Chapter 13
Big Models: From Beijing to the Whole China

13.1 A Golden Era of Big Models

Applied regional/urban models have attracted extensive attention from researchers in recent decades. Regional models are used for regional analysis and simulation at a macro-geographic level, such as for a collection of cities or a whole country, which generally involve a variety of spatial analysis approaches and statistical methods. On the contrary, urban models rely more on modeling and simulation approaches (Batty 2009). They are commonly used for understanding and predicting urban systems through abstracting and generalizing different components of a city. Urban models were firstly developed in the early 1950s and experienced several phases as they developed and evolved. Figure 13.1 presents the development line of urban models from static to dynamic models. The dynamic models further include top-down differential equation-based models and currently prevailing bottom-up models using cellular automata or agent-based approaches. The spatial unit of urban models is also in a transition from a larger geographical unit such as a large grid or a zone to a smaller unit such as a block, a parcel, or even a building (Hunt et al. 2005; Wegener 2004). Generally, these two types of models are utilized separately. According to existing research on applied regional/urban models, they are rarely used simultaneously or synthetically.

In practice, the existing applied regional/urban models can fall into two clusters based on their geographical scale and spatial unit. One is a fine-scaled model for a small area, e.g. part of a city or an entire city. The modeling spatial unit can be a parcel, a block, or a small cell. The other is a model for a large area, such as a region or an entire country. The modeling unit can be a county or a super cell. Because there is a general tradeoff between the spatial extent and the resolution due to the data paucity, it is hard to develop a model that can be applied to a large geographic extent but with a small spatial unit (see Fig. 13.2).

Y. Long, Z. Shen, *Geospatial Analysis to Support Urban Planning in Beijing*, GeoJournal Library 116, DOI 10.1007/978-3-319-19342-7_13

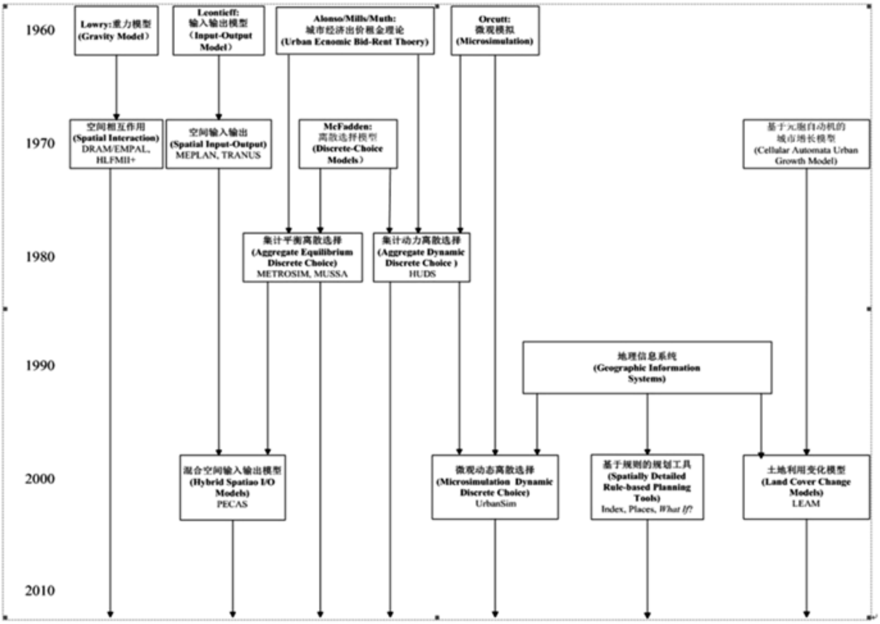


Fig. 13.1 The development line of Applied Urban Models (Adapted from Paul Waddell, Dynamic Microsimulation: UrbanSim, Webinar 5 of 8-part TMIP, Webinar series on land use forecasting methods)

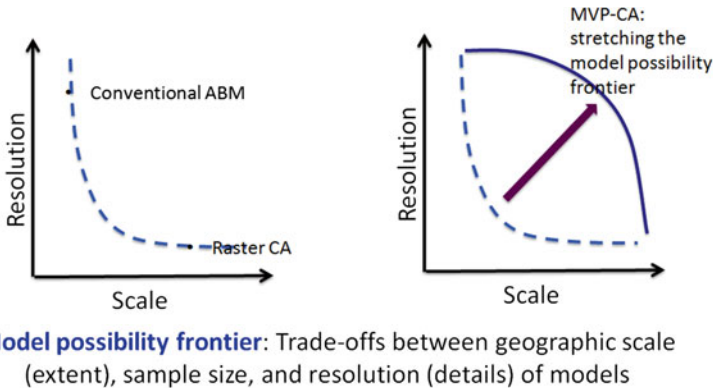


Fig. 13.2 Conventional models vs. "big models" (MVP-CA, a mega-vector-parcels cellular automata model, is our first big model for simulating urban expansion at the parcel level for all Chinese cities)

To the best of our knowledge according to extensive literature review, fine-scale applied urban models for a large area have been rare in academic research. As explained earlier, this stems largely from the lack of data and computation capacity limitation, which are particularly common in the case of China. In addition, collecting fine-scale data for feeding models in medium- and small-sized cities is often constrained by poorly developed digital infrastructures in developing countries. This condition, to some degree, has obstructed the progress of fine-scale urban simulation for a large area in developing countries in general and in China in particular. Overcoming data shortfalls has become the top priority for fine-scale urban simulation in developing countries, even in some developed countries, to support policy making.

