# Simulating Emergent Urban Form Using Agent-Based Modeling: Desakota in the Suzhou-Wuxian Region in China

Yichun Xie,* Michael Batty,† and Kang Zhao‡

*Department of Geography and Geology, Eastern Michigan University
†Centre for Advanced Spatial Analysis, University College, London
‡Department of Computer Science, Eastern Michigan University

We propose that the emergent phenomenon know as "desakota," the rapid urbanization of densely populated rural populations in the newly developed world, particularly China, can be simulated using agent-based models that combine bottom-up actions with global interactions. We argue that desakota represents a surprising and unusual form of urbanization well-matched to processes of land development that are locally determined but moderated by the higher-level macroeconomy. We develop a simple logic that links local household reform to global urban reform, translating these ideas into a model structure that reflects these two scales. Our model first determines the rate of growth of different spatial aggregates using linear statistical analysis. It then allocates this growth to the local level using "developer agents" who determine the transformation or mutation of rural households to urban pursuits based on local land costs, accessibilities, and growth management practices. The model is applied to desakota development in the Suzhou region for the period 1990 to 2000. We show how the global rates of change predicted at the township level in the Wuxian City region surrounding Suzhou are tempered by local transformations of rural to urban land uses which we predict using cellular automata rules. The model is implemented in the *RePast 3* software and is validated using a blend of data taken from remote sensing and government statistical sources. It represents an example of generative social science that fuses plausible behavior with formalized logics matched against empirical evidence, essential in showing how novel patterns of urbanization such as desakota emerge. *Key Words: agent-based modeling, desakota, emergence, lower Yangtze River Delta, rural-urbanization.*

Rapid urban change often leads to patterns of morphology, which are surprising in that they are unanticipated, often counter to what is expected. Regeneration and redistribution taking place in the industrial city, for example, have led to increasing specialization that manifests itself in phenomena as diverse as the "edge city" and patterns of segregation that form spontaneously among different population groups quite content to live side by side, notwithstanding mild preferences for adjacent locations. These patterns are often described as "emergent," reflecting processes that act from the bottom up, producing growth and change that are organic and unplanned in their genesis. A particularly clear example is associated with the urbanization in some newly developed countries, particularly in East Asia. There, rural landscapes, usually within the hinterlands of large cities, are rapidly urbanizing, not through rural depopulation to the cities with their subsequent peripheral growth, but through a process of change in which a majority of the rural population are transforming their lifestyles and activities into urban pursuits in situ. In these situations, the long-standing migration of the population to large cities which has historically marked third-world urbanization is much less significant than the transformations that are taking place as the rural population becomes urban without substantial movement to the cities. This phenomenon is called "desakota."

Desakota is a pattern of settlement characterized by an intensive mixture of agricultural and nonagricultural activities that reveals itself as a close "interlocking" of villages and small towns (Lin 2001). These patterns are neither urban nor rural, but have demonstrated features of both. The term desakota was first used by McGee (1989, 1991), who identified these morphologies with the Bahasa Indonesian word "desakota" from the words for village (*desa*) and town (*kota*). He says "These zones are characterized by high population densities, rapid growth of nonagricultural activities, labor mobility, occupational fluidity, and intense mixture of land with agriculture, cottage industries, industrial estates, suburban developments and other uses" (McGee 1991, 16–17).

Desakota is an emerging urban form that interlocks bottom-up rural urbanization with top-down urban

*Annals of the Association of American Geographers,* 97(3), 2007, pp. 477–495 
Initial submission, August 2005; revised submission, May 2006; final acceptance, August 2006
Published by Blackwell Publishing, 350 Main Street, Malden, MA 02148, and 9600 Garsington Road, Oxford OX4 2DQ, U.K.

expansion (Lin 1997, 2001; Wang 1997, 1998). It occurs in traditional agricultural areas and brings urban functions to rural villages and towns. Physically, desakota consists of noncontinuous impervious patches that are numerous and adjacent but small in size (Xie et al. 2006), located in rural areas distant from cities and town centers. They are the direct evidence of spontaneous economic growth motivated by village residents and farmers. This bottom-up impetus represents China's economic vitality and is a primary factor sustaining China's continued rapid economic growth (Marton 2000). Moreover, much of China's desakota growth has taken place in economically advanced regions of the East Coast, particularly where the influence of market reforms and globalization has been most strongly felt (Wang 1998). The economic push and pull from large cities are key driving forces to sustain desakota development, which completes a dynamic cycle of top-down urban expansion and bottom-up rural urbanization.

In one sense, it is easy to see why this pattern of growth characterizes rapid urbanization in places like China. Rural life has formed the bedrock of Chinese society for many thousands of years, revealing itself in a dense polynucleated quilt of villages and small towns with close economic links to the larger cities. Unlike the wholesale movement from the countryside to the towns in the United Kingdom and other countries in the nineteenth century, modern technologies now make it possible to urbanize in situ, so to speak, with the network of social and economic connections associated with an urban society already largely in place. Some argue that understanding this rural-urban nexus and its new landscape is a key to understanding China's tremendous social and economic transformation (W. Tang and Chung 2000). Somehow these patterns are representative of China's extraordinary economic vitality and provide clues to its continuing social and political stability in the face of great economic upheavals (Lieberthal 1995). However the phenomena is by no means confined to China; an equivalent of desakota has existed in parts of urban Europe for the past half century as urban growth has been accelerated on the polycentric network of towns and cities established more than 500 years ago (Kloosterman and Musterd 2001).

Emergence is a much more difficult concept to explain than to illustrate. One way to proceed is to build models of such phenomena whose fundamental entities or objects, sometimes called "agents," interact with one another from the bottom up (Parker et al. 2003). The key to understanding emergent phenomena is to fully understand the way the model's agents influence one another, usually over multiple time periods and across extended spaces, where surprising patterns often emerge as a consequence of nonlinear interactions between agent behaviors through positive feedback. This is the conventional wisdom underlying the rationale for complex systems modeling. Once a satisfactory understanding of such emergence has been gleaned, then it is an open question whether the phenomenon is still to be called "emergent." Moreover, in terms of urban growth and form, purely bottom-up explanations are unlikely to reflect the range of processes and agents that generate such spatial organization (Urry 2003), and therefore any model of this process must reflect the local and the global.

In the case of desakota, efforts to explain such phenomenon are reflected in at least two schools of thought. The first emphasizes the role of rural areas as the locations for development and gives priority to rural urbanization. Since Deng Xiaoping's "reform and opening-up" policy, central control by the Chinese government on rural areas has been relaxed and local cadres have assumed responsibility for many resources and institutions in the countryside. Townships and village officials have sought to replace declining state revenues with taxes and fees on local industries and have promoted and subsidized collective and commercial enterprises, an approach that has been widely adopted since the 1980s. This viewpoint argues that, although large metropolitan cities may provide markets and new technologies, much of the energy and drive for production is not demand-driven but comes from rural peasants and local cadres seeking to improve their lives. This is truly a bottom-up process reflecting local action (W. Tang and Chung 2000).

In contrast, the top-down view highlights the contributions of China's largest cities and coastal trade zones that appear to have reinvigorated and internationalized China's economy, culminating in its recent entry into the World Trade Organization (X. Li and Yeh 1998). This viewpoint argues that it is only the metropolitan regions that have supported the conditions for China's social and economic transformation to a modern economy consistent with competitive labor markets, high worker mobility, and free trade (Yeung and Zhou 1991; Yao 1992). In fact, McGee's (1991, 1998) model of desakota is a hybrid, drawing on elements from both approaches, wherein he implies that the resultant landscapes are based on industrialization in rural areas but consistent with a "friction of space" that privileges certain locations up to 200 km beyond the largest cities or between adjacent metropolitan areas (Oi 1999).

Desakota has been quite widely studied in a qualitative sense, but to date the phenomenon has been mainly identified and analyzed in descriptive terms, focusing on how the transformation of China in terms of the global economy and its internal restructuring has sped up the pace of this kind of urbanization. There have been attempts to simulate incremental urban change in rural areas using mainly physical models such as those based on cellular automata developed by X. Li and Yeh (2000) for the Pearl River Delta. There have been attempts at measuring the resultant morphologies, which show particular patterns of fragmentation (see Sui and Zeng 2000), and there are approaches to detecting differences between rural and urban in urbanizing regions using ideas from fuzzy sets (Heikkala, Shen, and Yang 2003). Xie et al. (2005) and Xie et al. (2006) have explored how these processes have resulted in loss of agricultural land and changes to the ecological balance. To date, however, there have been no attempts to simulate the way in which developers and entrepreneurs engage in the process of land development, which is central to the way rural activities are transformed to urban. We redress this in this article by explaining the evolution of desakota using an agent-based model that is embedded within a land development process driven both from the top down and the bottom up. We argue that desakota regions emerge from a combination of (i) behaviors toward the land and housing markets that reflect State and City policies that are instituted from the top down, and (ii) developer, entrepreneur, and consumer behaviors that respond to local conditions from the bottom up. Indeed, like A. Li et al. (2005), we argue that agent-based modeling should not be restricted to processes simulating growth and change from the bottom up.

In the next section, we describe how the processes that lead to desakota can be simulated by a spatial logic that meets both local and global conditions and constraints in the City of Wuxian, which surrounds Suzhou City in the lower Yangtze Delta about 100 km northwest of Shanghai. We outline the formal structure of the model used to transform the landscape surrounding big cities into desakota, emphasizing the way top-down processes of social and economic development interact with developer-agent behavior from the bottom up, thus initiating various feedback effects that determine the spontaneous transformation of land uses. We then describe the data we have for five-year periods from 1990 to 2000, showing how these data can be used to estimate rates of urban change for the twenty-seven townships that comprise the region and that determine the controls on overall growth that take place over the observed period. We outline the way the model works at a fine spatial scale in the cells that agents occupy in making the transformation from rural to urban. We show how well this model simulates the observed trajectories of urban change from 1990 to 2000 and then indicate how we can use the model to make forecasts for the middle range until 2010 and beyond. Our emphasis on using the model in prediction is to show how agents operating spontaneously at the fine spatial scale are influenced by and influence policy at the global level, which is governed by the actions of policymakers in the townships. We then conclude with ideas for further research and a brief commentary on the suitability of this approach for explaining unusual spatial patterns such as desakota.

