CHAPTER 17

Spatio-Temporal Difference in Model Outputs and Parameter Space as Determined by Calibration Extent

C.K. Dietzel

CONTENTS

17.1 INTRODUCTION

During recent years, models of land use change and urban growth have drawn considerable interest. Despite past failures in urban modeling (Lee, 1973, 1994), there has been a renaissance of spatial modeling in the last two decades due to increased computing power, increased availability of spatial data, and the need for innovative planning tools for decision support (Brail and Klosterman, 2001; Geertman and Stillwell, 2002). Spatial modeling has become an important tool for city planners, economists, ecologists and resource managers oriented towards sustainable development of regions, and studies have attempted inventories and comparisons of these models (Agarwal et al., 2000; EPA, 2000). These new models have shown potential

0-8493-2837-3/01/$0.00+$.50

in representing and simulating the complexity of dynamic urban processes and can provide an additional level of knowledge and understanding of spatial and temporal change. Furthermore, the models have been used to anticipate and forecast future changes or trends of development, to describe and assess impacts of future development, and to explore the potential impacts of different policies (Pettit et al., 2002; Verburg et al., 2002). Because many of these models are being used to provide information from which policy and management decisions are made, it is important that modelers have a clear understanding of how the geographic extent at which they are calibrating and modeling influences the forecasts that their models produce. This is directly linked to a larger geographic issue in modeling. Can large-scale (geographic extent) models accurately forecast local growth compared with smaller-scale applications, or should state/nation/global modeling be done at a local level and then aggregated to create a more realistic view?

The concept of geographic extent and how changing it alters a model's parameter space, and subsequently model outputs and forecasts is not something that has been studied extensively. In this work, extent is defined as the geographic boundary of a system. Generally speaking, when the extent of a geographic system is changed, so should the statistics that describe the system, and the interactions that take place within that system. Only in the case of a repeating pattern with small-scale structure, such as a checkerboard or white noise, is this not true. In the case of urban models, the effects of geographic extent on model calibration and outputs may be overlooked or brushed aside due to two constraints that inhibit this type of modeling in general: (1) many times researchers struggle to get the necessary data to run the models, at any spatial extent; and (2) as the spatial extent of data gets larger, the computational time increases, sometimes in more of an exponential manner than a linear one. These two issues have prohibited urban modelers from addressing sufficiently the issue of geographic extent and how it relates to urban model output, but even more importantly, how calibration at different extents can impact model forecasting and final outputs.

AQ: Straatman et al., 2003 is not in ref list; is this 2004 as in ref. list, please check

With any model, there is an explicit need to calibrate and parameterize the model to fit the dataset. Recently, modelers have increased their focus on the calibration phase of modeling to gain a increased understanding of how models, in particular cellular automata, work (Abraham and Hunt, 2000; Li and Yeh, 2001a; Silva and Clarke, 2002; Wu, 2002; Straatman et al., 2003). The calibration phase of modeling is one, if not the most important, stage in modeling because it allows the fitting of model parameters to the input data, to be further used in forecasting. Failure to calibrate a model to the input data results in an application that is not robust or justifiable. While these efforts have focused on refining the calibration process and the definition of parameters for these models, none of them have focused on how calibration at different spatial resolutions changes the parameter set and the model outputs. For all models, the "best" parameter space is defined as the area or volume of parameters that is searched to find the parameter set. The parameter set is then the "best-fit" set of parameters that describe the behavior of the system within the framework of the model. The parameter space is defined as an area or volume depending on the number of individual parameters. If there are only two parameters, then the parameter space in an area; any more than two then it is a n-dimensional volume. A better understanding of how spatial resolution changes a model's parameter set, and hence its outputs, is an important area of research, especially when many of these models are used in the decision-making process.

Should modelers take into account that an urban area may be a transition zone between two metropolitan areas or is influenced by a larger region in the calibration process? And how does incorporating the influence of these areas change the parameter space, and hence the spatial output of the model? Inclusion of outside influential areas into local urban models is not a new idea (Haake, 1972), but the study of how their inclusion changes the parameter space of current models may be. Advances in computing, especially the advent of parallelization and the cost effective strategy of "clusters" have significantly deflated the geocomputational cost of modeling larger spatial areas at fine spatial resolutions, so inclusion of possibly influential, but outside, areas is not as much of a taxing task as it once was. Capitalizing on these advances, this research focuses on the relationship between spatial extent and parameter space, and how calibration of an urban cellular automata model at varied spatial extent can allow for forecasts that are more typical of the local–regional interactions taking place.

AQ: Haake, 1972 is not in ref list please check

The SLEUTH urban model is a cellular automaton model that has been widely applied (Esnard and Yang, 2002; Yang and Lo, 2003; also refer Chapter 16 by Clarke, and Chapter 18 by Goldstein) and has shown its robust capabilities for simulation and forecasting of landscape changes (Clarke et al., 1997). The model makes use of several different data layers for parameterization, for example, multi-temporal land use and urban extent data, transportation network routes and digital elevation model data. Application of the model necessitates a complex calibration process to train the model for spatial and temporal urban growth (Silva and Clarke, 2002). This chapter documents work done on the role that geographic extent plays in the calibration of urban models by working with SLEUTH. A large geographic area was calibrated and modeled at three different geographic extents: global (extent of the entire system), regional, and county. The derived parameters were then used to forecast urban growth from 2000 to 2040. The results from the model forecasts were then compared to determine the extent that calibration at different geographic extents had on model forecasts. This analysis was then used to examine some general considerations about the geographic extent over which urban models are calibrated and used, and the implications that this has for using these models to evaluate policies and scenarios.

