GISTAM 2017
3rd International Conference on Geographical Information Systems Theory, Applications and Management
Porto, Portugal
27 - 28 April, 2017

Improving SLEUTH Calibration with a Genetic Algorithm

Keith C. Clarke
Dept. of Geography, University of California, Santa Barbara,
Santa Barbara CA 93106-4060 USA
kclarke@geog.ucsb.edu

UCSB geography

What is SLEUTH?

- Cellular automaton model of urban growth and land use change
- Over 100 applications documented in the LUC literature, 20 years of discussion forum, many YouTube How-To's
- Model supported by USGS, EPA, NSF and others over 4 versions since 1998
- Key publications have 2300 citations according to Google Scholar

A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area KC Clarke, S Hoppen, L Gaydos Environment and planning B: Planning and design 24 (2), 247-261	1332	1997
Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore KC Clarke, LJ Gaydos International journal of geographical information science 12 (7), 699-714	1034	1998

The 4 review papers

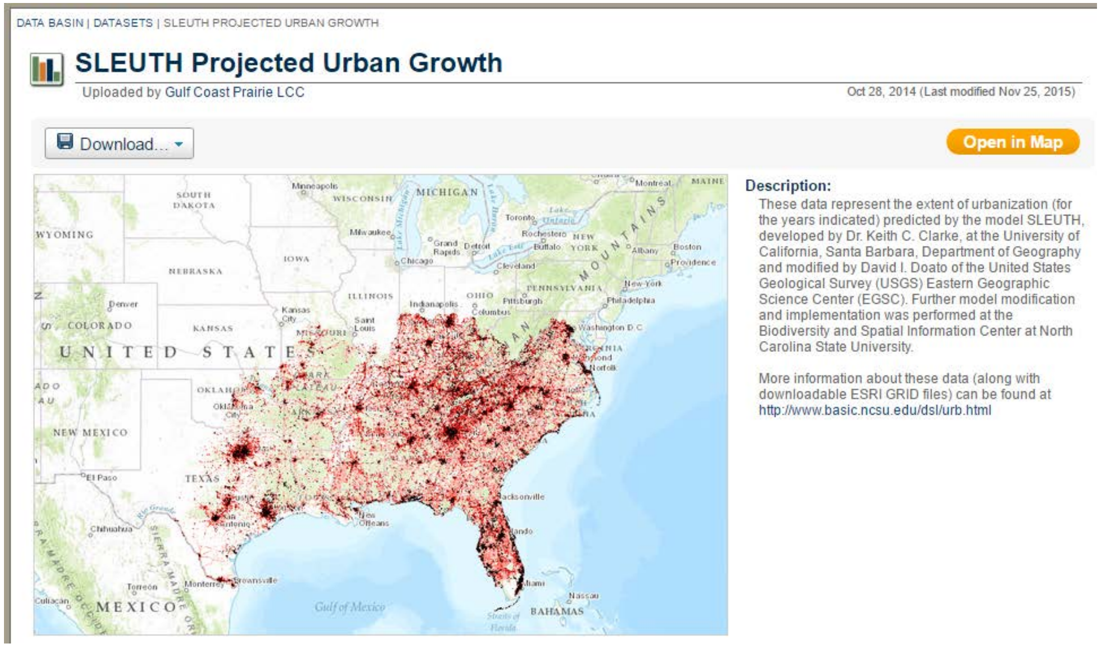
- Chaudhuri, G. and Clarke, K. C. (2013) The SLEUTH Land Use Change Model: A Review. *International Journal of Environmental Resources Research,* 1, 1, 88-104.
- Clarke, K.C. (2008) *Mapping and Modelling Land Use Change: an Application of the SLEUTH Model*, In Landscape Analysis and Visualisation: Spatial Models for Natural Resource Management and Planning, (Eds. Pettit, C., Cartwright, W., Bishop, I., Lowell, K., Pullar, D. and Duncan, D.), Springer, Berlin, pp 353-366.
- Clarke, K.C. (2008) *A Decade of Cellular Urban Modeling with SLEUTH: Unresolved Issues and Problems*, Ch. 3 in Planning Support Systems for Cities and Regions (Ed. Brail, R. K.) ,Lincoln Institute of Land Policy, Cambridge, MA, pp 47-60.
- Clarke, K. C, Gazulis, N, Dietzel, C. K. and Goldstein, N. C. (2007) *A decade of SLEUTHing: Lessons learned from applications of a cellular automaton land use change model*. Chapter 16 in Fisher, P. (ed) Classics from IJGIS. Twenty Years of the International Journal of Geographical Information Systems and Science. Taylor and Francis, CRC. Boca Raton, FL. pp. 413-425.

