Beijing City Lab

Long Y, Liu X J, 2013, Automated identification and characterization of parcels (AICP) with OpenStreetMap and Points of Interest. Beijing City Lab. Working paper # 16

Automated identification and characterization of parcels (AICP) with OpenStreetMap and Points of Interest

Ying Long[1,2], Xingjian Liu[3]

1. Beijing Institute of City Planning, China
2. University of Cambridge, UK
3. University of North Carolina at Charlotte, USA

Automated identification and characterization of parcels (AICP) with OpenStreetMap and Points of Interest

Abstract:
Aiming at the paucity of urban parcels in developing countries in general and China in particular, this paper proposes a method to automatically identify and characterize parcels (AICP) with ubiquitous available OpenStreetMap (OSM) and Points of Interest (POIs). Parcels are the basic spatial units for fine-scale urban modeling, urban studies, as well as spatial planning. Conventional ways of identification and characterization of parcels rely on remote sensing and field surveys, which are labor intensive and resource-consuming. Poorly developed digital infrastructure, limited resources, and institutional barriers have all hampered the gathering and application of parcel data in developing countries. Against this backdrop, we employ OSM road networks to identify parcel geometries and POI data to infer parcel characteristics. A vector-based CA model is adopted to select urban parcels. The method is applied to the entire state of China and identifies 82,645 urban parcels in 297 cities.

Keywords: Open street map (OSM), Point of Interest (POI), land parcel, automatic generation, urban planning

Introduction

Land parcel data are one of the cornerstones of contemporary urban planning (Cheng et al. 2006). Using as an analytical tool, parcels are the basic spatial units of urban models, as for example the latest urban simulation models are oftentimes vector-based and capture parcel-level dynamics (Pinto 2012; Stevens and Dragicevic 2007). More importantly, normative planning and policies are performed on parcels, ranging from devising master and detailed urban plans, to strategic plan implementation, and to policy effects evaluation (Frank et al. 2006; Jabareen 2006; Alberti et al. 2007).

Whereas parcel data for the developed world are generated by robust digital infrastructure and supplemented by open data initiatives (e.g., OpenStreetMap), researchers still lament the difficulty of attaining parcel data for developing countries. For example, the best available parcel map for China's capital Beijing – supposedly one of the most technologically advanced and rapidly developing cities in the erstwhile Third World – was dated in 2010 (Beijing Institute of City Planning 2010). In addition, collecting parcel data in medium- and small- sized cities in China is constrained by poorly developed digital infrastructures. That goes without saying that complete parcel-level features (e.g., land use type, urban functions, and development density) do not exist for many occasions. In addition to hard infrastructures, soft institutions have also created barriers for Chinese urban planners' access to parcel maps. For instance, our interviews with 57 planning professionals[1] reveal that access to existing parcel maps held by local planning bureaus/institutes is extremely restrained, as parcel maps are tagged as confidential within the current Chinese planning institutions. In summary, parcel data for the developing world are oftentimes outdated, limited in geographical scopes, and do not contain much necessary information other than basic parcel geometry. This condition has limited the progress of quantitative urban studies, urban planning compilation as well as urban management in developing countries in general and in China in particular.

As parcel data are moving towards the central stage of urban planning (Cheng et al. 2006), the lack of parcel data for cities in the developing world would constrain our ability to trace urban changes at high spatial-resolution, hinder the compilation and implementation of detailed urban plans, and rule out the possibility of adopting contemporary parcel-based urban management. Built on manual interpretation of remote sensing images and field surveys, conventional ways of generating parcel data are time-consuming, expensive, and labor-intensive (Erickson et al 2013). Thus many developing countries do not have the necessary capital and resources to produce parcel data in the conventional fashion. Overcoming such "data desert" scenario seems to be of high priority for urban planning in developing countries.

Against this backdrop, we propose a method for automatic identification and characterization of parcels (AICP), based on freely-available Open Street Map (OSM) and crowd-sourced Point-of-Interest (POI) data. The proposed method could (1) provide quick and robust delineation of land parcels; and (2) generate a variety of parcel level attributes, allowing for the examination of urban functions, development density and mixed land uses. We illustrate the usefulness of our method with 297 cities in China. Drawing upon entirely on open data, the methods developed in this paper can be easily extended to other cities in the developing world. The next section reviews progress in obtaining parcel-level geometry and features, followed by an elaboration of methods and the case study. We conclude with a discussion of the strength, limitations, as well as future applications of our method.

[1] We have interviewed 57 planning professionals in China (23 of these 57 professionals are affiliated with research institutes and universities; 21 with foreign planning/architect firms such as AEOCOM and Atkins; and 13 with domestic planning institutes and firms). Professionals often rely on manual digitalization of land use maps (often in image format as vector data files would not be released). This process is extremely time consuming and often produces land parcel maps of less desirable quality.

Background

Parcel boundaries and their features are conventionally identified through manual interpretation of high-resolution remote sensing images, topographic maps, building maps and field surveys. Manual operations are resource and time consuming (for example, it would take an experienced operator 3-5 hours to identify and infer land use for 35-50 urban parcels covering the area of one square kilometers); producing inconsistent data (as data quality largely depends on the experience and proficiency of individual practitioners); and not suitable for longitudinal updates and comparison. Volatile urban development (e.g., gentrification and urban sprawl) in many developing countries has also added difficulty to update existing parcel maps. Still, compiled parcels generally lack parcel level information, such as density and land use mix. As a case in point, data about parcel density for Beijing, China after 2010 could not be obtained, and are limited to the area within the sixth ring road (approximately 13.8% of the whole Beijing Metropolitan Area).

Attempts have been made to identify automatically parcel geometrics. For example, Yuan et al. (2012) proposed a raster-based approach for parcel delineation based taxi trajectories and POIs. However, Yuan et al. (2012) omitted road space in the delineation of parcels, the raster-based nature of the method generates heavy computational burden, severely limiting the method's applicability to large geographical areas. Meanwhile, Aliagaet al. (2008) presented an algorithm for interactively synthesizing parcel layouts for to-be-developed areas according to the structure of real-world urban areas. This study is limited by the fact that it does not account for parcel characteristics, and performs parcel subdivision within predefined blocks, instead of identifying blocks from the data.

In light of this situation, OSM has been proposed as a promising candidate for quick and robust delineation of parcels (Haklay and Weber 2008; Ramm 2010). As one of the most successful volunteered GIS projects, OSM provides street network data for a wide array of cities (Goodchild 2007; Sui 2008). Jokar Arsanjani et al. (2013a) predict that the data coverage and quality of OSM will continue to be improved in the coming years. More specifically, OSM data quality in well-mapped and oftentimes large cities is on par with that of topographic maps (Girres and Touya 2010; Haklay 2010; Over et al. 2010). The growth of OSM in developing countries has been encouraging, as the volume of OSM data in China has experienced a nine-fold increase during 2007-2013 (Figure 1).

Several preliminary studies suggest that OSM road networks are useful in identifying urban structures. For example, Hagenauer and Helbich (2012) extracted urban built-up areas from OSM, and Jiang and Liu (2012) identified natural grouping of city blocks based on OSM data. Existing analyses using OSM focus more on deriving universal laws and social physics (Jiang and Liu 2012) rather than producing data products for urban planning and studies. In a similar vein, Jokar Arsanjani et al. (2013b) identified land-use patterns for central Vienna, Austria (roughly 32 km^2) using OSM. Whereas Jokar Arsanjani et al. (2013b) introduced a volunteered geographic information based approach to generate land-use patterns, their approach focuses on developed countries with high-accuracy OSM data, could be extended to generate additional parcel features.

In addition to parcel geometries, planning practices also require parcel features such as urban functions and development density. There is a rich literature on inferring land use from remote sensing images (Kressler et al. 2001; Herold et al. 2002). However, as discussed previously, remote sensing images are not suitable for large scale parcel-level analysis, due to *inter alia* data availability and the sheer amount of resources required. Although some automatic or semi-auto techniques have been developed to address urban land-use classification (Herold et al. 2002; Pacifici et al. 2009), it is still difficult to identify certain land use types such as high-density residential areas and commercial areas from remote sensing images. More importantly, remote-sensing based methods often treat parcels as having homogenous land use types, and do not allow for quantitative analysis of mixed land use. More recently, researchers have inferred human use of urban space with human mobility data, such as smart card records (Long et al. 2013), mobile phone data (Soto and Frias-Martinez 2011; Toole et al. 2012), and taxi trajectories (Liu et al 2012; Yuan et al. 2012). Nevertheless, human mobility data are hard to obtain as they often involve privacy issues as well as profit-seeking data holders (Beresford and Stajano, 2003). Such data paucity greatly undermines the wide applicability of human-mobility based methods for large geographic regions.

