# Exploring Complexity in a Human–Environment System: An Agent-Based Spatial Model for Multidisciplinary and Multiscale Integration

Li An,* Marc Linderman,* Jiaguo Qi,** Ashton Shortridge,** and Jianguo Liu***

*Department of Fisheries and Wildlife, Michigan State University
**Department of Geography, Michigan State University
***Department of Fisheries and Wildlife and Center for Systems Integration and Sustainability, Michigan State University

Traditional approaches to studying human–environment interactions often ignore individual-level information, do not account for complexities, or fail to integrate cross-scale or cross-discipline data and methods, thus, in many situations, resulting in a great loss in predictive or explanatory power. This article reports on the development, implementation, validation, and results of an agent-based spatial model that addresses such issues. Using data from Wolong Nature Reserve for giant pandas (China), the model simulates the impact of the growing rural population on the forests and panda habitat. The households in Wolong follow a traditional rural lifestyle, in which fuelwood consumption has been shown to cause panda habitat degradation. By tracking the life history of individual persons and the dynamics of households, this model equips household agents with "knowledge" about themselves, other agents, and the environment and allows individual agents to interact with each other and the environment through their activities in accordance with a set of artificial-intelligence rules. The households and environment coevolve over time and space, resulting in macroscopic human and habitat dynamics. The results from the model may have value for understanding the roles of socioeconomic and demographic factors, for identifying particular areas of special concern, and for conservation policy making. In addition to the specific results of the study, the general approach described here may provide researchers with a useful general framework to capture complex human–environment interactions, to incorporate individual-level information, and to help integrate multidisciplinary research efforts, theories, data, and methods across varying spatial and temporal scales. *Key Words: Agent-based modeling, complexity theory, multidisciplinary and multiscale integration, household socioeconomics and demographics, giant panda conservation.*

Complex human–environment interactions have been increasingly attracting the attention of researchers with different backgrounds and research purposes. On the one hand, characterizing the environment and the complex role that human actions play within it is challenging. This is partially due to the inherent complexity of the processes. The accumulated impact of individual decisions made by dozens, hundreds, or millions of people is the immediate cause of human-induced environmental change. On the other hand, these individual actions are shaped by the particular social, political, economic, and environmental frameworks within which they occur. These frameworks change through time as conditions change. Furthermore, the imprint of these activities varies throughout space and across different spatial scales.

The science of complexity has provided key theoretical contributions and techniques for environmental modelers wrestling with these challenges (Flake 1998). It is concerned with the manner in which fundamental processes can lead to emergent phenomena or behaviors in complex adaptive systems (CAS), focusing on many kinds of complexities such as hierarchical structures, feedback, self-organization, scaling, and time lags (Malanson 1999). Levin et al. (1997) provide a broad overview of approaches to considering complexity in ecosystem modeling. They describe several advantages:

- Incorporation of substantial local and individual characteristics
- Recognition of the stochastic nature of complex systems
- Explicit characterization of the impact activities at one scale have on patterns at another

However, the challenge is also a technical or implementational one. How can researchers integrate data and models to deal with these complex processes? A variety of approaches have been adopted; a general overview of these was recently published in this journal (Parker et al. 2003) and is only briefly reviewed here. Geographers and other human–environment modelers often turn to Geographic Information Systems

*Annals of the Association of American Geographers,* 95(1), 2005, pp. 54–79 
Initial submission, October 2003; revised submission, June 2004; final acceptance, July 2004
Published by Blackwell Publishing, 350 Main Street, Malden, MA 02148, and 9600 Garsington Road, Oxford OX4 2DQ, U.K.

(GIS) to assist in data management and modeling of spatially explicit variables. GIS is a powerful tool to capture, store, manipulate, and analyze spatial data, and it has been extensively used in studying human–environment interactions. The data models employed by common GIS are inherently static, however; they do not handle time well, nor do they capture functions or dynamic processes effectively (Peuquet 1999). Environmental modelers usually resort to externally implemented methods to handle advanced modeling problems (although interesting integrated implementations exist, e.g., the PCRaster platform discussed by Wesseling et al. 1996). These methods include multivariate spatial models (e.g., Seto and Kaufmann 2003), Markov chain analysis (e.g., Brown, Pijanowski, and Duh 2000; López et al. 2001), and cellular automata (e.g., Batty, Xie, and Sun 1994; Li and Yeh 2002; Malanson 2002). In particular, cellular automata (CA) models have been shown to be powerful in modeling many ecological processes because, as a bottom-up approach, they have a better capacity than GIS overlay or map algebra functionality to capture and represent local interactions that give rise to global complex patterns (e.g., Li and Reynolds 1997; Clarke and Gaydos 1998). CA models, however, face challenges in simulating human decision making and capturing feedback elegantly (Parker et al. 2003).

It may be worth highlighting a few of these diverse approaches employed by scientists with similar research problems and data to those of the present study. These studies, like our own, seek to model relationships between social factors and landscape change. Pan et al. (2004) report on work regressing landscape pattern metrics with data collected in farmstead surveys in Ecuador. Geographic factors, including extent and spatial scale, are shown to play an important role in model results. Walsh et al. (1999) demonstrate the effect of spatial resolution on multiple-regression models relating population per unit of cultivated land to six independent social and physical variables for a region in Thailand. Both studies integrated social data collected with surveys with land-cover and other spatial data. The same Ecuadorian study site employed by Pan et al. (2004) was also subject to a cellular-automata-based model in another study (Messina and Walsh 2001). In that model, land-cover change over time is modeled using patterns observed in data collected in five time intervals over three decades. Rules identifying the probability of cell state change based on its neighbors are developed, leading to predictive models of land-cover dynamics. While the goals and data of these different studies are similar, very different methods are employed. The first two use multivariate techniques to identify overall relationships between land cover and human processes. The third employs a bottom-up approach in which individual cells change state over time according to probabilities associated with their properties and those of neighboring cells.

Agent-based modeling (similar to individual-based modeling in many ecological studies) is another bottom-up methodology that has been specifically employed to deal with complexity, especially when coupled with GIS. The research reported in this article utilizes this approach. Agent-based modeling (ABM) predicts or explains emergent higher-level phenomena by tracking the actions of multiple low-level "agents" that constitute, or at least impact, the system behaviors. Agents usually have some degree of self-awareness, intelligence, autonomous behaviors, and knowledge of the environment and other agents as well; they can adjust their own actions in response to environmental changes (Lim et al. 2002; Parker et al. 2003). The concepts underlying ABM are similar to those of the object-oriented programming (OOP) paradigm in computer science, and ABM models frequently employ object-oriented programming languages like C++ and Java.

Unlike procedural programming, for which data and operations on the data are separated, object-oriented programming groups operations and data (or behavior and state) into modular units called objects and lets the user combine objects into a structured network and form a useful program (Larkin and Wilson 1999). Figure 1(a) is an illustration of an object with operations (called methods) and data bound together. The strengths of OOP lie in its modularity, software reusability, and its separation between interface and implementation. Modularity reduces programming complexity by dividing code into relatively separated "parts" or modules, each with different functional focuses. Software reusability

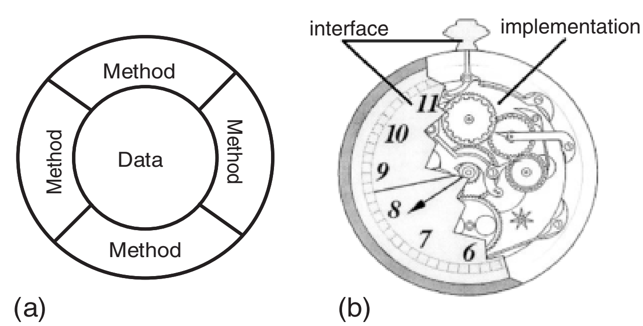


**Figure 1.** (a) An object in object-oriented programming with data and operations (methods) combined and (b) an interface and implementation of object-oriented programming (NeXT Software, Inc. 1992).

means that one piece of code, when defined and tested, can be reused as many times as possible. The separation between interface and implementation hides technical details inside the system surface, such as the parts in a clock and how these parts interact with each other (Figure 1(b); NeXT Software, Inc. 1992). The "implementation" feature (technical details) makes the system work well. A user-friendly interface running above this provides simple data input, output, and display functions so that other objects (or users) can call or use them.

A set of agent-based modeling tools with the above OOP features is readily available for use, e.g., Swarm, RePast, NetLogo, Ascape, and StarLogo.[1] Swarm may be one of the most powerful and comprehensive toolkits (Najlis, Janssen, and Parker 2001). Swarm is a software package for multiagent simulation of complex systems publicly available under GNU licensing terms. Originally developed at the Santa Fe Institute and now maintained by the Swarm Development Group, Swarm decomposes emergent phenomena at a certain level into attributes and actions of collections (swarms) of concurrently interacting agents at lower levels. Swarm allows for a hierarchical structure for agent organization and management, which means a higher-level agent in the hierarchy can include and manipulate a number of lower-level agents and their actions (Minar et al. 1996).

ABM alone does not address complexity (e.g., spatial heterogeneity, structural hierarchy) well, however. A growing number of efforts to integrate ABM with traditional approaches such as equation-based models and GIS have been made in the environmental modeling arena (e.g., Bian 1997; Berger 2001; Gimblett 2002; Jiang and Gimblett 2002), but relatively few studies have been implemented to examine the complex manner in which the accumulation of individual decisions, as affected by social/political factors and economic conditions, may affect the biophysical environment across a range of spatial and temporal scales. Here we address this critical topic with the following specific objectives: (1) linking spatial patterns and temporal processes by capturing complexity (e.g., heterogeneity, nonlinearity, feedback, and time lag) via ABM in a coupled human–environment system; (2) constructing a framework to integrate data and/or methods across disciplines, spatial/temporal scales, and aggregation levels; and (3) providing an effective policy-analysis tool for biodiversity conservation in relation to low-level anthropogenic (e.g., household life history) and environmental (e.g., spatially varying forest volume and growth rate) characteristics and relationships.

This article is fundamentally about integration: the integration of social and environmental drivers, the integration of diverse modeling techniques, the integration of fundamental knowledge generation with the policy implications of that knowledge, and the integration of data about processes operating across a range of spatial and temporal scales. It addresses these issues within an explicitly geographical framework and takes advantage of spatial tools that provide a practical modeling environment for conducting this complex analysis. We recognize that the project's technical architecture might divert attention from the project's specific purpose: to represent relationships between giant panda habitat loss in the Wolong Nature Reserve (China) and local household dynamics as affected by the social, economic, and political context. For this reason, a variety of validation efforts are employed to assess the fidelity of model results and their sensitivity to input parameters.

## Methods

### Study Area

An excellent site to conduct research with the above objectives is Wolong Nature Reserve in China (Figure 2) for the following reasons: (1) it is recognized as a globally significant biodiversity conservation site; (2) much is

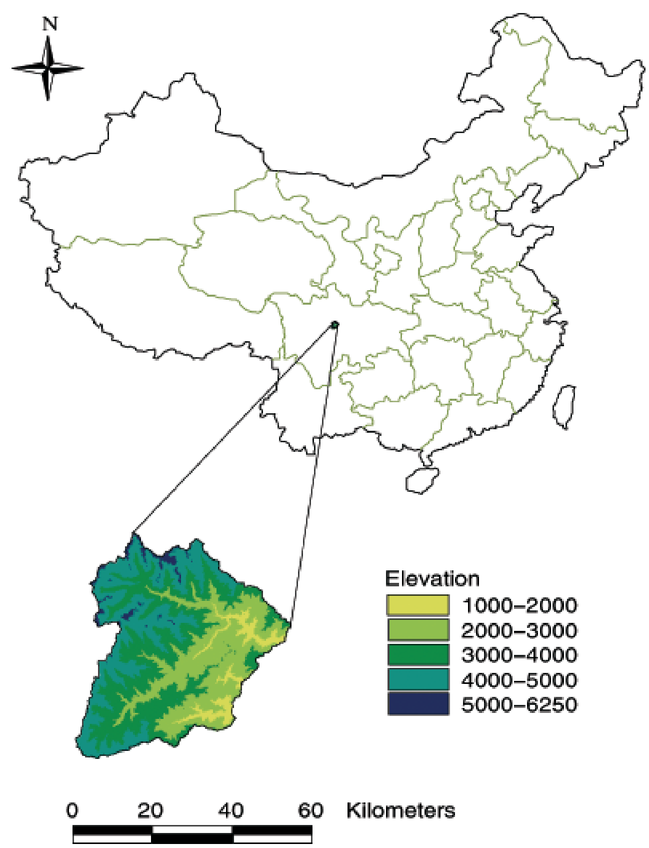


**Figure 2.** The location and elevation (m) of Wolong Nature Reserve in China.

known about the biology and physical environment of the giant panda; (3) human impact on panda habitat is a serious problem in the reserve; and (4) the complexity of the problem necessitates an integrated approach. We discuss each of these points in turn in the following paragraphs.

First, there is broad consensus that the Wolong Nature Reserve is of global significance. Established in 1975 for conserving the endangered giant panda (*Ailuropoda melanoleuca*), the reserve is within one of the twenty-five global biodiversity hotspots (Myers et al. 2000). Over 2,200 animal/insect species and more than 4,000 plant species (Wolong Administration 1987) cohabit with the giant panda in a diverse biophysical environment occupying approximately 2,000 $km^2$. Species richness, an important indicator of biodiversity, is directly linked to availability of associated habitat types based on the utter dependence of organisms on an appropriate environment (Ehrlich and Wilson 1991), especially for contexts associated with island biogeography. A famous example is the species–area relationship, i.e., $S = cA^z$, where S is the number of species that occur in a region with area A, and c and z are relevant constants (MacArthur and Wilson 1967). Thus, protecting habitat is a necessary step toward conserving any single type of organism. The implications for Wolong are clear: conserving panda habitat means that both the internationally renowned and endangered giant panda and the less widely known species that live within and comprise that habitat can be protected.

Second, we have collected extensive socioeconomic and environmental (e.g., remote-sensing) data about the Wolong area through our intensive fieldwork from 1998–2002. Extensive research efforts have been invested on giant panda biology, ecology, and habitat studies, such as the relationship between the giant panda and bamboo forest, the canopy cover that serves the giant panda as shelter, and the understory bamboo that serves as a primary food source (e.g., Schaller et al. 1985; Liu, Ouyang, Tan, et al. 1999).

