

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 Beijing Institute of City Planning, Beijing 100045, China

2 The Martin Centre for Architectural and Urban Studies, Department of Architecture,
University of Cambridge, Cambridge CB2 1PX, UK
longying1980@gmail.com

Xingjian Liu

3 Department of Geography and Earth Sciences, University of North Carolina at Charlotte,
NC 28223-0001, USA
xliu36@uncc.edu

(The first version, 24 November 2013)

Abstract: Parcels are the basic spatial units for fine-scale urban modeling and urban studies, as well as carrying out spatial plans. It is both time and money consuming to prepare parcel maps, which makes some cities in developing countries do not have well-prepared parcel maps for academic research and planning practices. In this paper, we proposed an automated approach for generating parcels and inferring their density, function (a proxy for land use) and land use mix using ubiquitous available OpenStreetMap (OSM) and Points of Interest (POIs). OSM was used to generate parcels and POIs within or close to each parcel were applied for inferring its density, function and land use mix. Vector cellular automata models were developed for each city to select urban parcels from all generated parcels. The method for automated identification and characterization of parcels was used in China with 82,645 urban parcels generating in 297 cities. The final results have been validated via comparing with urban area derived from DMSP/OLS and parcels prepared by planners in Beijing.

Key words: parcel; OpenStreetMap; POI; cellular automata

Online visualization:

[https://a.tiles.mapbox.com/v3/jianghaowang.gbjn3899/page.html?secure=1 - 5/37.125/107.754](https://a.tiles.mapbox.com/v3/jianghaowang.gbjn3899/page.html?secure=1-5/37.125/107.754)

1 Introduction

Land cover maps interpreted from remote sensing images have been extensively applied in strategic-level urban studies and urban plans. Existing land use and density maps for parcels, as the basis for carrying out master plan, zoning, journey analysis and site selection, are the most important maps for carrying out contemporary urban planning (Cheng et al., 2006). Parcels can help people clearly understand the complex characteristics of a city region. Parcel level maps, however, are not easy to be established by using remote sensing images. Cities in developing countries, e.g. middle and small cities in China, do not have existing parcel maps due to poor digital infrastructure, thus hindering the progress of urban plan compilation and management, as well as fine-scale urban studies. In addition, each parcel in the widely established land use maps is generally associated with a unique land use. Every parcel, however, is mixed in terms of land uses and associated urban functions, which has seldom been taken into account in planning practices and urban studies due to lacking of

necessary information on how mix a parcel is (Long and Liu, 2013). Considering aforementioned backgrounds and the time consuming process for parcel maps, we propose an automated approach for generating parcels and inferring their density, function (a proxy for land use) and land use mix using ubiquitous available OpenStreetMap (OSM) and Points-of-Interest (POIs) in this paper. OSM would be used to generate parcels and POIs within or close to each parcel would be applied for inferring its density, function and land use mix.

Conventionally, the geometrical boundaries of parcels are manually generated based on high-resolution remote sensing images (e.g. LIDAR and QuickBird) or relief maps. The process is quite time consuming and the quality of parcel boundary and their attributes (density, function and mix degree) heavily depends on a producer's experience. In addition, extra efforts have to be put on upgrading the already established map every several years. Less attention was paid to automatically parcel generation with an exception of Yuan et al. (2012), which proposed a raster-based approach for map segmentation and inferred their function using taxi trajectories and POIs. As the first research on parcel generation and function inference, there are several points deserve more attention. First, road space as the boundary for delineating parcels is not reflected in their generated parcel map. Second, the method is not suitable for large-scale geographical space since it is raster based. Third, density and land use mix function are not accounted. These would be reflected in our research using OSM and POIs for parcel generation and attribute inference.

In addition to parcel generation, inferring parcel attributes are also crucial for the further application of parcel data. There are considerable research on identifying land use using high-resolution remote sensing images, like Barnsley and Barr (1996). This is not suitable for a large scale, e.g. the whole country, due to data availability and time and money required. Recently, some research focused on inferring land use using human mobility information, like bus smart card records, mobile phone, and taxi trajectories (Long et al, 2013; Soto and Frías-Martínez, 2011; Toole et al, 2012; Yuan et al, 2012). To the best of our knowledge, the ubiquitous and light-weight (in terms of both data preprocessing and inferring model) POIs and check-ins are absent from existing literature for inferring the parcel attributes, like density, land use, and land use mix.

As one of most successful volunteered GIS, OSM provides street networks and is currently available in most cities, small or big (Goodchild, 2007; Sui, 2008). Increasing data coverage and quality of OSM provides it possible to be more realistic (see Figure 1 for the increasing contributions of OSM in China). This trend is expected to continue in the future (Jokar et al. 2013 proposed a cellular automata based approach for predicting future OSM coverage). Data quality of OSM has attracted several attentions, and the results show that OSM has good data quality comparable with that of survey map in some big cities (Girres and Touya, 2010; Haklay, 2010; Over et al., 2010). OSM could be a promising data source for delineating parcel boundaries (Haklay and Weber, 2008; Ramm, 2010). Transportation networks (streets, roads, highways, railroads) stored in OSM partition urban areas and are possible to be used for delineating parcels¹. There are several studies on using OSM for extracting urban area. E.g., Hagenauer and Helbich (2012) extracted urban areas from OSM and found the quality vary a lot across cities. Jiang and Liu (2012) used OSM to extract city/rural blocks and generate natural cities using blocks. Jiang et al (2013) identified urban area using OSM. According to the review, there is no existing research on using OSM for generating fine-scale parcels.

¹ Buildings in OSM are very few in China and not possible to aggregate them into parcels.

