The Challenges of Non-Stationarity, Multi- and Equi-Finality in Validating Land Change Models

Dan Brown

Environmental Spatial Analysis Laboratory
School of Natural Resources and Environment
University of Michigan

University of Michigan

Goals of Land Change Modeling

- Encode our knowledge of process
- Help test pattern-process links
  - Can help us examine feedbacks between ecosystem structure/function and human actions
- Test alternative futures under various hypotheses, policies, practices, and incentives
- Make projections of future landscape patterns

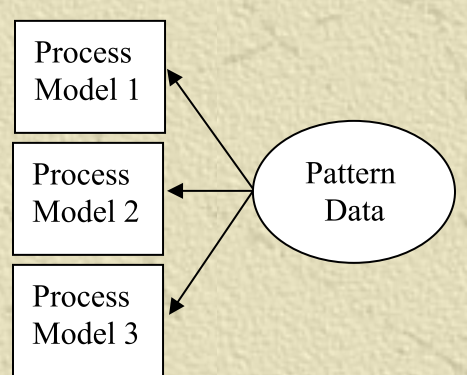
Non-Stationarity in Predicting

- Means that some aspect of the process *structure* or *parameters* are not constant over time or space.
- Models assume some degree of stationarity
  - Strong stationarity – “business as usual”
  - Controlled stationarity – scenarios
  - Weaker stationarity – adaptive, learning agents
- Even if model predicts well over one period, a non-stationary process (e.g housing market collapse) means that we know little about how it will do the next period.

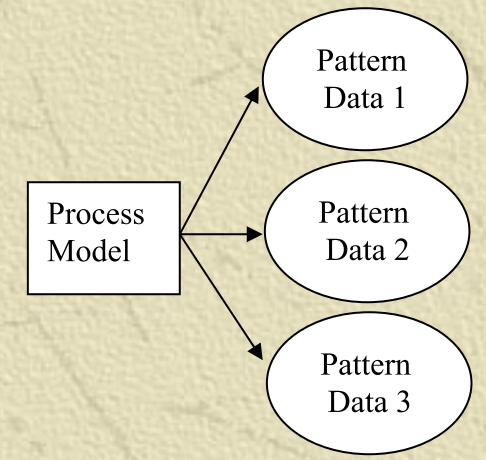
Explaining Pattern-Process Relationships

- Two other problems make explanations difficult.