Third, the rural human population in the reserve threatens panda habitat. Although Wolong enjoys high domestic standing as a "flagship" reserve in China with considerable domestic and international financial and technical support, the reserve also supports a substantial human population that is growing rapidly, with an even more rapid increase in the number of households (Liu, Daily, et al. 2003). The population (approximately 4,400 local residents in 2000) is comprised of four ethnic groups: Han, Tibetan, Qiang, and Hui, following a traditional rural lifestyle (Liu, Ouyang, Tan, et al. 1999; An et al. 2001). In spite of the enormous time, energy, and increasing difficulty involved in collecting fuelwood (mostly due to the shrinking forest area and the extremely rugged mountainous terrain), the majority of households in Wolong cut wood from the surrounding forests to cook and heat their homes. Although electricity is available in the reserve,[2] only a small proportion of the households use electricity for cooking and heating; the primary use of electricity is for lighting and electronic appliances (An et al. 2001, 2002). Assuming that an average household consumes 15 $m^3$ of fuelwood per year (An et al. 2001) and that an average hectare of beech (*Fagus*), oak (*Quercus*), birch (*Betula*) and poplar (*Populus*) forest contains 80 $m^3$ of fuelwood (Yang and Li 1992), then a $90 \times 90$ m pixel of mixed forest can sustain one household's fuelwood demand for about four years.

Despite abundant economic incentives (e.g., a lower agricultural tax) and policies (e.g., prohibiting some tree species from being harvested) implemented by the reserve administration, the past two decades have still witnessed a continued increase in annual fuelwood consumption (from 4,000 $m^3$ to 10,000 $m^3$ over the past two decades), contributing to a total reduction of over 20,000 ha of panda habitat (Liu, Ouyang, Taylor, et al. 1999). Degradation of forests comprising panda habitat undoubtedly accounts for part of the documented decrease in the Wolong panda population in recent decades: from 145 individuals in 1974 (Schaller et al. 1985) to 72 in 1986 (China's Ministry of Forestry and World Wildlife Fund 1989). This degradation may arise from some combination of ineffective enforcement of existing policies, the common-property nature of the forests, and the difficulty in monitoring, given the rugged landscape of the reserve. All these factors make biodiversity of the Wolong Nature Reserve highly sensitive to human activities and policy changes and threaten the long-term viability of the ecosystem to support wild populations of giant panda and other coexistent species.

Last, the complexity underlying various ecological, socioeconomic, and demographic processes has necessitated interdisciplinary research involving a range of spatial and temporal scales. Piecemeal approaches fail to account for many important factors; nonspatial approaches cannot characterize the critical role location and relative position play for interactions between humans and the environment; evaluating habitat in single time periods ignores the dynamic nature of the processes that drive habitat destruction. Existing research efforts to characterize the relationship between panda habitat, fuelwood consumption, and human socioeconomic/ demographic factors (e.g., An et al. 2001, 2002, 2003; Linderman et al. forthcoming) are inadequate because they account for too few aspects of Wolong's human–environment system.

For instance, the fuelwood model (An et al. 2001) links the household-level fuelwood demand to household demographic and socioeconomic factors, but it is basically aspatial and unable to link household fuelwood demand to its impact on the landscape. The electricity-demand model suffers from the same limitation in that it links the probability of switching to electricity to a set of socioeconomic, demographic, and geographic factors (An et al. 2002), but cannot identify the impact of this change to specific locations, or to the forests of Wolong. In addition, household demographics are exogenous in these models; no attempt to characterize the dynamic nature of the household was made. An, Mertig, and Liu (2003) develop a household formation model to address this deficit. This model attempts to characterize the movement of young adults from their parental homes to establish their own households in the Wolong region. However, it does not address the spatial location of households whose demands for fuelwood and electricity are determined as mentioned above. More importantly, these models alone still cannot effectively address the complexity in this coupled human–environment system: nonlinear interactions, cross-scale (spatial and temporal) data, feedback, and time lags between different subsystems. The model described in the subsequent paragraphs is designed to overcome these limitations and to accomplish the objectives discussed in the introduction. It is intended to be a comprehensive tool to (1) employ a valid representation of the Wolong landscape and incorporate its spatial heterogeneity; (2) incorporate the demographic dynamics of households and individuals situated within this landscape; and (3) link fuelwood/electricity demand to the changing characteristics of the landscape and households as established in (1) and (2).

## Conceptual Model

With an excellent study site and a wealth of data, we have developed an Integrative Model for Simulating Household and Ecosystem Dynamics (IMSHED). As in many other studies (e.g., Deadman et al. 2001; Liu et al. 2003; Rindfuss et al. 2003; Walsh et al. 2003), households are chosen to be the fundamental unit for local people's decision making and behavior in relation to consumption and production of local resources. The conceptual framework is illustrated in Figure 3. The model consists of three major components: household development, fuelwood demand, and fuelwood growth and harvesting. Each of these components is addressed in turn within the remainder of this subsection.

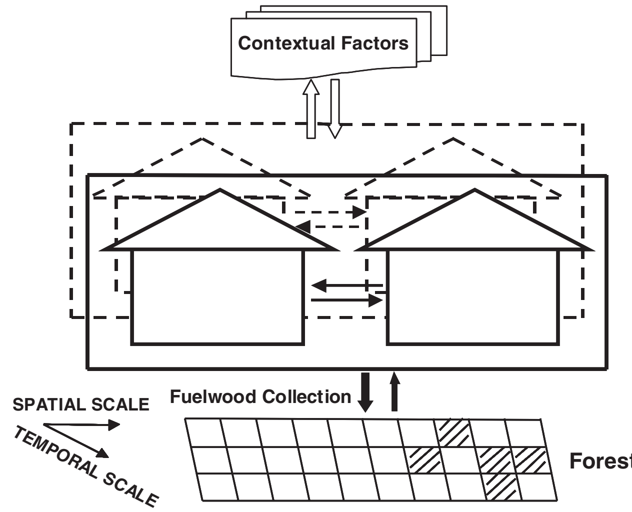


**Figure 3.** The conceptual framework of IMSHED.

First, the houses within the dashed box represent households in Wolong at a particular point in time, while the solid houses represent households in the same landscape but at a later time. Household change during this interval might involve one or more of many events: households may increase or decrease in size, dissolve, or relocate; new households may be initiated as individual persons go through their life history, the details of which are illustrated in the section "Demographic Submodel." The formation of a new household, which depends upon the actions of young adults who have reached marriageable age, is explained by a set of psychosocial factors using a structural equation modeling approach based on our 220-household in-person interview data (An, Mertig, and Liu 2003). The findings from this model and our in-person interviews indicate that the intention to leave the parental home and establish a new home (the variable "leave-home intention" in Tables 2 and 3) is determined or influenced by his/her sibling status (whether he/she has siblings, and whether he/she is the youngest son or daughter (Figure 5(b)), the availability of land, and the behavior of parents and peers. This intention is the major predictor of new household formation for young adults after marriage.

Second, at each snapshot over time, the fuelwood demand of a household may be modeled as a function of household size, whether there is a senior person (60+ years old) in the household, and the area of cropland based on our extensive fieldwork and interview data (An et al. 2001). Presence of seniors is positively correlated with fuelwood demand because Wolong households with seniors use more firewood for longer heating time and higher home temperatures throughout the winter. Cropland area is also positively correlated with fuelwood demand because larger acreages are associated with hog farming. These farms grow potatoes and corn, the bulk of which are cooked in fuelwood stoves and fed to pigs. This phenomenon may arise from a traditional belief that cooked food makes pigs grow faster and healthier. The pork or bacon thus obtained is consumed and often sold to tourists and local restaurants for cash. This explains why there is a link between cropland area (for corn and potatoes) and fuelwood demand. The probability of switching from fuelwood to electricity for each household is predicted by a number of socioeconomic and demographic factors using a discrete choice logit model (An et al. 2002), in which a household's decision to switch to electricity or continue to use fuelwood is regressed against the age, gender, and education of the household head, annual household income, electricity price, outage frequency, voltage levels, and a few other factors.

Last, the forests on the landscape, given no human interference, grow and die by themselves. The interactions between humans and the environment are realized through fuelwood collection, as shown by the two lower vertical arrows in Figure 3. Harvesters from local households, given a certain amount of fuelwood demand derived from the model described above, travel to the most convenient set of locations (pixels in a raster grid) to cut fuelwood. Increasing distance for fuelwood collection may, in turn, reduce local households' fuelwood demand and encourage the substitution of electricity. Important physical and social factors (context factors in Figure 3), including distance, elevation, policy decisions, and law enforcement, exert impacts on many processes such as demand for fuelwood and electricity.

## Major Agents/Objects

Major agents/objects include individual persons, households, pixels (square grid cells representing homogenous units of the landscape), and some management agents helping us manage various objects or tasks (e.g., a list containing many agents of the same type). We only describe the major agents, starting from definitions of the corresponding classes. A collection of management agents handle mostly technical details in Swarm, and are not further elaborated on here.

***Person.*** This class includes attribute variables such as personal ID, age, ID of the household that she/he belongs to, education level, gender, personal IDs of his/her mother and father, and his/her marital status. Also, the Person class has a few variables associated with childbirth: birth plan (how many children this person would have), birth interval (number of years between two consecutive children), marriage year (the year the person gets married), birth year (the year the person gives birth to a child), and first-child interval (the time between the marriage and the birth of her first baby). We will discuss the use of these variables in the section "Demographic Submodel."

The actions (called methods in Java) include: give birth, die, grow, marry, move out of a household, move into a household, and cut fuelwood. Some other detailed actions (e.g., set the value for an attribute variable) specific to Java-Swarm programming are not discussed here.

***Household.*** This class includes attribute variables such as household ID (consistent with that defined in the Person class), $x$ coordinate, $y$ coordinate, cropland area, household income, electricity price, outage level, voltage level, location of the household (Wolong Township for 0 and Gengda Township for 1; this is consistent with the dummy variable of location in the econometric model of An et al. 2002), distance of fuelwood transportation, and probability of switching from fuelwood to electricity. All the variables needed for predicting electricity and fuelwood demand are defined here because these demands are determined at the household level.

The actions in this class include formation of a new household, dissolution of a current household (when the number of people who belong to this household goes to zero), or an increase or decrease in household size (i.e., number of people in a household). We assume that when a new household is established, the area within 90 m (a parameter) around it has to be deforested and becomes nonhabitat for the pandas. This parameter is set in accordance with the fine spatial resolution that will be discussed in the data section. As a general rule, the value should be a multiple of the finest spatial resolution employed (90m; another resolution is reported on in this study) unless only the pixel containing the household is to be deforested.

***Pixel.*** The Pixel class contains all the information necessary for simulation of landscape changes. It contains attribute variables such as the $x$ coordinate, $y$ coordinate, elevation, slope, land-cover type, forest age, and forest volume (for non-forest pixels, this volume is automatically set to zero). Methods for the Pixel class include land-cover change (primarily from forest to nonforest), forest age increase, and volume growth (in forest pixels with tree species). We assume that forest volume reduction is primarily caused by fuelwood collection, because other factors such as forest fires and timber cutting are rare in the study region (M. Liu, personal communication).

## Data Collection, Preparation, and Integration

The performance and application of any spatial model depend substantially on the data available for parameterization, calibration, and validation. Our data fall into the three categories of spatial environmental data, demographic data, and socioeconomic data.

***Spatial environmental data.*** We have assembled a wealth of empirical data on both social and environmental factors and have built a database in both Microsoft Access and GIS (ArcInfo). Because remote sensing can provide views of the processes under study with adequate spatial extent, information detail, and temporal frequency (Herold, Goldstein, and Clarke 2003), we have used two time steps of high-quality, remotely sensed satellite images: Landsat TM (1997) and IKONOS (2000). We have conducted a supervised classification of the 1997 data based on 126 sample plots of $60 \times 60$ m, and validated them using a reserved set of sixty-three sample plots, resulting in an overall classification accuracy greater than 80 percent (shadow areas were classified as unclassified). We also have developed a digital elevation model of $30 \times 30$ m resolution, interpolated from one-hundred-meter-interval contour maps using the topogrid interpolation method in Grid ArcInfo. Based on a set of 313 Global Positioning System (GPS) points measured throughout the reserve, the DEM has a vertical accuracy of less than 50-m root mean square error (RMSE) and a standard deviation of approximately 37 m (in some most rugged areas the difference could be as large as 200 m).

Spatial coordinates of current household locations were obtained with Global Positioning System (GPS) measurement, and image analysis. The following steps illustrate the manner in which household locations were identified and linked with remotely sensed data and information from household surveys: (1) In the summer of 2001, we measured a total of fifty-nine households using a Trimble GPS unit with real time differential correction (Omnistar), for which positional accuracy was estimated to be within 2–3 m. Four IKONOS satellite images (with 1 m resolution) that cover most of the area of Wolong with human settlements were also obtained. (2) We printed out a set of IKONOS-derived maps with spatial resolution chosen so that households and their interrelationships were most easily identifiable. Using these maps, we visited each household and collected the demographic information described in the agents section above and linked that information to their spatial locations. (3) Using the coordinates of the fifty-nine households and fifty-five control points collected in the summers of 1998–1999 (Linderman et al. 2004) as control points, we georeferenced the four IKONOS images and recorded the coordinates of all the identifiable households. For households not identifiable in the IKONOS images, we used GPS to measure their coordinates. (4) Using the names of household heads as unique identifiers, we linked the demographic and socioeconomic data with the locational data (coordinates) on the IKONOS maps. Because nearly all Wolong houses are surrounded by their apportioned land, and land cannot be sold or traded in China (see the section "Demographic Submodel"), we did not record homestead field boundaries but assumed they were located immediately around the recorded house locations. For details of these processes, see Liu, An, et al. (2003).

Choosing an appropriate spatial resolution is a key challenge for this modeling effort. There is a trade-off between fidelity of spatial representation (an overly coarse cell resolution may mask some spatial variations) and efficiency (halving the resolution quadruples the amount of data and thus increases storage requirements and model execution times). We identified two resolutions to employ: 90 m and 360 m. For submodels requiring extensive human demographic factors (population size, number of households), the coarser (360 m) resolution was used. For submodels requiring or characterizing landscape characteristics (e.g., forest growth, distribution of panda habitats), the finer (90 m) resolution was used. Both resolutions were generated by resampling the 30 m raster DEM, slope, and Landsat TM-derived land-cover data; the methodology is reported by Linderman et al. (2004). Land-cover data were collapsed to nonforest (0), deciduous forest (1), conifer forest (2), and mixed forest (3). Processing took place in Erdas Imagine and ESRI ArcGIS, after which elevation, slope, and land-cover data were converted to ASCII text format for input to the Java-Swarm IMSHED model.

The fuelwood volume in each pixel is estimated according to the dominant tree species in that class. Class 1 consists of beech (*Fagus*), oak (*Quercus*), birch (*Betula*), and poplar (*Populus*), and the volume range is 60 to 100 m3/ha, with ages ranging from fifty to one hundred years old. Class 2 consists of fir (*Abies*), pine (*Pinus*), and spruce (*Picea*), and the volume is from 200 to 400 $m^3$/ha with ages ranging from 40 to 110 years old. Class 3 could be a mixture of any of these species and other woody ground cover; we set its volume from 125 (the average of lower bounds of Classes 1 and 2) to 250 $m^3$/ha (the average of upper bounds of Classes 1 and 2); and the age is set to be from forty to ninety years (Yang and Li 1992; Ouyang et al. unpublished data; Linderman et al. 2004).