17.2 CALIBRATING URBAN AUTOMATA

Modeling geographic systems using cellular automata (CA) models is a recent advance relative to the history of the geographic sciences (Silva and Clarke, 2002). Tobler (1979) was the first to describe these models in geography, briefly describing five land use models that were based on an array of regular sized cells, where the land use at location i, j was dependent on the land use at other locations. Applying this method of modeling to urban systems for planning applications was recognized early (Couclelis, 1985), and application of these models has proliferated in the last decade (Ward et al., 2000; Li and Yeh, 2001b; Almeida et al., 2003), including the development of SLEUTH. While the models themselves have proliferated, work on

AQ: Torrens and O'Sullivan, 2001 is not in ref list please check

AQ: Batty et al., 1999 is not in ref list please check

the calibration phase has lagged, and the need for more robust methods for calibrating and validating CA models has been noted (Torrens and O'Sullivan, 2001).

Model calibration has become an increasingly important consideration in the development phase of modeling (Batty and Xie, 1994; Landis and Zhang, 1998). Yet, the high flexibility in rule definition used in CA modeling, and application of these rules to manipulate cell states in a gridded world, makes parameter estimation a more difficult process (Wu, 2002). For CA models where the transition rules consist of equations for calculating future state variables, they generally consist of several linked equations for each land use, and these are complexly linked, so calibrating a model may require the fitting of tens, if not hundreds of parameters (Straatman et al., 2004). The general difficulty in finding the "golden set" of parameter values of CA is due to the complexity of urban development (Batty et al., 1999). Methods for calibration such as the use of off-the-shelf neural network packages have been suggested by some (Li and Yeh, 2001a), but some have argued that these sort of methods produce a "trained" model and not one that has intrinsic meaning in terms of known geographic principles (Straatman et al., 2004). Due to these difficulties and the parameters of CA models being dependent on the transition rules for the model, there has been little research on the parameter space or sets of urban cellular automata and how they are related to the geographic extent of calibration, although the work that has been done on calibration provides a starting point for looking at how the parameter space can be approached. This is in contrast to work on CA in computer science, where rules and parameters impacts on behavior have been studied exhaustively.

17.2.1 SLEUTH Calibration

Calibration of SLEUTH produces a set of five parameters (coefficients) which describe an individual growth characteristic and that when combined with other characteristics, can describe several different growth processes. For this model, the transition rules between time periods are uniform across space, and are applied in a nested set of loops. The outermost of the loops executes each growth period, while an inner loop executes growth rules for a single year. Transition rules and initial conditions of urban areas and land use at the start time are integral to the model because of how the calibration process adapts the model to the local environment. Clarke and Gaydos (1998) describe the initial condition set as the "seed" layer, from which growth and change occur one cell at a time, each cell acting independently of the others, until patterns emerge during growth and the "organism" learns more about its environment. The transition rules that are implemented involve taking a cell at random and investigating the spatial properties of that cell's neighborhood, and then urbanizing the cell, depending on probabilities influenced by other local characteristics (Clarke et al., 1997). Five coefficients (with values 0 to 100) control the behavior of the system, and are predetermined by the user at the onset of every model run (Clarke et al., 1997; Clarke and Gaydos, 1998; Candau, 2000). These parameters are:

1. *Diffusion* — Determines the overall dispersiveness nature of the outward distribution.
2. *Breed Coefficient* — The likelihood that a newly generated detached settlement will start on its own growth cycle.

3. *Spread Coefficient* — Controls how much contagion diffusion radiates from existing settlements.
4. Slope Resistance Factor — Influences the likelihood of development on steep slopes.
5. *Road Gravity Factor* — An attraction factor that draws new settlements towards and along roads.

These parameters drive the four transition rules which simulate spontaneous (of suitable slope and distance from existing centers), diffusive (new growth centers), organic (infill and edge growth), and road influenced (a function of road gravity and density) growth.

By running the model in calibration mode, a set of control parameters is refined in the sequential "brute-force" calibration phases: coarse, fine, and final calibrations (Silva and Clarke, 2002). In the coarse calibration, the input control data are resampled to one quarter of the original size (i.e., 100 m is resampled to 400 m), and then a Monte Carlo simulation of a broad range of parameters are tested for their fit in describing the input data. The results of the calibration run are then analyzed to narrow the range of tested parameters, based on metrics that describe spatial characteristics of the calibration runs against the input control data, specifically using the Lee–Sallee metric because of its "spatial matching" of the control data, although there has been some suggestion that other metrics can be used (Jantz et al., 2002). This metric is a shape index that measures the spatial fit between the model's growth and the known urban extent for the calibration control years. Upon narrowing the range of parameters based on the metrics, the original input data are resampled again, but to one half of the original size (i.e., 100 m is resampled to 200 m), and simulated over the narrowed range of parameters. Again, the results are analyzed, and the range of parameters narrowed. This final set of parameters is simulated with the full spatial resolution original data. The resultant parameters are then used to forecast urban growth. One of the drawbacks of the "brute-force" calibration is the massive amount of computational time required to calibrate the model. Clarke and Gaydos (1998) report using several hundred hours of CPU time to calibrate data for San Francisco, CA. The computational time required to calibrate the SLEUTH model has not decreased significantly with advances in computers, and has led to the search for other methods of model calibration, including the use of genetic algorithms, which have been tested (refer to Chapter 18 by Goldstein).