Equi-finality: The Inference Problem



Multi-finality: The Predictability Problem



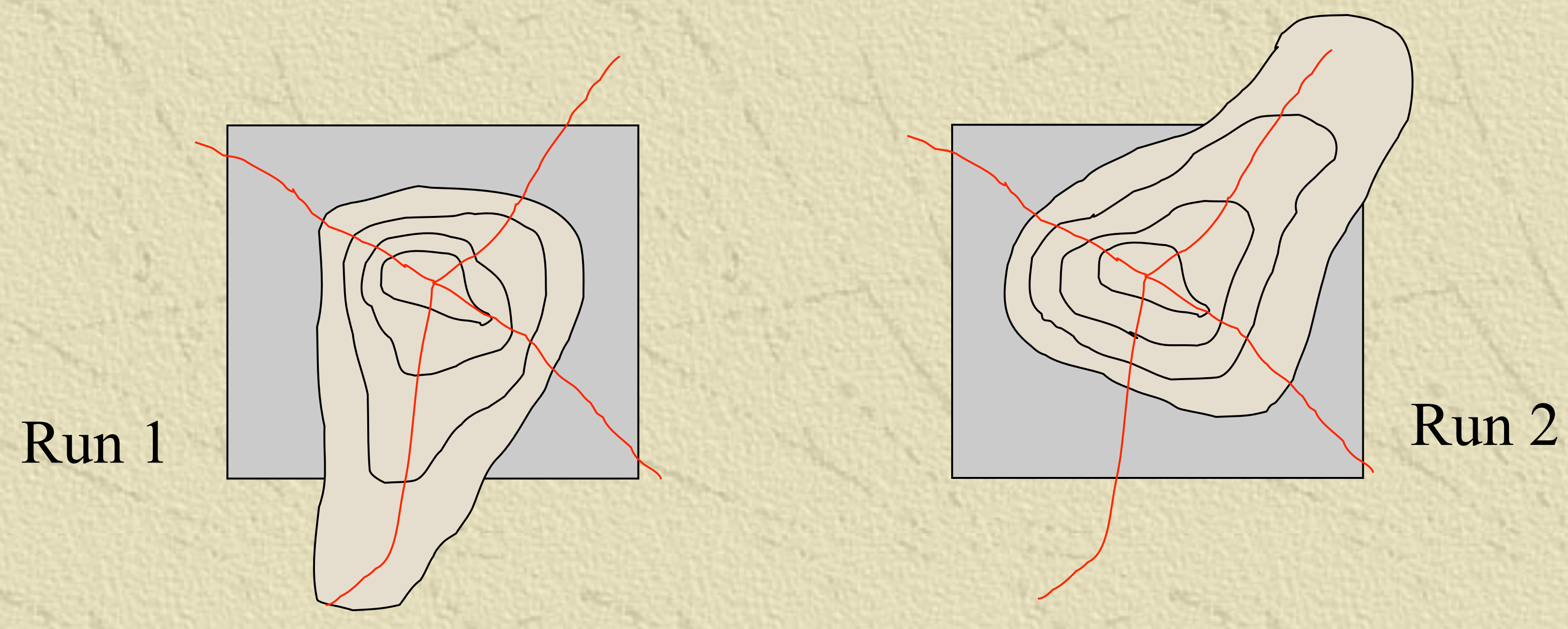
Unpredictability affected by…

- Stochastic processes
  - Processes in a model represented with random elements
- Path dependence
  - Feedbacks reinforce early steps in pattern evolution, potentially locking in certain patterns and making others impossible.

- With path dependence, uncertainty associated with stochastic processes can be magnified.
- Alternatively, path dependence can reign in stochastic processes, reducing the total range of possible outcomes.

How path dependence affects pattern

- In a path dependent system, small perturbations at early iterations can lead to big differences in outcomes.



Descriptions of Model Accuracy

- Predictive (Pattern) Accuracy
  - Aggregate-pattern validation
  - Spatial validation
- Process Accuracy
  - Structural validation
    - requires a structural model

Aggregate Pattern Validation

- Involves comparing distribution of aggregated measures of pattern across realizations with the observed pattern value at a given time.
- Example: spatial pattern metrics
  - Largest Patch Index (LPI)
  - Mean Patch Size (MPS)
  - Edge Density
  - Mean Nearest Neighbor (MNN)
- Only requires that patterns, not locations, are correct.

Spatial Pattern Validation

- Involves comparing model output on a location-by-location basis with a reference map at a given time.
- Example: Error Matrix, Kappa and variants

		Class in Reference Map		
		Developed	Undev	
	Developed	37	2	39
Class Predicted	Undev	3	48	51
by Model		40	50	90
			total correct =	85
			percent correctly classified (PCC) =	94.4

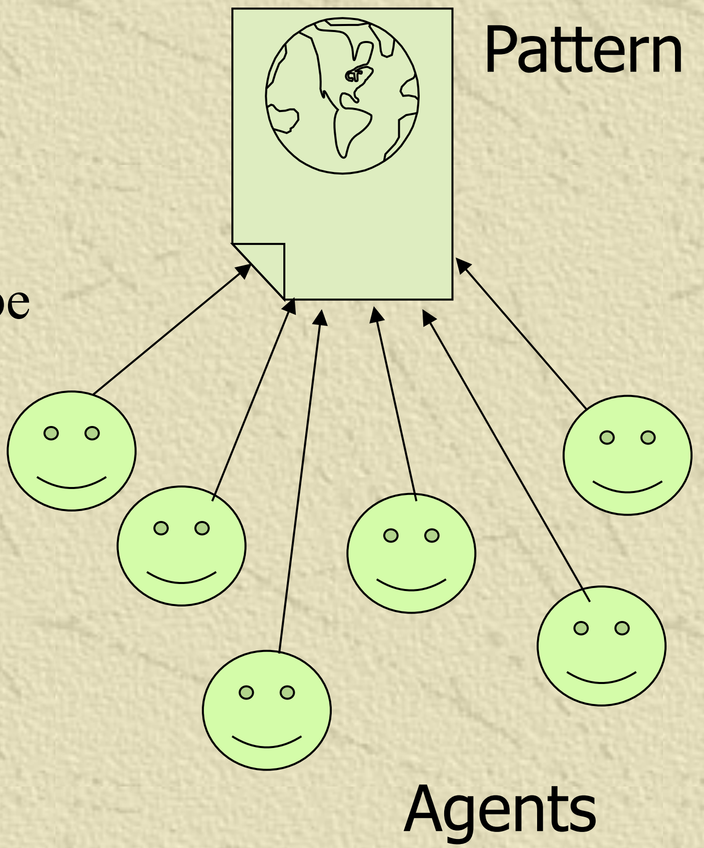
- Requires that our explanations of land change be specific enough to allow location-specific predictions.