Some Issues with SLEUTH

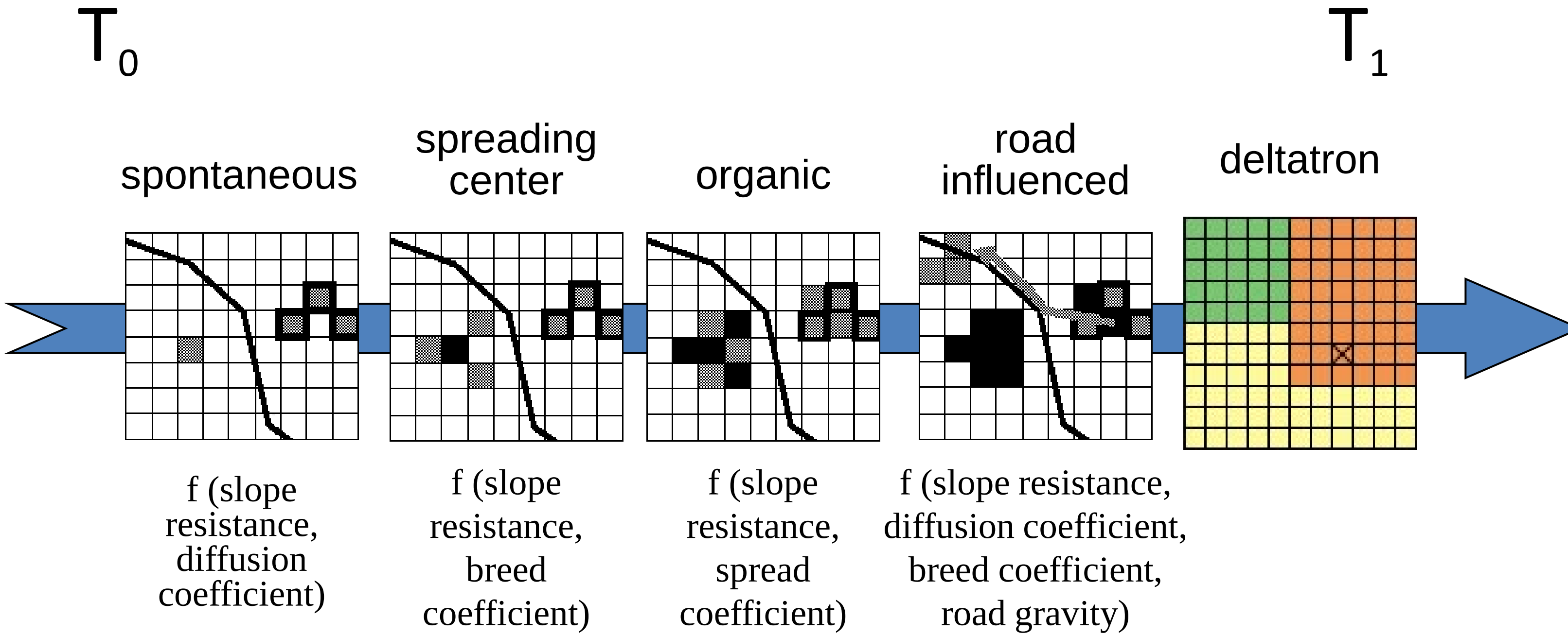
- Requires brute force calibration using hindcasting
- Expect non-linear feedbacks, no linearity
- Some applications have taken 6 months of CPU time
- First used MPI and parallel processing (3.0) on supercomputers
- Many time saving adaptations, e.g. SLEUTH-R
- Problem is that 5 behavior parameters with integer values 0-100 need to be tested = 101^5 combinations