***Demographic data.*** Socioeconomic data were obtained from a range of sources. Government data included the 2000 Population Census data of Wolong (Wolong Administration 2000), and the 1996 Wolong Agricultural census data (Wolong Administration 1996). Survey data were collected for approximately 1,000 households in Wolong, of which 220 were face-to-face interviews. Survey information included household economic status, social network (kinship relationship), and attitudes toward such issues as fertility. All these individual-based data, arranged by household, include personal ID, ID of the household that the person belongs to, gender, age, kinship relation to the household head, and other attributes of the "Person" and "Household" classes. These data cover all people in the reserve.

Each person (an object of Person class) keeps his/her father and mother IDs as attributes. In case the person's father or mother is unknown, dead, or not in the reserve, the value for the associated ID is set to zero. By doing so, we keep the kinship relations clear, which makes the simulations as realistic as possible. For instance, a brother in the single male list cannot "marry" his sister in the single female list by mistake because two people who share the same mother ID, father ID, or household ID are not allowed to "marry" each other.

These data can be employed to identify likelihoods for important household state changes: they are used to derive values for the parameters described in the following section. For example, consider the probability of in- and out-migration of young people from their parents' homes. Females emigrate from Wolong (0.28 percent); females immigrate to Wolong through marriage (0.19 percent); males emigrate from Wolong (0.043 percent); males immigrate to Wolong through marriage (0.043 percent). Since female in- and out-migration is relatively common, these parameters are employed in the demographic model. Parameters like these may not lead to changes in the amount of panda habitat over short time periods, but effects could be substantial at later times, and we examine their significance later in the section "Complexity Exploration."

***Socioeconomic data.*** The 1996 and 2000 demographic data sets identified in the previous section also contain some useful socioeconomic information, such as the cropland area for each household, which is very important in determining household fuelwood demand (An et al. 2001). However, primary socioeconomic data were obtained from our interviews of the 220 households. The primary data include current electricity prices, outage frequencies, and voltage levels, which are used in computing the probability to switch from fuelwood to electricity (see the section "Electricity demand"). Other questions in the same interview sessions have led to very useful information about what factors affect young adults' decisions about leaving their parental homes and establishing their own households after marriage. Those rules (to be described later in the section "Immigration and local movement through marriage"; also see Figure 5) about where to live after marriage are mainly based on these data.

## Demographic Submodel

All individual-based data were entered into an Access database and exported as text into IMSHED. The model keeps track of the life history of individuals (objects of Person class) as follows: persons may give birth or be born, die, get married, and move into or out of a household (subsequently, into or out of the reserve in some cases) through marriage. Out-migration occurs when local residents move out of the reserve, immigration occurs when people move into the reserve from other places, and local movement occurs when local residents move within the reserve (primarily through marriage). Households are affected by these changes: they increase or decrease in size, new households form, and some existing households dissolve.

***Death and out-migration.*** The death of each person is simulated through a random process. The likelihood of death for a person in a given year is in accordance with his/her age—0.00745 for people aged 0–5, 0.0009 for people aged 6–12, 0.00131 for people aged 13–15, 0.00196 for people aged 16–20, 0.00291 for people aged 21–60, and 0.05354 for people older than 60 (An et al. 2001). If a number drawn from a uniform distribution is less than the mortality rate associated with the person's age, he/she dies (as person on the left in Figure 4), and his/her spouse (if he/she has one) changes her/his marital status to "without spouse" while switching to the single male (or female) group; otherwise he/she survives the year.

Out-migration in Wolong is of two types: move-out through education and move-out through marriage;[3] the latter distinguishes between males and females. If a person survives, the model checks his/her age. If the age is between sixteen and twenty and the random number generator creates a number smaller than the college attendance rate for people in this age group (0.0192 for each of the five years; see An et al. 2001), then he/she goes to college and leaves the reserve (exits from the simulation in IMSHED) permanently. Otherwise, the person remains in the household for that year. The rationale for doing so is that nearly all of Wolong young

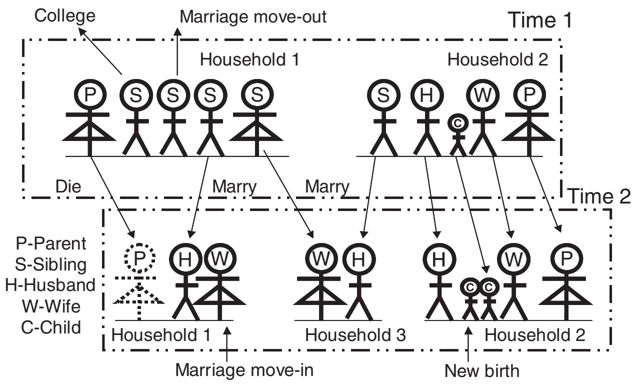


**Figure 4.** An illustration of individual-based demographic simulation.

people who leave for the college find work in cities after graduation and do not return to the rural life of the reserve. For each single person above twenty-two (the minimal age for marriage by law), he/she will move out of the reserve (exit from simulation in IMSHED) if the random number generator creates a number smaller than male/female's move-out through marriage probability (0.043 percent, and 0.28 percent, respectively).

***Immigration and local movement through marriage.*** Immigration is restricted due to Wolong's stand as a nature reserve for panda conservation. The only legal way for people outside the Reserve to move in and obtain permanent residence licenses (*Hukou*) is through marriages with local people. A very important decision associated with both immigration and local movement is to determine whether a newly married couple will initiate a new household or not. This is important in IMSHED because the efficiency of fuelwood consumption differs among households as household sizes change (Liu, Daily, et al. 2003). The following situations are included in IMSHED: (1) A local male brings an outside female into Wolong through marriage, and the decision process is illustrated in Figure 5(a). (2) A local female brings an outside male into Wolong through marriage. The decision process is similar to that in (1). Based on the findings of An, Mertig, and Liu (2003), the decision of whether to initiate a new household for these two people is: if (a) the female has no sibling, or (b) though she has siblings, all of them are females, and she is the youngest among them, then her husband and she will remain in

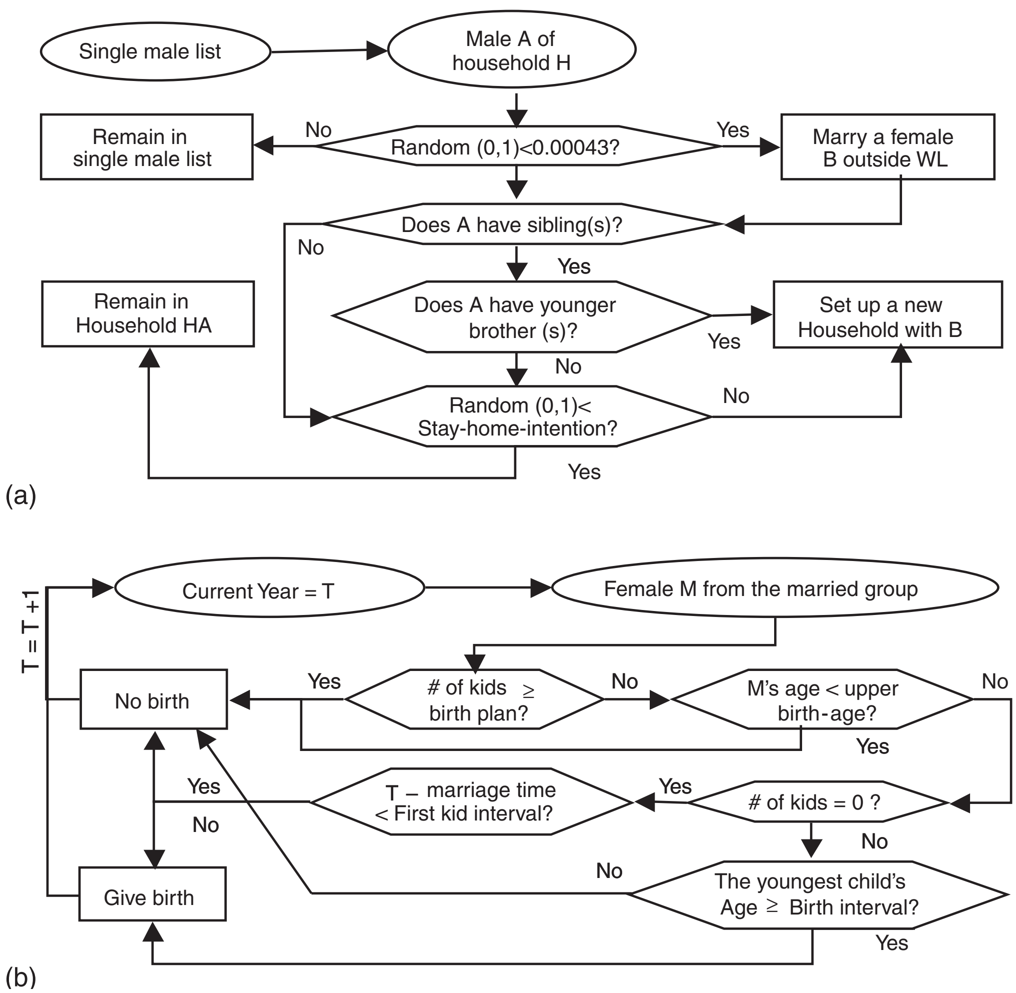


**Figure 5.** (a) Processes determining whether or not a new household is initiated when a local single male brings an outside female into Wolong through marriage. (b) Processes determining whether and when a household produces a child.

her original household with a probability of 0.58, a parameter subject to change. Otherwise, they will initiate a new household. (3) A local male marries a local female: When two local singles get married, if the husband (a) has no siblings, or (b) has only female siblings (sisters), or (c) is the youngest among male siblings, then the couple lives in the husband's original household with a probability of 0.58; otherwise check the sibling status of the wife. If she (a) has no siblings, or (b) she has only female siblings (sisters) and is the youngest among them, then the couple lives in the wife's original household with a probability of 0.58. Otherwise, the couple initiates a new household.

Based on the age and marital status of a person, a person is assigned to one of the four groups (lists in Java): (1) young group (all unmarried males and females less than twenty-two years old), (2) single male group (all single males over twenty-two years old, including the males whose spouses died), (3) single female group (all single females over twenty-two years old, including the females whose spouses died), and (4) married group (all females and males who have spouses with them). For example, if a male in the young group reaches twenty-two, he will move to the single male group; at some time, if he gets married, he will move to the married group. However, if for some reason his spouse dies, he moves back to single male group again and has the potential to get remarried, but the chance of doing so decreases as his age increases. The following equation is used to represent this relationship in accordance with our field observations:

$$\text{Rate of marriage at age } X = 0.35/(X-30)^{0.4}. \qquad (1)$$

***Childbirth.*** The event of childbirth only happens to females in the married group. For easier explanation, suppose that the woman under consideration is called M (indicating mother). As mentioned in the introduction of the Person class, each person has a birth plan that is used to set the number of children he/she may have. As indicated by Liu, Ouyang, Tan, et al. (1999), the number of children for each couple is 2.5. We use a binomial random variable Y to assign the number of children that M would have (Figure 5(b)). Since most families do not have more than five children (Wolong Administration 2000), we assume she would have 0, 1, 2, 3, 4, or 5 children with the probabilities of 0.03125, 0.15625, 0.3125, 0.3125, 0.15625, and 0.03125. The cumulative probabilities are 0.03125, 0.1875, 0.5, 0.8125, 0.96875, and 1, which are used to set probability intervals later. This is based on the classic problem of flipping a coin $n$ ($n = 5$ here) times and observing the number of heads above Y, where the probability of success (observing heads up) is 0.5, and Y is a random variable that could take values from 0, 1, . . ., to 5. From the binomial distribution, the average of Y is $n \times p = 5 \times 0.5 = 2.5$ (number of children per mother). The probabilities are computed by the following equation, where $p$ is the probability of "success."

$$\text{Prob.}\,(Y = y) = \binom{N}{y} \times p^{y} \times (1-p)^{(N-y)}. \qquad (2)$$

As an alternative childbearing model (overloading in object-oriented programming), we simply randomly choose a number between two integer bounds with equal probability. For other parameters, we set their values based on our field observations. Birth interval (age difference between two consecutive children) is randomly chosen between 1 and 6 years because the observed average birth interval is around 3.5 years. The first-child interval (the time between marriage of a couple and birth of their first child) is set to be 1 or 2 years with even probabilities. All these parameters are subject to change for different purposes, such as sensitivity or uncertainty analysis and policy design and test. We employed this model to simulate the childbearing for each female in the married group as illustrated in Figure 5(b), where the above parameters (e.g., birth plan, upper birth age, and birth interval) all affect her childbirth decision.

***Household dynamics.*** In accordance with all possible events for each individual, a household may decrease or increase in size, be initiated, or dissolve. When a new household is initiated, it is randomly assigned a site that is within a certain distance from its parental or original household, subject to two topographical restrictions of slope less than 37 degrees and elevation less than 2,610 m (He, Bearer, and Liu unpublished field data). This distance is controlled by a parameter with the default of 800 m based on our field observations. The new household is assigned a portion of the land from its parental household in proportion to its size rather than carved out from untenured land, which is based on the current Chinese land system. Farmers only have usufruct, and land can neither be traded nor developed without government permission because, based on China's constitution, the government and the collective organizations (quasi-governments) hold title to all land. The household responsibility system implemented in the late 1970s or early 1980s (the time for Wolong) assigned a certain amount of land to each rural household based on a set of criteria including household size and land quality, which has endured almost unchanged in spite of shifts in household sociodemographic factors such as household size. For this reason, new households founded by young married couples are often located in the immediate vicinity of one of the parental households. This fact explains why we did not map property boundaries for each household and why we did not employ a model to simulate land use change.

## Landscape Submodel

There are several landscape-oriented components of the model; one is a forest growth model that is used to determine fuelwood volume, two are concerned with path identification and selecting forested pixels for harvest, and one is a conventional GIS/cartographic model to identify panda habitat. Each will be discussed in turn. In all cases, the basic unit is the pixel, with resolutions of 90 m and 360 m, depending on the requirements of the specific model.

***Forest growth.*** Due to data limitations, we only consider forest growth using the simplified forest cover classification scheme presented in the section "Data Collection, Preparation, and Integration." According to Yang and Li (1992), the growth model for Class 1 is set to be 0.6, 0.8, and 1.0 $m^3$/ha/year if the forest is younger than twenty years, between twenty and eighty years, and older than eighty years. For Class 2, the rate is set to 2.0 $m^3$/ha/year regardless of the age. For Class 3 (mixed of Classes 1 and 2), the rate is set to 1.5. The maximal volumes for these three classes are set to be 350, 400, and 300; growth rate is set to zero when the volume of a pixel reaches its upper boundary.