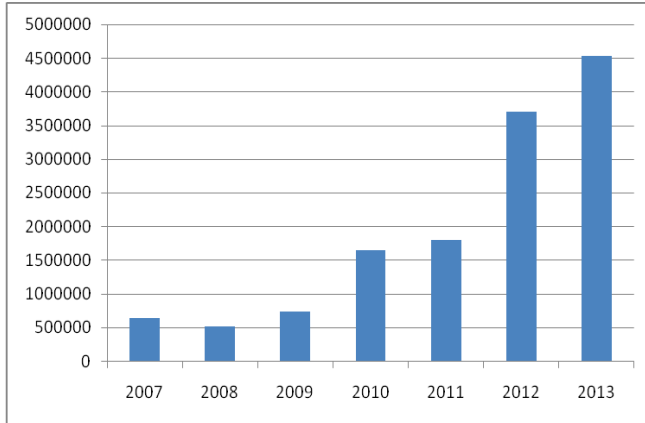


Figure 1 Increasing contributions of OSM in China (summarized from the OSM dataset in China, as of Oct 5, 2013)

In this paper, we would use roads in OSM to partition regional space and generate parcels. In addition to generated boundaries of parcels, we need open data source for the parcels' attributes inference. Since well-classified POIs provided by online map providers and reflecting human activities are available in most cities worldwide, we would use POIs for density inference, urban function and land use mix evaluation. The types of POIs generally include residence, working place, government, green space, clubs, and transportation centers, which could be regarded as a proxy of land use and urban function. This provides it possible for inferring the generated parcels' attributes. The full process is automatic and could be easily adapted to annually update existing parcel map supported with upgraded OSM and POIs. We would compare with generated parcels using OSM with the survey map and the parcel map generated by planners by hand. This paper is structured as follows. Section 2 describes the datasets used in this paper. The methods and their results are introduced in Section 3 in detail. We would compare the generated parcels together with their attributes with existing data source. Section 5 and 6 discuss the results as well as potential applications and make concluding remarks of this research, respectively.

2 Data

2.1 Administrative boundaries of all Chinese cities

We tested our approach in the whole China, the biggest developing countries in the world. Data infrastructure in China varied a lot across various levels of cities. The big cities like Beijing, Shanghai and Guangzhou have well-established urban GIS storing sufficient GIS layers e.g. existing and planned land use maps for conducting urban studies and carrying out city plans. Small cities in remote area of China could not even provide a digital city map to researchers and planners. In addition, data holders in big cities were always reluctant to provide well-established GIS data to researchers and planners in the context of data constraint. In this regard, establishing a parcel-level map for Chinese cities would be promising and contribute to researchers and planners, as well as inter-city studies.

Although parcels delineated in this paper cover the whole China, urban or rural, we focused on urban parcels for which we would assign density, function and land use mix. In this condition, confirming the administrative boundaries of all Chinese cities is necessary. The

whole territories of China could be divided into two types, rural area in counties (*Xian*)², and cities³. Totally 654 cities in China, ranging five levels, namely municipalities directly under the Central Government (MD, 4), sub-provincial cities (SPC, 15), other provincial capital cities (OPCC, 16), prefecture-level cities (PLC, 251), and county-level cities (CLC, 368) (for details on complicated Chinese administrative levels see http://en.wikipedia.org/wiki/Administrative_divisions_of_China), were selected for this study⁴. During the recent twenty years, some counties were upgraded as districts in prefecture-level cities or county-level cities. We generated the city boundaries based on the county boundaries released in 1990s, and the total urban built-up area of each city in 2012 was then linked with the boundary based on the latest Chinese City Construction Yearbook in 2012⁵. The urban land totals would be used for selecting urban parcels from all generated parcels using OSM in the following section.

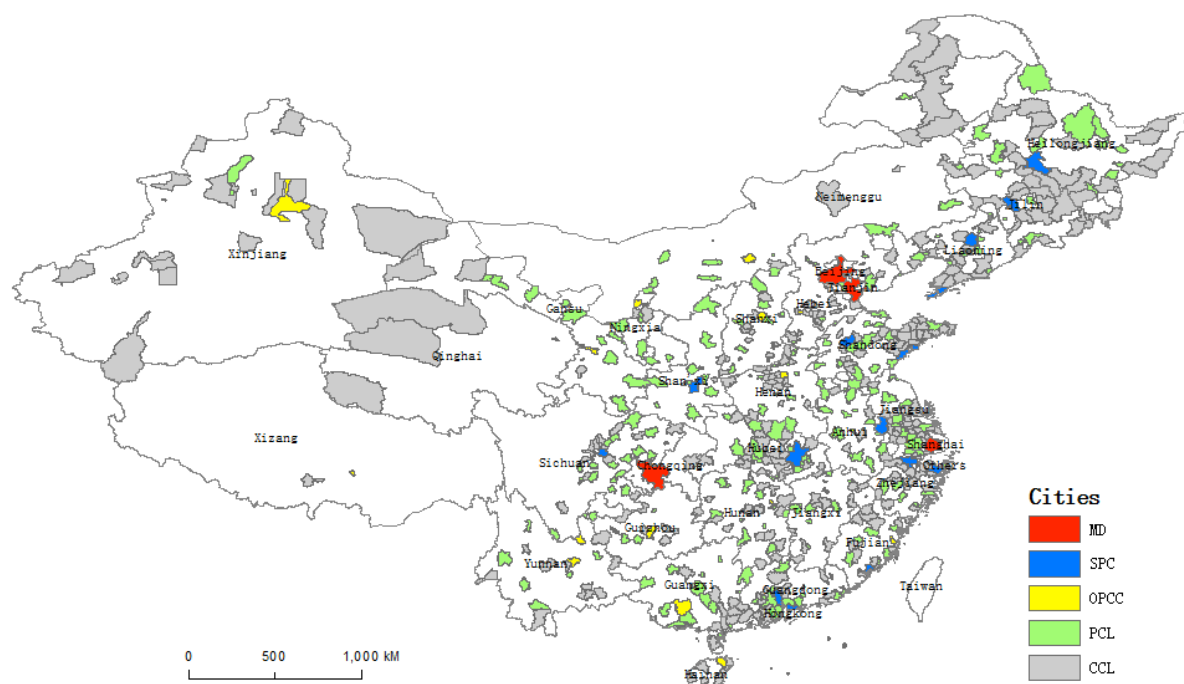


Figure 2 Administrative boundaries of Chinese cities

2.2 Road networks in ordnance survey of China

We obtained the ordnance survey map of China in 2011 with detailed road networks from a local road navigation firm based in Beijing. Both urban streets and regional roads were included in the data. Almost all detailed road networks in various levels were included in this data according to our examination via comparing with Google Maps and Baidu Map (a main online map service provider, also the dominating online search engine in China,

² The central town of a county, generally the administrative units locate, is only regarded as a town, rather than a city in this paper, also in China.