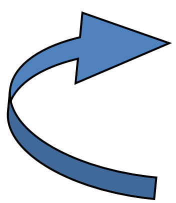
Behavior parameters

- Diffusion
- Breed
- Spread
- Slope resistance
- Road gravity



Behavior Rules

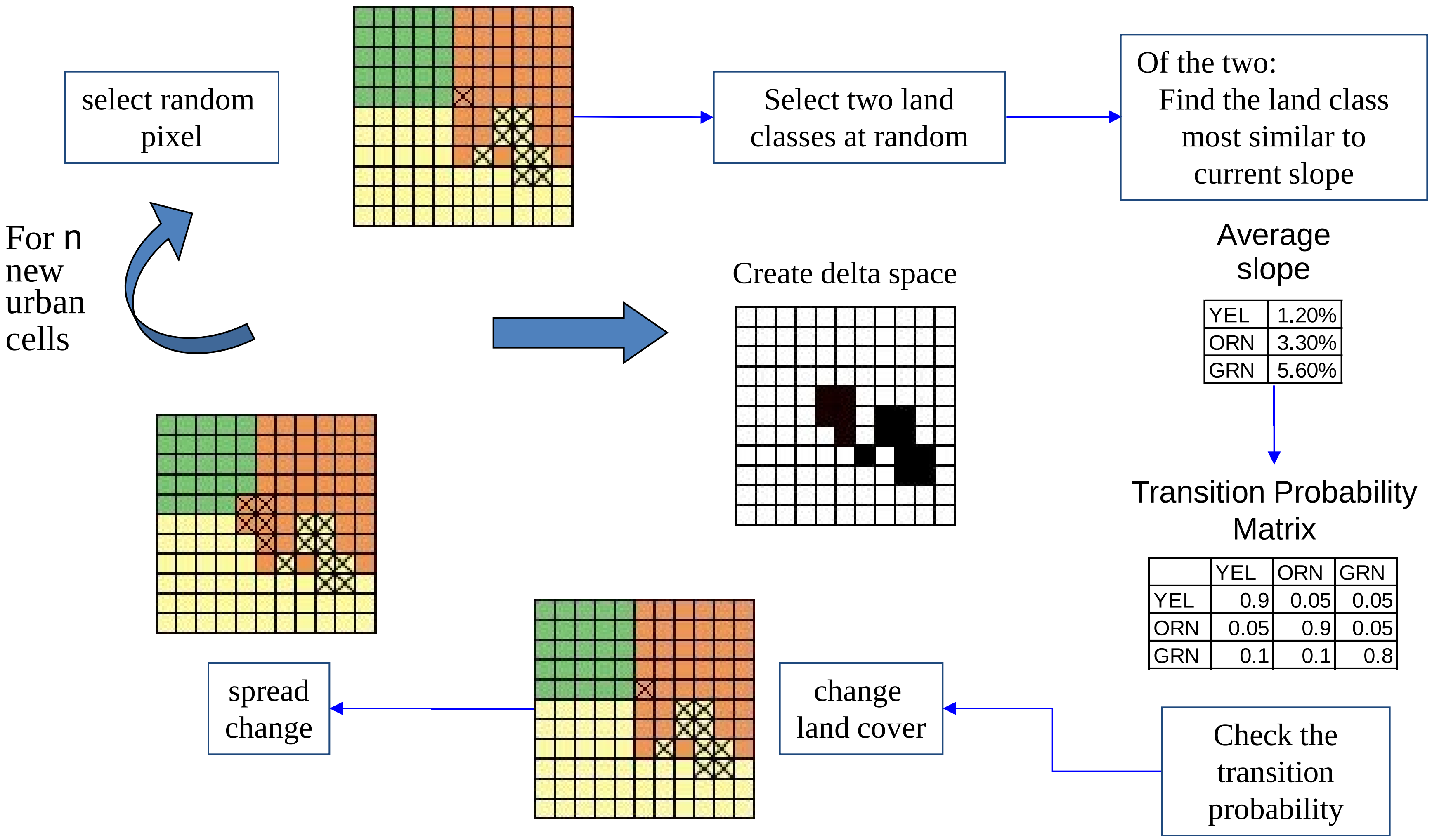




For i time periods (years)