Approaches to Structural Validation

- Statistical methods (e.g., econometric panel models) to estimate unbiased structural parameters
- Explicitly testing the influence of different model assumptions on its performance
- Evaluation of predictive accuracy for multiple, unrelated aspects of the system (e.g., land use map and income distribution)
  - Pattern-oriented modeling (Grimm et al. 2005)
- Explore structure in outcomes to understand
  - parameter sensitivity over time (Ligmann-Zielinska and Sun 2010)
  - range of of outcomes across multiple runs of a model (Brown et al. 2005).

Empirical Data in ABMs

- Agent actions aggregate to produce patterns.
- Need data on:
  - agents to support building the model.
  - aggregate patterns that can then be reserved for validation.
- Contrasts with approaches that perform calibration within the model based on observed patterns.



Empirical Support of Agents

- Social surveys
- Participant observation
- Field/lab experiments
- Companion (participatory) modeling
- Spatial/statistical inference

Robinson et al. 2007. *Journal of Land Use Science.*

Management Regimes

- We interviewed 25 exurban residents to ascertain their land management behaviors on several dimensions, e.g.,
  - Frequency of mowing, pruning
  - Fate of leaves
  - Irrigation and fertilizer
- Based on analysis of these data, we identified five management regimes that reflect observed variations. Choice of these is related to subdivision, neighbors, preference.
  - Conservationists
  - Watering Conservationists
  - Neatnik
  - Waterer
  - Infrequent Waterer

Nassauer et al., MS in Preparation

Neighbor Effects

Web-based survey (2005) asked ~500 residents preferences for

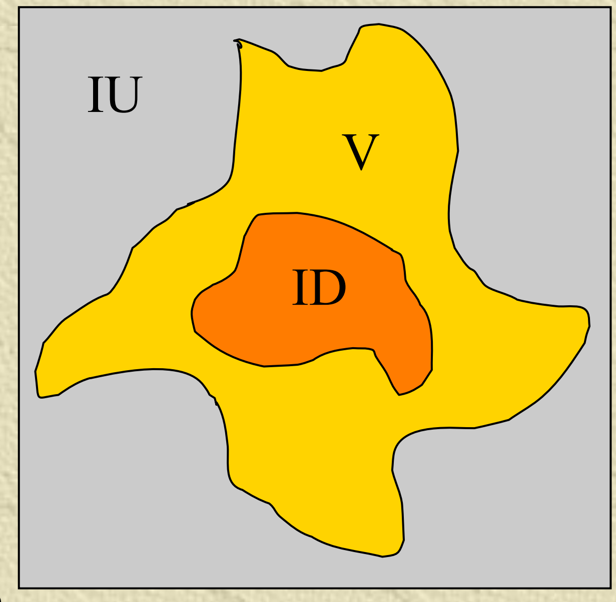
- Regional open-space availability and type
- Neighborhood open-space
- Yard-scale designs

- Results show strong *neighbor effect;* residents select yard designs based on what they see neighbors doing.