To this end, we argue that online POI data provide an alternative data source for characterizing parcels. The strength of POI data includes (1) containing sub-parcel level business information, which could serve as proxies for land use and urban functions; (2) being freely available from online mapping and cataloguing service providers; (3) having a nearly global coverage; and (4) having high spatial (e.g., geo-coded business locations) and temporal (e.g., routinely updated by service providers) resolutions. With all these advantages, it is surprising that few studies have tapped POI data's potential in characterizing parcel features.

Therefore, aiming to improve the identification and characterization of fine-scale urban land parcels, we introduce an automatic process using open data. OSM data are used to identify and delineate parcel geometries, while crowd-sourced POIs are gathered to infer land use intensity, function, and mixing at the parcel-level. We emphasize that our empirical framework is (1) fully automatic and use open data, allowing for the incorporation of other data sources (e.g., taxi trajectories and mobile phone data); (2) produce not only parcel geometry and land use types but also useful parcel-level information such as land use mix; (3) is applicable to large geographic areas, while most previous studies are limited to small areas; and (4) enables routine updates and free release of urban parcel data for China.

Data and methods

Data

Administrative boundaries of Chinese cities

Our analysis focuses on a total of 654 cities in China (Figure 2)[2], ranging across five administrative levels: namely municipalities directly under the Central Government (MD, 4 cities), sub-provincial cities (SPC, 15), other provincial capital cities (OPCC, 16), prefecture-level cities (PLC, 251), and county-level cities (CLC, 368) (Ministry of Housing and Urban Development, MOHURD, 2013; see Ma, 2005 for more details about the Chinese administrative system). As a city *proper* in China contains both rural and urban land uses, we narrow our analytical scope onto legally defined urban land within city *propers* and use administrative boundaries of urban lands to carve out OSM and POI data layers. In addition to administrative boundaries, we also gather information about total build-up area of individual cities in 2012 (MOHURD, 2013), which will be used in the urban parcel identification process.

OSM in China

OSM road networks for China were downloaded on October 5, 2013. We also gather the ordnance survey map of China at the end of 2011 with detailed road networks to verify results produced by OSM data. The OSM dataset contains 481,647 road segments (8.0% of that of the ordnance survey map) of 825,382 kilometers (31.5% of the ordnance survey map). Furthermore, road networks in OSM and the ordnance survey map are overlaid for a visual inspection of data quality (Figure 3). Although capturing a portion of the ordnance survey data, OSM data cover most urban areas in

[2] Sansha in Hainan and Beitun in Xinjiang appearing in MOHURD (2013) were not included due to spatial data availability. Taiwan was not included in all analysis and results in this paper.

China, especially large cities (Figure 3), and are potentially useful for identifying urban land parcels. The implication of OSM data quality will be elaborated in the final section.

POIs

A total of 5,281,382 POIs are gathered from and geocoded by business cataloging websites. The initial twenty POI types are aggregated into eight more general groups (Table 1): Commercial sites account for most POIs, followed by business establishments, transportation facilities, and government buildings. POIs labeled as "others" are used in estimating land use density, but removed from land use mix analysis as this type of POIs with mixed information are not well organized and classified according to our review. We also employ manual checking of randomly sampled POIs to ensure the data quality. Our empirical framework is extensible in the sense that POI counts can be replaced by other human activity measurements, ranging from the more conventional land use cover derived from remote sensing images to ubiquitously available online check-in service data (e.g., Foursquares).

Other data

DMSP/OLS (1-km spatial resolution; Yang et al. 2013) and GLOBCOVER (300-m spatial resolution; Bontemps 2009) remote sensing images are also obtained for model validation, as we will benchmark parcels identified with our empirical framework with those identified from remote sensing images, although the spatial scale varies between each other. In addition, manually generated parcel data for Beijing is gathered from BICP.

Methods

Delineating parcel boundaries

The working definition of a parcel is a continuously built-up area 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 OSM road data are merged as line features in a single data layer; (2) individual road segments are trimmed with a threshold of 200m to remove hanging segments; (3) individual road segments are then extended on both ends for 20m to connect adjacent but non-connected lines; (4) road space is generated as buffer zones around road networks. A varying threshold ranging between 2-30 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 (Figure 1); and (6) a final step involving overlaying parcel polygons with administrative boundaries to determine whether individual parcels belong to a certain administrative unit. Parameters used in these steps are determined pragmatically with topological errors of OSM data in mind.

Calculating density for all parcels

We define land use density as the ratio between the counts of POIs in/close to a parcel to the parcel area[3]. 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[4]. 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.

[3] POIs within the buffered road space were accounted by their closest parcels in our experiment.
[4] 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.

Selecting urban parcels

The next step identifies urban parcels from all generated parcels. The total urban land of individual cities was gathered from MOHURD (2013). We employ a vector-based constrained cellular automata (CA) model to identify urban parcels in individual cities[5] (Zhang and Long, 2013). More specifically, we use the CA model to predict possibility of being urban for individual parcels, and the total urban land is used as constraints for aggregated amount of urban parcels.

In the CA model, each parcel is regarded as a cell in CA, and the cell status was 0 (urban) or 1 (non-urban). The CA model essentially simulates the urban development. On the onset of the simulation, all cells are set to be rural. During each step during the simulation, whether a parcel is converted to "urban", i.e., the probability of being urban, depends on two aspects (Li and Yeh, 2002): Firstly, the proportion of neighboring parcels that are urban. In our empirical operationalization, the neighborhood of a parcel includes all parcels within a 500 m radius; and secondly, individual parcels' intrinsic attributes such as size, compactness, and the POIs density. These three attributes are combined using a logit function to influence individual parcels' probability of being urban (Wu, 2002). We then multiply the two aspects (neighborhood and intrinsic attributes) to determine whether the final probability is over a predefined threshold. In other words, a parcel surrounded by many urban parcels and with a high intrinsic probability would have more chance to be identified as an urban parcel in the simulation. Figure 4 provides a visual illustration of our CA model, where the final probability for being selected as an urban parcel for parcels A, B and C is 0.6 (0.8*6/8), 0.3 (0.6*4/8), and 0.225 (0.9*2/8) respectively. With a threshold of 0.5, the only parcel that would be selected as urban in our simulation is parcel a. The model stops at the iteration when the total area of urban parcels reaches total urban land.

For calibrating the weights for constraints, we conducted logistic regression on the existing parcels manually prepared by planners in the city of Beijing (12,183 km^2; Yanqing and Miyun counties in the Beijing Metropolitan Area are not included). Each parcel was regarded as a sample, and totally there were 125,401 samples (among them 57,817 urban parcels). The whole precision of logistic regression was 74.2%. The logistic regression results shown in Table 2 were applied in constrained CA models for all cities in China[6]. Our constrained CA model was meanwhile used in Beijing for model validation. The overall accuracy of 78.6% in terms of parcel count indicated the applicability of our CA model in identifying urban parcels from all parcels generated in a city.

Inferring dominating urban function and land use mix for selected urban parcels

Urban function for individual parcels is identified by examining dominant POI types within the parcels. A dominant POI type within a parcel is defined as the POI type that has accounted for more than 50% of all POIs within the parcel. For example, if 31 out of 60 POIs within a parcel are labeled as "business establishment", the urban function for that parcel will be assigned as "business". Note that not all parcels would have a dominant urban function.

As a supplement measurement for the dominating function, we computed a mix index to denote the land use mixed degree (Frank et al. 2004). The mixed index (M) of a land parcel is calculated as $M = -\mathrm{sum}(p_i * \ln p_i)$ ($i = 1, ..., n$), where n denotes the number of POI types, and pi is the proportion of POI type i among all POIs in the parcel. This index has been used before to better understand evolving

[5] Each city has its own constrained CA model for identifying urban parcels.
[6] We admit the heterogeneity of weights in various cities, however we do not have existing parcels for other cities at the time of this research.

travel mode choice and public health outcomes, as well to study changing senses of community (Manaugh and Kreider2013).

Validation

Our parcel identification and characterization results are validated at two spatial levels: At a first more fine spatial scale (i.e., parcel level), we compare the geometry and attributes of urban parcels generated by our program with those identified manually in the conventional approach. Due to data availability, this fine scale comparison is only performed for the city of Beijing (i.e., the aforementioned BICP data). Since urban parcels for Beijing was collected in 2010 with a total urban area of 1677.5 square kilometers, we re-ran the constrained CA model in Beijing using this total number and regenerated urban parcels for Beijing[7]. In order to remedy the limited availability of manually collected parcel data, we perform a second validation with full geographic coverage at the aggregated level (i.e., regional level). In this second validation, we focus on the statistical distribution of parcel geometries, and compare urban parcels identified from OSM and Ordnance Survey maps. To ensure the comparability of urban parcels from different approaches, we use road networks from the ordnance survey map (ORDNANCE) in place of OSM roads and re-run our program to identify urban parcels. As Ordnance Survey data reports the actual roads, thus according to our working definition of parcels, parcels generated with ORDNANCE data should correspond to real-world parcels. In other words, parcels generated based on ORDNANCE data are used to benchmark the validity of OSM-based product.