Deltatron Land Cover Model

Phase 1: Create change



Average slope

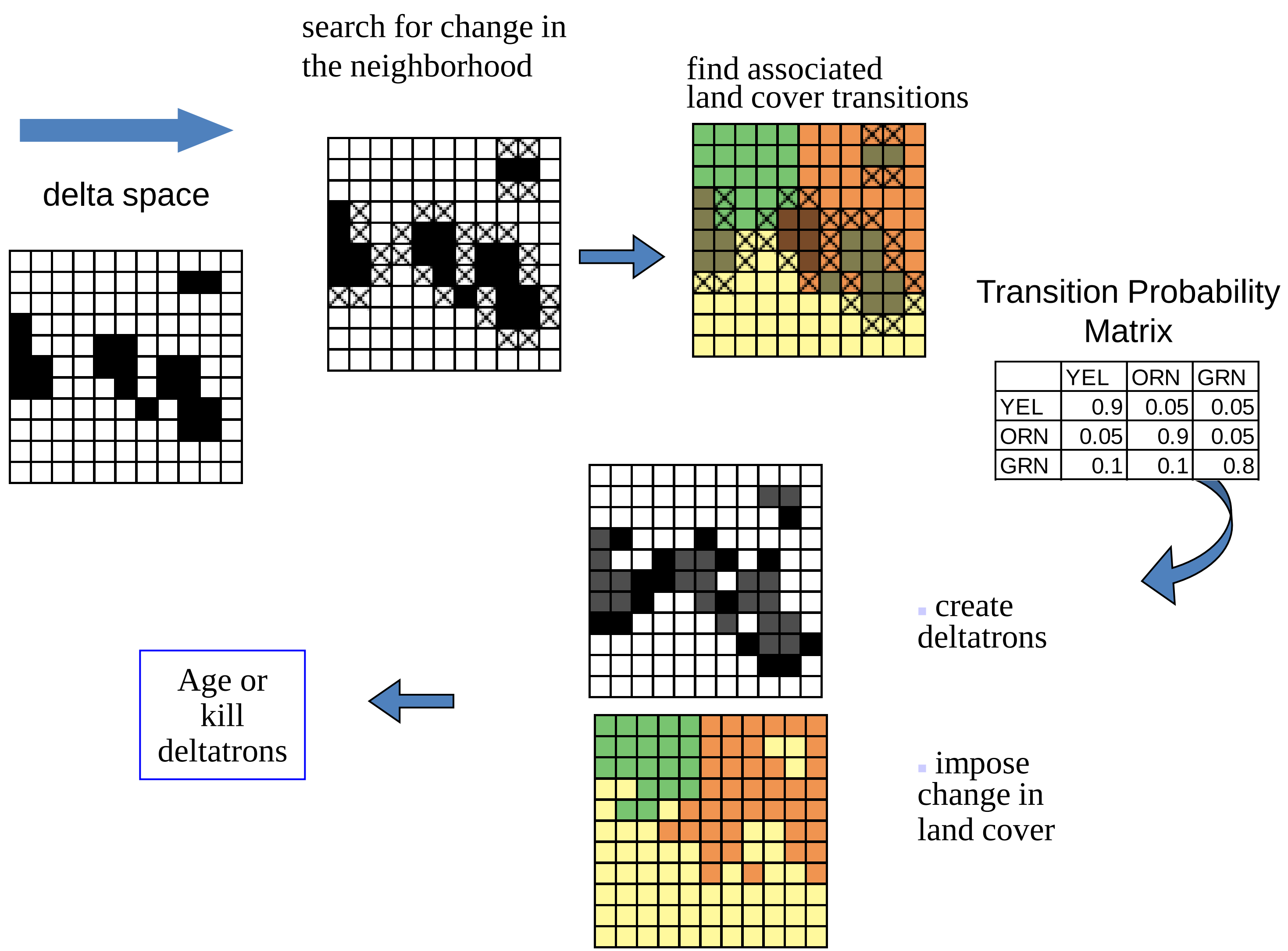
YEL	1.20%
ORN	3.30%
GRN	5.60%

Transition Probability Matrix

	YEL	ORN	GRN
YEL	0.9	0.05	0.05
ORN	0.05	0.9	0.05
GRN	0.1	0.1	0.8

Deltatron Land Cover Model

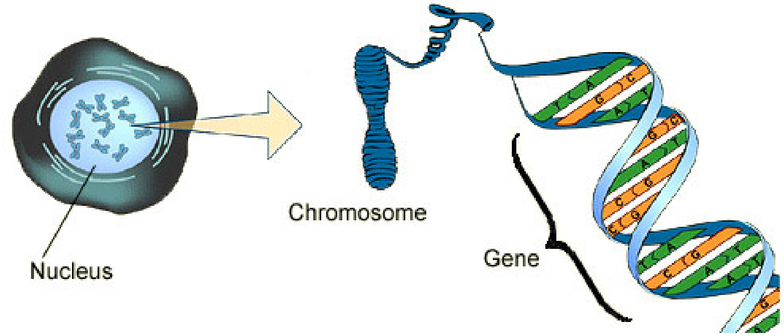
Phase 2: Perpetuate change



Transition Probability Matrix

	YEL	ORN	GRN
YEL	0.9	0.05	0.05
ORN	0.05	0.9	0.05
GRN	0.1	0.1	0.8

A new solution



