Destruction
- Make Prescriptive and Descriptive Statements about Processes
- Goal is to Minimize Ecological Damage
Two Types of Models

**Physical Analog Models**
- Examples - Markov Random Field, Diffuse Limited Aggregation, Correlated Percolation
- Based on well understood formal systems
- Hard to translate into 'real world'

**Agent-Based Models**
- Examples – Schelling's Tipping model of Segregation, Otter's ABLOoM model, Our model
- Easy to Incorporate New Ideas
- Ontology understood by Policy Planners
- Not well understood
Validation of ABMs

- A model is valid if it can correctly answer questions it was designed to answer (Casti, 97)
- Two Methods of Validation
  - Matching model outputs to measured variables
    - Micro-details
    - Macro-level patterns
  - Matching component structures and interactions
Difficulties in Validation

- Validation is Hard
- Positive Feedbacks
- Path Dependence
- Extreme Sensitivity to Initial Conditions
- Unpredictability of Agent Adaptation
- Micro-details often impossible to match
Sluce's Solution

- Our goal is not to match micro-level detail
- We build our model from “first principles”
  - Heterogeneity
  - Bounded Rationality
  - Correspondence between virtual agents and real agents
- Matching of Macro-Level Patterns
  - Zipf's Law
  - Clark's Law
Zipf's Law

- Zipf (49) showed that there is a power law relationship between city populations and their rank.
- Contemporary research has shown this also is true between frequency of developed clusters and size.

\[ N(A) \approx A^{-r} \]

- Universally \( r \approx 2 \), \( A \) is the size of a cluster, \( N(A) \) is the frequency of that size.
Clark's Law

• Clark (51) showed that as the radius of a circle around a city increases the density of development decreases exponentially

\[ y \approx Ae^{-bx} \]

• y is the density, x is miles from city center, b and A are constants
• The constants vary for different areas and times,
Overview of SOME
(Sluce's Original Model for Exploration)

• Three Main Components
  • Environment
  • Agents
  • Agent Interaction
• Modular Structure
• Landscapes are archetypal and GIS-based
• Implemented in Swarm
Environment

- Lattice (Variable; 301x301)
- Initial Distribution of Service Centers (1)
- Standard Characteristics (5-10 underlying maps)
  - Natural Beauty (exogenous; normal distribution from [0,1]; spatially autocorrelated)
  - Distance to Service Centers, nearest 8 service centers are used (endogenous)
- And others...
- Many Output Variables (30-50 outputs per step)
  - Clustering Statistics
  - Radius vs. Density Statistics
  - And many more....
Agents

- Residents
  - Several Preferences (Ideal and Weight)
    - Natural Beauty (0.5)
    - Distance to Service Centers (0.5)
  - Multiplicative utility model
- Service Centers
  - Follow Last Resident in
  - Every 100 Residents, 1 Service Center enters
Agent Interaction

• Residents enter every time step (10)
• Look at random locations (10)
• Choose the location with the highest utility for their preferences
• Corresponding endogenous variables are updated
Valid Results: Zipf's Law

$r$ of 2.067 is within the bounds found in empirical data.
Valid Results: Clark's Law

Empirical data shows b to vary quite a bit, but is usually much larger than 0.0069
Results of Validation

• Model matches closely with empirical Zipf's Law data
  • Our results through are only over a few decades
• Model matches Clark's Law relationally
  • However slope is different than empirical models
• Further validation is warranted
Future Work

- **Validation on a Real Landscape**
  - Scio Township and Washtenaw County
  - Satellite Data from 1978, Parcel Data from 1950s
  - Examine Relationship between amount of Information and Predictability (Graceful Decay)

- **Other Pattern Analysis Metrics**
  - Variant versus Invariant Regions, Certainty of Development
  - Kappa Statistics and Information Gain
  - Cross-Correlation of Development
Any Questions?
sluce@umich.edu