## A Logic for Modeling Spontaneous Urban Change in China

Urban development everywhere is influenced by decision making at multiple levels and scales. But for desakota in China in general, and development in the Suzhou City region in particular over the period from 1990 to 2000, we simplify the chain of development decisions to two levels, local and global. The global level is reflected in aggregate social and economic factors that pertain to districts or townships within the region, and that are used to define instruments that steer development to favorable locations consistent with regional and national economic policy. The local level involves the decisions that households in rural villages and small towns make to realize local-scale economies through the transformation of their activities from rural to urban pursuits or to a mix of these. We call this "two-front growth" because it combines two different policymaking levels, both of which contribute to the simultaneous development of urban and rural areas by fusing "city-leading-county" initiatives in the cities with the "household responsibility system" that has been introduced in the countryside. City-leading-county initiatives are not only geared to transforming collapsing State-owned enterprises into private ownership but often generating State-sponsored investments that reflect China's growing international trade and investment through the gateway cities. These new developments are initiated by local enterprises adjacent to large cities with foreign investment, often appearing like "flying intruders" in what was once farming land (Wei 2002).

In contrast, at the more local level since the early 1980s, introduction of the household responsibility system has dramatically changed rural areas through the decollectivization of agriculture and a return to family-centered crop production. The household responsibility system has provided strong incentives for rural towns and villages to diversify and grow their economies by developing nonagricultural enterprises. In general, this kind of rural urbanization often involves small-scale, individual, privately-owned, nonagricultural land use, termed "rural construction" in official Chinese statistical yearbooks (Wuxian City Statistical Bureau 2001), with most construction registered as housing. But the functional uses of such rural construction are diverse; many individual houses are in mixed use, based on small factories, craft and other retail shops, restaurants, and related privately-owned and operated businesses. It has been argued that this bottom-up impetus is core to China's economic vitality and is a primary factor sustaining China's continued economic miracle (Marton 2000). Agent-based modeling is an ideal way of encapsulating this kind of institutional policymaking with local physical responses in terms of land development and we will fashion our model around this logic.

We have abstracted this process in Figure 1, where we show the key feedbacks that appear to be plausible drivers of the processes that determine the transformation of the rural landscape into desakota in the Lower Yangtze Delta. At this point, we provide justification of our approach to understanding desakota using agent-based models as follows. As Page (2003) so cogently argues: "our models become better, more accurate, if they make assumptions that more closely match the behavior of real people . . . ," and to this end we consider the processes described in Figure 1 to be close to those we observe, albeit in a somewhat aggregate manner. We will further translate this by approximating the outputs of these processes by data when we come to validate the model in a later section, but in terms of verifying the model structure, we consider this structure to be consistent with the wide literature on desakota that has appeared so far. Before we specify the model formally, we need to describe this structure in somewhat more detail.

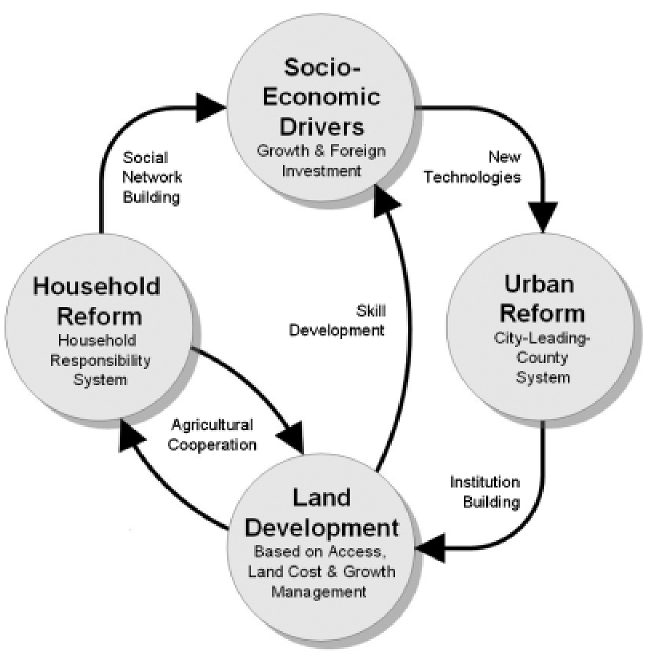


**Figure 1.** The logic of desakota.

In abstracting in this manner, we assume that distinct objects of interest can be defined as agents—whether literally as individuals and/or households, or somewhat more metaphorically as townships, districts, policy instruments, and the like. We also assume that the landscape on which and about which agents make their decisions is geographical in the traditional sense of the map. Agent-based models essentially simulate processes in which agents interact with each other but also with the landscape, where the assumption is that all possible feedbacks between landscape and agent can, in principle, take place (Batty 2005). In this context, we have already defined two levels, the local and the global, and we can thus define two types of agents and two types of landscapes associated with each of these levels. In terms of agents, we define "developer agents" who act on households, foreign investors, or even the State at the local level, and at the global level, we assume that the agents are townships. The associated landscapes are both geographical, with the global being the twenty-seven townships that exhaust the space of the Suzhou region and the local landscape being a regular grid of cells, which is the most neutral way of defining a geography wherein no individual location has an a priori advantage over any other.

These definitions map onto Figure 1 in the following manner. In general, the socioeconomic drivers of change are determined at the global township level where various policy instruments are exercised in terms of urban reform. At the township level, regional and national policies are determined, and in general this fixes the rate of growth, at least in the medium term. At the local level, household reform enables individuals and families to transform their lives by adopting urban pursuits, both to attract development and to initiate it themselves. In this sense, developer agents act as "probes" that condition households to respond to local conditions such as the costs and benefits of various types of accessibility as well as the cost of land and top-down policies for growth management. The interactions between these levels are of course critical and Figure 1 implies a degree of asymmetry in the processes just explained. Essentially the outer loop in Figure 1 represents a slow process in which socioeconomic conditions determine urban policy, which in turn provides the conditions for development to which households in the rural areas respond. In turn, these households initiate urban development on a much faster cycle than the outer loop implies. This means that at the local level, development takes place in the comparatively more stable context of wider global policy and economy but can be more volatile due to feedbacks posed by local conditions. The development that occurs then changes the socioeconomic conditions to which policy will respond in the longer term. Flow charts such as Figure 1 could imply that everything is connected to everything else, but in this context the chart shows that the global level responds more slowly than the local level. We will incorporate this feature somewhat bluntly in our model by computing rates of change at the global level over longer time periods that are used to activate the developer agents that initiate urban development at the local cellular level.

## A Formal Statement of the Model

Many kinds of agents take part in the transition from rural to urban growth, but in this model (and at the risk of gross simplification) we define two distinct types: developer agents such as individual entrepreneurs, small corporations, town- and village-owned enterprises, and privately-operated businesses at the microlevel, and township agents at the macrolevel. Meso-level actors such as town-village developers, municipality developers, external developers including investors from abroad and other China municipalities, as well as national policymaking agents, are lumped into the township agents in the current model but in the future versions these will be disaggregated. Formally, these two kinds of agents are defined in terms of the set of townships $Z_k$, indexed spatially by location as $k = 1, 2, \ldots, K$, and the set of developer agents, which we index as $j = 1, 2, \ldots, J$. The developer agents move on a landscape of cells that we index as $i = 1, 2, \ldots, I$. The township agents are immobile and are directly associated with their equivalent geographical space; that is, each township $k$ occupies an equivalent space $k$ where the number of townships is much less than the number of cells—that is, $K \ll I$. There is also a strict nesting of cells within townships, that is $\sum_k \sum_{i \in Z_k} \delta_i = I$, where the Kronecker delta $\delta_i$ simply indicates that we count $i$ as 1 if it is part of the township $k$. Just as cells are nested in townships, the micro-time periods over which development takes place from $t$ to $t+1 \ldots$, are nested within more macro-time periods from time $T$ to $T+1$, such that $\Delta T = [T+1] - [T] = \tau$, where $\tau$ is the number of micro-time instants associated with the change between $t$ and $t+\tau$ used to simulate local land development.

The model is specified at two levels. The key driver at the global level is a function that determines the rates of change in each of the townships measured by changes in households which can be converted into developable units. The rate of change in $k$, $R_k(\Delta T)$, is defined from the function $f(\bullet)$, which is specified as

$$R_k(\Delta T) = f\{X_k^1(T), X_k^2(T), \ldots\}, \tag{1}$$

where $X_k^\ell(T), \ell = 1, 2, \ldots, L$ are socioeconomic drivers associated with economic development and regional policy appropriate to the township level. Equation (1) is in fact the basis for the estimation in a later section of the importance of exogenous variables to the rates of change fitted using linear regression methods, with these rates determining the amount of growth over the macro-time period $\Delta T$. To generate a total for the end of such a time period, they are applied straightforwardly to the total households (as developable units) in $k$, $P_k(T)$, as

$$P_k(T+1) = [1 + R_k(\Delta T)]P_k(T). \tag{2}$$

To anticipate the lower-level cellular allocation, then the total households allocated at time $T+1$, $P_k(T+1)$ will always be the sum of the household population $p_{ik}(T)$ at the lower level; that is, $P_k(T+1) = \sum_{i \in Z_k} p_{ik}(T)$, where the households have already been aggregated over the number of time periods $\tau$.

From equations (1) and (2), total households can be counted at any scale and over any time period, but in the model the rates of change are in fact applied at the local level where all allocation takes place. If we define the cumulative rate from equation (2) as

$$1 + R_k(T) = \frac{P_k(T+1)}{P_k(T)}, \tag{3}$$

then we can factor this rate defined in equation (3) into a rate per unit time period $\Delta t = [t+1] - [t]$ by discounting the cumulative rate as

$$\tilde{r}_k(t) = \left\{\frac{P_k(T+1)}{P_k(T)}\right\}^{\frac{1}{\tau}} = 1 + r_k(t). \tag{4}$$

When applied cumulatively to the population $P_k(t) = \tilde{r}_k(t)P_k(T)$, equation (4) updates the totals at each time period $t$ to meet the constraint that $P_k(t+\tau) = P_k(T+1)$.