- Many new machine learning algorithms can be used for approximating optimization
- Replaced automated calibration routine with a genetic algorithm
- Start with a random set of genes within a five parameter multi-group chromosome
- Use evolution (combination, cross-over, mutation, replacement, competition)
- Evolve until fitness metric (Optimal SLEUTH metric) no longer increases for the best gene and the chromosome

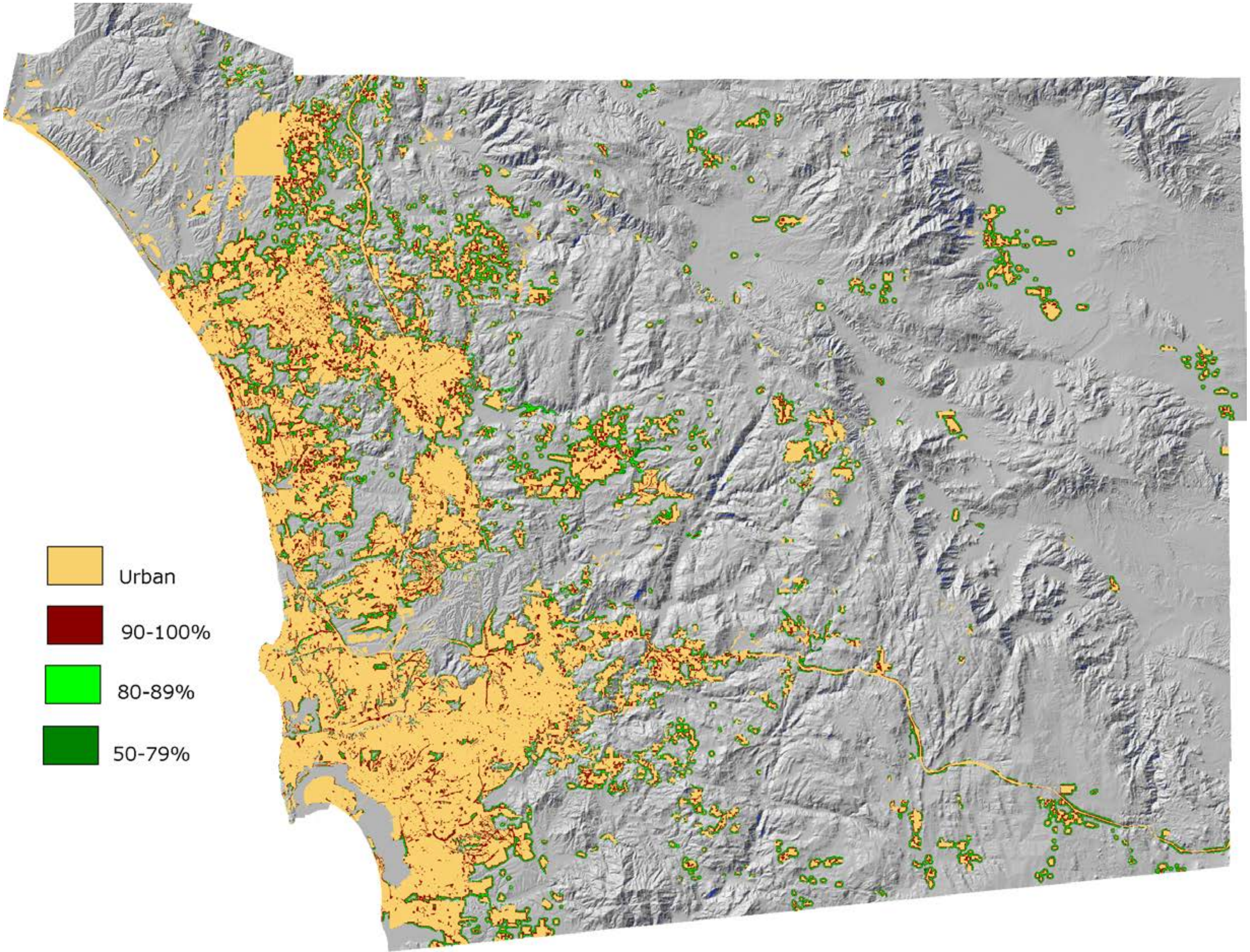
But, new control constants!