Additionally, as the errors in OSM-generated parcels may come from (1) errors in the raw OSM data; and (2) errors in our empirical framework, we attempt to single out pitfalls in our empirical framework and have cross-referenced ORDNANCE-based parcels with remote-sensing based parcels to demonstrate the capability of our empirical framework (Appendix 1).

Results

Parcel characteristics

We ran the proposed constrained CA model in all 654 cities. Our method generates exceedingly large parcels (i.e., individual parcels that would exceed the total urban area constraints) for cities with limited OSM data. We adopt a pragmatic threshold of ten parcels and deem the 297 cities with ten or more urban parcels as “successfully” processed by our algorithm (Figure 5). Due to the city’s sheer size, Chongqing was the only MD-level city absent from this group of successfully processed cities. All SPC cities, as well as half of the medium-to-small cities at the PLC and CLC levels have result in more than ten urban parcels.

A total of 232,145 parcels are identified for these 297 cities (Figure 5), and 82,645 out of all generated parcels are labeled as “urban”(total urban area 25,905 km^2). The average number of urban parcels for MD, SPC, OPCC, PLC and CLC cities are 1411, 407, 199, 79 and 26, respectively. As discussed previously, cities with more population and higher administrative ranks (e.g., Beijing, the national capital; Nanjing, a provincial capital; and Qingdao, a sub-provincial level city) tend to have more detailed OSM road network and subsequently greater number of parcels.

For all urban parcels, we calculate (1) land use density; (2) urban function; and (3) land use mix degree. Figure 6 illustrates the results for five representative cities. Density among parcels within a city or across cities could be compared in terms of inferred and standardized density attributes. Urban function and land use mix measurements point to substantive mixing of land use. More specifically, 55,728 (67.3%) out of the 82,645 urban parcels have “dominant” urban functions (Figure

[7] The urban area of the city of Beijing was 1445.0 km^2 in 2012, which was still less than the aggregated results (1677.5 km^2) using urban parcels in BICP. The data inconsistence between each other in China is not rare.

6), including 16,018 residential parcels, 16,381 commercial parcels, 18,351 firm parcels, and 10,018 government parcels. Moreover, the average land mix degree for all urban parcels in 297 cities is approximately 0.66 (with a maximum of 1).

Parcel validation

For validation at the parcel level, Table 3 summarizes the comparison of parcels generated by our approach and those contained in the BICP Beijing parcel data. Table 3 suggests that OSM-based approach generally produce larger parcels, due to the lack of information about tertiary and more detailed roads in the OSM dataset[8]. Nevertheless, the match degree (the total area of intersected urban parcels in both data divided by the total area of OSM-based urban parcels) is 71.2%, suggesting that both datasets largely capture the same geographic distribution of urban parcels and land use activities. In addition, we decompose the city of Beijing into sub-regions bounded by major ring roads, and calculate the proportion of parcels falling into individual sub-regions. The proportion of parcels falling into sub-regions between ring roads is consistent across both datasets. We also compare the size distribution of parcels, both of which are showing lognormal distribution with similar mean values.

Furthermore, density and urban functions of OSM-based urban parcels in Beijing are compared with other data sources. With the same OSM-generated parcel boundaries, we calculate development density for individual parcels (a total of 7,130 parcels[9]) based on (1) building information such as floor area gathered from BICP for 2008; and (2) POI data, as building information is the common data for inferring development density. The Pearson correlation coefficient between development densities calculated in two different ways is 0.858, suggesting that ubiquitously available POI data could be used as a proxy for urban density. As POI types and land use types in BICP data were not totally aligned with each other, we limited our comparison to OSM parcels with a dominating residence function and residential parcels in BICP. We overlaid residential parcels in OSM and BICP, and the overlapping area is 211.5 km^2 (56.3%out of total 375.6km^2 OSM-based residential parcels). In other words, the parcel level validation suggests that, despite only using online open data, our OSM-based approach often produce reasonably good approximations of data produced by and the conventional manual method.

Validation at the aggregated regional level is performed by comparing urban parcels generated by OSM and ORDNANCE in 297 cities where both datasets had included urban parcels (Table 4). Although the Ordnance Survey data were representing road networks at the end of 2011, in order to ensure comparability in terms of total parcel size, we use the 2012 total urban area in the parcel identification CA model. Table 4 suggests that OSM-based approach tends to generate parcels of larger size, again due to the relative sparseness of road networks in OSM data. The match degree between urban land by OSM and ORDNANCE was 58.1%, calculated as the ratio of the area of overlapping urban parcels to the area of all OSM-based urban parcels. When we disaggregated the overlapping results in each city level, the ratio for MC, SPC and OPCC was around 70% and the ratio for FLC and CLC was around 45%. This, following the comparison on road networks in both datasets in Figure 2, further proved the data completeness of OSM in big cities was much better than that in small cities in China.

---

[8] Parcels by ORDNANCE in Beijing were similar with those by planners in BICP in terms of parcel size.

[9] Density for BICP parcels is calculated based on floor space rather than POI. Floor space information was limited to the parcels within the six ring road of Beijing, which was used in comparison.

Conclusions

Aiming at the paucity of parcel data in cities of the developing world, our study proposes a novel and scalable empirical framework for the automatic identification and characterization of parcels using ubiquitously available OSM and POIs data. Our analysis represents a first attempt to use volunteered GIS data to identify and characterize urban parcels in China. Empirical results also suggest that OSM and POI could help to produce reasonably good approximation of parcels identified from conventional methods, thus making our approach a useful supplement

More specifically, the contribution of this paper lies in the following aspects: Firstly, we proposed a robust and straightforward approach to delineating parcels, identifying urban parcels, and characterizing parcel features, using ubiquitously available OSM and POI data. Secondly, we employ a novel approach that incorporates a vector-based cellular automata model into the identification of urban parcels. Thirdly, our approach has been applied to hundreds of cities in China, and could possibly be extended to generate parcel data for other developing countries. Our project is also part of the Open Data Initiative, as all our data products are freely available from the Internet.
The final product of our project is a dataset containing fine scale urban parcels with detailed features for 297 Chinese cities. This dataset can be applied to but not limited to the following aspects: Firstly, the dataset can be updated periodically and provides parcel maps for urban planning and studies in places where digital infrastructure development is weak. For example, official parcel data for Beijing are generally updated every three years and our approach would enable updating on a yearly basis to capture rapid growth in Chinese cities. Secondly, the dataset can serve as the base for emerging vector-based urban modeling, e.g., vector-based cellular automata models and agent based models (Stevens and Dragicevic 2007; Jjumba and Dragicevic, 2012). Urban parcels generated by our approach would enable us to establish large-scale urban expansion models for large geographical area (e.g., an entire nation) at parcel level. Such urban expansion models would open up new avenue for fine-scale regional growth management but were technically not possible without parcel data covering the same geographic extent. Our attempt to establish such parcel-level national-scale urban expansion model will be reported in a related paper. Thirdly, parcel attributes such as urban functions and land use intensity provide useful measurements for urban analysts to examine *inter alia* quality of life, urban growth, and land use changes (Frank et al. 2010). Though our dataset has been released for a very brief period of time, many planning and urban studies projected have reported to explore and utilize our parcel data. In fact, our parcel dataset has been downloaded more than 1500 times and received over 100 comments in its first week of release. Previously planning professionals in China have less chance to access land use data at such fine spatial scale. Fourthly, the generated parcels could be used as spatial units for incorporating other ubiquitous and spatially referenced data, e.g. check-ins, photos, and mini-blogs, as well as human mobility data like transportation smart card records, taxi trajectories and mobile phone traces. The estimation of inferred urban function, density and land use mix would be improved by integration different data sources.

Because the general limitations of using open data to study urban dynamics have been detailed elsewhere (Liu et al. 2013; Sun et al. 2013), we will conclude by noting limitations and possible future research avenues that are specific to our AICP framework. A first limitation of our approach is that OSM road networks are relatively sparse in many cities and lead to unrealistic large urban parcels. This deficiency is likely to be alleviated by the ever-increasing coverage and quality of OSM data in China (Figure 1). Techniques for parcel subdivision would be an alternative solution for generating more realistic urban parcels in small cities in China (Aliaga et al.2008). A second limitation is related to the use of POIs for estimating land use density. Our current approach focuses on the quantity rather than quality of individual POIs (e.g., a large department store and a small convenience store are treated equally). Possible improvements include the incorporation of online check-in data (e.g., Foursquares, and SinaWeibo – a Chinese equivalent of Twitter), taxi trajectories, and transportation smart card records to supplement inferring land use intensity. Lastly, more constraints like accessibilities to main roads and city centers, as well as exclusive development zones are expected to introduce into the constrained CA model used for identifying urban parcels to increase the overall identification precision.