In each macro time period $\Delta T$, the total change $\Delta P_k = P_k(T+1) - P_k(T)$ is broken into its finer temporal parts using equation (4), and each subtotal $\Delta P_k(t)$, $\Delta P_k(t+1)$, $\ldots$, $\Delta P_k(t+\tau)$, forms the control for the detailed urban development process at the cellular level. At this level, the variables that determine location are quite different from the global level in that it is accessibilities, land cost reflected through suitability, and growth management policies that determine the allocations. At this stage, we will define land suitability in the fine cell $i$ in township $k$ as $C_{ik}(t)$, accessibility to economic centers as $E_{ik}(t)$, and accessibility to transportation facilities as $T_{ik}(t)$. We also define a policy index $S_k(t)$, which is in effect a "Township Competition index" related to the rate of change in $k$, $R_k(T)$, $\forall_{i,t}$. This tempers the effects of accessibility and suitability with respect to the growth management and economic policies set at the township level with the index being set in proportion to the rate of growth of each township (see Xie et al. 2005). We will specify these variables in more detail when we validate the model, but in general these factors are used to determine a probability for development $\rho_{ik}(t)$, which is a form of utility given as

$$\rho_{ik}(t) = g\{T_{ik}(t), E_{ik}(t), S_k(t)\}. \tag{5}$$

In general, land is converted to urban uses by the developer agent $j$, who for each cell $i$ in township $k$ evaluates the probability of development, subject to the suitability of the land in question as reflected in the measure $C_{ik}(t)$. In principle, what each agent is doing is converting the land in question to an urban use, to $p_{ik}(t)$, by maximizing $\rho_{ik}(t)$ subject to the constraint posed by the land suitability $C_{ik}(t)$.

Because this process is implemented algorithmically in sequential form, the details differ from a pure optimization. As we will explain below, at the start of each macro-time period $T$ in the first micro-time period $t$, we set up a series of master agents that effectively seed the development process in the periphery of existing urban development. We define $\Delta P_k(t) = P_k(t+1) - P_k(t)$ such agents and we locate these agents so they occupy the cells $i$ with the highest probability for development $\rho_{ik}(t)$ using the standard random (Monte Carlo) mechanism used in such modeling (see Batty 2005). In fact, during this process, because land suitability is taken into account, developers will not develop a cell if the land suitability is less than a certain threshold $\Xi_k(T)$—that is, if $C_{ik}(t) < \Xi_k(T)$. The reason for this initial allocation step, which is different from the subsequent steps within the macro-time period, is that between 1995 and 2000 there was a strong shift in policy in this region and this needs to be reflected in the initial placement, as we will recount in the discussion that follows. We call this first process *random allocation*, but in subsequent time periods the master developer agents are used to "spawn" additional agents that add up to the total required in subsequent micro-time periods. These agents begin by considering development in the cellular neighborhood of each master agent activating a process we call *neighborhood allocation*. It is at this point that the probabilities defined in equation (5) are considered in neighborhood order; that is, the developer agent begins by considering cells in the immediate band of eight cells around the master agent—in the Moore neighborhood—and if no suitable cell is found, then the agent considers the next band of cells, and so on until a suitable cell is located. The reason for this somewhat convoluted process is to ensure that development remains "close" to existing development, which reflects the need for connectivity in the urbanizing system.

Once the process is concluded at the end of each micro-time period, new development changes the accessibility to transport infrastructure and economic centers as well as land suitability. In short, there are positive feedbacks initiated at this lower level between one time period and the next, as reflected in the direct feedback loop between developers and households in Figure 1. Formally, then,

$$\left.\begin{array}{l} T_{ik}(t+1) = z\{T_{ik}(t), p_{ik}(t+1)\} \\ E_{ik}(t+1) = q\{E_{ik}(t), p_{ik}(t+1)\} \\ C_{ik}(t+1) = h\{C_{ik}(t), p_{ik}(t+1)\} \end{array}\right\}, \tag{6}$$

and these changes defined in equation (6) reflect the transition of cells from urban or rural to desakota. Feedbacks at the higher level of course exist, although we have not implemented any so far due to the nature of the estimation (described in the next section). Moreover, there are many extensions that we might make to this model with respect to increasing the connectivity between the various elements. Nevertheless we consider that this captures enough of the desakota process to mirror the process of spontaneous development. In Figure 2, we illustrate the crucial steps in this simulation from which it should be clear to the reader how we might make additional connections and extensions to the model structure.

## Urban Change in the Suzhou-Wuxian Region

Suzhou-Wuxian is situated at the intersection of two major economic belts defined around the Coastal and Yangtze River Delta concentrations. The wider region is under the jurisdiction of Suzhou Municipality, which contains six county-level cities: Changshu, Zhangjiagang, Kunshan, Taicang, Wujiang, and Wuxian, with the urban district (Shiqu) of Suzhou located within the administrative territory of Wuxian. Suzhou Municipality has a population of 5.84 million (Suzhou Municipality Statistical Bureau 2003). Figure 3 shows the location of the region centered on Suzhou City with the twenty-seven townships within Wuxian. Unprecedented changes in the local economy have taken place in recent

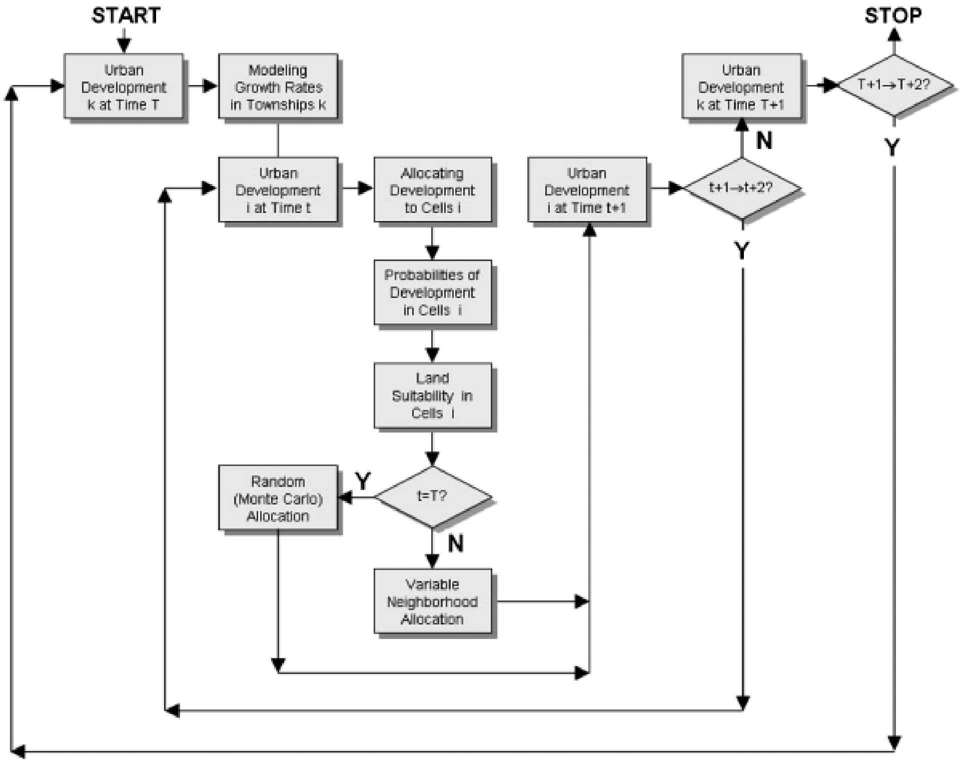


**Figure 2.** Sequence of operations in the model.

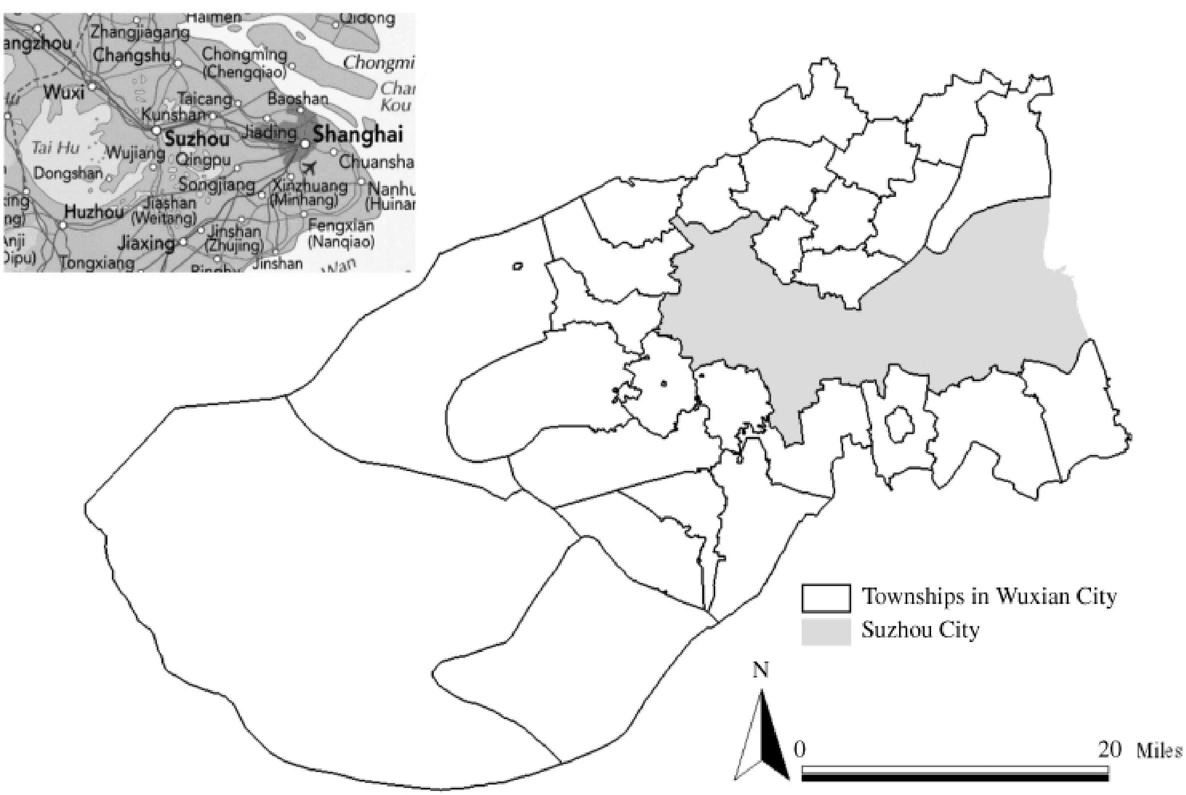


**Figure 3.** Suzhou, Wuxian, and the township level.

decades, and by the mid 1980s Suzhou provided a model for the development of rural industries based on diversified collective enterprises run by the local municipality model (K. Tan 1986). Two factors have dictated Suzhou's rural growth: its rich agricultural traditions and its network of preexisting towns and cities (Chen 1988; Marton 2000). Today agriculture still accounts for the employment of most of the rural labor force but provides less than 15 percent of the total value of production. More than 50 percent is provided by secondary industries, with another 30 percent by tertiary (service) activities. There is a demonstrable link between agricultural specialization, agricultural surpluses, and investments in nonfarm enterprises (Kirkby 2000), which is the economic basis of desakota. Moreover, due to loose and neglected protection of the environment, the consequences of uncontrolled rural industrialization are particularly serious. Rapid loss of farm land and environmental deterioration are typical concerns in Suzhou.