17.2.2 Study Area and Data

Using the San Joaquin Valley (CA) as a study area (Figure 17.1), input data for modeling urban growth using SLEUTH were compiled at 100 m spatial resolution (Table 17.1). Data sources for historical urban extent are listed in Table 17.1. Urban extent data for San Joaquin Valley for the years 1940, 1954, and 1962 were digitized from historical U.S. Geologic Survey 1:250,000 maps and based on air photo interpretation and supplemental ground survey information. Data from 1974 and later were captured directly from space-based remotely sensed imagery. The urban extent data for 1974 and 1996 were based on Landsat Multispectral

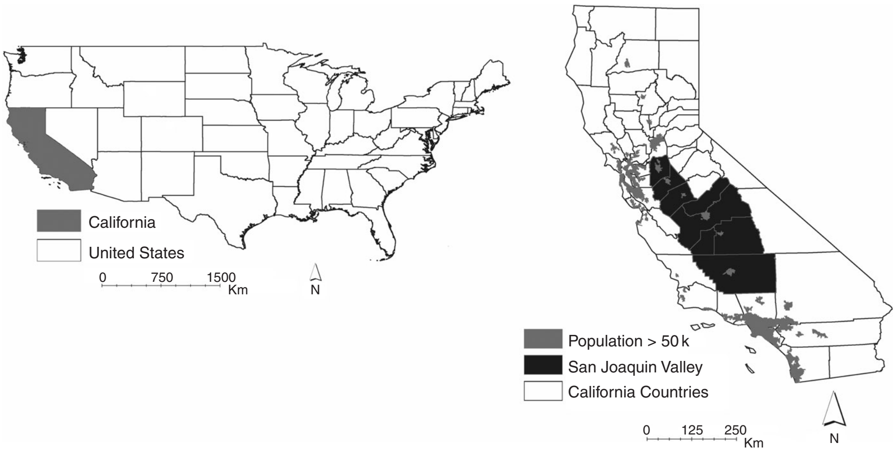

Figure 17.1 Location of California, population centers >50,000 people, and the location of the San Joaquin Valley.

Table 17.1 Sources, Description, and Resolution of Data Used in SLEUTH Modeling of the San Joaquin Valley (CA)

Data Layer	Source	Description	Spatial Resolution
Topography/Slope	USGS	30 m DEM	30 m
Land use	CA-DWR	Not used in this modeling	
Exclusion	CaSIL	Vector coverage of Federal and State owned land	N/A
Urban extent	USGS	Urban extent for 1940, 1954, 1962, 1974, 1996	100 m
	CA-FMMP	Vector coverage of developed land from 1984 to present in 2 year intervals	N/A
Transportation	CalTrans	Vector coverage of functionally classified roads from 1940 in 5 year increments	N/A

Scanner and Landsat Thematic Mapper mosaics compiled by the USGS Moffet Field, California office (http://ceres.ca.gov/calsip/cv/). Additional data for 1984, 1992, 1996, and 2000 were obtained from the California Farmland Mapping and Monitoring Program (CA-FMMP) that utilized aerial photography as a base mapping source (http://www.consrv.ca.gov/DLRP/fmmp/). The 1996 CA-FMMP data were merged with the USGS data to create a composite image of growth. Urban extent through time was treated as a cumulative phenomenon so that each time period built on the previous one, and urban extent was not allowed to disappear once it was established. Use of these data led to control years of 1940, 1954, 1962, 1974, 1984, 1992, 1994, 1996, and 2000. Pre-1974 data were all from cartographic sources, and the remainder were from satellite imagery and high definition aerial photos. All data processing was accomplished within a geographic information system (GIS) environment.