*** ***

Parcel data generated from OSM and POIs are freely available from the Beijing City Lab (Data15, http://longy.jimdo.com/data-released/) and visualized online at (https://a.tiles.mapbox.com/v3/jianghaowang.gcnng3cg/page.html?secure=1#5/36.014/105.996).

It is worth noting that these online visualizations serve as crowd-source validations for our methods (Fritz et al., 2012). Geometric and thematic errors of individual parcels are identified by data users with local knowledge, and used to fine-tune our CA model.

Acknowledgments: The first author acknowledges the financial support of the National Natural Science Foundation of China (No.51078213) and China Scholarship Council (No.201209110066). The authors thank Kang Wu (Capital University of Economics and Business, China), Dong Li (China Academy of Urban Planning & Design), Mingrui Mao (Beijing Institute of City Planning), Jianghao Wang (Chinese Academy of Sciences), and Qizhi Mao (Tsinghua University) for their generous help and comments in data collection, pre-processing, and methodology development. The usual disclaimer applies.

References

Alberti M, Booth D, Hill K, Coburn B, Avolio C, Coe S, Spirandelli D,2007, "The impact of urban patterns on aquatic ecosystems: an empirical analysis in Puget lowland sub-basins" Landscape and urban planning **80(4**) 345-361

Aliaga D G, Vanegas C A, Beneš B 2008, "Interactive example-based urban layout synthesis", In *ACM Transactions on Graphics (TOG)* ACM **27 (5)** 160

Barnsley M J, Barr S L, 1996, "Inferring urban land use from satellite sensor images using kernel-based spatial reclassification" *Photogrammetric Engineering and Remote Sensing* **62(8)** 949-958

Beijing Institute of City Planning, 2010 *Existing land use map of Beijing*. Internal working report

Beresford A R, Stajano F, 2003, "Location privacy in pervasive computing" *Pervasive Computing IEEE* **2(1)** 46-55

Bontemps S, Defourny P, Van Bogaert E, Arino O, Kalogirou V, Perez J R,2009 *GLOBCOVER: Products Description and Validation Report*

Cheng J, Turkstra J, Peng M, Du N, Ho P, 2006, "Urban land administration and planning in China: Opportunities and constraints of spatial data models" *Land Use Policy* **23(4)** 604-616

Erickson A, Rogers L, Hurvit P, Harris J, 2013, "Challenges and Solutions for a Regional Land Use Change Analysis" *Proceedings of ESRI*, Available online at http://proceedings.esri.com/library/userconf/proc06/papers/papers/pap_1472.pdf (accessed on 26 November 2013)

Feng Y, Liu Y, Tong X, Liu M, Deng S, 2011, "Modeling dynamic urban growth using cellular automata and particle swarm optimization rules" *Landscape and Urban Planning* **102(3)** 188-196

Frank L D, Andresen M A, Schmid T L 2004, " Obesity relationships with community design, physical activity, and time spent in cars" *American Journal of Preventive Medicine* **27(2)** 87-96

Frank L D, Sallis J F, Conway T L, Chapman J E, SaelensB E, Bachman W 2006, "Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality" *Journal of the American Planning Association* **72(1)** 75-87

Frank L D, Sallis J F, Saelens B E, Leary L, Cain K, Conway T L, Hess P M , 2010, "The development of a walkability index: application to the Neighborhood Quality of Life Study" *British journal of sports medicine* **44(13)** 924-933

Fritz S, McCallum I, Schill C, Perger C, See L, Schepaschenko D, van der Verde M, Kraxner F, Obersteiner M, 2012, "Geo-Wiki: An online platform for improving global land cover" *Environmental Modelling & Software* **31** 110-123

Girres J F, Touya G, 2010, "Quality assessment of the French OpenStreetMap dataset" *Transactions in GIS* **14(4)** 435-459

Goodchild M F, 2007," Citizens as sensors: the world of volunteered geography" *GeoJournal* **69(4)** 211-221

Haklay M, 2010, "How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets" *Environment and Planning B: Planning & design* **37(4)** 682

Haklay M, Weber P, 2008, "OpenStreetMap: User-generated street maps" *Pervasive Computing, IEEE* **7(4)** 12-18

Herold M, Scepan J, Clarke K C,2002, "The use of remote sensing and landscape metrics to describe structures and changes in urban land uses" *Environment and Planning A* **34(8)** 1443-1458

Jabareen Y R, 2006, "Sustainable urban forms their typologies, models, and concepts" *Journal of Planning Education and Research* **26(1)** 38-52

Jiang B, Liu X, 2012," Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information" *International Journal of Geographical Information Science* **26(2)** 215-229

Jiang B, Liu X, Jia T,2013 "Scaling of geographic space as a universal rule for map generalization" *Annals of the Association of American Geographers* (ahead-of-print)

Jjumba A, Dragićević S, 2012, ”High resolution urban land-use change modeling: Agent iCity approach” *Applied Spatial Analysis and Policy* **5(4)** 291-315
Jokar Arsanjani J, Helbich M, Bakillah M, Loos L, 2013, ”The emergence and evolution of OpenStreetMap: A cellular automata approach” *International Journal of Digital Earth* 1-30 (accepted)
Jokar Arsanjani J, Helbich M, Bakillah M, Hagenauer J, Zipf A, 2013, “Toward mapping land-use patterns from volunteered geographic information” *International Journal of Geographical Information Science* 1-15 (ahead-of-print)
Kressler F P, Bauer T B, Steinnocher K T, 2001, ”Object-oriented per-parcel land use classification of very high resolution images”, In *Remote Sensing and Data Fusion over Urban Areas*, IEEE/ISPRS Joint Workshop 2001 pp. 164-167 (IEEE)
Li X, Yeh A G O, 2002, “Neural-network-based cellular automata for simulating multiple land use changes using GIS” *International Journal of Geographical Information Science* **16(4)** 323-343
Liu Y, Sui Z, Kang C, Gao Y ,2013,” Uncovering patterns of inter-urban trips and spatial interactions from check-in data” arXiv preprint arXiv:1310.0282
Liu Y, Wang F, Xiao Y, Gao S ,2012,“Urban land uses and traffic ‘source-sink areas’: Evidence from GPS-enabled taxi data in Shanghai” *Landscape and Urban Planning* **106(1)** 73-87
Long Y, Han H Y, Yu X, 2013, “Discovering functional zones using bus smart card data and points of interest in Beijing“ Beijing City Lab, Working Paper # 11, http://longy.jimdo.com/working-papers/, accessed on 26 November 2013
Long Y, Liu X, 2014, “How mixed is Beijing, China? A visual exploration of mixed land use“ *Environment and Planning A* (Forthcoming)
Ma L J, 2005,”Urban administrative restructuring, changing scale relations and local economic development in China” *Political Geography* **24(4)** 477-497
Ministry of Housing and Urban-rural Development of the People’s Republic of China (MOHURD), 2013 *Chinese City Construction Statistics Yearbook 2012* (Beijing: China Planning Press)
Over M, Schilling A, Neubauer S, Zipf A, 2010,” Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany” *Computers, Environment and Urban Systems* **34(6)** 496-507
Pacifici F, Chini M, Emery W J, 2009,”A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification“ *Remote Sensing of Environment* **113(6)** 1276-1292
Pinto N N, 2012, “A cellular automata model based on irregular cells: application to small urban areas” *Environment and Planning B: Planning & Design* **37(6)** 1095-1114
Ramm F, Topf J, Chilton S, 2010 *OpenStreetMap: using and enhancing the free map of the world* (Cambridge: UIT Cambridge)
Soto V, Frías-Martínez E, 2011 “Automated land use identification using cell-phone records”, In *Proceedings of the 3rd ACM international workshop on MobiArch* ACM pp. 17-22
Stevens D, Dragicevic S, 2007, “A GIS-based irregular cellular automata model of land-use change” *Environment and Planning B: Planning & Design* **34(4)** 708-724
Sui D Z, 2008, “The wikification of GIS and its consequences: Or Angelina Jolie’s new tattoo and the future of GIS” Computers, *Environment and Urban Systems* **32(1)** 1-5
Sun L, Axhausen K W, Lee D H, Huang X, 2013, “Understanding metropolitan patterns of daily encounters” *Proceedings of the National Academy of Sciences* **110(34)** 13774-13779
Toole J L, Ulm M, González M C, Bauer D, 2012, ”Inferring land use from mobile phone activity”, In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing* ACM pp1-8
Wu F, 2002, “Calibration of stochastic cellular automata: the application to rural-urban land conversions” *International Journal of Geographical Information Science* **16(8)** 795-818
Yang Y, He C, Zhang Q, Han L, Du S, 2013,”Timely and accurate national-scale mapping of urban land in China using Defense Meteorological Satellite Program’s Operational Linescan System nighttime stable light data” *Journal of Applied Remote Sensing* **7(1)** 073535