The data for this model were derived from diverse sources. As population, household, and related socioeconomic data were not available at a scale equivalent to land parcels or even census blocks, data on urban and rural construction (which we assumed to be proportional to household change) were generated from remotely sensed imagery using Landsat (TM) images for 1990, 1995, and 2000. The pixel resolution of Landsat TM images is 30 m. This resolution suffices to differentiate urban construction patches since they are large and continuous. Patches of rural construction have a noncontinuous impervious surface and are small in size. Rural construction is a rather fuzzy type of land use, based on the official Chinese statistical yearbooks (Wuxian City Statistical Bureau 2001).

The functional uses of rural construction are diverse, but most rural construction is registered for housing. However, many individual houses have mixed uses, such as small factories, craft shops, retail shops, restaurants, and other village and privately owned and operated businesses. Moreover, the output values from these village and privately operated businesses account for increasingly larger portions of the total gross domestic production. Rural construction is the direct evidence of spontaneous economic growth inspired by village residents and farmers. At a regional scale, many studies have produced results promising enough to use Landsat data for mapping land use (Wilkie and Finn 1996; Reese et al. 2002). We assumed that data on urban and rural construction were proportional to household change and we generated this using Landsat TM images for 1990, 1995, and 2000. As population, household, and related socioeconomic data were not available at a scale equivalent to land parcels or even census blocks, we used township geography that is the finest administrative unit, on which the Chinese government collects socioeconomic statistics. We aggregated Landsat 30-m pixels over the township boundary data layer to tabulate areas of various land types and link the land use data to the townships' socioeconomic data used to simulate causal relationships.

These images were classified into a dozen land use categories used to derive the transition matrices indicating the amount of each land use that was converted to any other use during the two periods in question, 1990 to 1995 and 1995 to 2000: $T \rightarrow T{+}1$ and $T{+}1 \rightarrow T{+}2$. From these images, land parcel data were extracted and then converted to vector data sets, complemented by data associated with topography, geomorphology, vegetation, precipitation, and temperature used as the ancillary data in the interpretation process. Further details are given in Liu, Liu, and Deng (2002). The method adopted here to extract dynamic changes in the vector land use datasets was based on postclassification image comparison complemented by field sampling to ensure quality control in the resulting classifications (de Almeida et al. 2005). Control was executed by checking the identities and the boundaries of sample land use patches with manual adjustment to decrease the incidence of major errors. The TM images in 1995 were used to interpret the dynamic change vector data by comparing them with the vector data derived from interpreting TM images in 1990, and the same method was applied for the period 1995 to 2000. In fact, over both time periods comparing 1990 with 2000 data, the overall classification accuracy of measured areas of all land types was about 97 percent (Liu, Liu, and Deng 2002), which gave us a high level of confidence in the extracted change data and its allocation into land use categories.

We will not show the complete transition matrix among the thirteen categories of land use that we have extracted at these three dates, for our focus is not on how particular land uses are transformed into one another per se but on the impact of urban development on the range of land cover types. In Table 1 we show the changes from an aggregated set of classes into urban and rural construction (which we take to be urban/household unit development in this context). The total construction was significantly reduced during the period from 1995 to 2000 (19 percent) and in particular urban construction was dramatically trimmed down to around 64 percent. In contrast, rural construction increased 61 percent in 1995–2000 compared to the period 1990–1995. This conversion confirms that the policy enacted

**Table 1.** Percentage conversion of land use to urban and rural construction over the two macro-time periods

| Land use cover type | 1990–1995 | 1995–2000 |
|---|---|---|
| Converted hectares of urban and rural construction | 4,228 ( = 2692 + 1536) | 3,433 ( = 963 + 2470) |
| Shrub and loose forest | 0.09 | 2.08 |
| Other forest including orchards | 0.2 | 3.37 |
| Highly-covered grassland | 0 | 1.24 |
| Lake, reservoir, and pond | 1.17 | 1.88 |
| Shoal | 0 | 0.03 |
| Hill and plain paddy field | 50.84 | 77.13 |
| Hill and plain drylands | 46.69 | 10.85 |

in 1995 to protect agricultural land and curb the overheated economy had a notable impact on the desakota process through subsequent land use changes and patterns. Moreover, there has been a sharp increase in land use being converted from paddy fields and a consequent drop in conversion from drylands between the first and second time periods. This indicates that the preferred land supply—drylands—for urban development, has been severely diminished, which in turn has forced people to take more productive farmland—paddy fields. Although strict measures for controlling investment and urban expansion are noticeable, it has been hard for government to discourage rural residents from building more spacious houses due to increasing affluence.

Paddy fields and drylands completely dominate the process of land development, forming some 98 percent of the entire land use change in the first period and 88 percent in the second period. As conversion from paddy fields is the largest category in both periods, we can also examine the extent to which paddy fields are converted to other land uses. As shown in Table 2, urban and rural construction still dominate, taking some 77 percent and 74 percent of paddy field land in the two respective time periods, with factory and transportation uses taking 4 percent and 7 percent, respectively. The only other substantial transition is from paddy field to reservoirs and ponds, which simply indicates traditional changes in this kind of wetland agriculture with no real significance for urbanization.

**Table 2.** Percentage conversion of paddy fields to other land uses over the two macro-time periods

| Land use cover type | 1990–1995 | 1995–2000 |
|---|---|---|
| Lost hectares of paddy fields | 3816 | 3498 |
| Dense forest | 0.12 | 0.00 |
| Shrub forest | 0.03 | 0.00 |
| Sparse forest | 0.19 | 0.00 |
| Orchard | 0.14 | 0.48 |
| Dense grassland | 0.05 | 0.83 |
| River | 0.12 | 0.00 |
| Lake | 2.66 | 0.00 |
| Reservoir and pond | 14.02 | 17.73 |
| Shoal | 0.01 | 0.00 |
| Urban construction | 39.33 | 32.96 |
| Rural construction | 38.10 | 40.91 |
| Large factory and transportation | 3.51 | 6.93 |
| Plain dry land | 1.72 | 0.16 |

It is not meaningful in such a large region to examine absolute volumes of change. Suffice it to say that Shuzhou City's population was some 0.84 m (million) in 1990 and this grew to 1.11 m in 2000, with Wuxian City, the surrounding region, falling from 1.12 m to 0.96 m people during this time. In fact in 1995, due to boundary changes, the urban district of Suzhou incorporated five townships from Wuxian (Shenpu, Weiting, Kwatang, Xietang, and Fengqiao along with their 170,175 residents). However, the key point is that the entire region grew only slowly from 1.97 m in 1990 to 2.03 m in 1995 to 2.07 m in 2000. As a matter of fact, Suzhou Municipality recorded negative natural population growth rates in recent years ( − 0.28 percent in 1998, − 0.22 percent in 1999, 0.58 percent in 2000, − 0.39 percent in 2001, and − 0.27 percent in 2002), which partially explains the slow population growth in Suzhou (Suzhou Municipality Statistical Bureau 2003).

Nevertheless, there has been dramatic urbanization of Wuxian during this period, which is quite evident from analysis of the imagery. In terms of land area, some 60 percent of Wuxian is permanent lake. Paddy field is the next largest use at about 25 percent of the area in 1990, but paddy fields have reduced by 5 percent in each of the five-year periods (dropping by roughly 1 percent per year) and now constitute some 22 percent of the region. This loss has been taken up by rural construction, which was 3 percent of the region's area in 1990 and 4 percent in 2000, growing by some 25 percent in the first period and 46 percent in the second. Urban construction (within Wuxian) grew even more dramatically by some 18 percent in the first period and a staggering 240 percent in the second. The scale of this growth is quite characteristic of desakota with the boundary between what is defined as rural and urban becoming entirely blurred (Heikkala, Shen, and Yang 2003). Figure 4 shows the distribution of land uses taken from the remote imagery for 1990, 1995, and 2000, the differences between each of these dates, and the difference from 1990 to 2000, which indicates the degree of overall change. It is from these different maps that we compute the urban change used to drive the model from the township level.

## Estimating Rates of Urban Change

The rates of change at the township level that are applied to the total land use change extracted from the remote imagery at 1990 and 1995 are generated from a linear statistical model whose independent variables are based on socioeconomic data reflecting economic conditions and policy imperatives. Two rather different