***Path finding.*** Finding the path to collect fuelwood is one of the primary processes in landscape simulation. Aside from the land apportioned to a household (usually adjacent to the household), the vast amount of rural land in China (including forests) is accessible to the public unless otherwise specified or regulated. Although Wolong has some habitat regulation policies, as mentioned in the section "Study Area," their implementation was ineffective, and most forest could be regarded as an open resource (later, we use a parameter "house buffer distance" to represent a fuelwood restriction policy; see the section "Model Test"). Therefore, we simply consider the effects of topography and distance on the selection of routes to forested pixels, which may be visited by multiple households. In Figure 6(a), the household in the lower right corner needs to decide where to cut a certain amount of fuelwood, which has been determined by a number of socioeconomic and demographic factors and the probability of switching to electricity. Here we use a set of artificial intelligence rules in accordance with our in-site observations and interview data, such as the rule of limited viewing scope to be illustrated next.

To illustrate the function of the spatial fuelwood model, we consider the decision process underlying it (Figure 5(b)). (1) The fuelwood collector from the household has a limited geographical scope (the rule of limited viewing scope), so he/she only chooses among the forest pixels within the dashed window of size 5 × 5 (the window size is a parameter that may be adjusted by the user; 5 × 5 is only used for demonstration). (2) Within this window, only four pixels have forests, and the next step is to identify which pixel has the least cost to reach. Starting from the forested pixel, the fuelwood collector would deviate as little as possible from the direct path to each household pixel. We assume that his/her path-finding behavior is confined by the two southeast–northwest lines parallel to an assumed line cutting across the household and the pixel, where the distance between these two lines is a parameter. (3) Since he/she would not turn back while carrying a load of fuelwood, we assume that he/she goes northwest and

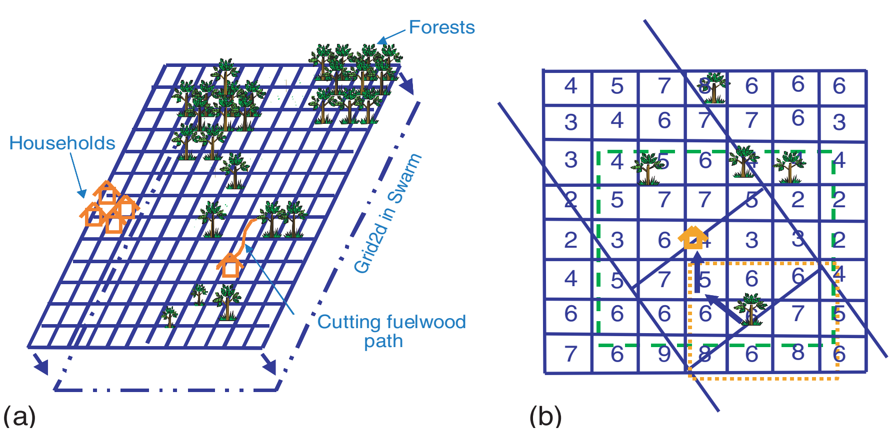


**Figure 6.** (a) An illustrative distribution of households and forests and (b) the procedure to find the least-cost path to collect fuelwood (the numbers in the pixels represent elevations).

does not go beyond the forest pixel; therefore, his/her path is also confined by the two lines that are perpendicular to the above two southeast–northwest lines. Within the area set by these four lines, he/she chooses the least-cost path. (4) Starting from the forest pixel, he/she chooses the pixel that has a lower elevation (or the one with the smallest elevation change if multiple pixels have lower elevations) and goes northwest as indicated by the arrow. For simplification, at some point if the household pixel is within one pixel to his/her standing pixel, he/she goes to the household directly. (5) He/she continues in this manner until the household pixel is reached.

Once the path is determined, its total length is calculated. Then it is adjusted by the slope between adjacent pixels along the route. The result indicates the cost of traversing the route between the home and a single forested cell. The same procedure is conducted on the remaining three forest pixels. Then, for that specific year, the collector chooses the pixel with the least-cost distance.

***Fuelwood site selection.*** For each household, we identified all of the forest pixels within a certain buffer distance (3,600 m as default based on unpublished field data from He, Bearer, and Liu) and put them on a list. We then calculate the cost distance between each forest pixel and the household using the method described above. We then group all these forest pixels into three categories of pixels that are 1,080 m, between 1,080 and 2,160 m, and over 2,160 m from the household, respectively. These three categories are chosen because interview data confirmed that 48.1 percent, 27.3 percent, and 24.6 percent, respectively, of the households collected fuelwood at sites corresponding to the above distances (He et al. [section II] unpublished field data). Therefore, if the random number generator creates a number smaller than 0.481, between 0.481 and 0.754, or greater than 0.754, the household will collect fuelwood in sites corresponding to the above distances.

We also apply artificial intelligence to the household under consideration: once it selects a pixel to collect fuelwood in a given year, it returns to the same pixel next year as long as the forest is still available. Doing so not only matches our field observations but also saves computer memory and time in computing and running the program, as fuel collection sites can be saved as an attribute of the household. Once this pixel is deforested, the household identifies a neighboring forested pixel.

The household "remembers" this distance (annually updated based on the location of the forest pixel), which affects its fuelwood demand by altering the household's perceived fuelwood collection proximity. Based on the travel distance employed in the previous year, proximity is set to be one of the three levels (short, moderate, and distant), corresponding to less than $X$ m, between $X$ m and $2.5X$ m, greater than $2.5X$ m, respectively ($X = 800$ m is the model default). Perceived proximity affects electricity demand and therefore provides a feedback effect. As fuelwood becomes harder to collect (sites become more distant), fuelwood demand drops, and households may substitute electricity for heating and cooking. The impact of the feedback effect is tested by varying the threshold $X$, a parameter called *perceived threshold distance* later in this article.

***Habitat determination.*** To identify habitat, we use the criteria of Liu, Ouyang, Tan, et al. (1999): any pixel with an elevation between 2,250 and 3,250 m, slope less than 30 degrees, and containing canopy forest is viewed as potential habitat. We combine two classes "highly suitable habitat" and "suitable habitat" from Liu et al. (2001) into one category, "habitat," and exclude "marginally suitable habitat" (Liu et al. 2001) in our simulation because we want to provide a conservative estimate of panda habitat.

## Socioeconomic Submodel

***Potential fuelwood demand.*** Fuelwood consumption is calculated on a household basis in IMSHED. According to An et al. (2001), the fuelwood demand from a household can be modeled as a linear function of (1) household size, (2) whether a household has a senior person, and (3) area cultivated in corn and potato. (These two crops are usually grown together, and both are primarily used as fodder. This area is largely in proportion [60–80 percent] to the total land area obtained from the parental household.) The first two factors, which are characterized in the demographic submodel, are checked annually, while area under corn–potato cultivation changes only if a new household is initiated from the parental household, in which case land is partitioned proportional to the sizes of the new household and the resultant parental household.

***Electricity demand.*** The fuelwood demand model just described does not consider the probability of switching from fuelwood to electricity and is, thus, incomplete. This switch probability is determined by the age, gender, and education of the household head, household annual income, current electricity price, outage frequency level, voltage level, perceived distance of fuelwood transportation, and location of the household under consideration (An et al. 2002). The following equation quantifies this relationship:

$$\begin{aligned} &Prob\,(\text{switch}|\mathbf{x}_{i,}\, \mathbf{z}_i,\, \alpha,\, \beta,\, \chi) \\ &\quad = \exp\left[\alpha + \beta(\mathbf{x}_i^1 - \mathbf{x}_i^0) + \chi \mathbf{z}_i\right] / \\ &\qquad \left[1 + \exp\left(\alpha + \beta(\mathbf{x}_i^1 - \mathbf{x}_i^0) + \chi \mathbf{z}_i\right)\right] \\ &\quad = 1/\left[1 + \exp\left(-\alpha - \beta(\mathbf{x}_i^1 - \mathbf{x}_i^0) - \chi \mathbf{z}_i\right)\right]. \end{aligned} \tag{3}$$

Vectors $\mathbf{x}_i^1$ and $\mathbf{x}_i^0$ represent the hypothetical and current electricity conditions (price, outage levels, and voltage levels), respectively, and $\beta$ is the parameter vector associated with $\mathbf{x}_i^1$ and $\mathbf{x}_i^0$. Other nonelectricity factors, such as age and geographic locations, are described by the vector $\mathbf{z}_i$ with an associated parameter vector $\chi$. The coefficients ($\chi$) for the two dummy variables of low perceived distance and moderate perceived distance are $-1.24$ and $-0.34$ (An et al. 2002), indicating that as the fuelwood collection sites become far enough to change the household's current perceived distance (see the section "Landscape Submodel"), the switch-to-electricity probability will rise and the demand for fuelwood will accordingly decrease.

***Reduced fuelwood demand.*** When electricity is available, the ultimate fuelwood demand is computed as the fuelwood demand derived above times the probability that the household does not switch to electricity, which is 1 minus the probability of switching from fuelwood to electricity as computed above.

## Programming for Simulation

The model is programmed using Java-Swarm 2.1.1, a collection of software libraries developed by Swarm Development Group and briefly described in the Introduction. Swarm (Java version; it also supports Objective-C) provides many readily useable packages for Java programmers. In addition, by resorting to a few readily made application programming interfaces (API's), IMSHED provides a user-friendly and graphical interface to set parameters and run the program.

IMSHED also provides a batch mode without graphical interfaces, where command-line arguments are allowed from the Unix/Linux shell. The modeling environment employs a command-line-based experiment manager written in Perl (Perl is a high-level programming language particularly well-suited for tasks involving quick prototyping, system utilities, system management tasks, World Wide Web programming, and so on; see http://www.perl.com), which allows for efficient experiments by sweeping varying combinations of parameters designated in the Perl manager, leaving multiple runs progress unattended, and writing simulation results to designated directories.

## Model Test

Model test, a crucial step after model calibration, is subject to many theoretical and practical challenges. Though models of any complex open system (e.g., agent-based spatial models like ours) may not be truly verified and validated (Overton 1977; Oreskes, Shrader-Frechette, and Belitz 1994), we still follow the traditional terms of model verification and validation. Both of these involve fitting the model to data or theory, but verification checks for the proper functioning of the programming, while validation investigates the correspondence between the software model and the conceptual model (structural validation) and between model outcomes and empirical data (empirical validation; see Manson 2001).

Model verification includes progressive debugging (see the paragraph below) and uncertainty testing (Table 1). Debugging is progressive in that model construction and calibration run in parallel with debugging/verification processes. We begin with a very simple model, and then add and test new features or algorithms progressively until we are confident in moving on. Testing involves assessing output of a series of thirty runs over a span of twenty years. This span is chosen because it is long enough for teenagers at model initiation (1996) to grow and experience nearly all the major life-history events such as marriage and household development, but short enough so that some assumptions or parameters (except the one(s) being tested) can be reasonably left unchanged, since socioeconomic and ecological uncertainties increase as we attempt to model farther into the future.

Uncertainty testing consists of extreme tests and extreme combination tests, which are employed to determine if the model becomes corrupted at some stages or returns wholly unreasonable values, which may signify potential programming bugs or design flaws (Rykiel, 1996). The former refers to setting each major parameter to minimum and maximum feasible values, conducting thirty runs, and constructing envelopes at the 95 percent confidence level over twenty years for the number of households, population size, and habitat area. Extreme combination tests combine sets of values of the four most sensitive parameters (see Table 1, Note 1) and observing model behavior (see Table 4). For the sake of simplicity, we only choose either the minimal or the maximal values of each parameter in each combination.

**Table 1.** Model Test Methods

| | Instrument | Testing Stage | Contents | Criteria | Data source |
|---|---|---|---|---|---|
| **Verification** | Progressive building & debugging | Beginning-completion | | | |
| | Uncertainty test | Upon completion | Extreme tests | Theory, experience | Simulation results |
| | | | Extreme combination tests[(1)] | Theory, experience | Simulation results |
| **Validation** | Empirical validation | Upon completion | Demographic validation | $t$-test at 0.05 level | Independent government records |
| | | | Habitat validation | Change rate closeness | Independent results by other researchers |
| | Sensitivity analysis | Upon completion | See Table 3 | Experience | Simulation results |
| | Experience/expert opinion | Upon completion | Spatiotemporal pattern | Theory, experience, & expert opinion | Simulation results |

*Notes*:
(1) The selection of the variables for the combination tests depends on the results of the sensitivity analysis: the most sensitive factor in each of the four categories in Table 3 is selected.

Model validation includes empirical validation, sensitivity analysis, and experience/expert opinion validation (in relation to predicted spatial pattern, Table 1) (Parker et al. 2003). To validate a model empirically, one may employ either spatially independent data or temporally independent data. Spatially independent data are collected at the same time as those used to calibrate the model, but from a separate region, and are not used to calibrate the initial model. Temporally independent data are collected in the same region as those used to calibrate the model, but at a different (usually later) time. Our empirical validation includes demographic validation, which is concerned with comparing predicted populations from 1997–2003 to observed data for that time period, and with numbers of households over 1997–2000. Our predetermined criterion is to pass the two-sample (observed and predicted data) paired $t$-test at the 0.05 $\alpha$ level with a null hypothesis that the differences between the model predictions and real observations are zero. We also empirically validate the habitat model by comparing predicted habitat change with results from other researchers' independent studies.

Experience/expert opinion validation is concerned with the plausibility of the model output (Manson 2001), in particular with the spatial pattern of the habitat model. We construct the probability for each cell to be deforested and become nonhabitat, map the result, and consider the map's plausibility based on our field observations and expert opinion of a few local researchers. Sensitivity analysis considers the robustness of model results to relatively small changes in input parameters. A highly sensitive model is undesirable, given the uncertainty in model input. Sensitivity may be assessed by perturbing each major parameter by a certain magnitude (here, 50 percent), and calculating the sensitivity index (Jørgensen 1986) as:

$$S_x = (dX/X)/(dP/P) \tag{4}$$

where $P$ is the value of the independent variable, $dP$ is the value for a small change of $P$, $X$ is the value of the dependent variable, and $dX$ is the corresponding change in $X$ in response to the change in $P$.

## Simulation Experiments

We employ two types of model experiments: scenario analysis and complexity exploration. The objective of scenario analysis is two-fold. On one hand, we use it as a continuation of the model test process since unexpected outcomes may signal potential errors or bugs in the model; on the other hand, we want to provide policy makers some insights into possible outcomes under various practical conditions, as opposed to the extreme conditions investigated during validation. Here we are interested in discovering how population size, number of households, and panda habitat respond to varying conditions: (a) baseline scenario: employing status quo conditions; (b) conservation scenario: setting the sensitive factors and a few demographic factors to values that would presumably benefit panda habitat conservation; (c) development scenario: setting the factors to values that would presumably degrade panda habitat. We change each parameter in such a magnitude that would be (1) practical in the real world, e.g., fertility would be more likely to reduce to 1.5 in the conservation scenario than to 0; and (2) large enough to make a difference in model output based on our sensitivity test or field observations. For details of these scenarios, see Table 5.