³ A Chinese city covers both urban and rural lands.

⁴ Sansha in Hainan and Beitun in Xinjiang appearing in the Urban Construction Yearbook in 2012 were not included due to spatial data availability. Taiwan was not included in all analysis and results in this paper.

⁵ In this paper, Urban streets were removed in the total urban area for each city as we did not consider road parcels which remained blank in the generated parcel map. The ratio of road lands in urban built-up space was 13% in 2012.

<http://map.baidu.com>). Total road length was 2,623,867 km for 6,026,326 segments⁶ (see Figure 3). This data would be used to examine the data quality of OSM in various levels of Chinese cities.

2.3 OSM in China

The layer “roads” in OSM data in China downloaded from (<http://download.bbbike.org/osm/bbbike/Beijing/>) on Oct 5, 2013, was used for our study in this paper. There are 481,647 road segments (8.0% of that of the ordnance map) with a total length of 825,382 km (31.5% of that of the ordnance map). This showed that the completeness of OSM roads was around 30% and OSM focused on the main roads, rather than small ones. For evaluating the data quality of OSM in China, we overlapped roads in OSM and the ordnance map in terms of total road length in Figure 3. The results indicated that OSM in big cities was significantly better than that in middle and small cities. This reminded us that in the current stage, our parcels generated would have more realistic and potential to be used in planning practices in big cities although our approach applied to all cities when data quality of OSM becomes better.

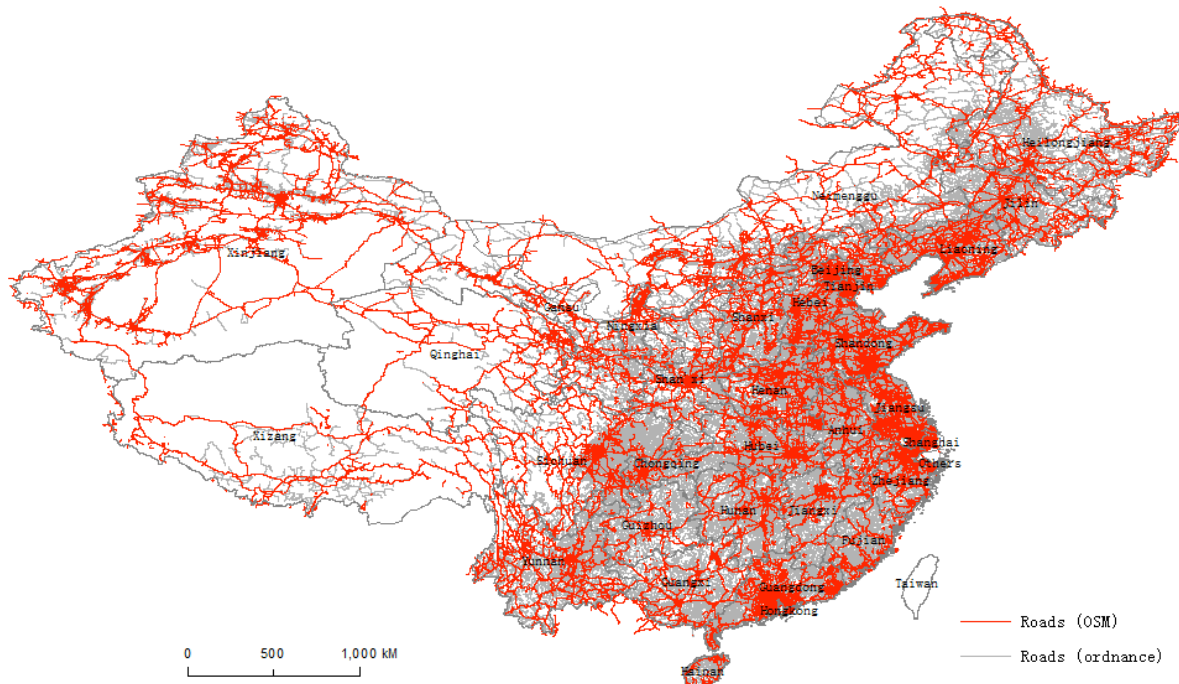


Figure 3 The comparison of roads in the OSM and ordnance map⁷

2.4 POIs

POIs used for inferring urban function for each established parcel were also obtained from the local navigation firm who provided us with the ordnance map used for comparing with OSM. Among 5,281,382 POIs, there were originally 20 types. For better projecting urban functions, we merged them into eight types as shown in Table 1. Commercial sites were

⁶ We would use the term “road” for all terms of transportation network elements, like streets and highways.

⁷ Roads in the ordnance map were partially hidden by roads in OSM. According to our careful check, almost all roads in OSM were meanwhile in roads in the ordnance map.

dominating among all merged types, followed by firms, transport facilities and government buildings. Note the "OTH" type POIs used in density estimation were not accounted in inferring urban function and calculating land use mix. All POIs concentrated in urban area according to our eye-ball check with Google Maps and Baidu Map. It was not easy to plot all POIs in a single map; we would reserve space for the generated results using OSM and POIs.