In this chapter, we propose a term, namely "big model" for the fine-scale urban simulation model of a regional area with a large geographical scale. Big model is defined as data-driven regional analysis and urban simulation tools involving a variety of modeling approaches in this chapter as a new type of research paradigm for urban and regional studies, thus overcoming the trade-off between simulated scale and spatial unit. More importantly, as our ability to collect, store, and process data has increased remarkably in recent years since the digital revolution, big models would provide us with new opportunities for better understanding how cities work. There are four major reasons making the widespread use of big models happen. (1) Today, big data, such as mobile traces, public transport smartcard records, online check-ins/points-of-interest, and floating car trajectories, are becoming pervasively available. The spread of mobile technologies and computing has made generating, tracking, and recording individual data as partial representation of daily life, greatly supporting the analysis and modeling with rich datasets. Some scholars even advocate that data are models themselves (Batty 2012). (2) Open access to data has been improved significantly as there have been calls for governmental transparency and accountability. For instance, people can access the dataset inventory of planning permits from the official website of Beijing Planning Commission, land transaction records from Beijing Land Bureau, and housing projects from Beijing Housing and Construction Commission. Generally, these records are associated with detailed project-level information, including fine-scale physical characteristics and urban development status. Supported by online geocoding services, these records can be utilized in big models in the form of point datasets. Without painstaking efforts towards an "open government", no such things would have been possible in China. (3) Computational capacity has been largely improved for running big models by means of techniques like parallel computation and Hadoop. (4) For those bottom-up simulation methods adopted by big models, such as cellular automata, agent-based modeling, and network analysis, they have evolved and matured, allowing more sophisticate and powerful application of big models. Therefore, we argue that big models will mark a promising new era for the urban and regional study field.

The purpose of this chapter is to summarize the progress of our existing research on the application of big models in China. The next section elaborates the basic ideas and characteristics of big models. Section 13.3 reviews the methodology development and several case studies in applying big models in various urban and regional researches. In the end, we conclude with a summary of our findings and suggest directions for further research.

13.2 Big Models: A Novel Research Diagram for Urban and Regional Studies

Big models have the following characteristics. First, they need large-scale geographic data including so called "big data" or large-scale "open data" for initialization. The data may be collected at the individual observation level or based on small spatial units. Second, both the existing inter-city and intra-city analysis methods can be integrated in big models (see Fig. 13.3 for an illustration of a big model combining inter-city and intra-city approaches). Third, the geographic scale, or extent of big models is generally larger than that of conventional models but with similar spatial units of simulation. For instance, quality-of-life (QOL) studies can draw conclusions on a city using data at the block/parcel level. But with big models, the analysis of QOL can be conducted to a larger geographic scale, such as for a region or an entire country, and still maintain the same spatial resolution. Fourth, for the same geographical area, a big model can maintain at a higher spatial resolution when compared to a conventional model. A good example is that, in a national-scale population density research, the conventional models may only be applicable at the county or city level, whereas big models driven by fine-scale datasets make it possible to address the issue at the sub-district, block or parcel level, thus helping bring out more meaningful implication for urban spatial planning and policy-making.

Big models can be applied in the following avenues. First, urban dynamics from cities of all sizes can be investigated and examined using big models. Currently, most of applied urban models (AUMs) can only be adopted in large cities where

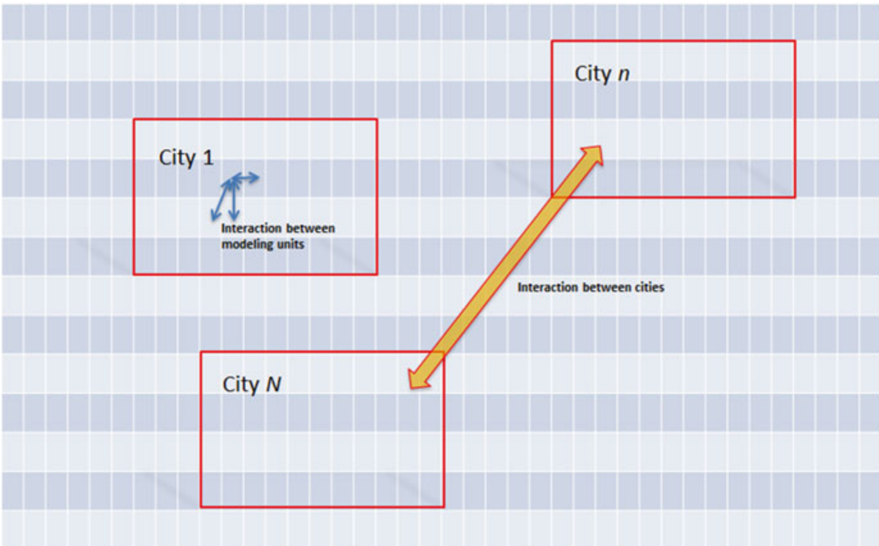


Fig. 13.3 An illustration of a big model integrating intra-and inter-city methodologies

data infrastructure and technical capacity are much better than those in middle- and small-sized cities in China. The introduction of big models could bridge the digital divide caused by data infrastructure. Second, focusing on individual data and fine-scale analyses and modeling, big models provide insightful solutions to various planning issues and contribute to a potential transition from a physical-concentration to a more collaborative and human-oriented planning process. Third, big models enable a variety of urban form and network indicators to be available and meaningful. These factors, combined with commonly adopted socio-economic aggregated indicators, can be adopted for inter-city analyses, which were particularly difficult previously due to lack of necessary road network and parcel geometries across many cities.

Simulating regional and urban dynamics using fine-scale and big models is advantageous as follows. (1) The adoption of local-level data with explicit geographical boundaries would be more appealing to local decision makers and citizens; (2) Land use regulations of spatial plan could be targeted directly at the fine spatial level. This would benefit those cities with limited capacity to analyze and forecast future development; (3) Such model can further be integrated with spatial interaction analysis (i.e. flows and networks).

13.3 Case Studies Using Big Models

Our efforts on the development and application of big models represent a first step towards a better understanding on cities using the emerging big data processing and analysis techniques in China. We outline our methodology development and research process of big models with several completed and ongoing research projects. As most of our case studies draw upon online data sources, the methodologies proposed in this chapter can be easily extended and applied to other cases.