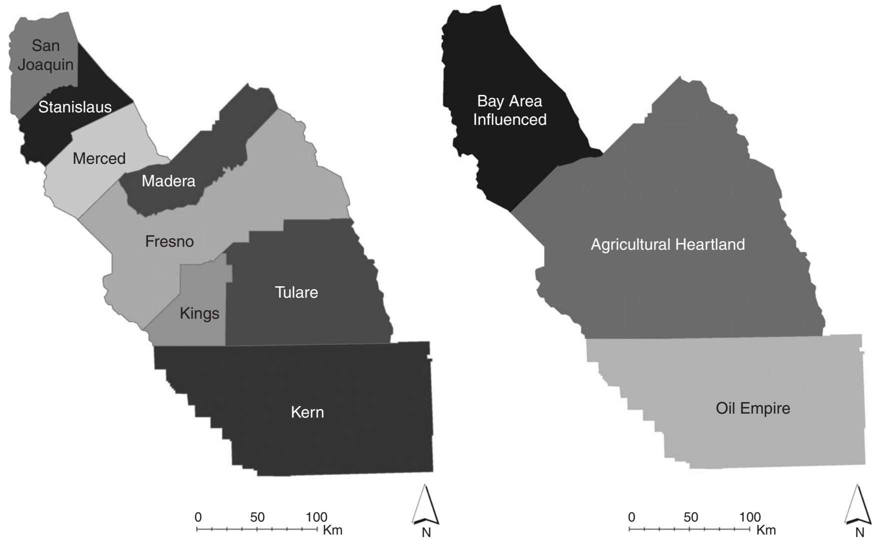

Figure 17.2 The eight independent counties in the San Joaquin Valley (left) and the three economic regions (right).

The San Joaquin Valley, while one large geographic and cultural area, is comprised of eight independent counties (Figure 17.2): San Joaquin, Stanislaus, Merced, Madera, Fresno, Kings, Tulare, and Kern. Additionally, these counties can be grouped into three distinct regions, based primarily on their economy (Figure 17.2). The Bay Area Region (San Joaquin, Stanislaus, Merced) is heavily influenced by the San Francisco Metropolitan Area economy and commuting patterns. Agriculture creates the economic base of Madera, Kern, Kings, and Tulare counties, uniting them as the Agricultural Heartland; and oil dominates Kern County which doubles as a county and region due to its unique natural resource and commuter patterns with the Los Angeles Metropolitan Area.

Input data for SLEUTH were calibrated for the San Joaquin Valley as a whole, each of the three regions, and each of the eight counties. The resulting parameter spaces were then used to forecast urban growth from 2000 to 2040, and using both tabular and graphical outputs, the results of the forecast were compared. Growth for San Joaquin, Madera, and Kern counties was then forecast using the global (San Joaquin Valley), regional, and county parameter sets to determine how that area grew specific to the others. These counties were chosen because they represent the typical historical growth trends in their respective economic region.

17.3 CALIBRATION AND FORECASTING RESULTS

The SLEUTH model was calibrated using the typical routine with self-modifying parameters values as described in Silva and Clarke (2002), and the data layers in Table 17.1. Calibration of the San Joaquin Valley and subsequent datasets resulted

Table 17.2 Resultant Growth Coefficients for the 11 Geographic Extents Calibrated using the SLEUTH Model

Area	Extent	Coefficients				
		Diffusion	Breed	Spread	Slope	Road
San Joaquin	County	2	2	54	1	3
Stanislaus	County	2	7	54	29	100
Merced	County	2	2	41	35	15
Madera	County	2	2	25	83	21
Fresno	County	2	5	58	41	52
Kings	County	2	2	45	1	2
Tulare	County	2	2	32	41	2
Kern	County/Region	2	2	58	46	31
Bay Influenced	Region	2	4	47	30	3
Agricultural Heartland	Region	2	2	45	36	41
San Joaquin Valley	Global	2	2	83	10	4

in a parameter set describing the growth of each area that was different for all areas included in this study (Table 17.2).

San Joaquin, Madera, and Kern counties were forecast using the global parameters along with their respective region and county parameters. Total urban area and new urban growth over time under each of the parameter sets were plotted for these three areas (Figure 17.3).

Total urban area from 2000 to 2040 in San Joaquin County was greater when the local county parameter set was used in forecasting compared with that of the Bay Influenced region and the global San Joaquin Valley (Figure 17.4). This is further supported by the total hectares of new urban growth for the county parameters set than the region and valley parameters (Figure 17.3), yet the difference between the three was slight over the forty year forecast. Total urban area is predicted to cover 125,000 ha under the county parameter set, 127,000 and 135,000 ha under the region and valley parameters.

Growth trends in Madera County under the three parameter set forecasts were opposite of San Joaquin County. Using the global San Joaquin Valley parameters to forecast future growth produced a county that grew three times faster than when the local county parameters were used, and twice as fast when the regional Agricultural Heartland parameters were used in forecasting (Figure 17.5). Total urban area under the county parameters was 36,000 ha, opposed to the 60,000 and 90,000 predicted using the region and valley parameters.

Kern County, a unique area that itself is a region, had growth curves that were lower than those of the global San Joaquin Valley (Figure 17.6). The number of hectares of urban growth is predicted to be 305,000 for the county and region parameters, and 395,000 under the valley parameters. This difference between the total urban area forecast by the different parameter sets was not as great as was found for Madera County, but larger than the minute differences found in San Joaquin County.