Yuan J, Zheng Y, Xie X, 2012, "Discovering regions of different functions in a city using human mobility and POIs", In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* ACM pp 186-194

Zhang Y P, Long Y, 2013, "Urban Growth Simulation Using V-BUDEM: A Vector-Based Beijing Urban Development Model" The conference of Spatial Planning and Sustainable Development, Beijing

Table 1 POIs types and aggregated information

Type	Abbreviation	Count
Commercial sites	COM	2,573,862
Business establishment	FIR	677,056
Transport facilities	TRA	561,236
Others	OTH	534,357
Government	GOV	468,794
Education	EDU	285,438
Residence communities	RES	167,598
Green space	GRE	13,041

Table 2 Logistic regression results for the BICP parcels

Name	Coefficient	S.E,	Sig.
a_0	1.562	0.033	0.000
a_1	-0.234	0.003	0.000
a_2	34.192	0.347	0.000
a_3	0.005	0	0.000

Table 3 Comparison of selected urban parcels in BICP and OSM in Beijing (R=ring road)

Parcels	**Parcel count**	**Average size (ha)**	**Overlapped with BICP**	**Spatial distribution (in terms of area, km^2)**					
				Within R2	**R2-R3**	**R3-R4**	**R4-R5**	**R5-R6**	**Beyond R6**
OSM	7,130	17.2	1194.2 km2 (71.2%)	42.5	74.0	113.4	263.5	666.5	519.9
BICP	57,818	2.9	-	48.6	69.7	99.8	229.5	687.9	544.4
OSM/BICP	0.12	5.93	-	0.87	1.06	1.14	1.15	0.97	0.95

Table 4 The comparison of urban parcels in OSM and ORDNANCE for 297 cities

Data	**Urban area (km^2)**	**Parcel count**	**Average parcel/patch size (ha)**	**Intersected with ORDNANCE (km^2)**
OSM	25,905	82,645	31.3	15,053
ORDNANCE	25,670	260,098	10.0	-

Figure 1 Increasing data volume in OSM-China (accessed on Oct 5, 2013; Unit: data points).

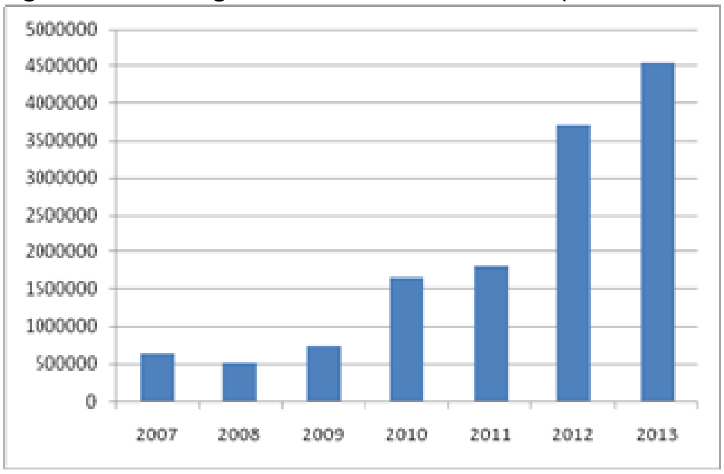


Figure 2 Administrative boundaries of Chinese cities

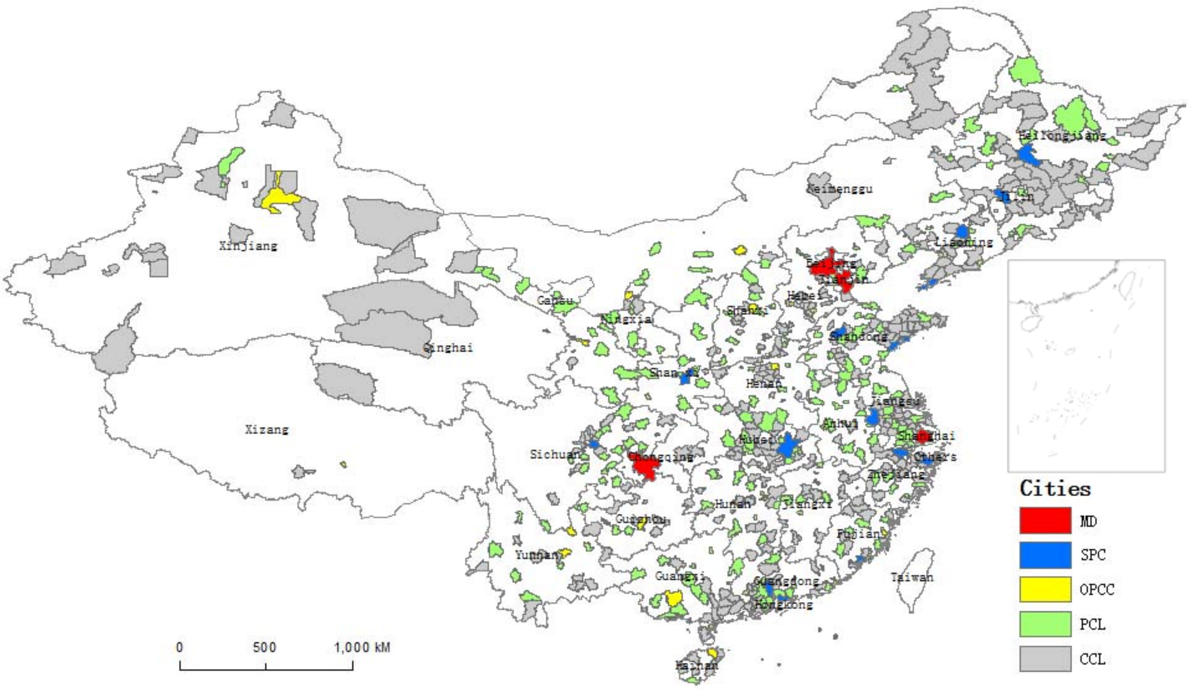
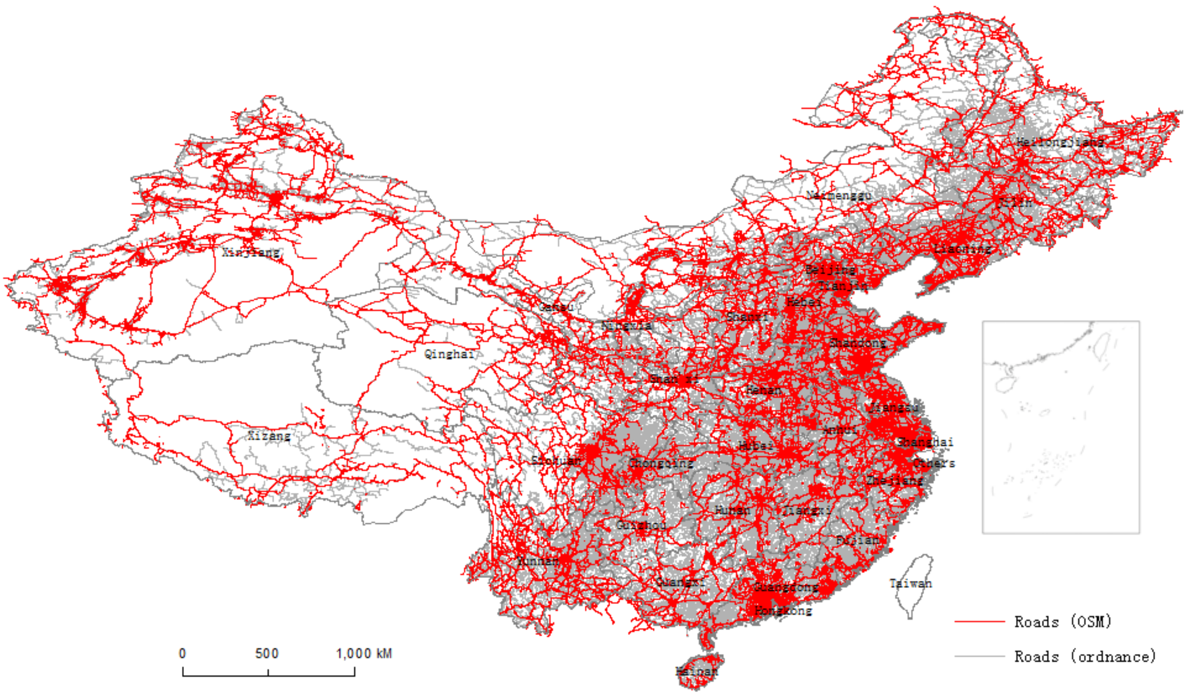


Figure 3 The comparison of roads in the OSM and ordnance map[10]



[10] Roads in the ordnance map were partially covered by roads in OSM. According to our careful check, almost all roads in OSM were meanwhile in roads in the ordnance map.