The use of these highly divergent scenarios may provide some insights into the possible trajectories of panda habitat change, and its consequent effect on the likelihood of giant panda survival.

In complexity exploration, we are mainly interested in using the model to find and test particular features of complexity, including the impact of time lags, feedback effects, and nonlinearity on key processes. We hypothesize that: (1) most of the demographic factors have substantial time-lag effects, and these effects escalate over time; (2) most of the spatial variables have complex nonlinear behavior due to human feedback and behavior adjustment. We illustrate (1) by focusing on the intention of young people to leave their parents' homesteads (a parameter called "leave-home intention" in Tables 2 and 3). This variable is important because it represents lifestyle changes, which affect the efficiency of resource utilization (Liu, Daily, et al. 2003). For (2), we use perceived threshold distance and house buffer distance (for definitions, see Table 2) to test how projected panda habitat changes when these two variables take either the minimal value, the value corresponding to the first quartile ($\frac{1}{4}$(max − min)), the median value, the third quartile ($\frac{3}{4}$(max − min)), or the maximum value. The perceived threshold distance is an indicator of local households' own perceptions of the ease of fuelwood collection. A fuelwood collection site that is within this distance from a specific household is considered to have a "short" proximity (see the section "Fuelwood site selection"). The larger the value, the more likely a household views the current fuelwood collection distance as a "short distance." This variable may be affected by many other sociodemographic or psychological factors. For example, an increase in a specific household's annual income may lead it to value leisure time more highly, thereby decreasing the perceived ease of fuelwood harvesting and reducing fuelwood demand. The house buffer distance could be viewed as a policy control: a zero distance represents enforcement of no-cut regulations against fuelwood collection, while a very long distance (e.g., 7,200 m in Table 2) represents no or little restriction.

## Results

We present the results in three sections. The first section shows the outcomes of the model test efforts, including both model verification and validation. The second section reports on results of the three scenarios described previously: the baseline scenario, the conservation scenario, and the development scenario. The third section illustrates the patterns of complexity detected by the model simulations.

### Model Test

We verify our model in two steps: (1) extreme value tests and (2) extreme combination tests. The outcome of these tests is reported in Table 2. The model behaves as expected under the two extreme values of each variable. For instance, when the parameter "leave-home intention" is set to be 0 (indicating that all young adults remain in their parents' home after marriage), the final model reports total habitat of 580.78 km$^2$. When set to 1.0 (indicating that all young adults leave their parents' home and establish their own households after marriage), habitat area falls to 569.90 km$^2$. This may be caused by the great difference in the number of households: approximately 1,600 as opposed to 860 (Figure 8(a)); the human population is identical under both scenarios.

The 95 percent confidence envelopes become increasingly wide for all three variables as the model projects into the future, indicating increasing uncertainty in the prediction. For the number of households (Figure 8(a)), the envelope for the value of 1.0 (leave-home intention) is much higher than that for the baseline, which, in turn is higher than that for the value of 0.0. Because all the adults remain in their parental homes and do not establish their own houses (leave-home intention = 0.0), there is relatively little uncertainty over time as the total number of households slightly decreases. The population size dynamics do not differ among the three situations (Figure 8(b)), with all three envelopes nearly overlapping each other. This is because the leave-home intention parameter only affects the likelihood for young adults to establish their own households, rather than population size. The area of panda habitat (Figure 8(c)) decreases over time in all scenarios, but the rate varies logically between the three situations: results using a value of 1 show the most rapid decrease, as far more households are established and consume substantially more fuelwood.

Before reporting the results of the extreme combination tests, we report sensitivity test outcomes (listed as part of model validation in Table 1) because we use them to identify the most sensitive variables for the extreme combination tests. Table 3 lists the sensitivity of model parameters in four key groups. The most sensitive parameters are age at marriage in the family-planning category (1.73 percent of sensitivity, see Table 3), leave-home intention in the migration group (−1.02 percent over thirty years), price change in the electricity group (−3.27 percent), and house buffer distance in the spatial group (2.36 percent). These variables are examined in greater detail in subsequent extreme combination tests.

**Table 2.** Extreme Test Design and Results

| | Parameters | Definition | Default Value(1) | Min, max | $\text{Habitat}_{\text{min}}$, $\text{Habitat}_{\text{max}}^{(2)}$ |
|---|---|---|---|---|---|
| Family-planning | Max first kid interval | Maximal time between marriage and birth of the 1st child (years) | 4 | 1,<br>20 | 576.33 (0.71)<br>578.40 (0.75) |
| | Max birth interval | Maximal time (years) between births of consecutive children | 6 | 1,<br>20 | 575.94 (0.95)<br>576.07 (0.86) |
| | Upper birth age | The upper age that a woman gives birth to child | 50 | 30,<br>60 | 575.16 (0.84)<br>574.52 (0.89) |
| | Marry age | The age of first marriage | 22 | 18,<br>40 | 575.55 (0.65)<br>581.52 (0.74) |
| | Fertility | # of children a woman may give birth to during her lifetime. | 2.0 | 0,<br>20 | 577.63 (1.05)<br>575.16 (0.81) |
| Migration | College rate | Ratio between the number of people who go to college and the total number of people between 16–22 at a year. | 1.92% | 0.0%<br>100% | 575.81 (1.06)<br>580.48 (0.99) |
| | Leave-home intention | Probability that a "parental-home dweller" (3) leaves parental household and set up his/her own. | 42% | 0.0%<br>100% | 580.61 (0.96)<br>569.85 (0.67) |
| | Female marry-out rate | Ratio of the females between 22–30 who moved out of Wolong by marriage to all the females between 22–30 at a year | 0.28% | 0.0%<br>100% | 574.90 (0.93)<br>579.44 (0.70) |
| | Male bring-female-in rate | Ratio of the males between 22–30 who bring females into Wolong by marriage to all the males between 22–30 at a year | 0.19% | 0.0%<br>100% | 574.78 (0.89)<br>571.02 (0.92) |
| Electricity | Outage change | Change of outage levels (0 for low, 1 for medium, 2 for high) | 0 | – 2,<br>2 | 582.42 (1.14)<br>570.11 (0.80) |
| | Voltage change | Change of voltage levels (0 for low, 1 for medium, 2 for high) | 0 | – 2,<br>2 | 565.96 (0.81)<br>574.65 (1.42) |
| | Price change | Price change (Yuan) | 0 | – 0.50<br>0.50 | 597.72 (0.42)<br>558.32 (1.24) |
| Spatial | Perceived threshold distance | Distance (m) within which the perceived fuelwood collection distance is low | 800 | 0<br>8,000 | 579.31 (0.83)<br>565.06 (1.15) |
| | House buffer distance | Maximum distance (m) within which households collect fuelwood | 3,600 | 0<br>7,200 | 603.68 (0.45)<br>584.50 (1.44) |
| | Children-parent house distance | Maximum distance (m) between households of children & parents | 800 | 0<br>7,200 | 580.87 (0.78)<br>567.52 (1.41) |

*Notes:*

[1] The default values for the associated variables based on field observations. Habitat area ($km^2$) under the default values at year 2016 is 575.49 (0.69). Minimal and maximal values for the associated variables used to test the model.

[2] The habitat area ($km^2$) under minimal and maximal values for the associated parameter. The numbers in the parentheses are standard errors.

[3] An adult child who remains in his/her parental home after marriage.

The model runs through all the extreme combination tests (i.e., a total of sixteen tests: two extreme values for each of the four variables, Table 4) as well and gives reasonable results. Model runs are reported for a string of four letters, such as "*bbss*"; *s* symbolizes a minimal value ("small"), while "*b*" symbolizes a large ("big") value; "*bsss*" means a combination that takes big, small, small, and small values for the first to fourth variables in the combination. The outcome habitat area ranges from 607.18 $km^2$ in such combinations as "*ssss*" and "*bsss*" to 562.98 in $km^2$ in a combination of "*sbbb*." These combination tests may also signify the relative importance or contribution of each of the variables in affecting panda habitat change. For instance, when the electricity price change is "*s*" (–0.5 Yuan, or a 0.5 Yuan decrease), the value of the house buffer distance ("*s*" or "*b*") does

**Table 3.** Sensitivity Tests for Model Input Parameters (Habitat Area Projected in 2016)

| | Parameters | Default Value | +50%[1] Perturbation | Habitat area ($km^2$)[2] | Different from baseline[3] | Sensitivity |
|---|---|---|---|---|---|---|
| Family-plan | Max first kid interval | 4 | 6 | 575.55 (0.63) | no | N/A |
| | Max birth interval | 6 | 9 | 576.72 (1.35) | no | N/A |
| | Upper birth age | 50 | 60[4] | 574.52 (0.63) | no | N/A |
| | **Age at marriage** | 22 | 33 | 580.48 (0.87) | yes | 1.73% |
| | Fertility | 2.0 | 3.0 | 575.16 (0.95) | no | N/A |
| Migration | College rate | 1.92% | 2.88% | 569.8 | yes | 0.06% |
| | **Leave-home intention** [5] | 42% | 63% | 565.29 | yes | − 1.02% |
| | Female marry-out rate | 0.28% | 0.42% | 569.43 | yes | 0.04% |
| | Male bring-female-in rate | 0.19% | 0.29% | 568.32 | no | N/A |
| Electricity | Outage change | 0 | 1 | 573.22 (0.88) | no | N/A |
| | Voltage change | 0 | 1 | 575.68 (0.68) | no | N/A |
| | **Price change** | 0 | 0.05 | 566.09 (0.88) | yes | − 3.27% |
| Spatial | Fuelwood-change distance | 800 | 1,200 | 574.65 (1.11) | no | N/A |
| | **House buffer distance (m)** | 3,600 | 5,400 | 582.29 (0.96) | yes | 2.36% |
| | Kid-parent house distance (m) | 800 | 1,200 | 574.52 (1.25) | no | N/A |

*Notes*:
[1]The perturbation range of 50 percent is determined in consideration of: (1) it should be relatively small (otherwise we can use extreme tests as in Table 2); and (2) the response magnitude of the habitat change should be large enough for our calculation. An alternative of –50 percent perturbation is not included simply for space consideration.
[2]The standard error is in the parentheses following each average value.
[3]We use two-sample paired *t*-test at the 0.05 level to test whether the predicted habitat area is different from the baseline value at year 20.
[4]Here only 20 percent perturbation because an upper birth age of 75 years old (a 50 percent increase) does not make sense in the real world.
[5]All the 4 parameters under the category "Migration" have insignificant impact on panda habitat over 20 years. The numbers reported here are simulation results over 30 years, and the sensitivity index is calculated using the amount of habitat over 30 years (568.20 $km^2$).

not make much difference; when it is "*b*" (0.5 Yuan), it makes a great difference, e.g., "*bsbs*" and "*bsbb*" give 607.18 and 574.78 $km^2$, respectively. This implies that, given an increase in the price of electricity, a hypothetical policy on forbidding fuelwood collection (here we do not consider its practicability because it is a model test)

**Table 4.** Extreme combination test design and results

| Marry age | Leave-home intention | Electricity price (Yuan) | Buffer Dist.[1] (m) | Average habitat area ($km^2$)[2] | | | |
|---|---|---|---|---|---|---|---|
| | | | | Year 5 | Year 10 | Year 15 | Year 20 |
| 18(s) | 0.0(s) | − 0.5(s[3]) | 0(s) | 607.18(0.00) | 607.18(0.00) | 607.18(0.00) | 607.18(0.00) |
| | | | 7,200(b) | 607.18(0.00) | 606.95(0.13) | 606.92(0.21) | 606.79(0.22) |
| | | 0.5(b) | 0(s) | 607.18(0.05) | 607.18(0.05) | 607.18(0.05) | 607.05(0.05) |
| | | | 7,200(b) | 601.47(0.34) | 588.64(0.96) | 577.11(1.26) | 571.54(1.21) |
| | 1.0(b) | − 0.5(s) | 0(s) | 603.03(0.74) | 601.34(1.05) | 599.79(0.84) | 598.23(0.89) |
| | | | 7,200(b) | 601.99(0.67) | 598.88(0.90) | 595.25(0.90) | 593.18(0.85) |
| | | 0.5(b) | 0(s) | 603.42(0.67) | 599.53(0.52) | 597.84(0.85) | 596.42(0.83) |
| | | | 7,200(b) | 596.16(0.62) | 582.29(0.57) | 571.67(0.98) | 562.98(1.07) |
| 40(b) | 0.0(s) | − 0.5(s) | 0(s) | 607.18(0.00) | 607.18(0.00) | 607.18(0.00) | 607.18(0.00) |
| | | | 7,200(b) | 607.18(0.00) | 607.05(0.08) | 606.92(0.17) | 606.53(0.18) |
| | | 0.5(b) | 0(s) | 607.18(0.00) | 607.18(0.00) | 607.18(0.00) | 607.18(0.00) |
| | | | 7,200(b) | 601.86(0.32) | 588.90(0.81) | 579.06(1.11) | 574.78(1.12) |
| | 1.0(b) | − 0.5(s) | 0(s) | 606.66(0.40) | 606.14(0.64) | 605.62(0.66) | 605.10(0.63) |
| | | | 7,200(b) | 605.75(0.27) | 604.45(0.42) | 603.29(0.50) | 600.83(0.93) |
| | | 0.5(b) | 0(s) | 606.27(0.36) | 606.14(0.41) | 605.49(0.60) | 602.12(0.59) |
| | | | 7,200(b) | 601.47(0.51) | 589.42(0.81) | 579.44(1.01) | 571.67(1.39) |

*Notes*:
(1) House buffer distance (see Table 2 for its definition).
(2) The standard error is in the parentheses following each average value.
(3) The letters "s" and "b" stand for "small" and "big", respectively, corresponding to the minimal and maximal values of each parameter.

will substantially reduce habitat loss compared to no or little enforced collection restriction (house buffer distance = 7,200 m). This is intuitively correct and consistent with our observations.

The model validation consists of three sections: empirical validation, sensitivity testing, and validation via expert opinion/corroboration. We reported the sensitivity test results earlier in this section. For empirical validation, we use two-sample paired *t*-tests to decide whether the predictions are acceptable in relation to field observations. The model passes such *t*-tests for both demographic variables (number of households and population size), resulting in two *p*-values of 0.89 and 0.88 (Figures 7(a) and 7(b)). In addition, our predicted annual population increase rate is 0.48 percent (a total of 9.50 percent over twenty years), while the same rate from 1982 to 1996 is 1.05 percent (Liu, Ouyang, Tan, et al. 1999). This decrease in the rate of population growth could be due to the strict population control in the 1990s (Liu, Ouyang, Tan, et al. 1999). Our predicted annual rate of increase for the number of households in Wolong is 1.18 percent (a total of 23.63 percent over twenty years, Table 3), which is greater than the increase in population. This is consistent with the pattern from 1975 to 1999, for which the number of households increased more rapidly than did the population size as reported by Liu, Ouyang, Tan, et al. (1999).