Using a GIS to overlay the model outputs, geo-algebra was done to determine the total area that was both unique and common between all parameter set forecasts for the year 2040 for each of the three areas (Figure 17.7). In concurrence with the total

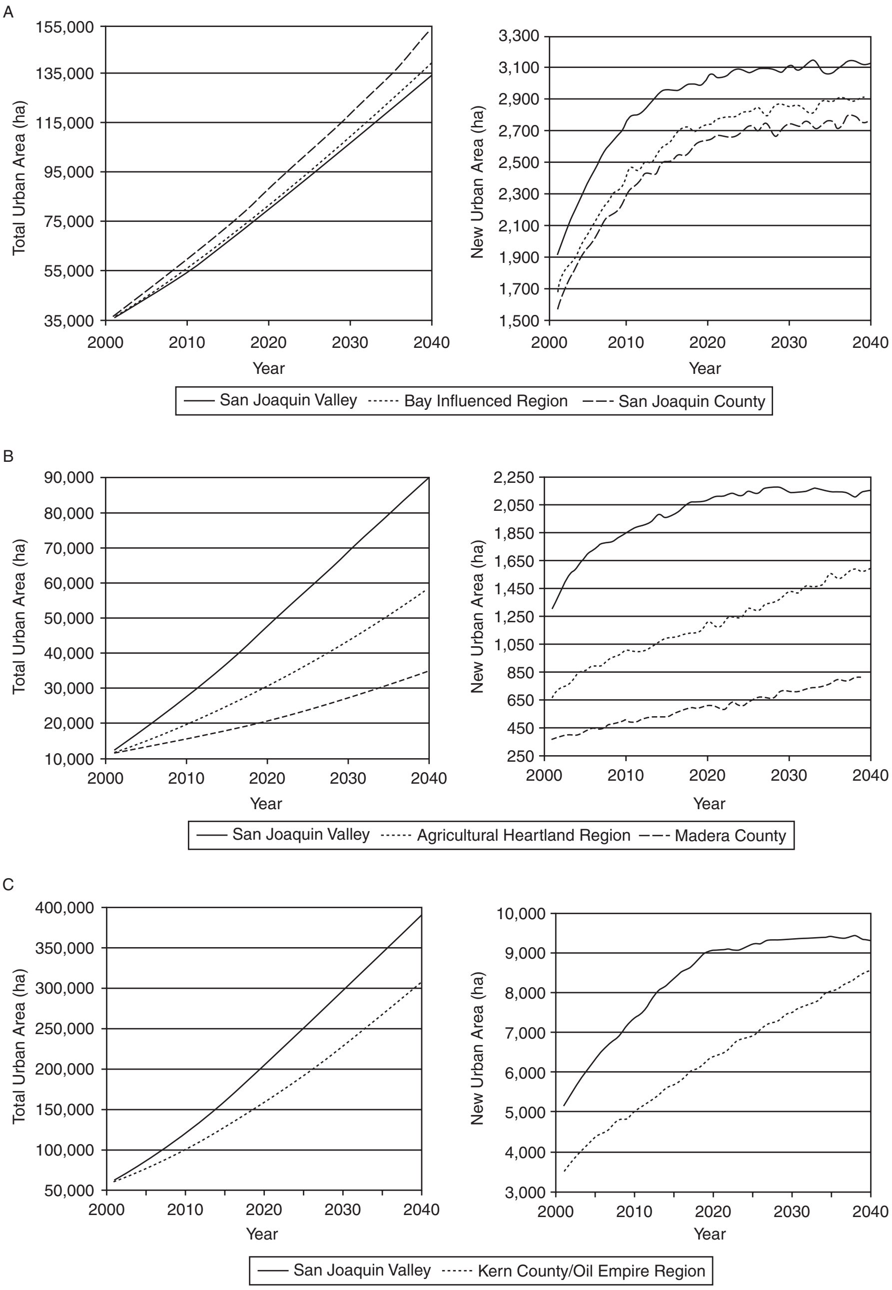

Figure 17.3 Total urban area and new urban area for each year of model projection (2000–2040) for San Joaquin (A), Madera (B), and Kern (C) counties. The model projections were made using the parameters derived from calibration of the entire San Joaquin Valley dataset, as well as those parameters derived from the calibration of the individual county, and the economic region that the county is part of.

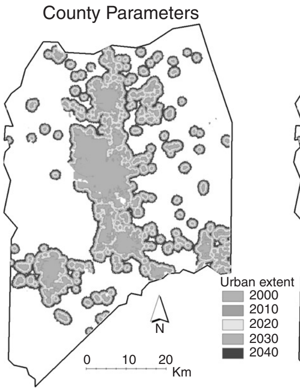

Figure 17.4 Predicted urban growth (2000–2040) for San Joaquin County using the county (left), region (middle), and valley (right) parameter sets.

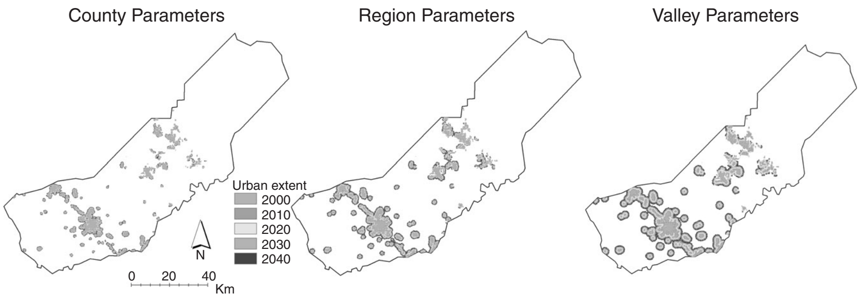

Figure 17.5 Predicted urban growth (2000–2040) for Madera County using the county (left), region (middle), and valley (right) parameter sets.