13.3.1 Mapping Urban Built-Up Area for All Chinese Cities at the Parcel/Block Level

Urban built-up areas play a strong role in representing urban spatial development for planning decisions, management, and urban studies. They not only illustrate spatial patterns but also reveal socio-economic characteristics within the built-up areas, e.g., population aggregation, social interaction, energy consumption, and land use efficiency, thereby reflecting how a city evolves in a complex manner (Batty and Ferguson 2011). Conventional methods for delineating urban built-up area from the top down have been applied in major cities around the world on a large scale. However, such methods cannot be applied to most of cities in developing countries due to lack of high-resolution data (Long et al. 2013). Moreover, the research approach of the existing methods for fine-scale studies is conditioned by the presence of data and study context and hence varies from case to case. Against this backdrop, an automatic bottom-up approach was developed in this chapter. Built upon morphological and functional characteristics determined by street network as well as point of interests (POIs), the proposed approach creates a unified way to define fine-scale cities of all sizes.

The definitions and measurements of urban built-up areas have been varied. Urban built-up areas in the United States are defined as Urbanized Areas (UA) in a typical administrative model for spatial statistics. A UA comprises one or more "central places" areas and the adjacent densely settled surrounding "urban fringe" areas, with a total population of 50,000 or more (Morrill et al. 1999). A counterpart in Japan is called "Densely Inhabited District" (DID). DID is a district which has a population density of more than 4000 people per km^2. Urban Areas (UA) in UK are derived from entities-built areas, where certain real-estate densities are detected through satellite images (Hu et al. 2008). On the other hand, socio-economic factors are also adopted to describe the actual urban areas, e.g. labor force markets and commuter sheds are utilized to represent Metropolitan Area (MA) (Berry et al. 1969). Urban built-up areas can be utilized for different purposes with respect to population characteristics, economic status, and built environments attributes.

There are many ways of recording and mapping urban built-up areas. From the perspective of capturing morphological characteristics, an increasing attention has been focused on remote sensing images and street network. Remote sensing and night-time satellite imaging help us gauge urban activity and measure the extent and shape of built-up areas through capturing land cover information and interpreting light data (He et al. 2006). Apart from that, a number of indicators of street network have been introduced to describe the spatial layout of the built environment and predict their correlation with social effects. Examples are street intersection density (Masucci et al. 2012), fractal indices (Jiang and Yin 2014), integration, and accessibility. In terms of the functional characteristics, socio-economic statistics such as demographic densities (Rozenfeld et al. 2009), effective employment density (SGS Economics and Planning 2011, 2012), and infrastructures accessibilities (Hu et al. 2008) have emerged as a standard method of defining urban statistical areas (US Census Bureau 2014).

Nevertheless, these aforementioned approaches have some drawbacks. Firstly, such methods cannot be applied to most of cities in developing countries due to lacking necessary data and fine digital equipment. Moreover, these existing methods still require multiple steps according to unique conditions if achieving a fine-scaled result is expected. Furthermore, these existing approaches seem to isolate the spatial characteristics and the functional ones; therefore the real urban activities seem to be absent in snapping the urban areas by existing methods.

In light of this situation, this chapter employs an automated framework – "automatic identification and characterization of parcels (AICP)" – that was proposed by Long and Liu (2014) to delineate urban built-up areas at the parcel level, based on increasingly standardized roadway asset data from ordnance surveys and crowd-sourced point-of-interests (POIs) data. Roadway data are used to identify and describe parcel configuration, and POIs are processed to infer the intensity, function, and mixing of land use and human activities.

The working definition of a parcel is a geographical entity bounded by roads. Identifying land parcels and delineating road space are therefore dual problems. In other words, our approach begins with the delineation of road space, and individual parcels are formed as polygons bounded by roads. The delineation of road space and parcels is performed as follows: (1) All roadway data are merged as line features in a single data layer; (2) individual road segments are trimmed with a threshold of 200 m to remove hanging segments; (3) individual road segments are then extended on both ends for 20 m to connect adjacent but non-connected lines; (4) road space is generated as buffer zones around road networks. A varying threshold ranging between 2 and 30 m is adopted for different road types (e.g., surface condition, as well as different levels of roads); (5) parcels are delineated as the space left when road space is removed; and (6) a final step involves overlaying parcel polygons with administrative boundaries to determine whether individual parcels belong to a certain administrative unit.

We regard POI density as the ratio between the counts of POIs in/close to a parcel to the parcel area. We further standardized the density to range from 0 to 1 for better inter-city and intra-city density comparison using the following equation: standardized density = log(raw)/log(max), where raw and max correspond to density of individual parcels and the nation-wide maximum density value.[1] We also note that other measures (e.g. online check-ins and floor area ratio) can substitute POIs and approximate the intensity of human activities.

A vector cellular automata (VCA) model is adopted to identify urban parcels from all generated parcels. In this model, each parcel is assigned a value of 0 (urban) or 1 (non-urban). Initially, all parcels are assumed to be rural. To determine the actual status of each parcel, we should take into account not only the individual parcel's intrinsic attributes, such as population density, neighborhood attributes, and some other spatial variables, but also the status of neighboring parcels. The model stops at the iteration when the total area of simulated urban parcels reaches total urban land.

We applied this approach to map city boundaries for all Chinese cities and compared them with urban areas identified by GLOBCOVER, DMSP/OLS and population density. The simulation process and results highlight our proposed framework is more straightforward, time-saving and precise than conventional methods (see Fig. 13.4 for the results in typical cities).

The contribution of this work lies in three major aspects: data, methodology, and innovation. Firstly, the final product of this project is a database containing urban built-up area maps with detailed parcel features for 654 Chinese cities. Featured by fine-scale parcel information, the detailed road network and POIs datasets consolidated in this research can be applied to support a variety of planning and urban

[1] The unit is the POI count per km^2. For parcels with no POIs, we assume a minimum density of 1 POI per km^2.