To corroborate our work with that of other experts, we compare our predictions about panda habitat dynamics with results from other researchers. According to Laurie and Pan (1991), the annual loss of forest area in Wolong was 2.5 km$^2$ prior to 1991. Because the ratio between the total area of habitat (607 km$^2$, see Figure 8(c)) and the total area of forest (1,249 km$^2$, calculated by adding all the cells of forest classes defined in Linderman et al. 2004) in 1997 was 49 percent, this forest loss of 2.5 km$^2$/year is largely equivalent to a 1.23 km$^2$/year loss of habitat if the same ratio applies. Our model shows that under the status quo, panda habitat will decline from approximately 607 km$^2$ to 576 km$^2$ from 1996 to 2016 (Figure 8(c)), which translates into an annual habitat loss of 1.55 km$^2$. We explain our slightly higher habitat-loss rate in this way: habitat is not evenly distributed in all types of forests; instead, it is located in forest areas within certain elevation and slope thresholds (see the section "Habitat determination"), to which people have easy and frequent access. Cutting trees in these areas would degrade the habitat more than for an average forest plot that might be less accessible. Therefore, using the ratio between the total area of habitat and the total area of forest (49 percent) may lead to a lower value than the true value.

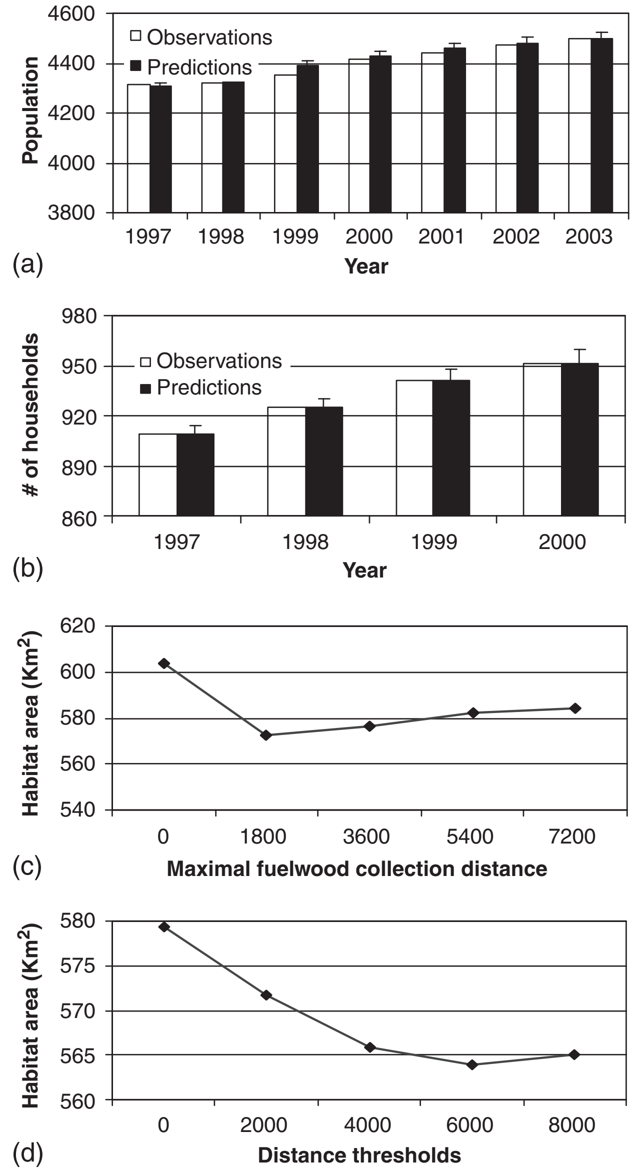


**Figure 7.** (a) Comparison between the observed and predicted population size ($t = -0.14$, $p = 0.89$). (b) Comparison between the observed and predicted number of households ($t = 0.16$, $p = 0.88$). (c) Habitat under varying maximal fuelwood collection distance (m). (d) Habitat under varying perceived distance for fuelwood collection easiness.

Last, we report our model validation by the spatio-temporal patterns. When time is controlled, the households occupy more land when leave-home intention is 1, which is more obvious in year 20 (Figure 9(b), top) than in year 10 (Figure 9(a), top). This is largely due to the model assumption that all young adults establish their own households and thus use more land. Consequently, habitat loss is more severe when leave-home intention is 1.0 than when it is 0.0, in particular for year 20 (Figures 9(b) and 9(d), bottom). From the spatial gradients, the most likely pixels for future households are

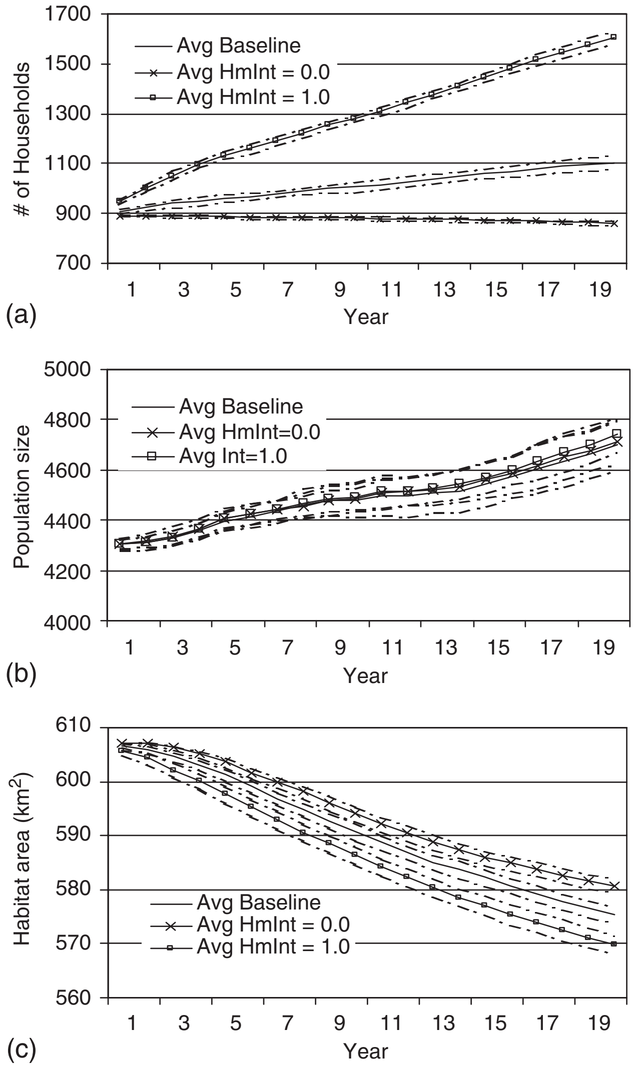


**Figure 8.** (a) The average numbers of households, (b) population size, and (c) predicted panda habitat and associated upper and lower 95 percent confidence envelopes for 1) the baseline simulation, 2) leave-home intention set to 0.0, and 3) leave-home intention set to 1.0.

close to those existing households. These are surrounded by more distant, less likely pixels, which are in turn surrounded by the least likely pixels. Farthest away are the totally unlikely ones. For habitat, we see a reversed trend: the least likely habitat cells are closest to households, then the less likely ones, and the most likely ones are the farthest. These phenomena are consistent with a diffusion model characterizing household choice in clearing land for construction and cutting fuelwood: begin with the nearest suitable sites, and then move outwards. Moreover, the model predicts the potential habitat gains or losses when leave-home intention equals 1.0 or 0.0. In the first case, habitat loss is increasingly large with respect to the baseline simulation as time progresses in the model (Figures 10(a) and 10(b)). In the second case, less habitat is lost relative to the baseline simulation (Figures 10(c) and 10(d)). These results agree with the experience of our own and local researchers (Shiqiang Zhou, personal communications[4]).

## Scenario Analysis

The conservation and development scenarios lead to great differences in the three key variables. The conservation scenario predicts approximately 600 km$^2$ of remaining panda habitat in 2016, the final model year. The development scenario estimates just 554 km$^2$ of remaining panda habitat. The conservation scenario predicts approximately 873 households, with a total human population of 2,611. The development scenario predicts much larger numbers of households and people: 1,602 and 6,305, respectively (Table 5).

The spatial patterns corroborate the above numerical changes in the number of households and population size. Under the development scenario, the households expand rapidly outward as time progresses, destroying some previous panda habitat. Under the conservation scenario, the land occupied by local households remains nearly unchanged, resulting in little habitat loss. Figure 11 gives a snapshot of panda dynamics under the conservation and development scenarios, which shows that the development scenario exerts much more severe impacts on panda habitat in parallel with a big cluster of households.

## Complexity Exploration

As shown in Figure 8(a), the numbers of households differ from each other increasingly over time when the leave-home intention takes 0, 0.42, and 1.0; the same is true for panda habitat (Figure 8(c)). The differences in panda habitat among these three scenarios become more significant over time: approximately from year 15, the lower bound of the 0.0 intention envelope is substantially higher than the upper bound of the baseline envelope; from year 11, the lower bound of the baseline envelope is substantially higher than the upper bound of the 1.0 intention envelope. This supports our first hypothesis that differences in initial demographic factors have large and escalating effects on model outcome.

The second hypothesis concerns nonlinearity in observed spatial patterns of household processes. Here we consider the impact of house buffer distance, which is a parameter representing a possible habitat protection policy to control how far people are allowed to search for fuelwood from their homes. As the house buffer distance rises, consequent total panda habitat area falls to a certain

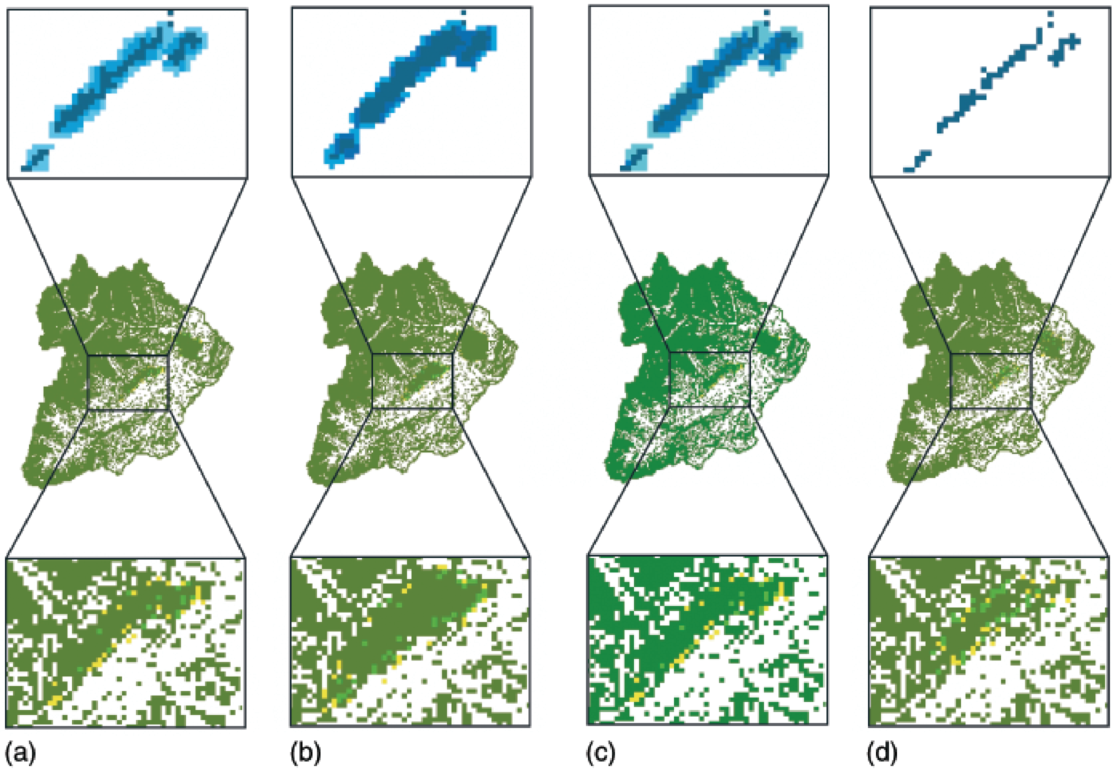


**Figure 9.** The gradients for household locations (top) and panda habitats (bottom) for model predictions employing (a) leave-home intention = 1.0 and time = 10; (b) leave-home intention = 1.0 and time = 20; (c) leave-home intention = 0.0 and time = 10; (d) leave-home intention = 0.0 and time = 20. For the household location gradients, the white, light blue, blue, and dark blue colors represent probabilities of <25%, 25%–50%, 50%–75%, and >75% for the occurrence of households. For the habitat gradients, the green, light green, yellow, and white colors represent probabilities of <25%, 25%–50%, 50%–75%, and >75% for the occurrence of habitat.

point (approximately 573 km$^2$), and then rises slowly (Figure 7(c)). We then consider a different, yet related, variable: perceived fuelwood distance. As the threshold for the perceived fuelwood collection distance rises, a similar pattern occurs, except that the changing point is much larger (approximately 6,000 m; Figure 7(d)).

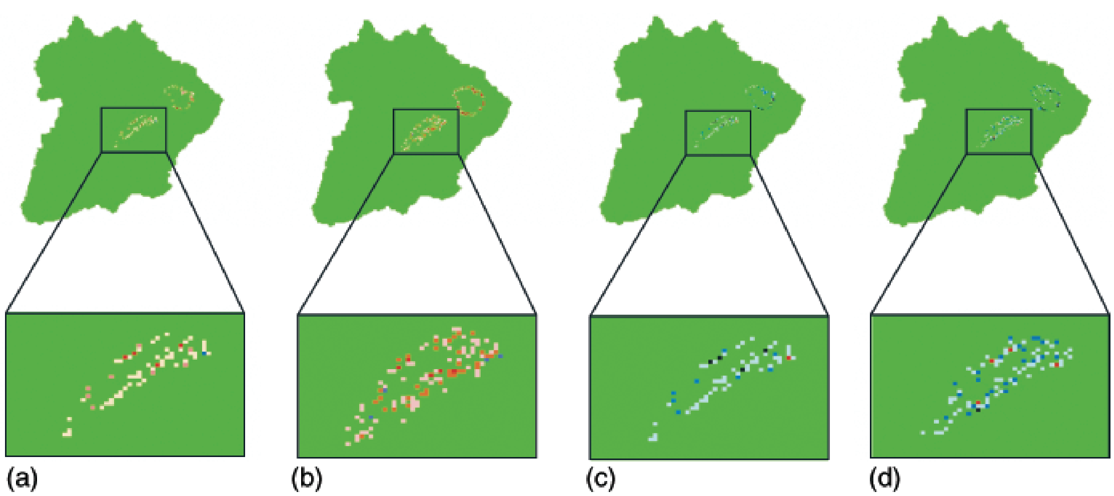


**Figure 10.** The habitat change gradients compared to the baseline simulations for various alternative scenarios: (a) leave-home intention = 1.0 and time = 10; (b) (b) leave-home intention = 1.0 and time = 20; (b) leave-home intention = 1.0 and time = 20; (c) leave-home intention = 0.0 and time = 10; (d) leave-home intention = 0.0 and time = 20. Shades of blue represent habitat gain with respect to the baseline, while red shades indicate habitat loss. The three levels of light yellow, pink, and red in (a) and (b) represent the probabilities of <25%, 25%–50%, >50% for habitat loss relative to the baseline, while the three levels of blue in (c) and (d) represent the probabilities of <25%, 25%–50%, >50% for habitat gain relative to the baseline.