Table 1 POIs types and aggregated information

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

3 Method and results

3.1 Delineating parcel boundaries

Parcel boundaries could be generated by OSM. Data preprocessing was necessary before delineating parcel boundaries using OSM. All transportation layers were merged into one layer GeoLines, in which the original layer names were retained for determining buffer distances. For tacking topological errors raised by contributors in OSM, the layer GeoLines was trimmed 200 m for eliminating hanging line segments, followed by extended with 20 m for connecting untouched lines. The processed layer Geolines was buffered in both sides of lines with the buffer width varying across original transportation layers according to Chinese actual road conditions. The buffer distance (one side) ranging from 2-30 m varied with the level of roads. We then erase buffered Geolines from the whole study area, namely the whole China territories (see Figure 1), to derive parcel boundaries, which was further overlaid with the administrative boundary of counties of China to distinguish parcels in each county. We finally generated 232,145 parcels in the whole China (see Figure 4).

3.2 Calculating density for all parcels

The density for all generated parcels could be estimated using various indicators, e.g. the density of POIs, online check-ins, road network as well as buildings (floor area ratio, FAR). These indicators could be used as the proxy for the density of human activity, both in the physical space and the virtual space. In this paper we would use the POIs density (the POIs count within each parcel divided by the parcel area) as the density of a parcel⁸. We further standardized the density to range from 0 to 1 for better inter-city and intra-city density comparison.

⁸ POIs within the buffered Geolines were accounted by their closest parcels in our experiment. We admit the urban activities over each POIs are not homogenous. We are crawling the whole country's check-ins for POIs used in this paper to eliminate this problem. We hope we could be able to use check-in density for evaluating urban density for each generated parcel at time of releasing the parcels online.

3.3 Selecting urban parcels

Urban parcels were selected from all generated parcels in Section 3.1 for inferring their urban functions and land use mix. The total urban land of each city region was estimated from the Chinese City Construction Yearbook in 2012 introduced in Section 2.1. Whether a parcel was urban or not within each city could be determined by allocating the total land area via considering the “potential” of a parcel to be urban. Since the urban land tended to be continuous in space, we used vector-based constrained cellular automata (CA) for allocating urban parcels in each city⁹ (Zhang and Long, 2013). In the proposed CA model, each parcel was regarded as a cell in CA, and the cell status was 0 (urban) or 1 (non-urban). At the very beginning of model simulation, all cells’ status was set as rural. Based on Li and Yeh (2002), the concept model of the proposed constrained CA was represented as $S_{ij}^{t+1} = f(S_{ij}^t, \Omega_{ij}^t, Con, N)$, where S_{ij}^t and S_{ij}^{t+1} are the states of a cell ij time t and $t + 1$, respectively; f is the transition function; Ω_{ij}^t is the neighborhood evaluation function; Con are the constraints on urban expansion; and N is the total number of cells.

Specifically, the probability of a cell ij changing its state from non-urban to urban at time t can be represented as $P_{ij}^t = (P_l)_{ij} \times (P_\Omega)_{ij} \times con(\cdot) \times P_r$, where $(P_l)_{ij}$ is a local potential representing the potential of a cell conversing from non-urban to urban state; $(P_\Omega)_{ij}$ is the state conversion potential of the cell within its neighborhood; $con(\cdot)$ is the restrictive condition for urban development; and P_r denotes the stochastic disturbance of any unknown errors (Feng et al., 2011). In this paper we did not consider the restrictive conditions and stochastic factors and the transition rule was then represented as $P_{ij}^t = (P_l)_{ij} \times (P_\Omega)_{ij}$. The local potential $(P_l)_{ij}$ can be determined through a set of factors using a logistic regression method

(Wu, 2002),
$$(P_l)_{ij} = \frac{1}{1 + \exp[-(a_0 + \sum_{k=1}^m a_k c_k)]}$$
, where a_0 is a constant, c_k is a spatial variable, and a_k is the estimated parameter/weight of c_k , and m is the count of spatial variables. In this paper, we selected three spatial factors, (1) the logarithm value of parcel size; (2) the compactness of a parcel; (3) the POIs density calculated in Section 3.2. The state conversion

potential of the cell within the neighborhood can be defined as
$$(P_\Omega)_{ij} = \frac{\sum con(S_{ij}^t = urban)}{n}$$
, where $con(S_{ij}^t = urban)$ represents the number of urban cells amongst the neighborhood of cell ij , and n is the count of cells in the neighborhood of cell ij . The parcels within 500 m were regarded as the parcel’s neighborhood in this paper. By comparing the global probability P_{ij}^t with a predefined threshold value P_{thd} in the range of $[0, 1]$, the model was

⁹ Each city has its own constrained CA model for allocating urban parcels.

then used to decide whether a non-urban cell can be converted to urban state at time $i + 1$:

$$S_{ij}^{t+1} = \begin{cases} \text{Urban} & \text{for } P_{ij}^t \succ P_{thd} \\ \text{NonUrban} & \text{for } P_{ij}^t \leq P_{thd} \end{cases}.$$