Figure 4 Examples of identifying urban parcels using constrained CA. The black border indicates a parcel's neighborhood, and the number in brackets reflects the parcel's intrinsic probability.

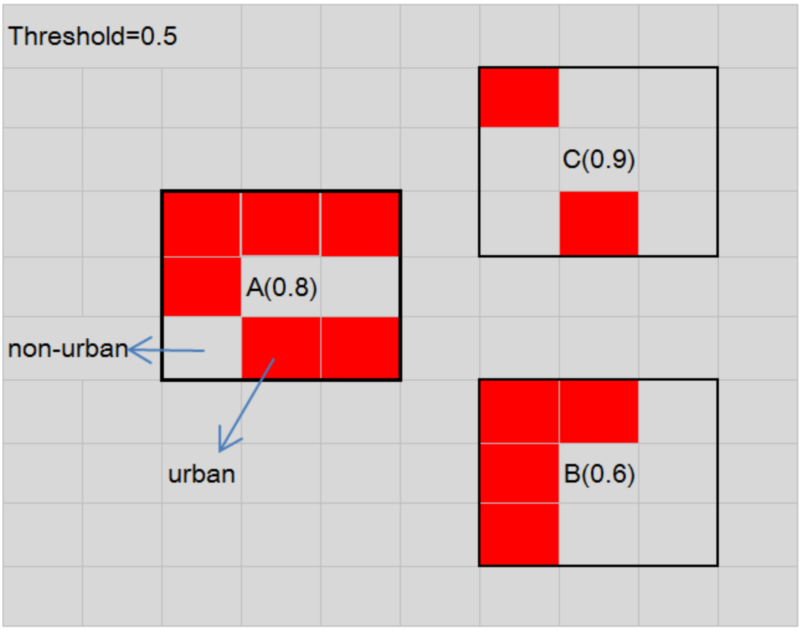
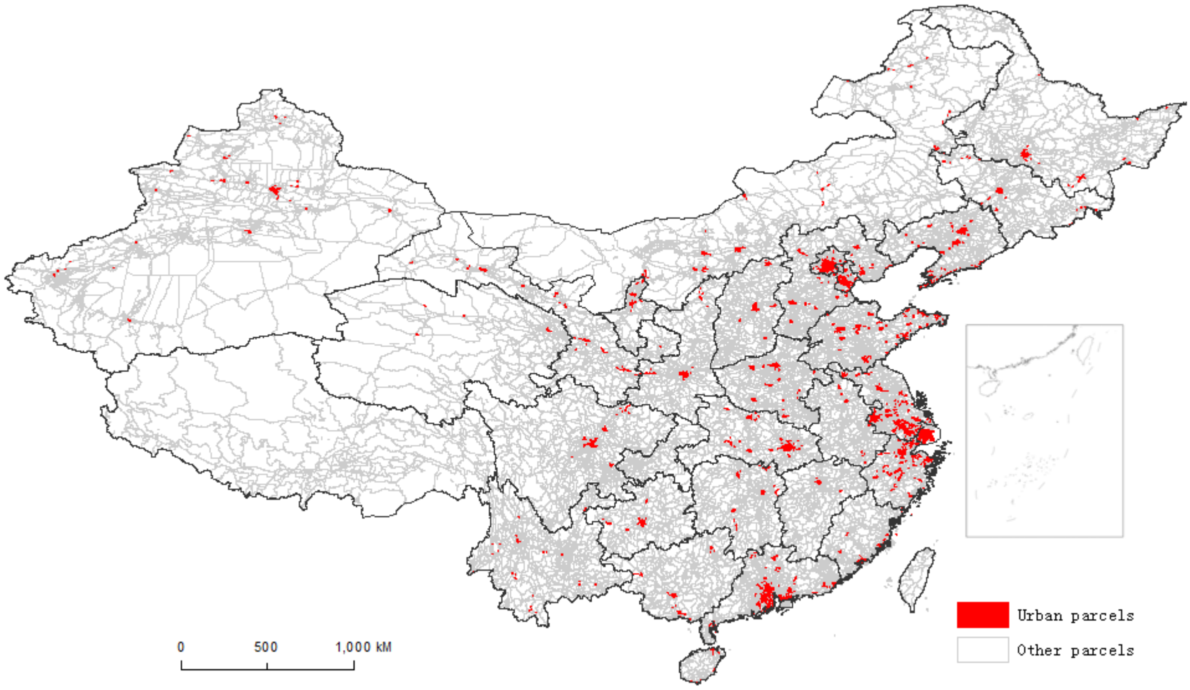
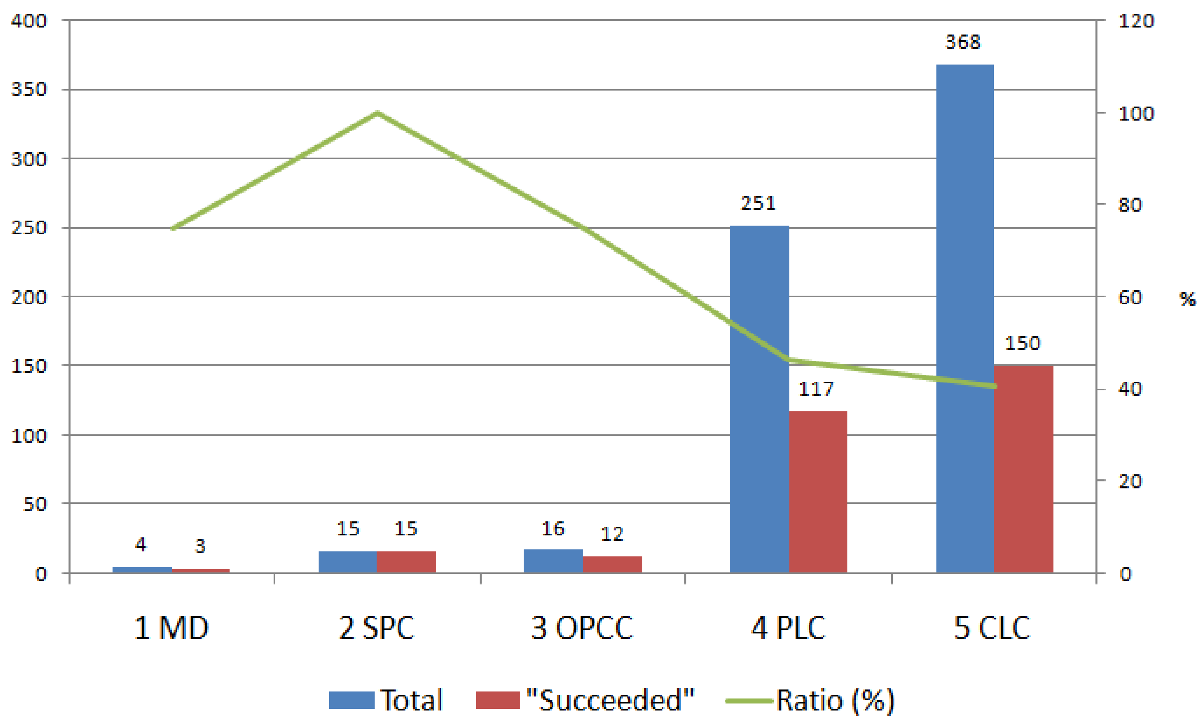


Figure 5 All generated parcels and urban parcels in China (a, spatial distribution; b, the profile of "successfully processed" cities)