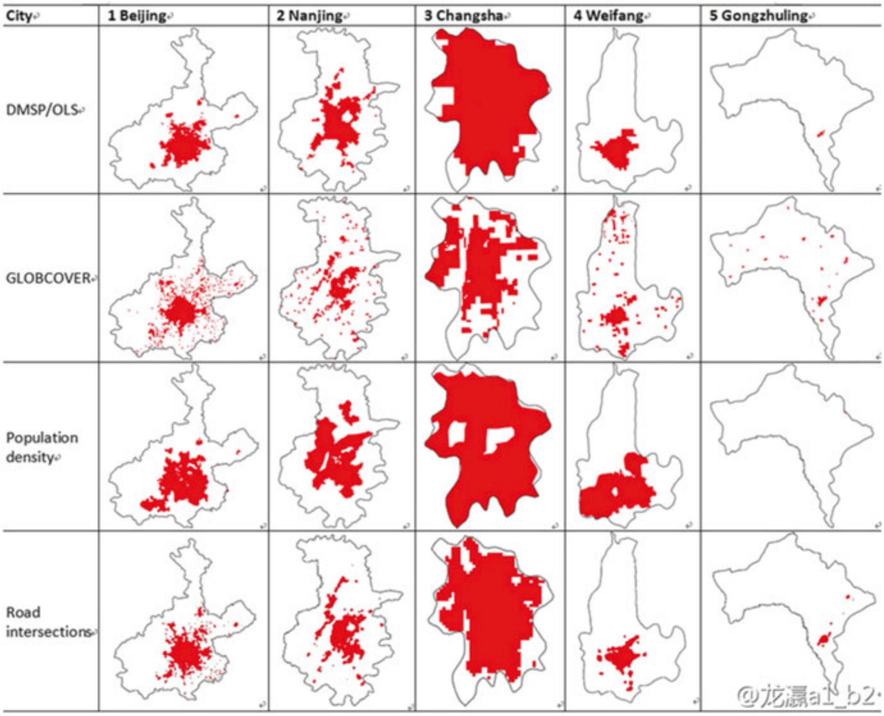


Fig. 13.4 Mapped urban areas in five typical Chinese cities by various methods

studies projects covering a wide range of geographic extent. Secondly, our research proposed a straightforward and consistent approach to identifying urban built-up areas across the country. Unlike previous methods that are somewhat laborious and subjective, our proposed methodology driven by VCA modeling is automatic, straightforward, and objective. The generated parcels can serve as basic spatial units for incorporating other high-resolution ubiquitous and spatially referenced data. In addition to the contribution of delineating urban built-up areas, this research also provides a robust framework for understanding complex urban system across cities from a bottom-up perspective.

13.3.2 Simulating Urban Expansion at Parcel Level for All Chinese Cities

China, as the largest developing country in the world, has experience rapid levels of urbanization in recent year since the introduction of Chinese Reform and Opening-up policies (Montgomery 2008; Liu et al. 2012). Featured by the history's largest flow of rural-to-urban migration and unprecedented economic growth, the urbanization process has shaped and transformed China from a rural to a more urban society. In light of this situation, increasing efforts on urban development assessment and management tools have been made in an attempt to promote a more sustainable development in China; among them are scenario-based urban simulation models (Zhang and Long 2013).

Large-scale simulation models are generally associated with large spatial units in space, like counties or super grids, sometimes reaching tens of square kilometers. Few applied urban models have the ability to pursue a large geographic scale extent with fine-level spatial units simultaneously due to data paucity and computation capacity limitation as discussed previously. Urban expansion simulation at a large geographic extent with a fine-scale (i.e. parcel scale) spatial unit could be promising for several reasons. Firstly, simulation and analysis at the parcel level would be more meaningful for local planners, decision makers, and residents to understand, administer, and monitor urban developments. Secondly, simulation modeling at the large geographic extent enables those administrative entities who have limited capacity to analyze and forecast the urban growth taking place within their boundaries by their own to have an insight on overall urban development scenario within the region and to gauge their growth and take action properly. Also, such simulation models make inter-city comparison possible.

In this section, we developed a mega-vector-parcels cellular automata model (MVP-CA) for simulating urban expansion in the parcel level for all 654 Chinese cities. Three modules, the macro module, the parcel generation module, and the vector CA module, were included in the MVP-CA, as shown in Fig. 13.5.