Details: Nassauer et al. 2009. *Landscape and Urban Planning*

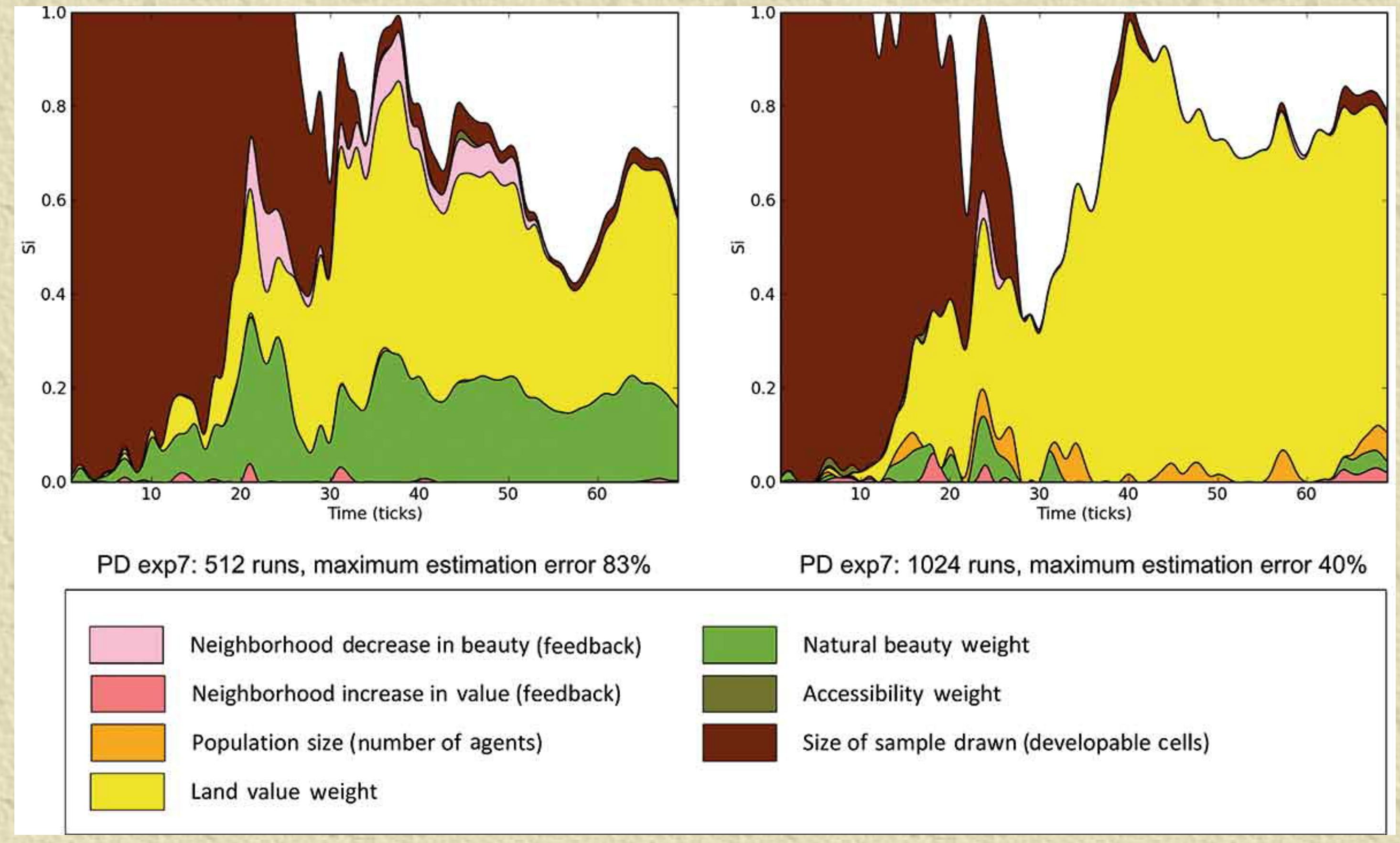
Variant-Invariant Approach

- An alternative approach to spatial validation.
- Identify the area that all (or most) of the runs predict the same (*the invariant region*)
- Identify the areas in which model outcomes vary from run to run (*the variant region*)
- Calculate size of V
- Compare accuracy with that expected (i.e., random) across entire map and within variant region



V – variant
ID – Invariant developed
IU – Invariant undeveloped

Time Dependent Global Sensitivity Analysis



Ligmann-Zielinska and Sun 2010. *IJGIS*.