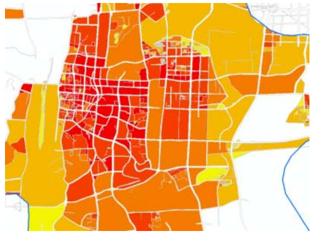
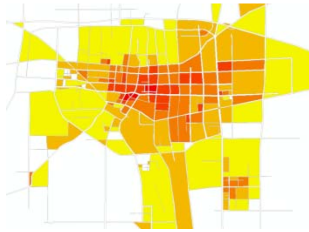
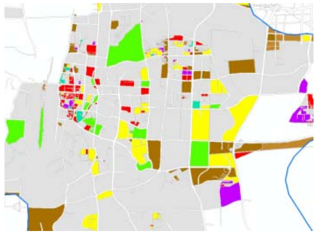
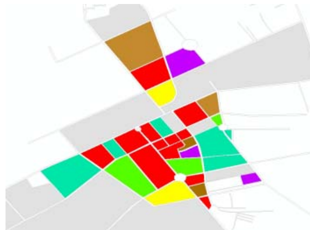
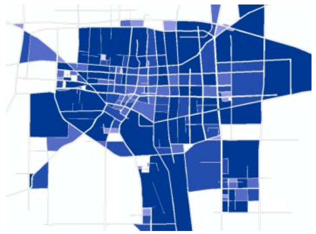


(a)



(b)

Figure 6 The generated parcels and their attributes in typical cities of China

	Beijing	Nanjing	Changsha	Weifang	Gongzhuling
City level	1 MD	2 SPC	3 OPCC	4 PLC	5 CLC
Density Low High					
Function COM EDU FIR GOV GRE RES TRA Not classified					
Land use mix Low High					

Appendix 1: Comparing urban parcels by the ordnance survey map and other data sources

We derived parcel boundaries and selected urban parcels from road networks in the 2011 Ordnance Survey (termed "ORDNANCE") in the whole China. Urban parcels in ORDNANCE were generated and selected using the same parcel generation and selection methods like “OSM”. Among 1,184,524 parcels generated, we successfully selected 350,102 urban parcels in 627 cities in ORDNANCE (also for the year of 2011).

Since we did not have fine-scale land use in the whole China, we compared the urban parcels in ORDNANCE (limited to the 627 cities,) with the 300m-resolution urban area of China in GLOBCOVER (Bontemps, 2009), and 1 km-resolution urban area of China from DMSP/OLS in 2008 (Yang et al. 2013), in terms of urban land distribution. Considering the data completeness of the ordnance survey, we expected this would validate our methods for parcel generation and urban parcel selection, while eliminating the data quality influence of OSM. As indicated by the comparison results in Table a-1, an average urban parcel was around 300 m * 400 m, which was much smaller than a patch in GLOBCOVER or DMSP/OLS thus providing more details in the urban spatial distribution map. There were 21,553 km^2 urban lands in ORDNANCE (54.2%) intersected with those of DMSP/OLS. If assuming road spaces not accounted in ORDNANCE were all covered by DMSP/OLS, the intersected ratio would increase to 60.0%. In addition, the time gap between ORDNANCE and DMSP/OLS, in some degree, underestimated the intersected ratio, which might also be hampered by the inconsistent of spatial resolution of two datasets. In this regard, our methods for generating parcels and selecting urban parcels produce reasonably good results for Chinese cities according to the aforementioned evaluation.

Table a-1 The comparison of urban parcels/patches in various data for 627 cities

Data	**Year**	**Spatial resolution**	**Urban area (km^2)**	**Parcel/patch count**	**Average parcel/patch size (ha)**	**Intersected with ORDNANCE (km^2)**
ORDNANCE	2011	-	39746	350102	13.0	-
DMSP/OLS	2008	300 m	44720	1293	3458.6	21553
GLOBCOVER	2009	1 km	39389	12515	314.7	15206

We also found the comparison results between ORDNANCE and GLOBCOVER were not as promising as between ORDNANCE and DMSP/OLS, which might be raised from the inconsistence between DMSP/OLS and GLOBCOVER. The intersected area of the two datasets was 19,501 km^2, 49.5% of urban area in GLOBCOVER and 43.6% of urban area in DMSP/OLS. Note that when we directly compared urban parcels in OSM with DMSP/OLS, the overlap ratio was 61.4% (underestimated by road space and the time lag as discussed in the previous paragraph), slightly larger than that between ORDNANCE and DMSP/OLS.

Appendix: The city list with selected urban parcels

(Order by the urban parcel count)

City name	City level	Parcel count	Urban parcel count	City area (km^2)	Urban land area (km^2)
上海 Shanghai	ZXS	13253	13148	4931	2391
北京 Beijing	ZXS	13184	7127	11713	1229
深圳 Shenzhen	FSJ	4195	3686	1837	734
天津 Tianjin	ZXS	5732	2641	6816	614
武汉 Wuhan	FSJ	5432	2200	8284	687
大连 Dalian	FSJ	2974	1806	2344	325
南京 Nanjing	FSJ	3283	1673	4595	551
沈阳 Shenyang	FSJ	2922	1613	3329	388
广州 Guangzhou	FSJ	3091	1584	3198	558
杭州 Hangzhou	FSJ	2483	1516	3182	340
佛山	DJS	5124	1356	3677	121
珠海	DJS	1541	1347	1207	319
西安	FSJ	2323	1312	3534	271
长春	FSJ	2043	1276	3562	355
青岛	FSJ	1814	1254	2934	312
成都	FSJ	1895	1253	2164	431
长沙	SH	1143	1137	320	228
南昌	SH	1787	1129	481	182
厦门	FSJ	1996	1004	1391	229
江门	DJS	1951	988	1756	137
乌鲁木齐	SH	1878	974	16745	320
宁波	FSJ	1424	929	2268	288
郑州	SH	1482	893	1004	285
无锡	DJS	1759	859	1556	237
临沂	DJS	1948	757	1654	162
苏州	DJS	1411	750	4530	375
石家庄	SH	1027	726	301	184
中山	DJS	1817	713	1540	86
济南	FSJ	1083	654	3009	318
兰州	SH	631	614	1707	165
太原	SH	881	552	1409	239
呼和浩特	SH	1289	527	2022	182
哈尔滨	FSJ	931	516	7050	322

City name	City level	Parcel count	Urban parcel count	City area (km^2)	Urban land area (km^2)
威海	DJS	958	501	369	122
包头	DJS	1070	490	2225	163
烟台	DJS	815	479	2582	233
汕头	DJS	504	479	1947	187
海口	SH	767	358	2162	102
嘉兴	DJS	572	316	962	89
常州	DJS	762	302	1767	152
洛阳	DJS	305	293	442	154
温州	DJS	604	284	828	132
惠州	DJS	650	281	2587	185
唐山	DJS	673	279	3257	164
蚌埠	DJS	392	256	378	103
大理	XJS	589	250	1490	34
鞍山	DJS	356	246	608	131
荆州	DJS	450	238	1362	59
宜昌	DJS	255	238	4146	107
秦皇岛	DJS	382	237	378	84
酒泉	DJS	520	231	3637	33
义乌	XJS	474	229	1057	80
日照	DJS	383	224	1799	82
扬州	DJS	469	222	2244	108
贵阳	SH	470	218	2374	186
丹东	DJS	315	209	517	48
淮安	DJS	457	208	3172	188
西宁	DJS	315	205	319	65
鄂尔多斯	DJS	352	204	2175	135
上虞	XJS	577	202	1142	22
丽江	DJS	308	195	1275	17
三亚	DJS	455	193	1841	29
绵阳	DJS	329	192	1572	96
清远	DJS	201	190	578	51
儋州	XJS	235	189	3267	25
南宁	SH	326	181	6490	202
潍坊	DJS	267	181	1527	130
银川	SH	446	178	1568	116
镇江	DJS	409	178	1048	103
江阴	XJS	588	177	937	43
嘉峪关	DJS	335	177	1358	53
衢州	DJS	316	175	2357	55
鹤山	XJS	355	173	1049	22
宝鸡	DJS	185	173	3654	81
南通	DJS	341	171	1869	172
常熟	XJS	350	163	1194	73
湖州	DJS	300	163	1520	80
淄博	DJS	356	161	2935	200
孝感	DJS	306	160	733	26