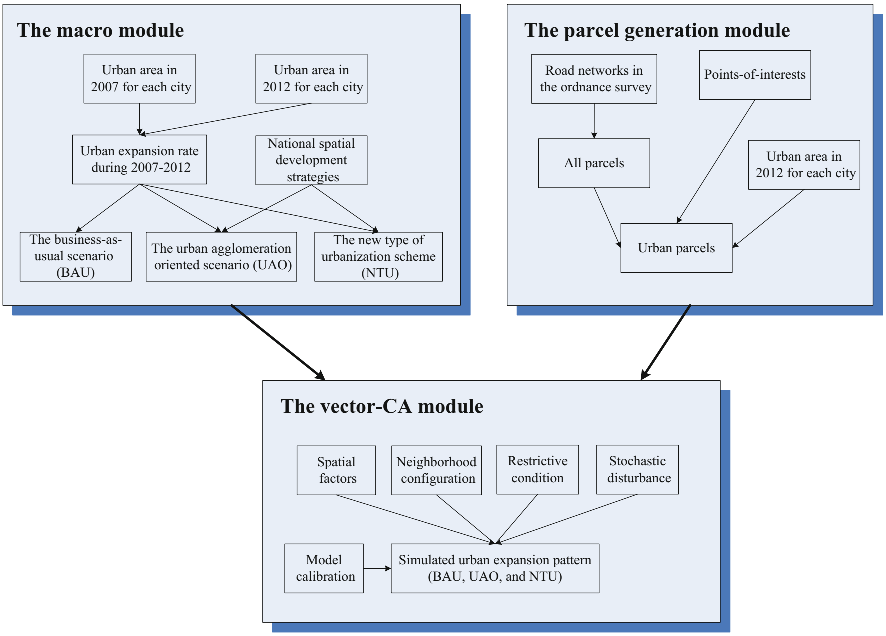


Fig. 13.5 The structure and flow diagram of MVP-CA

The macro module was responsible for setting urban expansion rate in the next five years for each city, taking into account historical urban expansion rate and national spatial development strategies. The parcel generation module was used for identifying existing urban parcels in 2012 using the framework of AICP (automatic identification and characterization of parcels) proposed by Long and Liu (2014). The vector CA module was applied for simulating urban expansion during 2012–2017. This module was examined using calibrated parameters abstracted from Beijing data. Three urban expansion scenarios – baseline, urban agglomeration, and new urban development- have been simulated during 2012–2017 by MVP-CA, respectively. The simulation results are shown in Fig. 13.6. We validated the simulation results by comparing the baseline scenario of Beijing with the results using a raster CA model BUDEM we developed previously.

As one of the first large-scale urban expansion models at the fine-scale for the whole China, our contributions of this chapter mainly lie in the following aspects. First, a vector-based cellular automata model was introduced for simulating urban expansion in a large geographical scale at the parcel level, which is rare in existing literature in the domain urban expansion modelling. Second, we proposed a solution for linking spatial development strategies with urban expansion via reflecting as the urban expansion speed of each city. This enables simulating macro policies in a very fine-scale through the channel of the MVP-CA model. Last, we simulated the near-future urban area for all Chinese cities in China, which, together with existing urban area, has already been shared online as an important data infrastructure for both practitioners and researchers.

13.3.3 Evaluating Urban Growth Boundaries for 300 Chinese Cities

Among the various urban growth management policies, urban containment policies have been widely adopted in an attempt to control the spread of urban areas, increase urban land use density, and protect open space (Nelson and Duncan 1995; Long et al. 2011). In general, urban containment policies seek to manage urban growth through at least three different types of tools – greenbelts, urban growth boundaries (UGBs), and urban service boundaries (USBs) (Pendall et al. 2002). UGB is one of the most widely discussed tools in the planning field. Through zoning, land development permits, and other land-use regulation tools, UGBs demarcate urban and rural uses and aim to contain urban development within the predefined boundaries (Pendall et al. 2002). In China, urban construction boundaries determined in master or detailed plans have been commonly recognized as Chinese/planned UGBs (Long et al. 2013), since they have a similar mechanism to UGBs in the U.S. as well as some other Western countries.

In China, conventional methods of delineating UGBs are based on planners' expertise and experiences; thus, they lack an adequate scientific basis and quantitative support. Consequently, the UGBs often fail to manage urban growth. According

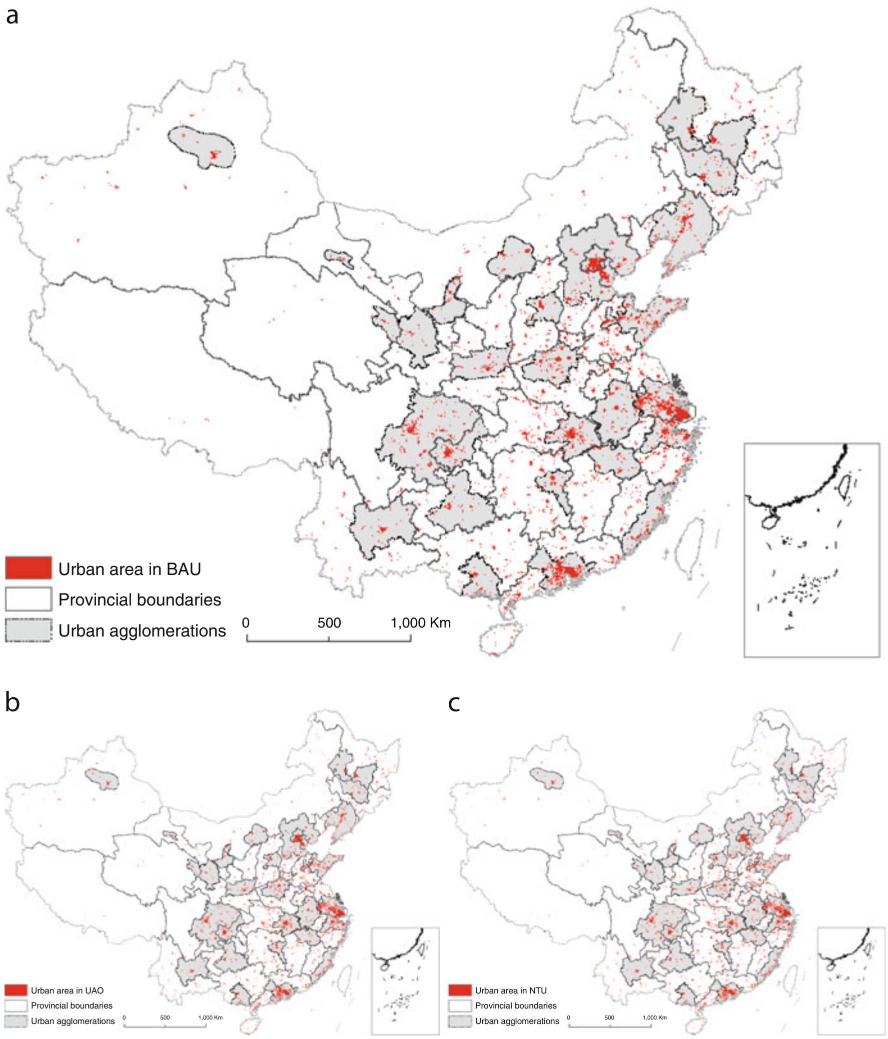


Fig. 13.6 Urban area of all Chinese cities (**a**), and urban expansion patterns of the entire China for three scenarios ((**a**) BAU, (**b**) UAO, (**c**) NTU)

to Han et al. (2009)'s study on the examination of the implementation of planned UGBs within the sixth ring road of Beijing using multi-temporal remote sensing images, more urban land developments were found outside than inside the UGBs during the previous two planning periods (1983–1993 and 1993–2005). Tian and Shen (2011) and Xu et al. (2009) also suggested that substantial urban development occurred outside of UCBs in Guangzhou and Shanghai in recent years. These findings were also supported by Long et al. (2012)'s research, which evaluated five master plans compiled and implemented in Beijing during 1958–2004. Though

Fig. 13.7 The profile of raw figures for planned UGBs (partially shown)

considerable progress has been made in revealing and quantifying the extent of urbanization and/or evaluating the urban policies' effectiveness on managing urban growth, we have found that most of them have been focused on a single city or region, and little work on the city-level comparison of the performance of UGB's implementation has been done.