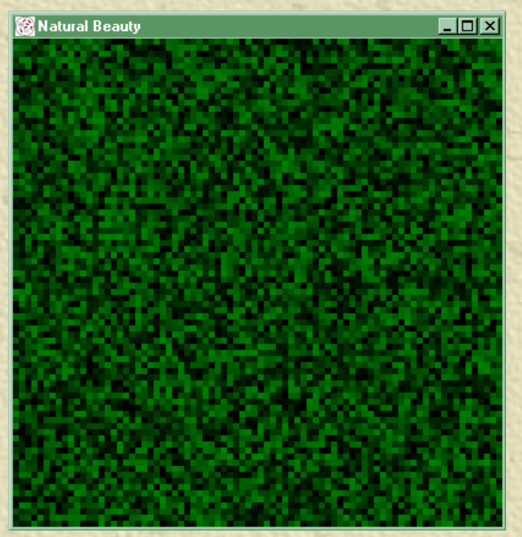
Agent-Based Models

ABMs are simulation models that can incorporate several **important processes** (i.e., they describe the *structure* of the system):

- heterogeneity among the actors
- interaction among actors at different scales (townships, developers, residents)
- feedback effects between land-use decisions and environmental characteristics
- measures of aggregate patterns of landscape change as a result of dynamic interactions of multiple actors.

SLUCE Agent-Based Model

- Implemented using objective-C and Swarm (www.swarm.org)
- Environment
  - Represented by lattice; each cell has
    - a value of aesthetic quality (q), assigned randomly or based on defined pattern
    - score for distance to service centers (sd), based on the sum of inverse distances to the nearest 8, updated at each step



Random pattern of Aesthetic Quality

ABM: Agents

- Each cell in lattice can accommodate many or only one resident or service center
- *residents*
  - have magnitudes of preference for
    - aesthetic quality (α_q)
    - nearness to SCs (α_{sd})
    - neighborhood density (α_{nd})
- *service centers* (SC)
  - initial SC located in center of map
  - one new SC created near location of each 100th resident - SCs follow residents (positive fb).

Red cells are service centers

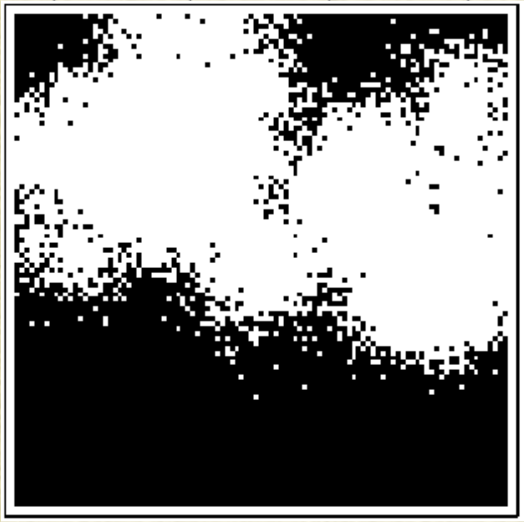
Experimental Approach

- Run the model 30 times for a given set of parameter settings.
- Select one run of the model as the reference map (i.e., to represent the "truth").
- Calculate aggregate and spatial validation statistics.
- Any differences area features of the process the model represents, rather than flaws.

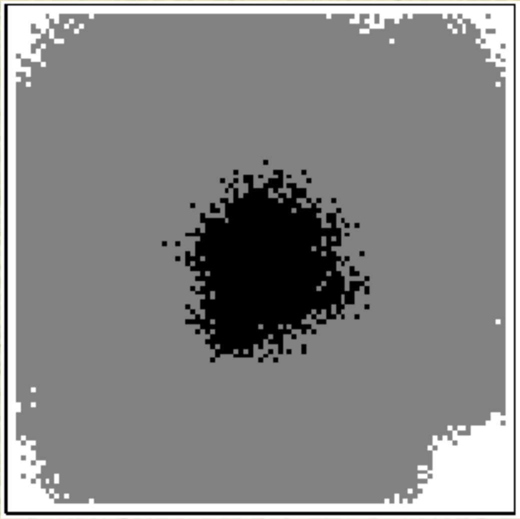
Reference: Brown, D.G., S. Page, R. Riolo, M. Zellner, and W. Rand. 2005. *International Journal of Geographical Information Science*, 19(2): 153-174.