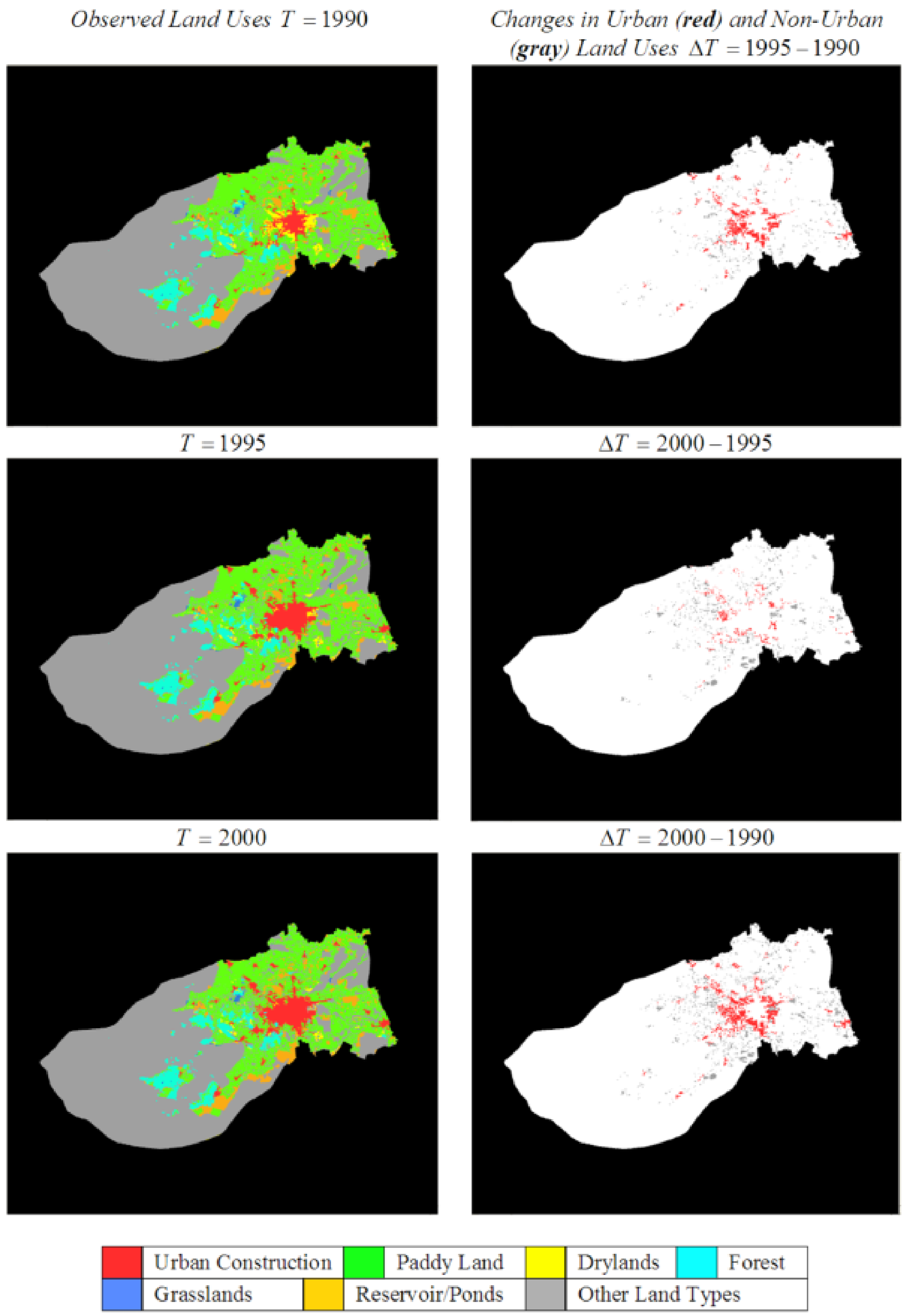


**Figure 4.** Land use types at 1990, 1995, and 2000 and first differences associated with modeling urban change.

models resulted from the estimations in each macro-time period: the first based on simple demographic variables, the second on new data reflecting income and tax. This is simply due to the stepwise procedure used to identify significant independent variables in the model rather any differences in data. The fact that the models from the two periods are different in structure reflects quite distinct differences between the political and economic regimes dominating development in the Suzhou region over the past decade. During this period, there was a strong shift to economic issues associated with income and taxation in contrast to the earlier period when demographic factors appeared stronger. In making forecasts, we will use the rate model from the second period because it has variables that can be more directly associated with policy.

In general the rate of urban change in township $k$, $R_k(\Delta T)$, defined above in equation (1), can be estimated from the following linear form:

$$R_k(\Delta T) = \alpha(T) + \sum_{\ell} \beta_k^{\ell} X_k^{\ell}(T) + \varepsilon_k(T), \qquad (7)$$

where $\beta_k^{\ell}$ are weights on $X_k^{\ell}(T), \ell = 1, 2, \ldots, L$, the $L$ independent socioeconomic variables defined at the township level $k$, $\alpha(T)$ is a constant, and $\varepsilon_k(T)$ are the associated error terms. In the period 1990–1995, a stepwise regression using the large data set in Table 3, which includes a series of population and employment-labor force variables, resulted in the following equation being judged to be the most parsimonious with the best fit of equation (7) for the period 1990–1995:

$$\left.\begin{array}{l} R_k(1990 \rightarrow 1995) = 15.68 + 0.627\, RP_k(1990) + 7.69 P_k(1990) \\ \qquad\qquad\qquad\qquad (\mathbf{0.89}) \quad (\mathbf{2.48}) \qquad\qquad (\mathbf{4.54}) \\ \qquad\qquad\qquad\qquad - 5.02\, E_k(1990) \\ \qquad\qquad\qquad\qquad (\mathbf{-3.49}) \end{array}\right\}, \qquad (8)$$

where $RP_k(1990)$ is the rural (nonurban) population, $P_k(1990)$ the urban population, and $E_k(1990)$ the employment (labor force) total, all at 1990. The $t$ statistics (**bold**) under each weight and variable make clear that the parameters $\beta_k^{\ell}$ are all significantly different from zero at the 5 percent level. The amount of variance explained by this equation is 72 percent, which is acceptable for driving the simulation from its start point.

In the second period, 1995–2000, other variables from the database, based on the same twenty key economic and demographic variables used in the 1990–1995 calibration, appeared. The dataset included output by employment sector, fixed asset values, incomes in different sectors, revenue, and taxable receipts. As in the first time period, we cycled through a stepwise regression procedure that ultimately converged on an equation with a larger and very different set of independent variables, much more related to policy instruments such as taxation. The form of equation (7) for the period 1995–2000 is

**Table 3.** Demographic and socioeconomic variables considered in the stepwise regressions

| Description | Unit[a] |
|---|---|
| Agricultural population | IND |
| Nonagricultural population | IND |
| Total population | IND |
| Land area in cultivation | MU |
| Total output value of agriculture, forestry, animal husbandry, and fishery | MY |
| Gross domestic product value | MY |
| Gross product value of primary and secondary industries | MY |
| Gross product value of tertiary industries | MY |
| Total value of fixed assets investment | MY |
| Total income of rural economy | MY |
| Income in agriculture, forestry, animal husbandry, and fishery | MY |
| Income in nonagriculture, nonforestry, non–animal husbandry, and nonfishery | MY |
| Total expense in rural economy | MY |
| Total income of the farmers | MY |
| Total value of industrial assets | MY |
| Net value of the fixed assets | MY |
| Number of factories | CNT |
| Number of employed people at the year's end | IND |
| Total tax value | MY |
| Sold ratio of the product value | Percentage |

[a]IND = individual count of all people; MU = 1 mu = 1/15 ha = 1/6 acre; MY = million Chinese yuan; CNT =count of all factories.

$$\left.\begin{aligned} &R_k(1995 \rightarrow 2000) \\ &\quad = 0.99 + 0.24\,TAX_k(1995) - 0.28\,INA_k(1995) \\ &\quad\;\; (\mathbf{0.17}) \quad (\mathbf{8.58}) \qquad\qquad (\mathbf{-4.65}) \\ &\quad\;\; + 0.17\,GDP_k(1995) - 0.14\,FA_k(1995) + 0.25\,RE_k \\ &\quad\;\; (\mathbf{4.35}) \qquad\qquad (\mathbf{-4.04}) \qquad\quad (\mathbf{2.37}) \end{aligned}\right\}, \tag{9}$$

where $TAX_k(1995)$ is the total tax levied, $INA_k(1995)$ is income in the nonagricultural sectors, $GDP_k(1995)$ is gross domestic product, $FA_k(1995)$ is the net value of fixed total assets, and $RE_k(1995)$ is the expenditure in the rural economy, all at 1995, and defined in each township $k$. The $t$ statistics (**bold**) under each variable imply that the $\beta_k^{\ell}$ parameters are all significantly different from zero at the 5 percent level. The variance explained by this equation is 88 percent, which is particularly high given the aggregation and uncertainties posed by the quality of the data.

Equations (8) and (9) are used in the global model, and in terms of the simulation they provide the parameters determining the overall rates of change from 1990 to 1995 and 1995 to 2000 in the twenty-seven townships. These are used to compute the rates input to equation (2), which in turn is used to factor the total urban change into its constituent components. The amount of the change is then allocated to the cells by the lower-level agent model. Note that the calibration of the global model is accomplished outside the overall model framework as represented in Figure 2.

Before we deal with more detailed questions of simulation, it is worth noting that the difference between the two major periods 1990–1995 and 1995–2000, which are used to calibrate two rather different aggregate models, is related directly to policy changes. A major policy change regarding the agricultural land protection was enacted in the mid-1990s. China's policy from 1990 to 1995 was prioritized to promote socioeconomic development and to improve living standards. As a result, rapid economic growth brought about rapid urban expansion with an increasing proportion of affluent rural residents consuming arable land for residence and services. Dramatic demands for construction resulted in land degradation, deforestation, habitat fragmentation, and biodiversity loss. The overheated economy forced the Chinese government to issue new policy directives in the mid-1990s (Jones 2002). The new farmland protection law stipulates tough rules on fixed-asset investment, urban construction, and farmland (paddy field) loss, but less strict regulations on marginal lands (dry plain and shrub lands, in particular) that are less productive (but environmentally sensitive).

Under these different policies, socioeconomic factors clearly interact in different ways with distinct changes in land use cover. From 1990 to 1995, rural and urban population and employment were the critical factors driving rural urbanization. However, in the period from 1995 to 2000 the most important drivers for rural industrialization and commercialization were the total tax levied in each township, income from the nonagricultural sectors, gross domestic product in each township, the net value of fixed total assets, and the expenditure in the rural economy. Production, revenue, profit, and efficiency thus play much more significant roles than the sheer sizes of population and employment.

## Agents in a Cellular Landscape: Simulating Growth and Change in Wuxian City

The cellular level used to allocate urban change generated at the global level is based on defining probabilities of transition from whatever use the cell has at the start of the simulation at time $t$ to urban use at time $t+1$. These probabilities were defined generically in equation (5), where it was argued that they depend on accessibility to economic activity in town centers $E_{ik}(t)$ and accessibility to transport infrastructure $T_{ik}(t)$. The other factor, which we have called the Township Competition index $S_k(t)$, is related to the rate of growth at the township level; it cements the local and global levels together not only by controlling the amount of growth but also by inputting the influence of the township on the local level. We will now detail how these factors are used to define the probabilities of urban change.

Economic accessibility is based on distance to town centers. It is determined through a GIS operation, buffering the town centers at five successive distances in 0.5 km increments (at 0.5 km, 1.0 km, 1.5 km, . . .) from their physical centroids and then recoding the distances as 1 ($<$0.5 km), 2 (0.5–1.0 km), 3 (1.0–1.5 km), and so on. We refer to the areas created as economic opportunity zones. Transportation accessibility to the main roads is computed by buffering at the same ranges of distance and then scoring the successive buffers in the same way.