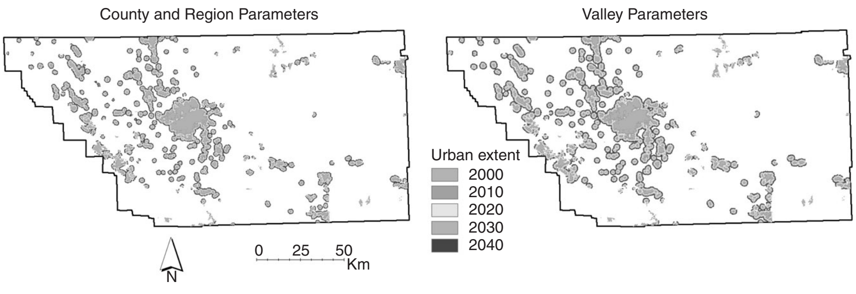

Figure 17.6 Predicted urban growth (2000–2040) for Kern County using the county and region (left), and valley (right) parameter sets.

urban area curves in Figure 17.3, the urban area common between growth forecasts using the county and regional parameters, compared with the valley parameters, were most similar in San Joaquin County. The total urban area common between the county and valley parameters was 89%, while there was 94% similarity between the

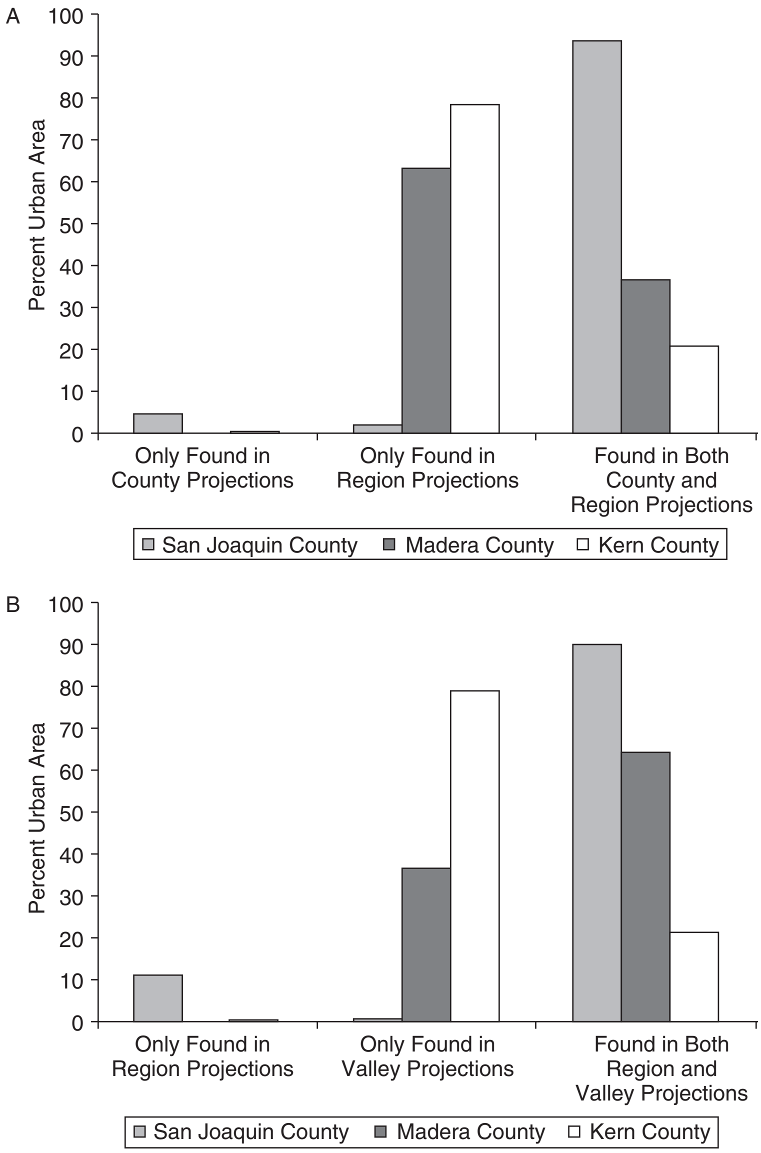

Figure 17.7 Spatial agreement between urban growth forecast (2000–2040) using the county and valley parameters (A) for San Joaquin, Madera, and Kern counties, and between forecast made for those same counties using the region and valley parameters (B). Urban area is measured in hectares.

region and valley parameter forecasts. The contrast between the three forecasts for Madera County was further demonstrated, and their was only a 36% spatial agreement between the county and valley forecasts, and 63% between the region and valley. Urban growth forecast for Kern County had a 79% spatial agreement between the county/region and valley forecasts.

17.4 DISCUSSION

Taking a hierarchical approach to modeling urban growth in three counties, each part of a different region in one large geographic area, resulted in a different parameter set for modeling at each level within the hierarchy. While the differences in the parameter sets were small between the diffusion and breed parameters, the differences in the spread, slope, and road gravity parameters were greatest, and probably had the largest impact on the differences in urban growth forecast under the different parameter sets. The total urban area forecast in San Joaquin County appeared to be similar under all three parameter sets, while Madera County growth differed by two or three times the amount forecast using the local county parameters. County and region parameters for Kern County produced urban forecasts that differed by $<20\%$, which was more than San Joaquin County, but much less than Madera County.