Base Case – Distance only

- Model has large stochastic component.
- Large V.
- Reproduces pattern statistics well.
- Overall accuracy of prediction, and especially that within V, not much better than random (1.13 and 1.05, respectively).



Ref Map



V-IV Map

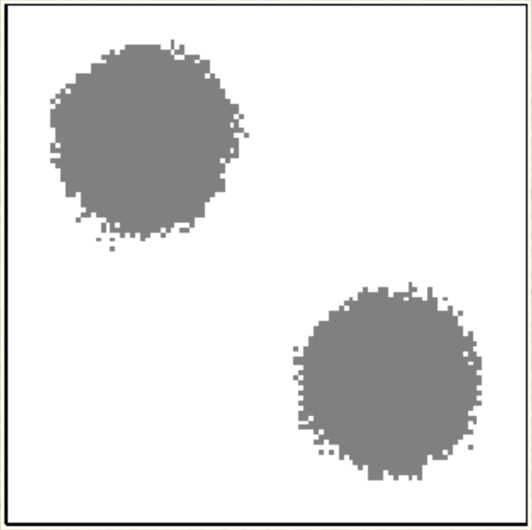
Hypothesis 3 and 4

- 3: Strong path dependence *can* lead to unpredictability.
- 4: It is possible to create a model that fits the reference map better than does the model that produced it.

Extreme Environmental Effect

- Two paths with prediction accuracy worse than random.
- Reference map has development following one of the paths.
- Overall accuracy is low and V large with low average accuracy.

Ref Map

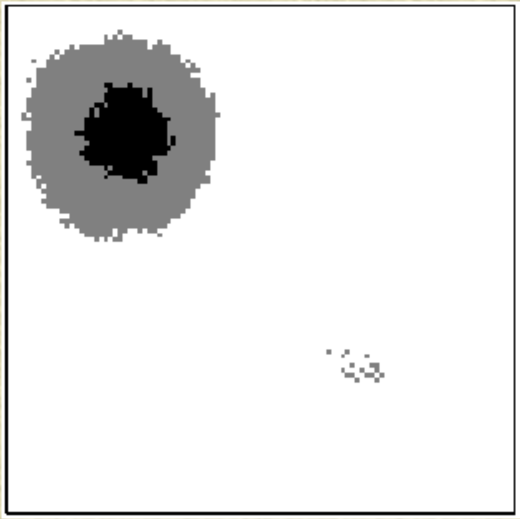


V-IV Map

Overfit Model

- Added preference for nearness to "lake," on left side of map.
- By "calibrating" the model to fit the reference pattern increased IV and its accuracy.
- Accuracy is nearly double the model that actually produced the reference map.