City name	City level	Parcel count	Urban parcel count	City area (km^2)	Urban land area (km^2)
台山	XJS	475	158	3074	25
瑞丽	XJS	638	157	1029	9
东营	DJS	288	156	2694	95
克拉玛依	DJS	269	153	8924	49
岳阳	DJS	215	150	1020	76
马鞍山	DJS	149	145	288	84
湛江	DJS	160	141	1355	81
文登	XJS	346	140	1840	38
襄樊	DJS	270	140	3603	94
营口	DJS	190	136	501	94
徐州	DJS	227	135	3000	190
喀什	XJS	212	131	94	46
黄山	DJS	238	130	2260	34
泰安	DJS	236	130	2027	101
南阳	DJS	195	128	1943	104
德州	DJS	135	127	254	82
信阳	DJS	360	126	3639	67
二连浩特	XJS	130	124	325	37
龙口	XJS	227	122	868	38
密山	XJS	628	119	7924	12
保山	DJS	214	115	5180	19
阜新	DJS	162	115	418	62
武威	DJS	220	113	6181	31
黑河	DJS	174	112	14553	21
桂林	DJS	145	110	562	56
库尔勒	XJS	315	107	8561	58
广元	DJS	237	107	5014	42
大丰	XJS	186	102	2300	16
乌兰浩特	XJS	234	101	752	41
鄂州	DJS	116	101	1587	53
增城	XJS	231	100	1703	8
开平	XJS	227	99	1629	32
廊坊	DJS	176	98	965	61
牡丹江	DJS	105	98	1363	69
东台	XJS	203	97	2264	29
格尔木	XJS	161	97	137357	27
大同	DJS	245	96	2041	98
盐城	DJS	203	96	1697	81
东港	XJS	223	95	2247	26
张掖	DJS	162	94	3842	30
东阳	XJS	275	93	1620	31
满洲里	XJS	239	93	705	23
普宁	XJS	102	91	1618	29
伊宁	XJS	142	87	128	32
石嘴山	DJS	138	87	1500	51
玉门	XJS	289	86	14365	16

City name	City level	Parcel count	Urban parcel count	City area (km^2)	Urban land area (km^2)
辽阳	DJS	133	86	567	95
金昌	DJS	129	86	1230	33
乌海	DJS	129	85	1543	49
蓬莱	XJS	185	83	1185	23
敦煌	XJS	158	83	27759	17
蒙自	XJS	115	83	2222	20
金华	DJS	185	82	2027	64
哈密	XJS	272	79	85978	35
衡水	DJS	166	78	581	36
富阳	XJS	306	76	1789	22
栖霞	XJS	154	76	1999	15
嵊州	XJS	143	76	1762	32
昆山	XJS	295	75	917	48
招远	XJS	113	75	1404	27
丽水	DJS	206	72	1182	30
平湖	XJS	278	71	518	18
扎兰屯	XJS	155	71	16921	16
楚雄	XJS	109	71	4539	30
曲阜	XJS	133	70	807	24
承德	DJS	82	70	648	54
阿勒泰	XJS	99	69	12775	14
北海	DJS	94	68	284	57
昌吉	XJS	264	67	9199	43
诸暨	XJS	203	66	2310	32
景洪	XJS	151	65	7427	22
商丘	DJS	118	65	1626	56
咸阳	DJS	228	64	512	60
兴化	XJS	196	64	2366	29
奉化	XJS	89	64	1217	28
辽源	DJS	71	64	231	39
珲春	XJS	107	62	5495	21
驻马店	DJS	94	62	102	49
晋中	DJS	153	61	1290	47
临沧	DJS	68	61	2768	19
石河子	XJS	99	60	556	27
太仓	XJS	158	59	644	39
高邮	XJS	103	59	1938	21
海宁	XJS	130	58	682	30
莆田	DJS	179	57	1928	46
文山	XJS	72	57	3084	23
海门	XJS	243	56	964	20
永安	XJS	125	56	2931	19
宿州	DJS	113	56	2923	57
乌苏	XJS	84	56	17507	16
凌海	XJS	73	56	2732	25
泰兴	XJS	206	55	1403	20

City name	City level	Parcel count	Urban parcel count	City area (km^2)	Urban land area (km^2)
博乐	XJS	117	54	8993	19
都江堰	XJS	105	54	1225	29
龙海	XJS	156	53	1341	15
荣成	XJS	156	53	1577	38
黄冈	DJS	80	53	411	28
吴忠	DJS	158	52	975	33
永州	DJS	101	52	3183	50
丹江口	XJS	60	52	3129	23
张家口	DJS	108	50	802	82
奎屯	XJS	82	50	1231	26
锡林浩特	XJS	60	50	15730	41
牙克石	XJS	165	49	27987	19
荥阳	XJS	146	49	908	24
舟山	DJS	131	49	877	46
韶关	DJS	127	49	3494	77
雅安	DJS	122	49	1102	18
海阳	XJS	107	49	1814	24
和田	XJS	459	48	56415	29
新郑	XJS	165	48	851	23
安宁	XJS	150	48	1324	19
许昌	DJS	69	48	99	63
平顶山	DJS	56	48	256	60
瓦房店	XJS	229	47	3551	35
即墨	XJS	100	47	1771	48
临安	XJS	84	47	3101	12
黄骅	XJS	222	46	2144	23
永康	XJS	180	46	1050	30
昌邑	XJS	74	46	1700	22
昭通	DJS	69	46	2244	27
百色	DJS	61	46	3821	34
钦州	DJS	116	45	4558	74
文昌	XJS	95	45	2385	19
温岭	XJS	82	45	893	30
姜堰	XJS	135	44	1180	19
普洱	DJS	105	44	4101	22
常德	DJS	92	44	2722	68
当阳	XJS	75	44	2125	14
邳州	XJS	71	44	2046	37
莱西	XJS	71	44	1572	27
铁岭	DJS	65	44	164	40
邵阳	DJS	62	44	314	53
寿光	XJS	186	43	2179	34
天水	DJS	148	43	5903	46
和龙	XJS	49	43	5136	11
宁安	XJS	48	43	7314	9
沧州	DJS	78	42	157	54

City name	City level	Parcel count	Urban parcel count	City area (km^2)	Urban land area (km^2)
德令哈	XJS	169	41	129303	18
榆林	DJS	147	41	6994	39
英德	XJS	135	41	5607	30
启东	XJS	124	41	1154	31
公主岭	XJS	66	41	4192	24
连州	XJS	91	40	2627	14
临海	XJS	90	40	2200	38
图木舒克	XJS	77	40	766	11
莱阳	XJS	72	40	1683	38
项城	XJS	68	40	1063	27
庄河	XJS	191	39	3743	32
龙岩	DJS	112	39	2676	36
迁安	XJS	67	39	1215	33
胶州	XJS	103	38	1279	41
青州	XJS	90	38	1533	40
新沂	XJS	72	38	1574	32
同江	XJS	60	38	6270	17
庆阳	DJS	45	38	979	20
如皋	XJS	43	38	1578	24
吐鲁番	XJS	142	37	17493	15
宜兴	XJS	106	37	2002	65
枣阳	XJS	67	37	3268	35
禹城	XJS	56	37	968	28
大安	XJS	163	36	4923	15
海城	XJS	175	35	2719	28
霸州	XJS	108	35	824	16
万宁	XJS	102	35	1877	12
福鼎	XJS	52	35	1452	20
清镇	XJS	41	35	1540	14
额尔古纳	XJS	116	34	28574	9
余姚	XJS	80	34	1367	41
朝阳	DJS	52	34	122	41
阿图什	XJS	128	33	18550	12
四会	XJS	85	33	1200	26
陆丰	XJS	65	33	1719	18
鹰潭	DJS	45	33	121	23
遵义	DJS	57	32	315	54
六盘水	DJS	51	32	2284	32
图们	XJS	48	32	850	8
广汉	XJS	50	31	574	37
福清	XJS	102	30	1441	26
塔城	XJS	61	30	4541	15
海林	XJS	146	29	9002	15
浏阳	XJS	86	29	4974	33
四平	DJS	62	28	421	47
鹿泉	XJS	193	27	615	16

City name	City level	Parcel count	Urban parcel count	City area (km^2)	Urban land area (km^2)
青铜峡	XJS	95	27	1860	30
诸城	XJS	42	27	2136	35
兰溪	XJS	36	27	1330	28
三河	XJS	95	26	611	17
阳泉	DJS	69	26	652	38
长乐	XJS	45	26	678	21
阜康	XJS	37	26	9811	9
邢台	DJS	30	26	122	59
普兰店	XJS	165	25	2768	30
乳山	XJS	92	25	1571	28
内江	DJS	61	25	1585	41
仙桃	XJS	41	25	2518	36
建德	XJS	90	24	2269	9
河源	DJS	47	24	101	27
遵化	XJS	46	24	1436	26
高碑店	XJS	31	24	663	13
龙井	XJS	29	23	3284	12
灵武	XJS	139	22	3706	11
句容	XJS	100	22	1355	20
新民	XJS	93	22	3307	22
根河	XJS	63	22	19949	11
洮南	XJS	59	22	6042	16
长治	DJS	48	22	355	48
东方	XJS	113	21	2294	19
定西	DJS	76	21	3710	15
辛集	XJS	38	21	936	26
崇州	XJS	32	21	1127	23
临夏	XJS	29	21	117	25
藁城	XJS	87	20	793	16
阿克苏	XJS	70	20	21194	32
桐乡	XJS	41	20	706	34
安顺	DJS	38	20	1682	30
孝义	XJS	32	20	943	19
巩义	XJS	31	20	1011	27