The differences in total urban area forecast using the different parameter sets was further illustrated by the spatial agreement between the forecast of the county and region parameters against the valley (Figure 17.7). Spatial agreement is defined as an area that forecast to be urban under the forecasts of each parameter set. San Joaquin County had the highest percent agreement between the county and valley parameters (89), and the region and valley parameters (94). The large difference in growth under the county and valley parameters for Madera County showed in the low 36% agreement between the county and valley parameters, and the 63% between the region and valley. Kern County, the county that was itself a region, had a 79% agreement. Both Kern and Madera counties, were areas where growth under the valley parameter set produced more urban area in 2040, had a large portion of urban growth that only occurred with the valley parameters. None of the parameter sets used in forecasting growth produced a substantial portion of urban growth under the county and region parameters that was not captured by the valley parameters.

17.5 CONCLUSIONS

Calibration and forecasting of a hierarchical system has provided insight into whether large-scale (geographic extent) models can accurately forecast local growth compared with smaller-scale applications. Using the San Joaquin Valley (CA) as a study area for this investigation provided examples that demonstrated that regional and global calibration and forecasting of models can provide similar outputs (up to 94%) to a local scale application. But there are also cases, such as Madera County, where large-scale (San Joaquin Valley and Agricultural Heartland scale) modeling was grossly different from the local modeling effort. The question is then, how to distinguish when it is appropriate to and when not to use global and regional modeling to look at local spatial phenomena, because it is the local scale at which most policy and land use decisions are made? This question is most likely dependent on the model being used and its parameters, since they are what determine the model output. One noticeable feature that may play a role in most models is topography. In SLEUTH, this is addressed by the "slope resistance" parameter, but it is undoubtedly used differently in many other models. Focusing on Madera County, the local parameterization of the model showed a slope resistance of 83 as opposed to 36 and 10 by the region and valley parameters. These lower slope resistance parameters allowed growth to occur where it would not normally have under conditions characterized by local parameters. This is different from the parameters from San Joaquin and its respective region, where local slope resistance was 1, regional was 30, and the valley parameter was 10; not as strikingly different as in the case of Madera County. Taking this into account, a general rule for determining whether a large-scale application can be useful for looking at a local area might be that if the topography is uniform across the entire geographic area, then the model is more likely to accurately capture local growth patterns at a larger scale. Under this line of thought, it would be possible to model many of the states and geographic regions in the United States, and look at county/metropolitan growth patterns, as well as countries like the Netherlands. But this will most likely not be the case, especially when the modeling is being done at the national, continental, and global scale. For these size applications it will most likely be necessary to model at a smaller state or region level, and then aggregate the results, producing a composite output that is more reflective of local behavior and growth.

The issue of whether large-scale modeling efforts can accurately forecast local growth compared with smaller-scale applications is inevitably also tied to spatial resolution and how changes in spatial resolution change the parameter space and model forecasts. Although this issue has not been addressed directly in this paper, the link between spatial resolution and geographic extent is one that should not be ignored. Coarser spatial resolution modeling will undoubtedly dilute model outputs at the local level, but may capture regional growth better. The converse may also be true, but these are areas of research that need exploration.

Modelers should continue to work to find the optimal geographic extent to model local urban growth, while still allowing the inclusion of influences from surrounding urban areas. While the role of geographic extent issues within the calibration and forecasting routine might appear to be a too finely detailed area to warrant extensive research efforts, this chapter has demonstrated that changes in geographic extent can create different model outputs, and in some cases, gross differences between results using large- and small-scale calibration. Future efforts should continue to research this area using current working models, building on the knowledge gained to further improve them, or create new ones that are more robust in forecasting the future. Only once the models, their behavior, and their proper use are fully understood with honesty (Chapter 16 by Clarke), will they be able to be used in an understandable and believable manner for application to the planning and decision-making process.

ACKNOWLEDGMENTS

The author would like to thank Keith Clarke for comments on the original draft as well as the three anonymous reviewers, and the Public Policy Institute of California for their funding of a larger project, from which data for this work were used.

REFERENCES

ABRAHAM, J. and HUNT J., 2000. Parameter estimation strategies for large-scale urban models. *Transportation Research Record*, **1722**, 9–16.

AGARWAL, C., GREEN, G., GROVE, M., EVANS, T., and SCHWEIK, C., 2000. *A Review and Assessment of Land-Use Change Models: Dynamics of Space, Time, and Human Choice* (Bloomington: U.S. Forest Service and the Center for the Study of Institutions, Population, and Environmental Change (CIPEC)).

ALMEIDA, C., BATTY, M., MONTEIRO, A., CAMARA, G., SOARES-FILHO, B., CERQUEIRA, G., and PENNACHIN, C., 2003. Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation. *Computers, Environment and Urban Systems*, **27**, 481–509.

BATTY, M. and XIE, Y., 1994. Modeling inside GIS: part 2. Selecting and calibrating urban models using Arc-Info. *International Journal of Geographic Information Systems*, **8**, 470–541.