Driven by our proposed urban growth simulation model and other relevant big models studies, we launched effort to create a systematic approach to horizontally examine and evaluate the effectiveness of UGBs across cities and regions. We collected raw planning drawing maps on planned UGBs in over 300 Chinese cities (see Fig. 13.7 for a partial sample of cities) and digitalized the boundaries in GIS to facilitate spatial analysis and statistics on these planned UGBs. After that, the planned UGBs of a city were overlaid and compared with the actual extent of urban expansion in the past years since the plan was first implemented, and the ratio of legal development to all urban development can be directly calculated to facilitate city-level comparison. Furthermore, the ambitious degree of each city can be inferred by dividing the actual extent of urban expansion by the planned-to-be-development land area.

Compared with previous studies on big models to date, this research generalizes the planned UGBs across cities and regions and helps make sense of differing results of urban development. In addition, it can provide an insight of the overall trend of urban development in China and thus would be useful for planners to evaluate, monitor, and manage urban planning efforts.

Meanwhile, the digitalized UGBs can also be used to supplement the MVP-CA urban expansion model for all Chinese cities (see our first case study in this chapter) as an institutional constraint, thus accounting for the simulation results. In addition, the project may help identify some universal law of governing the pattern of planned UGBs among all Chinese cities.

13.3.4 Estimating Population Exposure to PM2.5

Chinese cities have for many years suffered from air pollution, which has been a major downside to rapid economic growth and increased urbanization. Currently, few studies of air pollution have been conducted to assess population exposure to PM2.5 over large geographical areas and time periods in China. The existing studies mainly focus on air pollution's effects on health and ecosystems or relevant monitoring methods and measurement, but less has been done on the link between urban spatial structure and air pollution exposure, not to mention their spatiotemporal pattern.

In this study, we collected daily PM 2.5 concentrations during April 08, 2013 and April 07, 2014 from 945 monitoring stations in 190 cities across China.[2] The air quality data were acquired from China National Environmental Monitoring Center (http://www.cnemc.cn). These datasets enable us to understand the PM 2.5 concentration of each station all year round, and can be used as a key input for our estimation. Considering the sparse distribution of monitoring stations across China, we further used Moderate Resolution Imaging Spectroradiometer (MODIS) Aerosol Optical Depth (AOD) retrievals to supplement the PM 2.5 estimates on a daily basis. Demographic statistics were drawn from China's 2010 census data. The spatial distribution of population density across China was determined by geocoding population density of each sub-district based on Google Map API. In total, there are 39,007 sub-districts[3] in China, and the average population density for all sub-districts is 977 persons per km^2. Population have been divided into three age groups (age 0–14, age 15–64, and 65 years and older), with an aim to differentiate the exposure estimates for different sensitive groups such as children and seniors. It is worth mentioning that this is the first time to use sub-district population density for estimating human exposure to air pollution in China, whereas former studies were conducted at the county level at best.

The population exposure estimation involves three major steps. (1) Interpolate the PM2.5 concentration site data into surface data using both ground station-level data and MODIS ADO: PM2.5 concentration data were obtained from all air quality monitoring stations across the country and supplemented with MODIS ADO data. Using numerous spatial interpolation methods, the station-level data can be interpolated into surface data. The outcome of this step is the average daily PM2.5 concentration over the entire area and over time. (2) Estimating population exposure to PM2.5 for each sub-district. Based on interpolated PM2.5 data, a daily PM2.5 concentration above the national standard of 75 mg/m^3 is considered to be unhealthy and thus defined as "exposed". In this way, the total exposed days all year round of each sub-district can be estimated. Further, the exposure intensity for each sub-

[2] There are 657 cities in mainland China as of the end of 2012.

[3] There are three forms of township-level administrative units in China, sub-districts (*jiedao*), towns (*zhen*), and township (*xiang*). Jiedaos are mainly in city area. *Jiedao*'s counterparts in the rural area are towns and townships. Hereafter in this chapter, we use the term sub-district for representing all types of township-level administrative units in China.