Ref Map



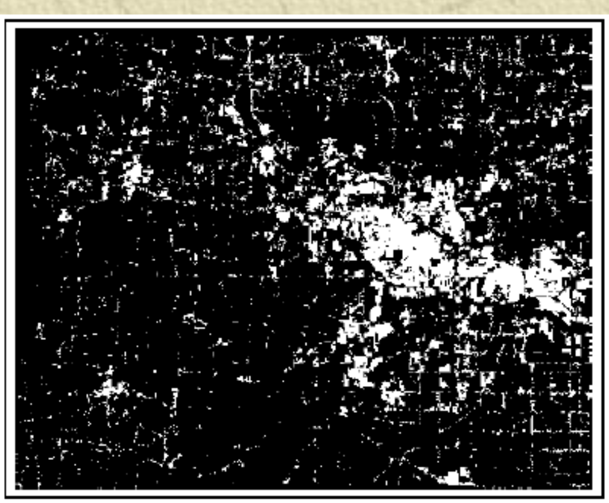
V-IV Map

Results III

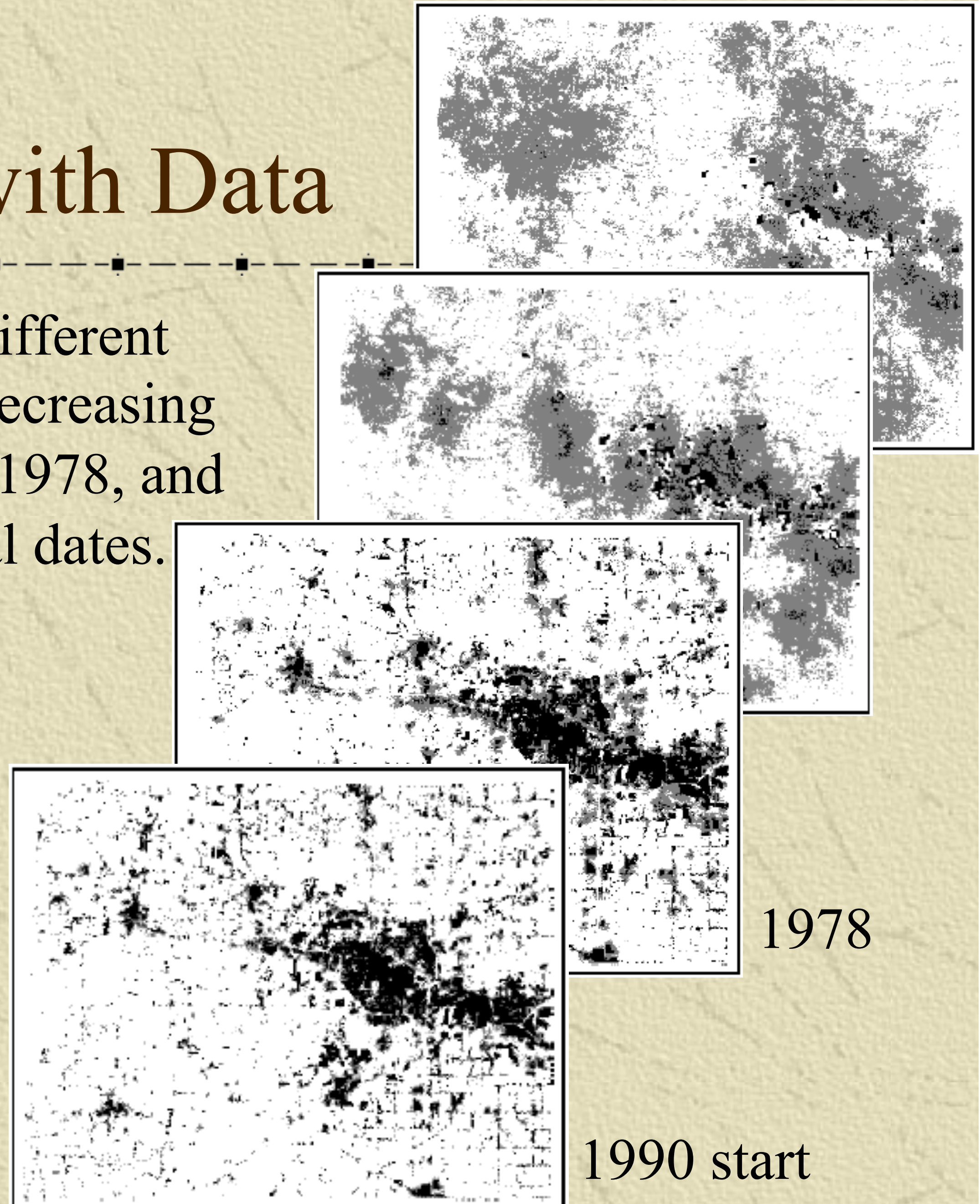
- Where path dependence is strong, a single realization (e.g., reality) may be insufficient to describe the possible outcomes.
- The model with the best prediction accuracy is not necessarily the model with the most accurate description of the process.
- By calibrating models to data we may be missing important opportunities to evaluate novel mechanisms to control patterns.