These simplifications are required so that the hundreds of thousands of interactions between cells in the model can be computed efficiently. The way this works in the model is as follows. When an agent considers accessibility to transport or to town centers, in the absence of any detailed information at the cellular level the model assumes that with decreasing distance to the town center, more and more economic opportunities are accumulated, and this is then weighted against the index score so that an opportunity surface is established. This is similar to gravitational notions in terms of intervening opportunities weighted against distance and it enables an accessibility score to be computed for each of the accessibilities in question. The Township Competition index is also computed by transforming township attributes into scores. Townships are sorted from high to low according to how many urban agents are associated with the urban change predicted for the township in question. We simply order the towns in terms of growth rates, from largest to smallest, and assign priority orders of 1 to the townships ranked from 1–5, 2 from 6–10, 3 from 11–15, 4 from 16–20, and 5 for townships ranked $>20$. The elements then used to compute the probabilities are dimensionless. The probability equation is set up in linear form as

$$\rho_{ik}(t) = \mu T_{ik}(t) + \lambda E_{ik}(t) + \psi S_{ik}(t), \qquad (10)$$

where each variable ranges from 1 to 5 with the entire range being between 1 and 15. The parameters $\mu$, $\lambda$, and $\psi$ are those that enable the model to be calibrated at the local level, a process we will describe shortly.

One last feature of the local allocation needs to be established before we briefly describe how we calibrate the model and then consider its use in forecasting. This involves land suitability, which we earlier formalized as a constraint on the optimization of the probability of urban development. In fact, we use a strict priority ordering for the transition of land to urban use. The urban agents will try to occupy dryland first, then paddy fields, forest, reservoirs and ponds, and finally grassland.

The land suitability process is initiated in the first micro-time period when the random allocation of master agents to locations is made. We outlined this process earlier but at this point we need to be crystal clear about how the whole microsimulation is implemented. This is at the heart of the process of generating spontaneous urban growth in the countryside, which is the essence of the way desakota emerges. Each cell is 100 m $\times$ 100 m and there are roughly 687,000 cells in total allocated to the twenty-seven townships. In each macro-time period of the simulation from $T \rightarrow T+1$, there are about 50,000 to 60,000 urban developer agents $\{j\}$ that roam the cellular landscape $\{i\}$ looking for cells to transform from rural to urban. These of course are discounted back to about 9,000 to 12,000 for each micro-time period $t \rightarrow t+1$. The transformation process that they initiate is different from the usual cellular automata model structure in which cells change state from rural to urban dependent on land suitability and accessibility rules, for the urban developer agents are essentially mobile. Strictly speaking, for a model to be agent-based, it must contain agents that can move, for if the agents are passive and simply in one-one correspondence with cells, then the agent layer is redundant (Batty 2005). In this case at the microcellular level, the agents act as "probes" searching the landscape for cells that are suitable for transformation from rural to urban and their movement on the landscape reflects the process of searching for suitable sites (cells).

As noted earlier, the process of allocation consists of randomly allocating the first round of agents to cells that have the highest probability of development based on equation (10), subject to the land suitability ordering, when $t = T$ or $t+\tau+1 = T+1$ (i.e., in the time period 1990–1991 or 1995–1996). These agents are the master agents; they then seed all subsequent allocations of agents using the neighborhood allocation rules in the remaining time periods of the microsimulation, 1992–1993, 1993–1994, and so on, until the end of the second macro-time period in 2000. In the second micro-time period, an appropriate number of new urban agents associated with the master agents in each township area, but at the cellular level, are generated and then allocated using the neighborhood allocation rules. This involves these new agents assessing the suitability and probability of land for urban development in the neighborhood of the master agents. If these agents are unable to find suitable land for conversion in these immediate neighborhoods (which will always be the case because the number of agents being generated is likely to far exceed the available cells in these restricted Moore neighborhoods), the search is widened and the agents move to the next band of cells, continuing in this way until all the agents associated with township in that micro-time period are allocated. At this point, these new agents become master agents seeding the next round of conversions in the next micro-time period until all the urban agents associated with that macro-time period have been allocated. This process is akin to a process of continual mutation of land uses until enough urban development has been generated to meet the control totals consistent with the rates of change that are simulated at the township level using the linear model.

The local level model, which is essentially the structure pictured in Figure 2, is implemented in the open source modeling language *RePast 3* (Collier, Howe, and North 2003; and http://repast.sourceforge.net/). Our implementation is unusual in that we have a very large cellular landscape and thousands of agents and it is one of the first "realistic" implementations of *RePast* for spatial agent-based simulation, as the results presented below will show. The details of the simulation process need not concern us, other than noting that there is another layer of time within the operation of the simulation, which is referenced as ticks. These ticks do not match the real times $t$ and $T$ for they are essentially used to track the movement of agents across the space as they search for suitable cells to transform and, as such, they reflect the various iterations that are used to achieve the control totals from the global level. It is also important to note that the search process for an urban agent is not confined to the cells associated with a particular township. The agents are free to search over the entire space. In fact, one major measure of fit that we use is to compute how many units of development are generated in each township.

We have run the model in two ways: first we calibrated the model from 1990 to 2000 in terms of the local parameters $\mu$, $\lambda$, and $\psi$ associated with the probability of development in equation (10); and second, having chosen these best parameter values, we ran the model from 2000 to 2010 using the global parameters from the regression model in equation (9). To calibrate the local level model, we should choose a range of values for each parameter and run the model over all significant combinations of values in these ranges.

For each combination, we compute the goodness of fit of the model in terms of the number and location of cell conversions from rural to urban land, and then choose that combination of values that is closest to what we observe in the overall period from 1990 to 2000. The values of the three parameters $\mu$, $\lambda$, and $\psi$ are arbitrary as the scoring used in forming their variables makes them comparable and thus it is their relative values that are important. This is a very standard method of searching for best-fit parameters in intrinsically nonlinear models, and one that goes back many years (see Batty 1976), although what we have actually done is to sample the phase space in a systematic way rather than sweeping the entire space in comprehensive fashion. In extensions to this model, we will selectively sample and search the space hierarchically as we have done in other agent-based models we have been working with (Xie and Batty 2005).

The goodness-of-fit criterion that we are currently using is based on the difference between the urban cells that have been converted from rural predicted by the model with respect to those observed from the remote imagery maps shown earlier in Figure 4. This is for the entire period from 1990 to 2000 but aggregated to the twenty-seven townships. Although the townships act as the control on total growth at the global level, urban development at the local, cellular level is not restricted to particular townships as we noted earlier; that is, agents are free to search the entire space. Formally this criterion is:

$$\Phi(\mu, \lambda, \psi) = \sum_{k=1}^{27} \left[ P_k(T \to T+2) - \sum_{\tau=1}^{10} \sum_{i \in Z_k} p_{ik}(\tau) \right]^2 \Bigg/ 27, \tag{11}$$

where the temporal summations are over the period from 1990 to 2000 and the spatial summations are over the number of cells in each township.

To get the best parameter values, we first developed a very crude sweep of the phase space, choosing four values of each parameter and focusing on the area of the space that seemed to yield the lowest goodness of fit. This was the area around $25 \le \mu \le 35$, $1 \le \lambda \le 10$, and $1 \le \psi \le 10$. We searched one dimensionally across each of these parameters, computing the goodness of fit, $\Phi$ ($\mu$, $\lambda$, $\psi$), through varying $\mu$, then $\lambda$, then $\psi$ with the respective other two parameters held constant in each case. In Figure 5 we show variations in the goodness of fit, which represent transects through the three-dimensional phase space. The best values ultimately identified were $\mu \sim 30$, $\lambda \sim 5$, and $\psi \sim 5$, and Figure 5 shows that these give a point of minimum difference between predictions and observations with respect to this local area of the phase space. Short of sweeping the entire phase space at this level of detail, which would involve running the model thousands of times, we consider this to be as good as we are likely to get at this initial stage. This may not be the optimum optimorum but we are also certain that there are no such global optima within phase spaces associated with models of this kind. Nevertheless we are confident that the parameter values identified produce good simulations, for these revealed patterns of growth are close to those that we observe over the calibration period 1990–2000.

It is important to examine these spatial results directly as statistics such as those in equation (11) are not spatially weighted in any way. When we look at the predictions from 1990–1995 and 1995–2000 in Figure 6, we see that the model produces rather plausible patterns of desakota, quite consistent with what we have extracted as observed urban development from the remote imagery. The patterns show that there is spontaneous growth into the hinterland of Suzhou in all townships as land is converted to urban. In fact, a much clearer picture of what is happening and the ability of the model to generate in situ growth and change is given in the snapshots from the animation that is shown in Figure 7 where we plot the actual change in urban development from 1990 to 2000. This is then extended as a forecast to 2010 using the 1995 to 2000 global parameters in fixing the amount of urban change over this future ten-year period.

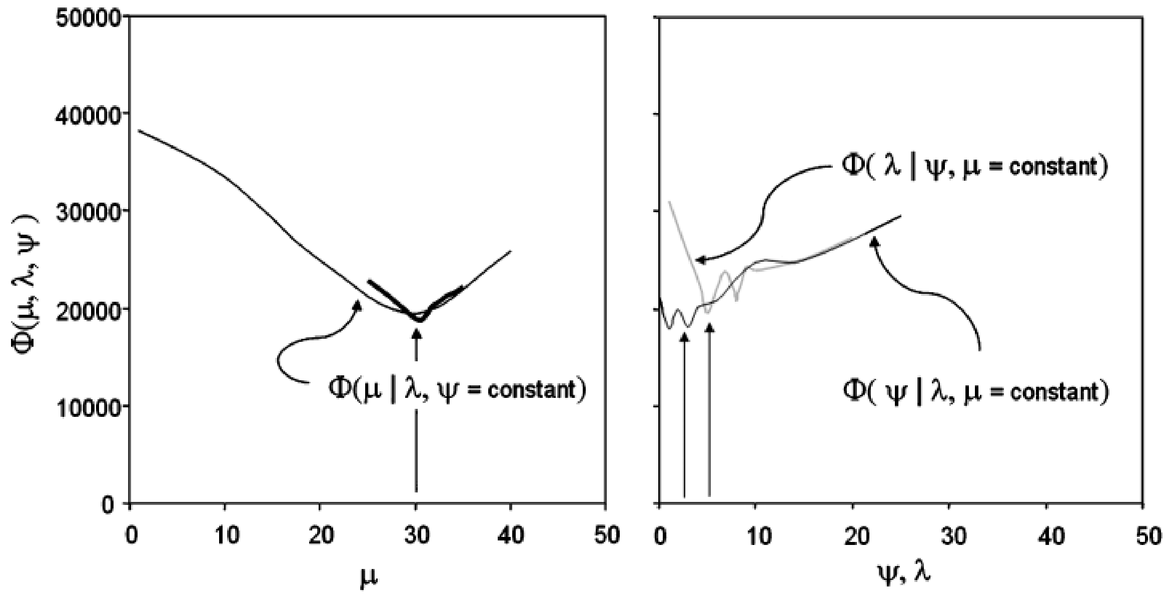


**Figure 5.** Optimizing the goodness of fit within the search space defined by local parameter values.

These figures speak for themselves in illustrating how urban development pops up all over the region, notwithstanding the modest population growth in the area which was very close to only 1 percent per annum through the decade from 1990 to 2000. The web site http://www.casa.ucl.ac.uk/desakota/ gives access to an animation of this process from which these snapshots are taken.