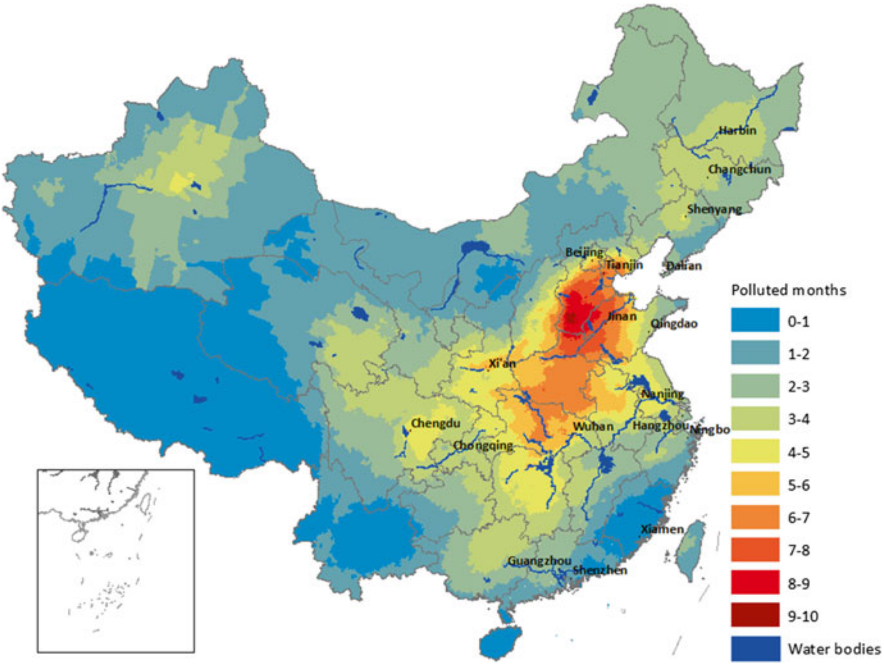


Fig. 13.8 The number of total exposed months for each sub-district in China

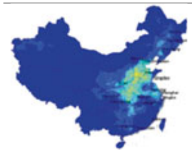
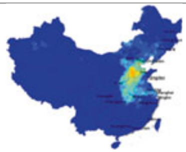
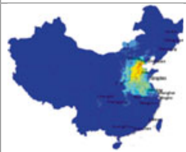
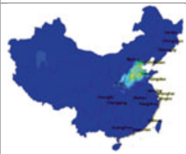
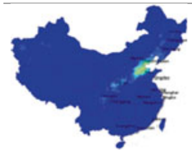
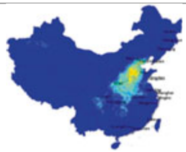
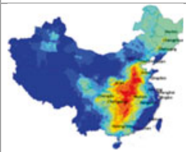
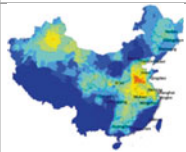
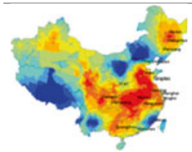
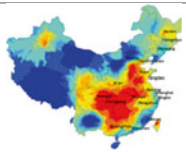
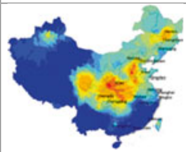
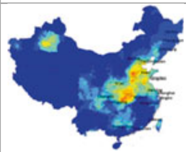
district can be calculated using the Equation: Exposure intensity = Population density * Exposed days. The greater exposed days or population density for a sub-district, the higher exposure intensity. This indicator reflects the strength of population exposure to PM2.5. The population density can be subject to specific sub-population groups for estimating the effects on members of sensitive groups. (3) Aggregating the estimated results spatiotemporally. To gain ideas on spatiotemporal pattern of population exposure to PM2.5, we can further aggregate the estimated results in both temporal and spatial dimensions. For the temporal dimension, the total number of exposed month can be calculated for each sub-district, thus presenting a big picture of population exposure to air quality over time. For the spatial dimension, the exposure of each city can be inferred by averaging the estimation results of all sub-districts in each city's administrative boundary.

The number of months subject to exposed condition across the entire country is presented in Fig. 13.8.

The daily exposure for each sub-district was further aggregated by each month. Table 13.1 displays the percentage of exposure days per month from April 2013 to March 2014.

The exposure intensities were obtained by multiplying population density for each sub-district with the estimated exposure days during the period. The final result is presented in Fig. 13.9. It is worth pointing out that the overall exposure intensity pattern generally coincides with the distribution of population density across the country.

Table 13.1 Exposed days in each month for each sub-district in China

Apr 2013	**May 2013**	**Jun 2013**	**Jul 2013**	**Exposed percentage in each month**
				.0 - .1 .1 - .2 .2 - .3
Aug 2013	**Sep 2013**	**Oct 2013**	**Nov 2013**	
				.3 - .4 .4 - .5 .5 - .6
Dec 2013	**Jan 2014**	**Feb 2014**	**Mar 2014**	
				.6 - .7 .7 - .8 .8 - .9 .9 - 1.0

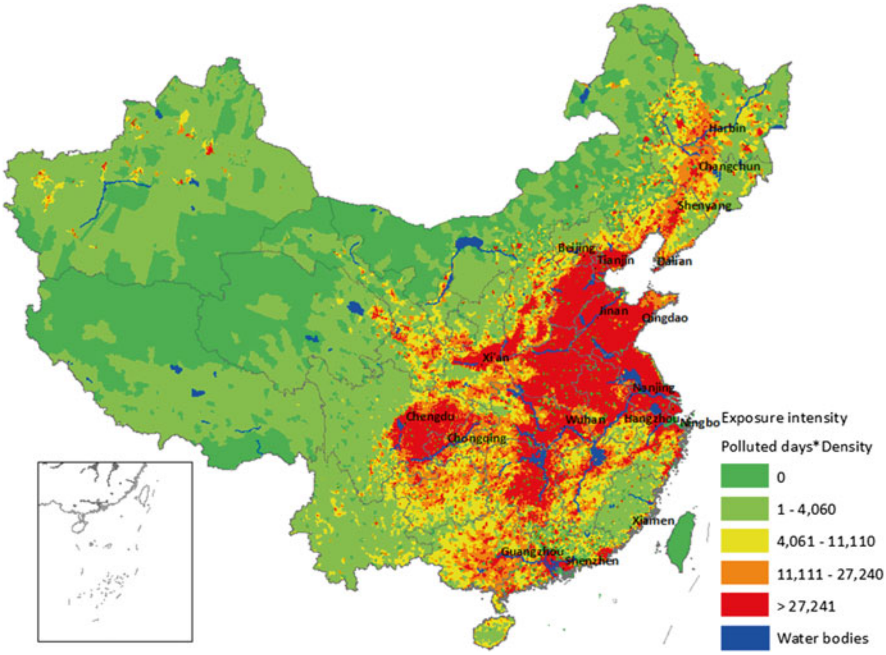


Fig. 13.9 Exposure intensity at the town level of China

13.4 Conclusions and Future Directions

This chapter has proposed the concept of big model as a novel research paradigm for regional analysis and urban studies. The concept, characteristics, and potential applications of big models have been elaborated. Meanwhile, we addressed several case studies to illustrate the progress of research and applications of big models, including mapping urban areas for all Chinese cities, performing parcel-level urban simulation, and several ongoing research projects. Most of these applications can be adopted across the whole country, and all of them are focusing on a fine-scale level, such as a parcel, a block, or a township (sub-district), which is quite different from the existing studies using conventional models. Believing that big models will mark a promising new era for the urban and regional studies in the era of new data environment, we hope our efforts on urban analytics and modeling in Beijing City will set new research agenda and inspire innovative ideas all over the country.