- Number of genes in the chromosome (Population)
- Number of offspring per generation
- Mutation rate
- Maximum number of generations
- Proportion of genes to be replaced per generation
- Maximum number of gene replacements

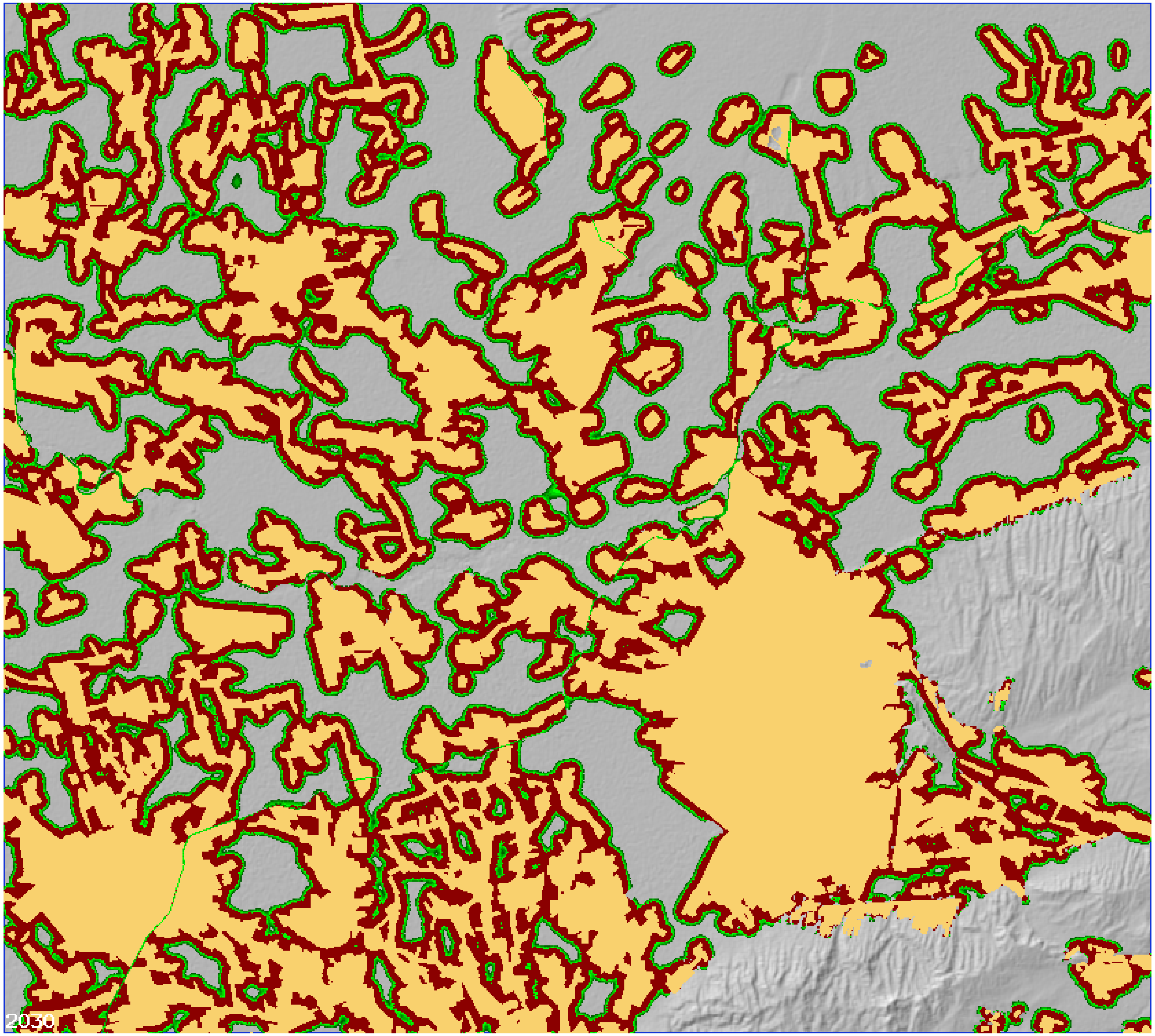
The Current Work

- Implement SLEUTH-GA (and post to web site)
- Choose best (San Diego, USA) and worst (Andijan, Uzbekistan) calibrations to date
- Calibrate using Brute Force
- Calibrate using GA, while monolooping through the 6 control parameters
- Determine best parameters and build into model as defaults

San Diego



Andijan



Results: Brute Force

Table 1: Brute Force Calibration Results. Values for constants are after calibration, with high and low coefficients in the top 8 solutions given, then after averaging to the last time period.

	San Diego	Andijan
Calibration period	1960-1999	1934-2013
Best OSM	0.7414836	0.0773797
Diffusion/derived	(100:98-100) 100	(63:60-63) 100
Breed/ derived	(97:97-99) 100	(100:85-100) 100
Spread/ derived	(25:24-25) 25	(1:1-2) 3
Slope/derived	(15:15-18) 1	(80:75-79) 1
Road gravity/ derived	(53:45-53) 53	(25:15-25) 38
Calibration time (s)	175589	440715

122.4 hours

Results: Best Gene

Table 3: Genetic Algorithm Calibration Results