## Conclusions: Findings and Next Steps

We have extracted a few findings about the desakota development in the Suzhou–Wuxian region in China. First, rapid rural urbanization and the emergence of market towns are apparent in the study area. In addition to the significant expansion of Suzhou City, many of the twenty-seven towns witnessed noticeable urban growth. Second, comparing the observed urban growth (Figure 4) with the growth predicted by the agent-based modeling (Figure 6) shows the observed desakota sites to be more aggregated and clustered and the predicted desakota areas to be relatively evenly dispersed among the twenty-seven townships. Though this finding may indicate that additional attention should be given to the calibration process, it has exposed the complex nature of spatial dynamics of desakota growth. As we pointed out earlier, key to desakota development is the integration of commercial and industrial functions in agricultural landscapes. In other words, desakota is associated with significant increases in nonagricultural population, rural economy revenue, and tax, which are the explanatory variables that we used here to estimate the growth rates of urban agents at the level of the township geography. Apparently the desakota phenomenon demonstrates a more complicated relationship that cannot be simply explained by the linear associations between these socioeconomic driving factors at the level of township aggregation. Nevertheless our findings do reveal the need for considering the township competitiveness from the regional perspective.

We have also illustrated a number of features about the agent-based models. First we have designed a relatively large-scale model of urban growth and change that is agent-based in terms of its local simulation, with the agents being used as probes to convert land use from rural to urban development. Agents are used as devices to search the space, and in future versions we will give these objects a more realistic form by dividing them into urban entrepreneurs/developers and farmers, elaborating the search process as one of profit maximizing where the various accessibility and land suitability attributes are considered as relevant to the market process (Xie and Batty 2005). Second, our use of agents in this fashion is rather innovative as the agents are designed to be "change agents" rather than literal households or individuals, thus enabling the dynamics of the agent-based software we have used to be configured not only for temporal change but for change associated with search and optimization within the model structure. Third, to our knowledge this is one of the first applications of agent-based modeling that uses large datasets fusing remotely sensed imagery with socioeconomic data.

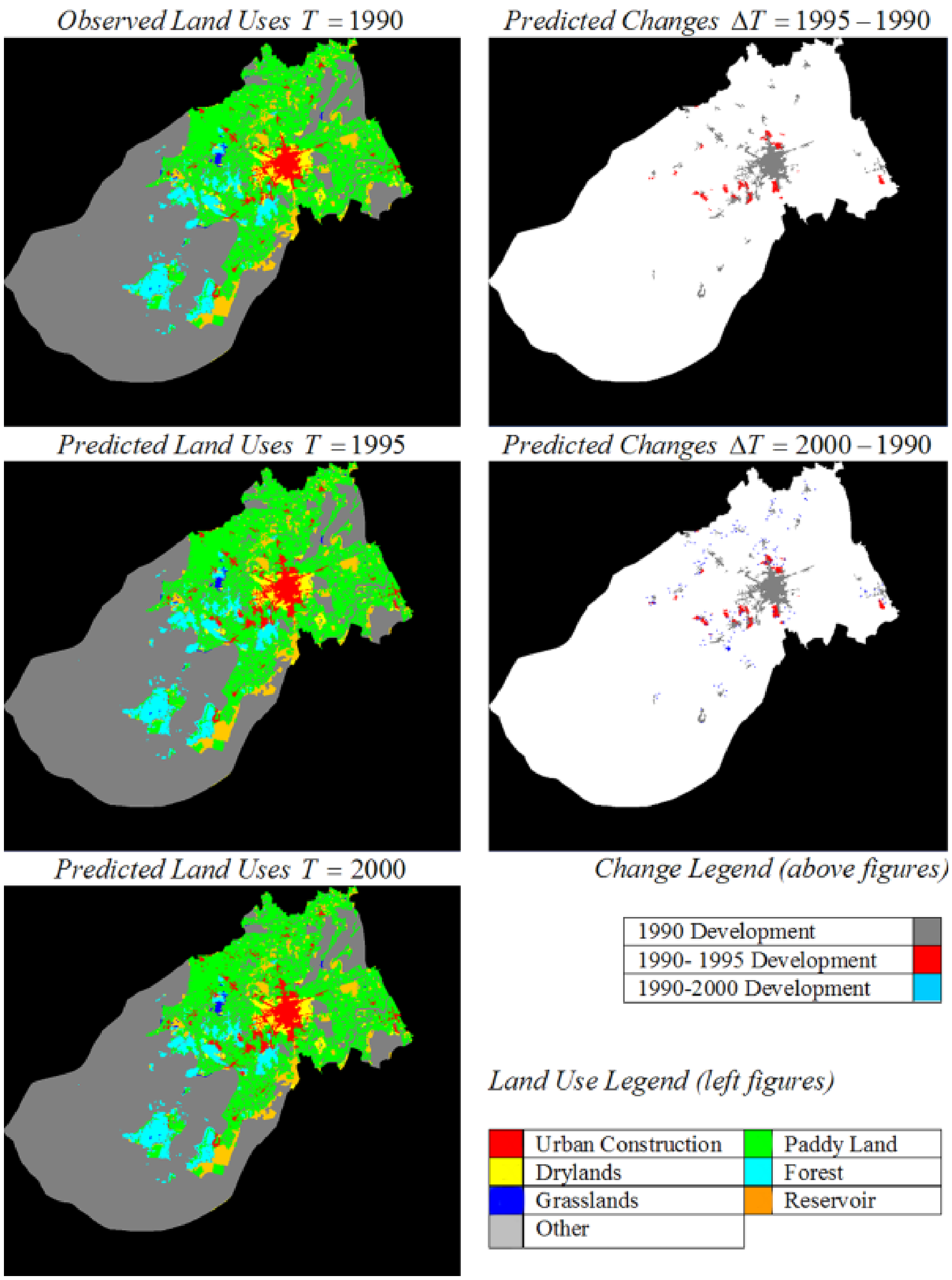


**Figure 6.** Predicted land use and changes in land use 1990 to 2000.

Fourth, we are firm in our belief that agent-based models do not simply apply to disaggregate systems but can be used to integrate different levels and scales essential to simulating what at first sight appears to be a bottom-up phenomenon such as desakota. Of course, enabling agents to search and to extract information beyond their locality is a challenging task. Our handling of the agents at the local neighborhoods and at the level of the townships in this article is a considered attempt to understand the interactions of geographic agents at different scales. However, as we pointed out earlier, the consideration of agents acting in a competitive environment both from regional and national perspectives in a world of economic globalization leads to more realistic simulations.

Many features of our simulations are rough around the edges and require considerable refinement. We are also aware that in calibrating such models to match real data, we are hardly testing the model in its widest sense. But this is little different from the movement in science and social science to embed plausible behavioral assumptions into our models, which we consider to be

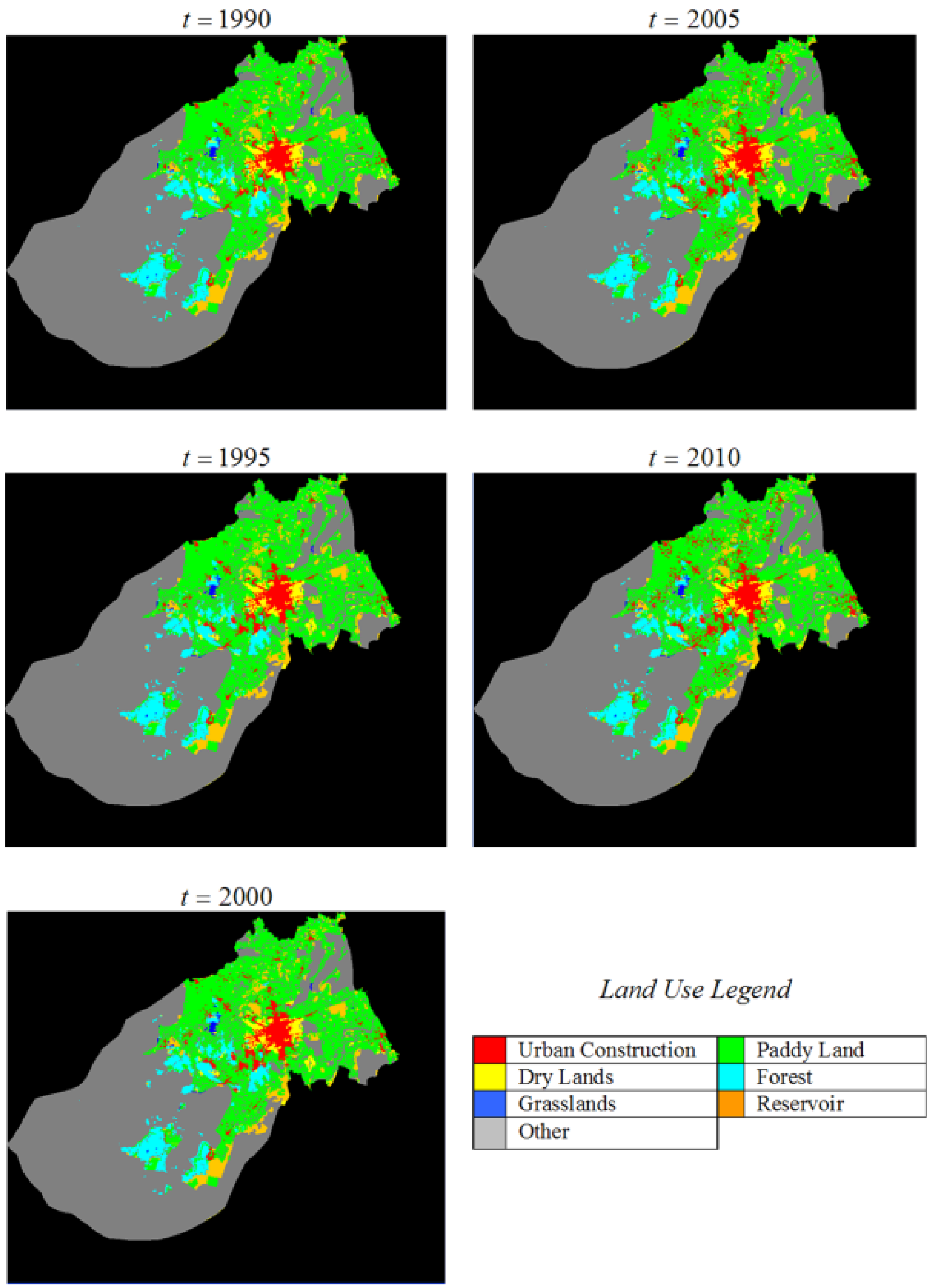


**Figure 7.** Snapshots of urban development through the micro-time periods 1990 to 2010 (For the complete animation, see http://www.casa.ucl.ac.uk/desakota/.)