There are several avenues on big models that deserve further studies. First, it is necessary to combine both intra-city and inter-cities methods in big models. Existing case studies in this chapter mainly rely on bottom-up intra-city approaches. City level connections are essential to be included in big models in the next step. For instance, a spatial equilibrium module considering the city level input-output would be helpful for the current MVP-CA model via addressing the interaction between cities. Second, more general theory on big models can be identified through more in-depth case studies analysis.

Acknowledgement We thank Ms Yichun Tu for her editing the language of this chapter.

References

Batty, M. (2009). *Urban modeling. International encyclopedia of human geography*. Oxford: Elsevier.

Batty, M. (2012). Building a science of cities. *Cities, 29*, S9–S16.

Batty, M., & Ferguson, P. (2011). Defining city size. *Environment and Planning B: Planning & Design, 38*(5), 753–56.

Berry, B. J. L., Goheen, P. G., & Goldstein, H. (1969). *Metropolitan area definition: A re-evaluation of concept and statistical practice 28*. Washington, DC: US Bureau of the Census.

Han, H., Lai, S., Dang, A., Tan, Z., & Wu, C. (2009). Effectiveness of urban construction boundaries in Beijing: An assessment. *Journal of Zhejiang University SCIENCE A, 10*, 1285–1295.

He, C., Shi, P., Li, J., Chen, J., Pan, Y., Li, J., Li, Z., & Ichinose, T. (2006). Restoring urbanization process in China in the 1990s by using non-radiance calibrated DMSP/OLS nighttime light imagery and statistical data. *Chinese Science Bulletin, 51*(13), 1614–1620.

Hu, Y., Wu, Z., Xiong, W., & Pan, C. (2008). Study of identifying urban built-up area: Taking Wuhan as an example. *City Planning Review, 32*(4), 88–92.

Hunt, J. D., Kriger, D. S., & Miller, E. J. (2005). Current operational urban land-use–transport modelling frameworks: A review. *Transport Reviews, 25*(3), 329–376.

Jiang, B., & Yin, J. (2014). Ht-index for quantifying the fractal or scaling structure of geographic features. *Annals of the Association of American Geographers, 104*(3), 530–540.

Liu, Z., He, C., Zhang, Q., Huang, Q., & Yang, Y. (2012). Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landscape and Urban Planning, 106*(1), 62–72.

Long, Y., & Liu, X. (2014). Automated identification and characterization of parcels (AICP) with OpenStreetMap and points of interest. arXiv preprint. arXiv:1311.6165.

Long, Y., Shen, Z., & Mao, Q. (2011). An urban containment planning support system for Beijing. *Computers, Environment and Urban Systems, 35*(4), 297–307.

Long, Y., Gu, Y., & Han, H. (2012). Spatiotemporal heterogeneity of urban planning implementation effectiveness: Evidence from five master plans of Beijing. *Landscape and Urban Planning, 108*, 103–111.

Long, Y., Han, H., Lai, S.-K., & Mao, Q. (2013). Urban growth boundaries of the Beijing metropolitan area: Comparison of simulation and artwork. *Cities, 31*, 337–348.

Masucci, A. P, Stanilov, K., & Batty, M. (2012). Limited urban growth: London's street network dynamics since the 18th century. arXiv preprint. arXiv:1206.5298.

Montgomery, M. R. (2008). The urban transformation of the developing world. *Science, 319*(8), 761–764.

Morrill, R., Cromartie, J., & Hart, G. (1999). Metropolitan, urban, and rural commuting areas: Toward a better depiction of the United States settlement system. *Urban Geography, 20*(8), 727–748.

Nelson, A. C., & Duncan, J. B. (1995). *Growth management principles and practices*. Chicago/Washington, DC: Planners Press/American Planning Association.

Pendall, R., Martin, J., & Fulton, W. (2002). *Holding the line: Urban containment in the United States*. Washington, DC: The Brookings Institution Center on Urban and Metropolitan Policy.

Rozenfeld, H. D., Rybski, D., Gabaix, X., & Makse, H. A. (2009). The area and population of cities: New insights from a different perspective on cities (No. w15409). National Bureau of Economic Research.

SGS Economics & Planning. (2011). Melbourne's economy: A stunning success or captured by complacency?, presentation at State Library of Victoria by Terry Rawnsley.

SGS Economics & Planning. (2012). Melbourne metro: Move to more productive jobs, draft version 2.0, 3 February 2012.

Tian, L., & Shen, T. (2011). Evaluation of plan implementation in the transitional China: A case of Guangzhou city master plan. *Cities, 28*, 11–27.

United States Census Bureau. (2014). History: Metropolitan areas. U.S. Census Bureau. http://www.census.gov/history/www/programs/geography/metropolitan_areas.html. Accessed 31 March.

Wegener, M. (2004). Overview of land-use transport models. *Handbook of Transport Geography and Spatial Systems, 5*, 127–146.

Xu, Y., Shi, S., & Fan, Y. (2009). Methodology of Shanghai city master planning in new position. *Urban Planning Forum, 2*, 10–15.

Zhang, Y. P., & Long, Y. (2013). *Urban growth simulation using V-BUDEM: A vector-based Beijing urban development model*. Beijing: The conference of Spatial Planning and Sustainable Development.