**Table 5.** Definition and results of conservation and development scenarios and simulation results over 20 years(1)

| Category | Variable | Conservation scenario | Development Scenario |
|---|---|---|---|
| Electricity | Price | 0.05 Yuan decline | 0.05 Yuan rise |
| | Outage levels | One level decrease | One level rise |
| | Voltage levels | One level increase | One level decline |
| Migration | Leaving parental home intention | 0.42 → 0.21 | 0.42 → 0.95 |
| | College rate | 1.92% → 30% (16–20 youth) | 1.92% → 0.0% |
| | Female marry-out rate | 0.28% → 20% | 0.28% → 0.0% |
| Family planning | Fertility | 2.0 → 1.5 | 2.0 → 5.0 |
| | Birth interval | 3.5 → 5.5 | 3.5 → 1.5 |
| | Marriage age | 22 → 28 | 22 |
| Fuelwood | Distance for demand change (m) | 800 → 0 | 800 → 8000 |
| Results | # of households | 873.00 (7.48) | 1,602.00 (12.12) |
| | Population size | 2,611 (27.51) | 6,305 (76.92) |
| | Habitat ($km^2$) | 599.92 (0.54) | 553.52 (1.13) |

*Notes*:
(1) The first numbers in the spaces are the default values in the model, and the second values are those used in the associated scenarios.

## Discussion

Though the agent-based model presented here shows great potential for addressing practical issues about panda protection, it has been developed to address more general, theoretically important issues such as integrating socioeconomics, ecology, and demography, understanding complexities in some coupled society–environment systems, and linking ABM and GIS to study spatiotemporal dynamics of land-use and land-cover changes. The following subsections discuss several of these issues.

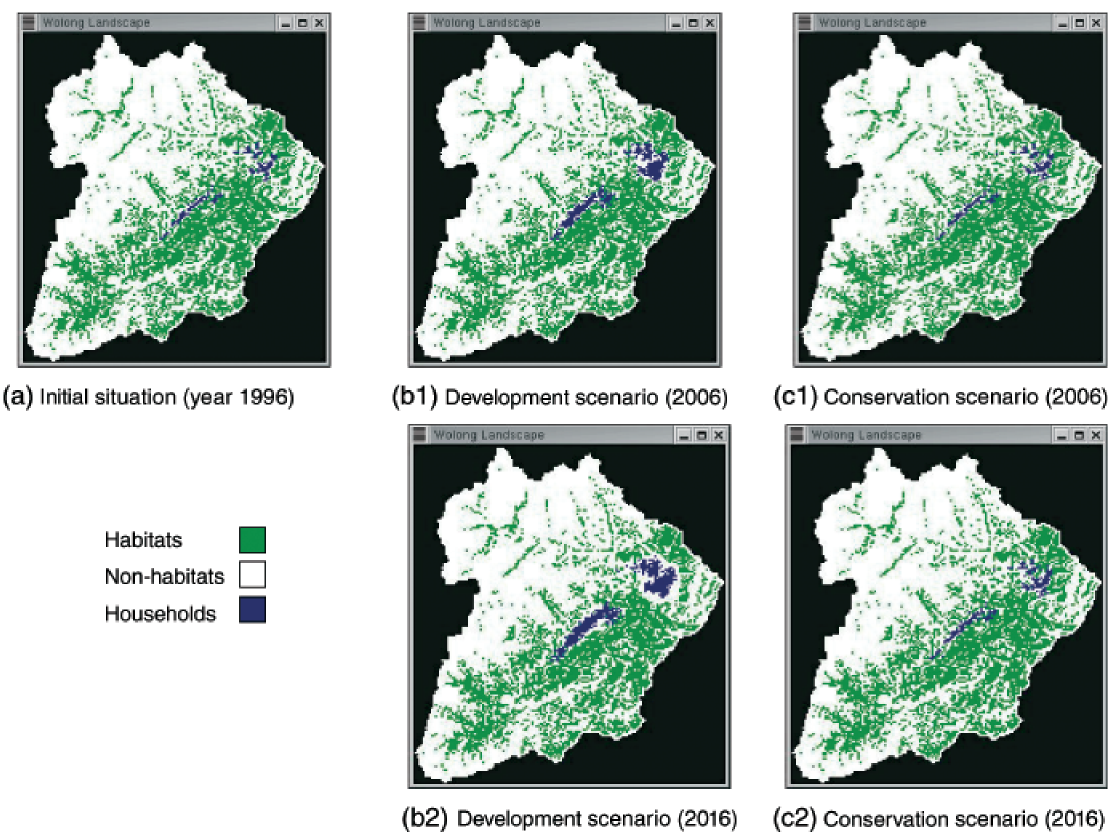


**Figure 11.** A snapshot of panda habitat fragmentation and household distribution in 1996 (a), 2006 (b1 for development and c1 for conservation), and 2016 (b2 for development and c2 for conservation).

## Model Results and Policy Implications

Model verification and validation turn out to be a theoretical and practical challenge in modeling complex systems (e.g., Manson 2001; Parker et al. 2003), and we address this challenge by designing a systematic strategy (see section "Model Test"; Table 1) and subjecting our model to this strategy. Though our model passes all the tests, it would be preferable to collect more detailed spatial data for forest volume and panda habitat in order to further validate the model. Greater detail would also facilitate the application of the model in balancing the needs for panda conservation and those for human well-being within the reserve.

This study used total area of potential panda habitat as a primary model output to assess the impact of different scenarios on the panda. Though a decrease of 5.27 percent (about 32 km$^2$) under the baseline situation (the status quo scenario) over twenty years may seem moderate, the spatial distribution and fragmentation of panda habitat should be of concern. Pandas appear to prefer areas that humans also tend to visit for fuelwood collection (Schaller et al. 1985; Liu, Ouyang, Taylor, et al. 1999). Therefore, the predicted decrease in potential habitat may occur in the panda's preferred habitat. Furthermore, employing total area alone does not address the problem of habitat fragmentation, which may render potential habitat pixels unusable for pandas. According to Schaller et al. (1985), a giant panda usually occupies a home range with an area of about 2 km$^2$. In our model, all pixels meeting the simple spatial criteria, including fragmented areas smaller than 2 km$^2$, are counted as panda habitats; therefore, it is likely that our model overestimates actual viable panda habitat. However, this approximation could be viewed from another perspective: within all the potential panda habitats (green blocks in Figure 11), there might be pandas; within all the areas that are counted as nonhabitats, there should be no pandas. A final spatial concern for the results is that stochastic environmental shocks, such as forest fires, could lead to a substantial sudden loss of panda habitat. This model does not account for such factors.

Model results under conservation and development scenarios indicate that human socioeconomic and demographic factors substantially affect panda habitat, but the impact of these factors takes time to manifest itself on the environment. Implementing policies that encourage family planning, human out-migration from the reserve, lifestyle change, or the increased use of electricity could result in subsequent preservation of panda habitat to varying degrees. For instance, the model predicts that an electricity subsidy of 0.05 Yuan could reduce total habitat loss by the year 2016 from around 32 km$^2$ to 18 km$^2$. If combined with other conservation activities, even more habitat could be spared. However, our results indicate that the environmental benefits of such policies—or the penalties should they not be implemented—are not immediately obvious.

The nonlinear and counterintuitive relationship between the amount of panda habitat and the house buffer distances as shown in Figure 7(c) may be caused by the following relationships: when this buffer distance is very small (even zero), people are allowed to harvest little or no fuelwood, and thus panda habitat is better preserved; when this buffer distance is very large, local households' fuelwood collection is scattered throughout a large buffer region, and some areas in this region may have a substantial forest volume or a rapid regeneration rate (capturing this type of uncertainty in the section "Spatial environmental data" is one strength of this agent-based model), so cutting wood may not cause severe habitat loss. If, however, this buffer distance is somewhere in between (around 1,800 m), local households' unrestricted fuelwood demand is satisfied through cutting all available wood in this small region (very likely going beyond its carrying capacity) and causing more habitat loss. As time goes on, the local households are likely to move outwards, as the pattern in Figures 10(a) and 10(b) shows.

Similarly, the parameter perceived threshold distance also leads to nonlinear changes in panda habitat (Figure 7(d)). This variable could be explained as the ease of adjusting household fuelwood demand based on the existing fuelwood collection distance: the longer the distance, the harder to reduce their fuelwood demand, thus more habitat loss caused by satisfying this demand. However, after a threshold (6,000 m), this effect diminishes and yields to other complexities (such as the above-mentioned habitat vs. buffer distance relationships).

## Methodology: Integration, Complexity, and Coupling of ABM and GIS

The IMSHED developed for this research has utilized data and models across scales, disciplines, and time periods. Data integration included working with data at different scales (e.g., individual-level data such as age vs. population data such as mortality rate), from different disciplines (e.g., ecological data such as forest regeneration vs. sociological data such as leave-home intention), and with varying degrees of uncertainty (e.g., accurate human individual demographic data vs. forest volume data with a wider range of uncertainty). Far more so than data, however, the combination of methods and models from different disciplines reflects the breadth of the human–environment modeling challenges. For instance, IMSHED employs a fuelwood demand model (An et al. 2001) and an econometric model for electricity demand (An et al. 2002) to compute fuelwood demand on a household basis (see the section "Demographic Submodel"). By integrating methods and empirical results concerning young adults' propensity to leave their parental homes and form their own households, IMSHED is able to project household dynamics and link them to panda habitat loss. In addition, ecological uncertainties or variability (e.g., variations in forest volume and regrowth rates) are taken into consideration through assigning varying values to the attributes of the forest pixel objects, which demonstrate the utility of object-oriented programming (OOP) for capturing environmental variability.

From the above analyses, we are confident that our modeling framework effectively integrates individual-level data and transdisciplinary models for projecting spatiotemporal changes of some key response variables. This challenge has been viewed as very significant for studying environmental sustainability (Clark 2002). This bottom-up approach "starts from the 'parts' (i.e., individuals) of a system and then tries to understand how the system's properties emerge from the interactions among these parts" (Grimm 1999). Consequently, this framework can efficiently deal with many research needs that traditional approaches may find difficult or impossible to deal with, and may provide more accurate predictions. However, this increased accuracy can only be found in the aggregated results, such as human population size or number of households. Some stochastic processes are simulated at the agent level (e.g., an individual person's leaving parental home decisions), and individual simulation runs are essentially single realizations of the process. These simulation runs are useful in predicting the overall number of households and the resultant spatial pattern, but whether a particular household will be established at a specific location is not a question that our model resolves to answer.

Developing and using ABM do not discredit the traditional state variable, statistic, or analytic approaches. On the contrary, in many situations, our framework uses these approaches because it is unnecessary or sometimes impossible to account for every detail of the agents under consideration. For instance, when computing the probability to switch from fuelwood to electricity, we use a logistic regression (see the section "Socioeconomic Submodel"), and obviously this regression model is an average trend derived from a number of households. It is important to balance between using outcomes based on individual agent actions and averaged trend data to find an appropriate level of resolution and aggregation in predictions. The choice depends on research needs, the applicability of individual data, and available resources (time, budget, and other conditions like computational power).

The complexities in many coupled society–environment open systems have been barriers for effectively studying and understanding such systems. By decomposing the population-level dynamics into life histories of all the individuals and characterizing the dynamics of all households in the landscape, it becomes easier to capture any time-lag effects of demographic factors. For instance, an increase of 0.5 in fertility may be "considered" by many couples planning to have children at an appropriate time (the model "knows" the time), and this consideration may lead to an increase in population size and number of households with a cumulative effect. This effect may not cause observable habitat degradation in ten years, but may do so in thirty years. In addition, with feedback (households decrease their fuelwood demand as collection distances rise) built into our model, the environment is not simply a passive cache of resources waiting to be developed; instead, its geography imposes opportunities and limitations on the human inhabitants. This leads to a more dynamic portrayal of the human–environment system and, we believe, more representative model results.

Finally, the integration of ABM and GIS in this study has allowed for further insights into the spatial trajectories of some key ecological or socioeconomic processes, such as the gradients for household locations and panda habitats in Figure 9. These insights are not only important in validating the model (abnormal spatial trajectories may signify potential bugs; see Parker et al. 2003), but also may be significant for panda conservation efforts because the trajectories provide policy makers information about where, when, and under what conditions panda habitat would be lost or conserved. However, this integration is still in its fledgling stage. Much readily usable functionality in GIS (such as finding the cost distance) has to be coded in Java Swarm by the authors, which sometimes becomes a heavy burden. As an alternative, some model outcomes have to be exported to ArcGIS for further spatial analysis (e.g., the gradient of households locations in Figure 9). Coupling environmental models and GIS has long been recognized as a key challenge (Goodchild, Parks, and Steyaert 1993; Wesseling et al. 1996), and our experience reflects this. A more specific issue for coupling ABM and GIS is the development of rule sets based on empirical metrics. Current metrics may not adequately reflect key aspects of the environment, and new landscape and/or spatial metrics (e.g., Herold, Goldstein, and Clarke 2003) may need to be developed to capture different spatial and temporal dynamics of landscape/habitat change.

In this study, we develop a framework to integrate geographical, ecological, socioeconomic, and demographic data into different levels or types of agents or objects (persons, households, pixels), incorporate some complex mechanisms (e.g., time lag, feedback), and project the spatial patterns of panda habitat extent over time. This framework enables us to study how changes in socioeconomic and demographic factors work in both straightforward and complex ways to affect panda habitat. This impact is characterized over time in a spatially explicit manner. Using this combined model has enabled us to develop a better understanding of the relationships between people and panda habitat in Wolong, which may, in turn, help to develop environmentally sound policies in the reserve. More broadly, we have provided a working example of a framework (including the tool) to explain or predict overall landscape patterns as a result of the actions of many agents. This framework is a powerful means for integrating data and models across varying scales and disciplines and shows promise for many human-environment studies.