Model Runs with Data

- Start model at four different dates, representing decreasing information – 1990, 1978, and two psuedo-historical dates.



1995 Reference Map



1978

1990 start

Questions

- How might we know when the process itself is unpredictable, thereby placing limits on model predictability and validation?
- Are there other approaches to structural validation?

Location Decision

- A number of residents is created during each discrete time step (based on growth rate). Initially set to 10 per step.
- Residents locate by:
  - selecting a # of cells (*numtests*) randomly – initially set to 15
  - moving to the cell that provides the highest utility - based on utility model.

$$u_{xy} = q_{xy}^{\alpha_q} \times sd_{xy}^{\alpha_{sd}} \times \left|1 - \left(\beta_{nd} - nd_{xy}\right)\right|^{\alpha_{nd}}$$

Black cells are residents

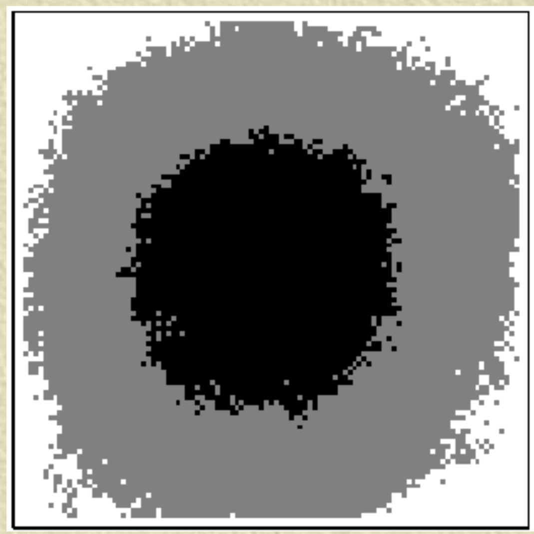
Hypothesis 1 and 2

- 1: More and stronger feedbacks will lead to more path dependence.
- 2: Spatial variability in the environment can both strength path dependence and reduce the number of possible paths.

Base + Agents want Density

- Including neighborhood density reduces V and improves accuracy overall and within V (1.46 and 1.11, respectively).
- All pattern metric values from the reference map were within distribution from model.

Ref Map

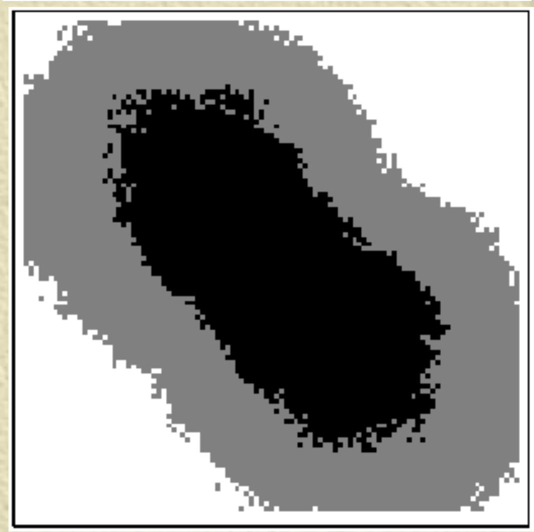


V-IV Map

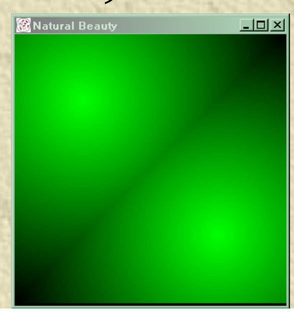
Base + Density + Environron.

- Variable environment further reduces V (i.e., number of available paths) further.
- Environment provides basis for more accurate predictions, overall and in V (1.55 and 1.13, respectively).

Ref Map



V-IV Map



Aesthetic Quality Map

Results 1 and 2

- Spatial pattern of reference map was always within the distribution of patterns in model.
- Density feedback and environmental variability both increased the degree of predictability of the model.
  - Environmental variability reduced the number of possible paths.