	San Diego	Andijan
Calibration period	1960-1999	1934-2013
Best OSM	0.729724	0.072920
Diffusion/ derived	(90: 79-90) 100	(54:53-94) 82
Breed/ derived	(23: 22-25) 26	(2:0-2) 3
Spread/ derived	(89:74-98) 100	(88:62-94) 82
Slope/derived	(13:2-32) 1	(70:9-70) 3
Road gravity/ derived	(19:19-98) 30	(47:14-47) 75
Calibration time (s)	55588	19866
Speed Up	3.16	22.18

5.52 hours

Best GA Parameters: 10-12 generations

Table 2: Genetic Algorithm Parameter Monolooping Calibration Results

City	San Diego	Andijan
Calibration period	1960-1999	1934-2013
Best OSM	0.72972	0.07292
Maximum # of evaluations	900	900
Population (Chromosome size)	55	55
Mutation Rate	0.13	0.13
Number of offspring	55	55
Replacement per generation	50	50
Calibration time (s)	55588	19866

Conclusion

- GA at least equal, and often superior calibration results
- Seems to lower modeling uncertainty
- Differences in the calibration parameter sets and forecasts are small
- CPU time for calibration was reduced by about a factor of 3 for San Diego and 22 for Andijan
- GA can provide a convergent set of genes that can be further optimized by a narrower brute force such as the range over the top 8 genes
- Code now available on SLEUTH website