## Acknowledgements

We thank Guangming He, Jinyan Huang, Shiqiang Zhou, Zhiyun Ouyang, and Hemin Zhang for their assistance in data acquisition. We appreciate the valuable comments from Dr. Michael Goodchild and the four anonymous reviewers. We are also indebted to financial support from the National Science Foundation (NSF), National Institute of Child Health and Human Development (NICHD), American Association for Advancement of Sciences, the John D. and Catherine T. MacArthur Foundation, and Michigan State University. Special thanks go to the online help from Swarm Development Group, in particular Dr. Paul Johnson at the University of Kansas.

## Notes

1. For link to these tools, go to http://wiki.swarm.org/wiki/Main_Page (last accessed on 17 June 2004).
2. Electricity is the readily available substitute for fuelwood in the Reserve, subject to government price control and some quality problems (An et al. 2002). Other energy sources such as coal, charcoal, biogas, and sun/wind power are not used and no market exists for them.
3. Though we observed some people who took temporary jobs in outside areas (primarily big cities), they still had their residence registration (known as *Hukou*) license in Wolong. More importantly, they often come back to Wolong during busy agricultural seasons and Chinese spring festivals and conduct resource-related activities such as fuelwood collection. Thus, they are not treated as out-migrants.
4. Shiqiang Zhou from the Wolong Nature Reserve is an experienced researcher with extensive knowledge in local biology, ecology, and socioeconomic and demographic situations. The authors have closely worked with him to collect the data, build the models (including earlier models as published by An et al. 2001, 2002, 2003), and discuss the model outcomes during 1998–2004.

## References

An, L., J. Liu, Z. Ouyang, M. A. Linderman, S. Zhou, and H. Zhang. 2001. Simulating demographic and socioeconomic processes on household level and implications on giant panda habitat. *Ecological Modeling* 140:31–49.

An, L., F. Lupi, J. Liu, M. A. Linderman, and J. Huang. 2002. Modeling the choice to switch from fuelwood to electricity: Implications for giant panda habitat conservation. *Ecological Economics* 42 (3): 445–57.

An, L., A. G. Mertig, and J. Liu. 2003. Adolescents' leaving parental home: Psychosocial correlates and implications for biodiversity conservation. *Population and Environment: A Journal of Interdisciplinary Studies* 24 (5): 415–44.

Batty, M., Y. Xie, and Z. L. Sun. 1994. Modelling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems* 23:205–33.

Berger, T. 2002. Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource-use changes, and policy analysis. *Agricultural Economics* 25 (2–3): 245–60.

Bian, L. 1997. Multiscale nature of spatial data in scaling up environmental models. In *Scale in remote sensing and GIS*, ed. D. A. Quattrochi and M. F. Goodchild, 13–26. New York: Lewis Publishers.

Brown, D. G., B. C. Pijanowski, and J. D. Duh. 2000. Modeling the relationships between land use and land cover on private lands in the Upper Midwest, U.S. *Journal of Environment Management* 59: 247–63.

China's Ministry of Forestry and World Wildlife Fund. 1989. *Conservation and management plan for giant pandas and their habitat*. Beijing, China.

Clark, W. C. 2002. Research systems for a transition toward sustainability. In *Challenges of a changing earth: Proceedings of the global change open science conference (Amsterdam NL, 10–13 July 2001)*, eds. W. Steffen, J. Jäger, D. Carson, and C. Bradshaw. Berlin: Springer-Verlag http://ksgnotes1.harvard.edu/BCSIA/sust.nsf/pubs/pub54 (last accessed 11 June 2004).

Clarke, K. C., and L. Gaydos. 1998. Loose coupling a cellular automaton model and GIS: Long-term growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science* 12 (7): 699–714.

Deadman, P., K. Lim, D. Robinson, E. Moran, E. Brondízio, and S. McCracken. 2001. LUCITA: multi-agent simulations of land-use change near Altamira, Brazil (Section 3.6). In *Agent-based models of land-use and land-cover change*, ed. D. C. Parker, T. Berger, and S. M. Manson (W. J. McConnell, Man. Ed.). Available at http://www.csiss.org/maslucc/ABM-LUCC.htm (last accessed 11 June 2004).

Ehrlich, P. R., and E. O. Wilson. 1991. Biodiversity studies: science and policy. *Science* 253:758–62.

Flake, G. W. 1998. *The computational beauty of nature: Computer explorations of fractals, chaos, complex systems, and adaptation*. Cambridge, MA: MIT Press.

Gimblett, H. R. 2002. Integrating geographic information systems and agent-based technologies for modeling and simulating social and ecological phenomena. In *Integrating geographic information systems and agent-based techniques for simulating social and ecological processes*, ed. H. R. Gimblett, 1–20. New York: Oxford University Press.

Goodchild, M. F., B. O. Parks, and L. T. Steyaert, eds. 1993. *Environmental modeling with GIS*. New York: Oxford University Press.

Grimm, V. 1999. Ten years of individual-based modeling in ecology: What have we learned and what could we learn in the future? *Ecological Modeling* 115:129–48.

He, G., S. Bearer, X. Chen, and J. Liu. Unpublished. The household fuelwood collection survey (Section II). 2003.

He, G., S. Bearer, and J. Liu. Unpublished. The household fuelwood collection survey (Section I). 2001–2002.

Herold, M., N. C. Goldstein, and K. C. Clarke. 2003. The spatio-temporal form of urban growth: measurement, analysis and modeling. *Remote Sensing of Environment.* 86 (3): 286–302.

Jiang, B., and H. R. Gimblett. 2002. An agent-based approach to environmental and urban systems within geographic information systems. In *Integrating geographic information systems and agent-based techniques for simulating social and ecological processes*, ed. H. R. Gimblett, 171–89. New York: Oxford University Press.

Jørgensen, S. E. 1986. *Fundamentals of ecological modelling*. Amsterdam: Elsevier.

Larkin, D., and G. Wilson. 1999. *Object-oriented programming and the Objective-C language*. Cupertino, CA: Apple Computer, Inc.

Laurie, A., and W. Pan. 1991. Working group report on giant pandas in the wild. In *Giant panda and red panda conservation workshop: Working group reports: National Zoological Park*, ed. D. Klerman and M. Roberts, 5–12. Washington, DC: Smithsonian Institution.

Levin, S. A., B. Grenfell, A. Hastings, and A. S. Perelson. 1997. Mathematical and computational challenges in population biology and ecosystems science. *Science* 275:334–43.

Li, H., and J. F. Reynolds. 1997. Modeling effects of spatial pattern, drought, and grazing on rates of rangelands degradation: a combined Markov and cellular automaton approach. In *Scale in remote sensing and GIS*, ed. D. A. Quattrochi and M. F. Goodchild, 211–30. New York: Lewis Publishers.

Li, X., and A. G. Yeh. 2002. Neutral-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science* 16 (4): 323–43.

Lim, K., P. J. Deadman, E. Moran, E. Brondizio, and S. McCracken. 2002. Agent-based simulations of household decision-making and land use change near Altamira, Brazil. In *Integrating geographic information systems and agent-based techniques for simulating social and ecological processes*, ed. H. R. Gimblett, 277–308. New York: Oxford University Press.

Linderman, M., L. An, S. Bearer, G. He, Z. Ouyang, and J. Liu. Forthcoming. Modeling the spatio-temporal dynamics and interactions of households, landscapes, and giant panda habitat. *Ecological Modelling*.

Linderman, M., J. Liu, J. Qi, Z. Ouyang, L. An, J. Yang, and Y. Tan. 2004. Using artificial neural networks to map the spatial distribution of understory bamboo from remote sensing data. *International Journal of Remote Sensing* 25 (6): 1685–1700.

Liu, J., L. An, S. S. Batie, R. E. Groop, L. Zai, M. A. Linderman, A. G. Mertig, Z. Ouyang, and J. Qi. 2003. Human impacts on land cover and panda habitat in Wolong Nature Reserve: linking ecological, socioeconomic, demographic, and behavioral data. In *People and the environment: Approaches for linking household and community surveys to remote sensing and GIS*, ed. J. Fox, R. R. Rindfuss, S. J. Walsh, and V. Mishra, 241–63. Boston: Kluwer Academic Publishers.

Liu, J., G. C. Daily, P. R. Ehrlich, and G. W. Luck. 2003. Effects of household dynamics on resource consumption and biodiversity. *Nature* 421:530–33.

Liu, J., M. Linderman, Z. Ouyang, L. An, J. Yang, and H. Zhang. 2001. Ecological degradation in protected areas: the case of Wolong Nature Reserve for giant pandas. *Science* 292:98–101.

Liu, J., Z. Ouyang, Y. Tan, J. Yang, and H. Zhang. 1999. Changes in human population structure: Implications for biodiversity. *Population and Environment: A Journal of Interdisciplinary Studies* 21:46–58.

Liu, J., Z. Ouyang, W. W. Taylor, R. Groop, and H. Zhang. 1999. A framework for evaluating the effects of human factors on wildlife habitat: the case of giant pandas. *Conservation Biology* 13:1360–70.

López, E., G. Bocco, M. M. Mendoza, and E. Duhau. 2001. Predicting land-cover and land-use change in the urban fringe: A case in Morelia city, Mexico. *Landscape and Urban Planning* 55:271–85.

MacArthur, R. H., and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton, NJ: Princeton University Press.

Malanson, G. P. 1999. Considering complexity. *Annals of the Association of American Geographers* 89: 746–53.

———. 2002. Extinction-debt trajectories and spatial patterns of habitat destruction. *Annals of the Association of American Geographers* 92:177–88.

Manson, S. M. 2001. Calibration, verification, and validation (Section 2.4). In *Agent-based models of land-use and land-cover change*, ed. D. C. Parker, T. Berger, and S. M. Manson (W. J. McConnell, mng. ed.). http://www.csiss.org/maslucc/ABM-LUCC.htm (last accessed 11 June 2004).

Messina, J. P., and S. J. Walsh. 2001. 2.5D morphogenesis: modeling landuse and landcover dynamics in the Ecuadorian Amazon. *Plant Ecology* 156:75–88.

Minar, N., R. Burkhart, C. Langton, and M. Askenazi. 1996. The Swarm simulation system: A toolkit for building multi-agent simulations (SFI Working Paper 96-06-042). http://www.swarm.org/archive/overview.ps (last accessed 11 June 2004).

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403:853–58.

Najlis, R., M. A. Janssen, and D. C. Parker. 2001. Software tools and communication issues (Section 2.3). In *Agent-based* *models of land-use and land-cover change*, ed. D. C. Parker, T. Berger, and S. M. Manson (W. J. McConnell, mng. ed.), 27–39. http://www.csiss.org/maslucc/ABM-LUCC.htm (last accessed 11 June 2004).

NeXT Software, Inc. 1992. *Object-oriented Programming and the Objective-C Language*. Online: http://www.gnustep.org/resources/documentation/ObjectivCBook.pdf (last accessed 11 June 2004).

Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 263 (4 February): 641–46.

Ouyang, Z., J. Liu, S. Zhou, and Y. Tan. Unpublished. The vegetation survey in Wolong Nature Reserve. 1998.

Overton, S. 1977. A strategy of model construction. In *Ecosystem modeling in theory and practice: an introduction with case histories*, ed. C. Hall and J. Day, 49–73. New York: John Wiley & Sons.

Pan, W. K. Y., C. M. Erlien, F. Baquero, S. J. Walsh, R. E. Bilsborrow, and B. G. Frizzelle. 2004. Farm-level models of spatial patterns of land use and land cover dynamics in the Ecuadorian Amazon. *Agriculture, Ecosystems, & Environment* 101 (2–3): 117–34.

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.

Peuquet, D. J. 1999. Time in GIS and geographical databases. In *Geographical information systems* (2nd ed.), ed. P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, 91–103. New York: Wiley.

Rindfuss, R. R., P. Prasartkul, S. J. Walsh, B. Entwisle, Y. Sawangdee, and J. B. Vogler. 2003. Household-parcel linkages in Nang Rong, Thailand: challenges of large samples. In *People and the environment: Approaches for linking household and community surveys to remote sensing and GIS*, ed. J. Fox, R. R. Rindfuss, S. J. Walsh, and V. Mishra, 241–3. Boston: Kluwer Academic Publishers.

Rykiel, E. J. , Jr. 1996. Testing ecological models: the meaning of validation. *Ecological Modeling* 90: 229–44.

Schaller, G. B., J. Hu, W. Pan, and J. Zhu. 1985. *The giant pandas of Wolong*. Chicago: University of Chicago Press.

Seto, K. C., and R. K. Kaufmann. 2003. Modeling the drivers of urban land use change in the Pearl River Delta, China: integrating remote sensing with socioeconomic data. *Land Economics* 79 (1): 106–1.

Walsh, S. J., R. E. Bilsborrow, S. J. McGregor, B. G. Frizelle, J. P. Messina, W. K. T. Pan, K. A. Crews-Meyer, G. M. Taff, and F. Baquero. 2003. Integration of longitudinal surveys, remote sensing time series, and spatial analysis: approaches for linking people and place. In *People and the environment: Approaches for linking household and community surveys to remote sensing and GIS*, ed. J. Fox, R. R. Rindfuss, S. J. Walsh, and V. Mishra, 241–63. Boston: Kluwer Academic Publishers.

Walsh, S. J., T. P. Evans, W. F. Welsh, B. Entwisle, and R. R. Rindfuss. 1999. Scale-dependent relationships between population and environment in northeastern Thailand. *Photogrammetric Engineering & Remote Sensing* 65 (1): 97–105.

Wesseling, C. G., D. J. Karssenberg, P. A. Burrough, and W. P. A. Van Deursen. 1996. Integrated dynamic environmental models in GIS: the development of a dynamic modelling language. *Transactions in GIS* 1:40–48.

Wolong Administration. 1987. *Wolong vegetation cover and plant resources*. Chengdu: Sichuan Science and Technology Press.

———. 1996. 1996 Agricultural Census, Wolong (unpublished, in Chinese).

———. 2000. *Wolong 2000 nationwide population census data*. (unpublished, in Chinese).

Yang, Y., and C. Li, eds. 1992. *Sichuan forests*. Beijing, China: Yunnan, China Forestry Press.

*Correspondence*: School of Natural Resources and Environment, University of Michigan, Ann Arbor, MI 48109-1115, e-mail: lian@umich.edu (An); Department of Fisheries and Wildlife, Michigan State University, 13 Natural Resources Building, East Lansing, MI 48824, e-mail: linderman@geog.ucl.ac.be (Linderman); Department of Geography, Michigan State University, East Lansing, MI 48824, e-mail: qi@msu.edu (Qi); Department of Geography, Michigan State University, East Lansing, MI 48824, e-mail: ashton@msu.edu (Shortridge); Center for Systems Integration and Sustainability, Michigan State University, East Lansing, MI 48824, e-mail: jliu@panda.msu.edu (Liu).