BRAIL, R. and KLOSTERMAN, R., 2001. *Planning Support Systems: Integrating Geographic Information Systems, Models, and Visualization Tools* (Redlands: ESRI Press).

CALIFORNIA FARMLAND MAPPING AND MONITORING PROGRAM, 2003. WWW document, http://ceres.ca.gov/calsip/cv/.

CANDAU, J., 2000. Calibrating a cellular automaton model of urban growth in a timely manner. Presented at the *4th International Conference on Integrating GIS and Environmental Modeling* (GIS/EM4). Banff, Alberta, Canada, September 2–8, 2000.

CLARKE, K. and GAYDOS, L., 1998. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographic Information Science*, **12**, 699–714.

CLARKE, K., HOPPEN, S., and GAYDOS, L., 1997. A self-modifying cellular automata model of historical urbanization in the San Francisco Bay area. *Environment and Planning B: Planning and Design*, **24**, 247–261.

CLARKE, K., PARKS, B., and CRANE, M., 2002. *Geographic Information Systems and Environmental Modeling* (Upper Saddle River: Prentice Hall).

COUCLELIS, H., 1985. Cellular worlds: a framework for modeling micro–macro dynamics. *International Journal of Urban and Regional Research*, **17**, 585–596.

EPA, U.S., 2000. *Projecting Land-Use Change: A Summary of Models for Assessing the Effects of Community Growth and Change on Land-Use Patterns* (Cincinnati: U.S. Environmental Protection Agency, Office of Research and Development).

ESNARD, A. and YANG, Y., 2002. Descriptive and comparative studies of 1990 urban extent data for the New York Metropolitan Region. *URISA Journal*, **14**, 57–62.

GEERTMAN, S. and STILLWELL, J., 2002. *Planning Support Systems in Practice* (Heidelburg: Springer Verlag).

JANTZ, C., SHEELEY, M., GOETZ, S., and SMITH, A., 2002. *Modeling Future Urban Growth in the Washington, D.C. Area. College Park, MD, U.S.A.* (College Park: University of Maryland Department of Geography, Mid-Atlantic Regional Earth Science Applications Center).

KLOSTERMANN, R., 1999. The What if ? Collaborative planning support system. *Environment and Planning B: Planning and Design*, **26**, 393–408.

LANDIS, J. and ZHANG, M., 1998. The second generation of the California urban futures model. Part 2: Specification and calibration results of the land use change submodel. *Environment and Planning B: Planning and Design*, **25**, 795–842.

LEE, D., 1973. Requiem for large-scale models. *Journal of the American Institute of Planners*, **39**, 163–178.

LEE, D., 1994. Retrospective on large-scale urban models. *Journal of the American Planning Association*, **60**, 35–40.

LI, X. and YEH, A., 2001a. Calibration of cellular automata by using neural networks for the simulation of complex urban systems. *Environment and Planning A*, **33**, 1445–1462.

LI, X. and YEH, A., 2001b. Zoning land for agricultural protection by the integration of remote sensing, GIS, and cellular automata. *Photogrammetric Engineering and Remote Sensing*, **67**, 471–477.

PETTIT, C., SHYY, T., and STIMSON, R., 2002. An on-line planning support system to evaluate urban and regional planning scenarios, in S. Geertman and J. Stillwell (Eds.), *Planning Support Systems in Practice* (Heidelberg: Springer-Verlag), pp. 331–347.

SILVA, E. and CLARKE, K., 2002. Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. *Computers, Environment and Urban Systems*, **26**, 525–552.

STRAATMAN, B., WHITE, R., and ENGELEN, G., 2004. Towards an automatic calibration procedure for constrained cellular automata. *Computers, Environment and Urban Systems*, **28**, 149–170.

TOBLER, W., 1979. Cellular geography, in S. Gale (Ed.), *Philosophy in Geography* (Dordrecht: D. Reidel Publishing Group), pp. 379–386.

U.S. GEOLOGICAL SURVEY, 2003. Preliminary assessment of urban growth in California's Central Valley. WWW document, http://ceres.ca.gov/calsip/cv/.

VERBURG, P., SOEPBOER, W., VELDKAMP, A., LIMPIADA, R., ESPALDON, V., and MASTURA, S., 2002. Modeling the spatial dynamics of regional land use: The CLUE-S model. *Environmental Management*, **30**, 391–405.

WARD, D., MURRAY, A., and PHINN, S., 2000. A stochastically constrained cellular model of urban growth. *Computers, Environment and Urban Systems*, **24**, 539–558.

WEGENER, M., 1994. Operational urban models: state of the art. *Journal of the American Planning Association*, **60**, 17–30.

WHITE, R. and ENGELEN, G., 1993. Cellular automata and fractal urban form — a cellular modeling approach to the evolution of urban land-use patterns. *Environment and Planning A*, **25**, 1175–1199.

WU, F., 2002. Calibration of stochastic cellular automata: the application to rural–urban land conversions. *International Journal of Geographical Information Science*, **16**, 795–818.

YANG, X. and LO, C., 2003. Modeling urban growth and landscape changes in the Atlanta metropolitan area. *International Journal of Geographical Information Science*, **17**, 463–448.