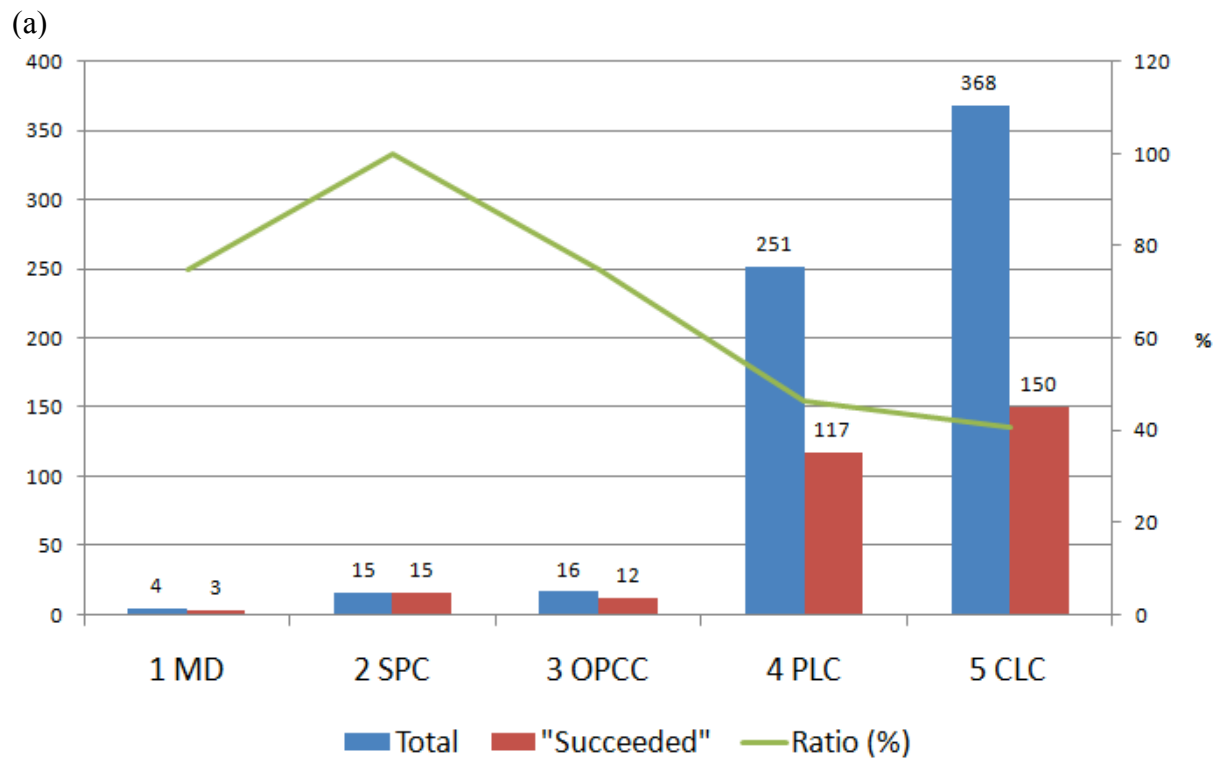
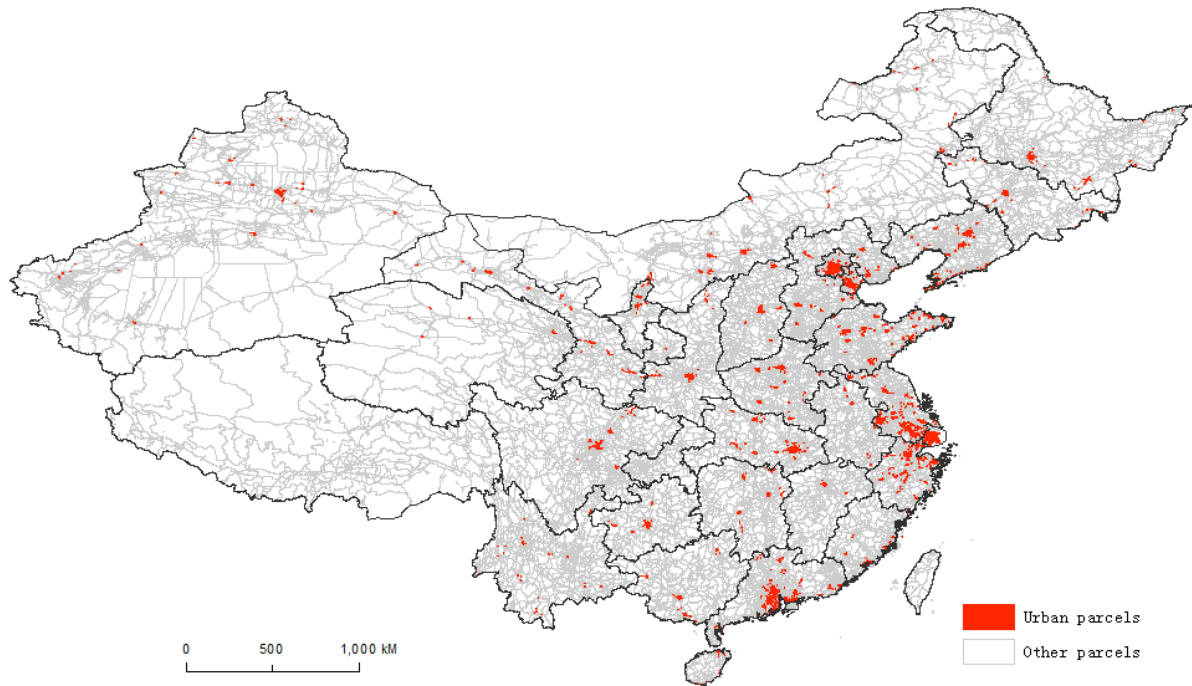
For calibrating the weights for constraints, we conducted logistic regression on the existing parcels manually prepared by planners in the city of Beijing (12,183 km², Yanqing and Miyun in the Beijing Metropolitan Area not included). Each parcel was regarded as a sample, and totally there were 125,401 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 China¹⁰. 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 selecting urban parcels from all parcels generated in a city region.

Table 2 Logistic regression results for the Beijing 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

We ran the proposed constrained CA model in all 654 cities separately. There were 357 cities having no accurate urban parcels meeting the total urban area constraint, due to over-large parcels in these cities. OSM data quality in these cities was very poor and only branch roads existed. More, we regard these cities with less than or equal to ten urban parcels as failed cities in selecting urban parcels. Finally, urban parcels in the 297 “succeeded” cities together with the identification ratio for each levels of cites were displayed in Figure 4. Only Chongqing among all MD cities was absent from "succeeded" cities, and all SPC cities were "succeeded". Over half of middle and small cities in the levels of PLC and CLC were also absent from the "succeeded" city list. For those "succeeded" cities, total urban parcels count was 82,645, 35.6% of all generated parcels in the 297 cities, and the average parcel count for each city was 278. Total urban area for the 297 cities was 25,905 km².

¹⁰ We admit the heterogeneity of weights in various cities. We do not have existing parcels in other cities in writing this paper.