important to explanation but are often, indeed usually, absent from more parsimonious model structures. The new quest for generative modeling in the social sciences is illustrated rather well in the model developed here (Epstein 1999). The idea that we need to demonstrate how our assumptions can generate plausible outputs is encapsulated in the idea of growing our systems through various forms of dynamics, temporal and otherwise. As Page (2003, 344) notes: "the generative claim that 'if you didn't grow it, you didn't show it' should be ignored at our peril . . ." The example of desakota is rather a good test bed on which to illustrate this argument.

We are planning a number of extensions of this model. We need to develop a much stronger link from the global to the local and vice versa which will be based on strengthening the feedback loop between development and socioeconomic drivers illustrated in Figure 1. We need to generate several different types of agents at different scales for the rural and urban regimes in our model and we need to consider transitions other than those between rural and urban. In this way, we plan to extend our model to examine impacts on the environment and the extent to which the kind of desakota appearing in China is sustainable. We intend to calibrate the model at the fine cellular scale and to develop strategies for multilevel calibration, a relatively new feature of agent-based modeling occasioned by our use of more than one scale of agent. And last but not least, we intend to improve the detailed measurement and simulation of accessibilities in the model relating the allocation process to capacities on land as well as its suitability.

## Acknowledgments

The authors wish to thank The Center for Ecological Research, Institute of Botany, Chinese Academy of Sciences, for partial financial support of this research through The One Hundred Scholars—Distinguished Overseas Scholar Funds, and Dr. Guangjin Tian at the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, for data preprocessing.

## References

Batty, M. 1976. *Urban modelling: Algorithms, calibrations, predictions*. Cambridge, U.K: Cambridge University Press.

———. 2005. *Cities and complexity: Understanding cities through cellular automata, agent-based models, and fractals*. Cambridge, MA: MIT Press.

Chen, H. 1988. Jiangsu expands its economic horizons. *Beijing Review* 23:18–25.

Collier, N., T. Howe, and M. North. 2003. Onward and upward: The transition to Repast 2.0. *Proceedings of the First Annual North American Association for Computational Social and Organizational Science Conference*. Electronic Proceedings, Pittsburgh, PA: 241–68; http://repast.sourceforge.net/ (last accessed 9 August 2005).

de Almeida, C. M., A. M. V. Monteiro, G. Câmara, B. S. Soares-Filho, G. C. Cerqueira, C. L. Pennachin, and M. Batty. 2005. GIS and remote sensing as tools for the simulation of urban land-use change. *International Journal of Remote Sensing* 26:759–74.

Epstein, J. M. 1999. Agent-based computational models and generative social science. *Complexity* 4:41–60.

Heikkala, E., T. Shen, and K. Yang. 2003. Fuzzy urban sets: Theory and application to desakota regions in China. *Environment and Planning B* 30:239–54.

Jones, S. 2002. Social constructionism and the environment: Through the quagmire. *Global Environmental Change* 12:247–51.

Kirkby, R. 2000. *Urban land reform in China*. Oxford, U.K.: Oxford University Press.

Kloosterman, R. C., and S. Musterd. 2001. The polycentric urban region: Towards a research agenda. *Urban Studies* 38:623–33.

Li, A., M. Linderman, J. Qi, A. Shortridge, and J. Liu. 2005. Exploring complexity in a human-environment system: An agent-based spatial model for multidisciplinary and multiscale interaction. *Annals of the Association of American Geographers* 95:54–79.

Li, X., and A. Yeh. 1998. Principal components analysis of stacked multi-temporal images for monitoring rapid urban expansion in the Pearl River Delta. *International Journal of Remote Sensing* 19:1501–18.

———. 2000. Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographic Information Science* 14:131–52.

Lieberthal, K. 1995. *Governing China: From revolution through reform*. New York: Norton.

Lin, C. 1997. Transformation of a rural economy in the Zhujiang Delta. *The China Quarterly* 149:57–80.

———. 2001. Evolving spatial form of urban-rural interaction in post-reform China: A case study of the Pearl River Delta. *The Professional Geographer* 53:56–70.

Liu, J., M. Liu, and X. Deng. 2002. The land-use and land-cover change database and its relative studies in China. *Journal of Geographical Sciences* 12:275–82.

Marton, A. 2000. *China's spatial economic development: Restless landscapes in the Lower Yangtze Delta*. New York: Routledge.

McGee, T. 1989. Urbanisasi or kotadesasi? Evolving patterns of urbanization in Asia. In *Urbanization in Asia*, ed. F. J. Costa, A. K. Dutt, L. J. C. Ma, and A. G. Noble, 93–108. Honolulu: University of Hawaii Press.

———. 1991. The emergence of desakota regions in Asia: Expanding a hypothesis. In *The extended metropolis: Settlement transition in Asia*, ed. N. B. Ginsberg, B. Koppel, and T. G. McGee, 3–25. Honolulu: University of Hawaii Press.

———. 1998. Five decades of urbanization in Southeast Asia: A personal encounter. In *Urban development in Asia: Retrospect and prospect*, ed. Y. Yeung, 55–91. Shatin, Hong Kong SAR: Chinese University of Hong Kong.

Oi, J. C. 1999. *Rural China takes off: Institutional foundations of economic reform*. Berkeley: University of California Press.

Page, S. E. 2003. Agent based models. In *The new Palgrave dictionary of economics and the law: Volume 2*, ed. P. Newman, 341–49. New York: Palgrave Macmillan.

Parker, D. C., S. M. Manson, M. A. Janssen, M. J. Hoffmann, and P. Deadman. 2003. Multi-agent systems for the simulation of land-use and land-cover change: A review. *Annals of the Association of American Geographers* 93: 314–37.

Reese, H. M., T. M. Lillesand, D. E. Nagel, J. S. Stewart, R. A. Goldmann, T. E. Simmons, J. W. Chipman, and P. A. Tessar. 2002. Statewide land cover derived from multiseasonal Landsat TM data: A retrospective of the WISCLAND project. *Remote Sensing of Environment* 82:224–37.

Sui, D. Z., and H. Zeng. 2000. Modeling the dynamics of landscape structure in Asia's emerging desakota regions: A case study of Shenzen. *Landscape and Urban Planning* 75:1–16.

Suzhou Municipality Statistical Bureau. 2003. *Statistical yearbook of Suzhou Municipality—2002*. Suzhou Municipality: Jiangsu Province.

Tan, K. 1986. Revitalized small towns in China. *Geographical Review* 76:138–48.
Tang, W., and H. Chung. 2000. Urban-rural transition in China: Beyond the desakota model. In *China's regions, polity, and economy: A study of spatial transformation in the post-reform era*, ed. S. Li and W. Tang, 275–308. Hong Kong: The Chinese University Press.
Urry, J. 2003. *Global complexity*. Cambridge, U.K.: Polity Press.
Wang, M. Y. 1997. The disappearing rural-urban boundary? Rural socioeconomic transformation in the Shenyang-Dalian region of China. *Third World Planning Review* 19:229–250.
———. 1998. *Mega urban regions in China*. Beijing: Ocean Press.
Wei, Y. 2002. Beyond the Sunan model: Trajectory and underlying factors of development in Kunshan, China. *Environment and Planning* A 34:1725–47.
Wilkie, D. S., and J. T. Finn. 1996. *Remote sensing imagery for natural resources monitoring: A guide for first-time users*. New York: Columbia University Press.
Wuxian City Statistical Bureau. 2001. *Statistical yearbook of Wuxian City: 2000*. Wuxian, Jiangsu Province, China: Wuxian City Government.
Xie, Y., and M. Batty. 2005. Integrated urban evolutionary modelling. In *Geodynamics*, ed. P. M. Atkinson, G. M. Foody, S. E. Darby, and F. Wu, 273–93. Boca Raton, FL: CRC Press.
Xie, Y., M. Yu, Y. Bai, and X. Xing. 2006. Ecological analysis of an emerging urban landscape pattern: Desakota, a case study in Suzhou, China. *Landscape Ecology* 21:1297–1309.
Xie, Y., M. Yu, G. Tian, and X. Xing. 2005. Socio-economic driving forces of arable land conversion: A case study of Wuxian City, China. *Global Environmental Change* 15:238–52.
Yao, S. 1992. *Zhongguo de Chengshi Qun* [The urban agglomerations of China] Hefei, PRC: The Science and Technology University of China Press.
Yeung, Y., and Y. Zhou. 1991. Human geography in China: Evolution, rejuvenation and prospect. *Progress in Human Geography* 15:373–94.

*Correspondence:* Department of Geography and Geology, Eastern Michigan University, 205 Strong Hall, Ypsilanti, MI 48197, e-mail: yxie@emich.edu (Xie); Centre for Advanced Spatial Analysis, University College, London, Gower Street, London, WC1E 7HB, England, e-mail: MBatty@geog.ucl.ac.uk (Batty); Department of Computer Science, Eastern Michigan University, Ypsilanti, MI 48197, e-mail: kzhao@emich.edu (Zhao).