(b)
Figure 4 All generated parcels and urban parcels in China (a, spatial distribution; b, the profile of "succeeded" cities)

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

The dominating urban function for identified urban parcels could be identified via analyzing each urban parcel's POIs composition. The urban function with a proportion over 50% in terms of POI count would be assigned as the dominating function for a parcel. E.g. there were 10 POIs in a parcel, and the "FIR" type POIs count was 6. In this case, the dominating urban function of this parcel would be "FIR (firms)". For this, all urban parcels were attached with the information of their urban function composition for identifying dominating urban function. in the 297 cities, only partial parcels (55,728, 67.3% of all urban parcels) were be assigned with a dominating urban function considering our definition on "dominating urban function". See Figure 5 for dominating urban functions for selected cities.

As a complementary for the dominating function of each urban parcel, we computed a mix index to denote the land use mixed degree (Frank et al., 2004). The mixed index (M) of a parcel is calculated as $M = -\sum_{i=1}^n p_i \times \log(p_i)$, where n denotes the number of POI types under consideration ($n=7$ in this paper), and p_i is the proportion of the i_{th} type POIs. This indicator, together with its dominating urban function, if any, helped us better understanding the urban function of a parcel in the background that almost all parcels are mixed. The calculating results for selected cities were shown in Figure5.

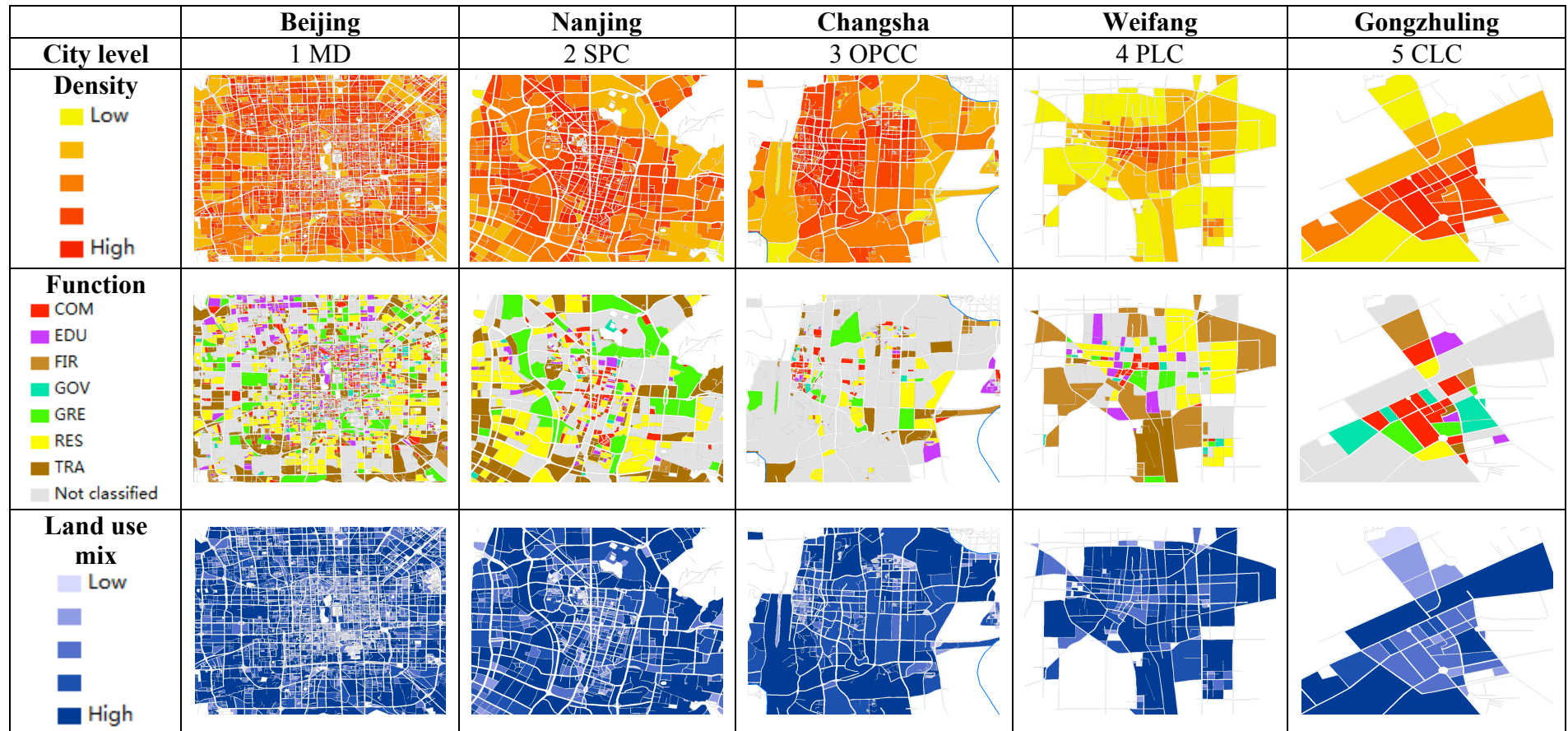


Figure 5 The generated parcels and their attributes in typical cities of China

4 Validation

We validated our established parcels in two levels, the regional level (geometry, attribute and density of parcels not accounted) and the parcel level.

4.1 *The regional level*

Since OSM data in China was not complete in terms of road coverage according to our preliminary evaluation in Section 2.3, we divided the validation in the regional level into two steps. First, we derived parcel boundaries and selected urban parcels from road networks in the 2011 ordnance survey (termed "ORDNANCE") in the whole China as detailed in Section 2.2. 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 3, 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 km² urban lands in ORDNANCE (54.2%) overlapping with those of DMSP/OLS. If assuming road space that was not accounted in ORDNANCE was all covered by DMSP/OLS, the overlapping ratio would increase to 60.0%. In addition, the time gap between ORDNANCE and DMSP/OLS, in the degree, underestimated the overlapped 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 were workable for Chinese cities according to the aforementioned evaluation. The generated urban pattern in this study could be upgraded via introducing more factors like main roads, city centers and exclusive development zones into the vector CA model in the near future.

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 inconsistency between DMSP/OLS and GLOBCOVER. The overlapped area of the two datasets was 19501 km², 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.

Table 3 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)	Overlapped 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

Second, we compared urban parcels between OSM and ORDNANCE in 297 cities where both datasets had selected urban parcels. Although road networks in both datasets were different (2013 and 2011 respectively), we selected urban parcels using the same total urban land for each city, thus making it possible to assuming urban parcels selected were for the year 2013. Moreover, the parcels in ORDNANCE were created using observed surveyed maps, which could guarantee parcels in ORDNANCE were all existing parcels in China. Since parcels in OSM and ORDNANCE were generated using the same approach, compare urban parcels derived from both datasets provided a new channel for evaluating the OSM data quality. The comparison results in Table 4 demonstrated that OSM has a much larger parcel size than that of ORDNANCE, which was generated using road networks with more details. There was 58.1% shared urban land by OSM and ORDNANCE. 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.

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)	Overlapped with ORDNANCE (km ²)
OSM	25905	82,645	31.3	15053
ORDNANCE	25670	260,098	10.0	-

4.2 The parcel level

In the parcel level, we compared the generated urban parcels using OSM and those established by planners in Beijing Institute of City Planning (BICP) using conventional approaches. Due to data availability, we only compared parcels in the city of Beijing with the total territorial area of 12,183 km², Yanqing and Miyun counties in the Beijing Metropolitan Area not included, in terms of average parcel size, spatial distribution as detailed in Table 5 as well as the parcel size distribution. Since urban parcels in BICP were drafted in 2010 with the total urban land of 1677.5 km², we re-ran the constrained CA model in Beijing using this total number and regenerated OSM for Beijing¹¹. The comparison results in Table 5 indicated that parcels by OSM were much larger than those drafted by planners, due to insufficient small roads in Beijing which subdivided parcels¹². The overlapping results between OSM and BICP was promisingly high as 71.2%, suggesting that the spatial pattern of urban parcels in

¹¹ The urban area of the city of Beijing was 1445.0 km² in 2012, which was still less than the aggregated results (1677.5 km²) using urban parcels in BICP. The data inconsistency between each other in China is not rare.

¹² Parcels by ORDNANCE in Beijing was similar with those by planners in BICP in terms of parcel size.

both datasets was similar which was further proved by the circular distribution comparison (see Table 5).

Table 5 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

We also compared the frequency distribution of size of all parcels in OSM and BICP as shown in Figure 6. All parcels in both datasets were in log-normal distributions with similar mean values, indicating that our approach for generating parcels using OSM was practical.

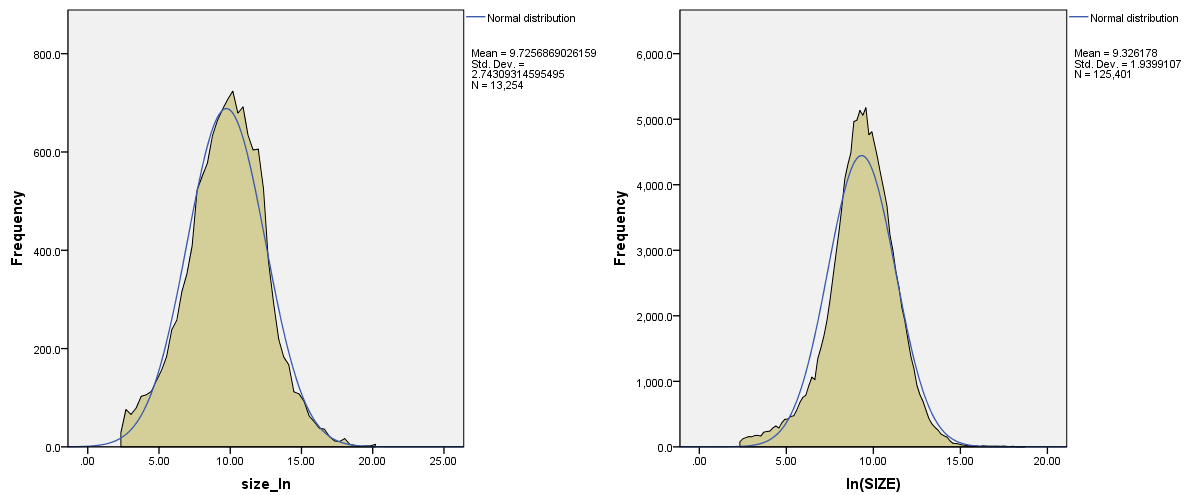


Figure 6 The parcel size probability distribution of all parcels in OSM (left) and BICP (right)

We further examined the inferred attributes of urban parcels in OSM. First, we examined the inferred density using POIs of urban parcels in OSM via comparing with floor space calculated from building footprints and their floor number¹³. The results showed that the correlation between POI count and floor space for each urban parcel in Beijing was positively correlated (Pearson correlation coefficient=0.858). This demonstrated that it was acceptable to use the ubiquitous and economic POIs as a proxy of urban density, in the background that the datasets of buildings were very costly in China. Second, since the classification of POI types and land use types in BICP were not totally consistent. We compared OSM parcels with dominating urban function of residence with the residential parcels in BICP. Lastly, there was no land use mix information for parcels in Beijing (even in the whole country). We therefore did not conduct the result validation in terms of land use mix.

¹³ Floor space information was limited to the parcels within the six ring road of Beijing, which was used in comparison. The sample size in the correlation analysis was 5523.

5 Discussion

5.1 *Potential bias and next steps*

While admitting the merits of this study, there are several limitations in the present study. First, some urban parcels by OSM were rather large due to not being enough road networks in OSM in some cities, although we have controlled the maximum size (10 km^2) of an urban parcel during selecting them. This condition is expected to be avoided with the increasing quality of OSM by more contributors in China (as demonstrated by Figure 1). Second, there was bias on using POIs for estimating a parcel's density; in the fact the activity strength among POIs might not be equal (e.g. a large residence community and a grocery in POIs). In the near future, we would introduce more data like check-ins for POIs, taxi trajectories and bus/metro smart card records into more realistic density estimating via considering flows and activities over space. Last, it would be possible to gain knowledge from other land-cover and land use datasets (mostly in a coarser spatial scale) for assigning urban function for urban parcels, rather than only using POIs that did not well coincident with the land use inventory in China. This study, nevertheless, provided an alternative framework, integrating ubiquitous available OSM and POIs, to automatically generating parcels and inferring their attributes. This framework is expected to be more efficient and accurate in the close future, on being fed with more fine-scale data related with human mobility and urban activities.

5.2 *Potential applications*

The established national-scale urban parcels with necessary attributes for 297 cities in China could be applied in but not limited to the following aspects. First, parcels reflecting the characteristics of living, working and entertainment etc provide researchers it possible to conduct large-scale inter-city urban studies using fine-scale data, which would be a new agenda and stimulates more interesting findings, e.g. urban morphology analysis for the whole country as well as inter-city comparison on quality of life (QOL). Second, the dataset provides practitioners (like planners) for planning compilation and urban spatial monitor, thus promoting the scientific degree of spatial plan of China. Urban spatial development is then possible to be revealed via comparing automatically generated urban parcels using OSM in various years. Third, we would be able to identify fine-scale urban built-up area for all Chinese cities (expandable to all world cities), thus possible to evaluate the data quality of OSM via comparing with real networks. Fourth, we could use generated parcels as a proxy of other ubiquitous data, e.g. check-ins, photos, mini-blogs, and aggregated information via allocation for accounting environmental impact, energy consumption, and carbon emission. Then we could finally extend the present dominating point-based analysis on these “big data” to polygon based. Lastly but not least, it would be possible to combine established parcels with parcel subdivision techniques and urban growth models (e.g. a rule-based vector CA based model) for simulating future urban parcels in the manner of the national-scale. We are now working on some aspects based on the outcomes of this paper. With the release of the established urban parcels using OSM, we expect it would attract more researchers to address some of but not limited to the aforementioned aspects.

6 Concluding remarks

In this paper, we developed a method for automatically generating parcels and inferring their attributes like density, urban function and land use mix using OSM and POIs. The whole research includes the following processes. First, parcels were delineated by roads in OSM via eliminating buffered roads from the whole study area. Second, urban parcels were selected from all generated parcels using a constrained cellular automata based model, considering the size, compactness, and location of parcels, as well as total urban area for each city. Third, density of parcels was derived via calculating the POIs count within each urban parcel. Fourth, the dominating urban function and land use mix of each urban parcel was identified via analyzing the urban functions of POIs within each parcel. The method was used in China for generating 82,645 urban parcels in 297 cities. The final results have been validated in several dimensions. First, urban parcels were compared with urban areas in other coarser data source. Second, they were compared with urban parcels generated from the survey map of China, and in Beijing we compare them, together with their function and density, with urban parcels manually prepared by planners. Finally, the inherent characteristics of generated parcels were analyzed in terms of the probability density of parcel size and compactness.

Our contributions of this paper mainly lie in the following aspects. First, we proposed a simple and straightforward framework for delineating parcels, selecting urban parcels and calculating their density, function and land use mix using ubiquitous available OSM and POI. This is a first test on using volunteered GIS data. The generated results could be updated at a year base using updated OSM and POIs. Second, a vector-based cellular automata model was introduced for selecting urban parcels among all generated parcels according totals from yearbooks. This extends the application domain of cellular automata. Third, land use mix has been calculated for each generated parcels supported by the diversity POIs, which is seldom calculated in the parcel level. Fourth, we generated the existing land use and density maps for over six hundred cities in China, which would be shared online (when the manuscript is published) as an important data infrastructure for both practitioners and researchers. Moreover, our new than Jiang and Liu (2012) lies in that, although we are using the similar process to generate parcels/blocks, our research focuses on how to selecting urban parcels and inferring their attributes, and they focuses on universal rules of blocks and natural cities as well as identifying centers.

Acknowledgments: The first author would like to acknowledge the financial support of the National Natural Science Foundation of China (No.51078213). Dr. Kang Wu at Capital University of Economics and Business is appreciated for his helping us preparing the boundaries of all 654 Chinese cities. Our thanks also go to Dr Dong Li at China Academy of Urban Planning & Design for generously providing us with the ordinary survey map of China, Mr. Mingrui Mao at Beijing Institute of City Planning for checking the data quality of generated urban parcels in Beijing, and Dr Jianghao Wang at IGSNRR, Chinese Academy of Sciences for visualizing the established urban parcels online at MapBox. We have released the generated parcels using OSM and POIs online for free downloading to everyone. For downloading and accessing online visualization of urban parcels of China, please go to the data-released channel of Beijing City Lab, <http://longy.jimdo.com/data-released/>.

References

1. 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.
2. Bontemps, S., Defourny, P., Van Bogaert, E., Arino O, Kalogirou, V., & Perez, J. R. (2009). GLOBCOVER: Products Description and Validation Report.
3. 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.
4. 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.
5. 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.
6. Girres, J. F., & Touya, G. (2010). Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*, 14(4), 435-459.
7. Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), 211-221.
8. Haklay, M., & Weber, P. (2008). Openstreetmap: User-generated street maps. *Pervasive Computing, IEEE*, 7(4), 12-18.
9. 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.
10. 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.
11. 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).
12. 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*, (accepted), 1-30.
13. 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/>)
14. Long, Y., Liu, X. (2013). How mixed is Beijing, China? A visual exploration of mixed land use. *Environment and Planning A* (accepted)
15. 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.
16. Ramm, F., Topf, J., & Chilton, S. (2010). OpenStreetMap: using and enhancing the free map of the world. Cambridge: UIT Cambridge.
17. Soto, V., & Frías-Martínez, E. (2011). Automated land use identification using cell-phone records. In *Proceedings of the 3rd ACM international workshop on MobiArch* (pp. 17-22). ACM.
18. 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.

19. 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 (pp. 1-8). ACM.
20. 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.
21. 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), 073535-073535.
22. 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 (pp. 186-194). ACM.
23. 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.

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