Research Paper

Understanding uneven urban expansion with natural cities using open data

Ying Long[a], Weixin Zhai[b,c], Yao Shen[d], Xinyue Ye[e,*]

[a] School of Architecture and Hang Lung Center for Real Estate, Tsinghua University, PR China
[b] Institute of Remote Sensing and Geographic Information System, Peking University, PR China
[c] Department of Geography and Earth Sciences, University of North Carolina at Charlotte, United States
[d] University College London, United Kingdom
[e] Kent State University, United States

ARTICLE INFO

ABSTRACT

The last several decades have witnessed a rapid yet uneven urban expansion in developing countries. The existing studies rely heavily on official statistical yearbooks and remote sensing images. However, the former data sources have been criticized due to its non-objectivity and low quality, while the latter is labor and cost consuming in most cases. Recent efforts made by fractal analyses provide alternatives to scrutinize the corresponding "natural urban area". In our proposed framework, the dynamics of internal urban contexts is reflected in a quasi-real-time manner using emerging new data and the expansion is a fractal concept instead of an absolute one based on the conventional Euclidean method. We then evaluate the magnitude and pattern of natural cities and their expansion in size and space. It turns out that the spatial expansion rate of official cities (OCs) in our study area China has been largely underestimated when compared with the results of natural cities (NCs). The perspective of NCs also provides a novel way to understanding the quality of uneven urban expansion. We detail our analysis for the 23 urban agglomerations in China, especially paying more attention to the three most dominating urban agglomerations of China: Beijing-Tianjin-Hebei (BTH), Yangtze River Delta (YRD) and Pearl River Delta (PRD). The findings from the OC method are not consistent with the NC method. The distinctions may arise from the definition of a city, and the bottom-up NC method contributes to our comprehensive understanding of uneven urban expansion.

1. Introduction

Most of the urban expansion studies use remote sensing images or statistical yearbooks, while emerging new data and open data provide novel opportunities for understanding urban expansion. In this paper, we will examine national wide urban expansion using open data as well as referring to the idea of "Natural City" (NC) which means spatially clustered geographic events, such as the agglomerated patches aggregated from individual social media users' locations (Jiang & Miao, 2015; Long, 2016). Natural city may include natural town as well. NCs are a product of the bottom-up process in terms of data collection and geographic units or boundaries, which can be extracted from RS images, GPS and location-based social media data (Jiang & Miao, 2015). Fractal geometry is the base of NCs. A fractal is a mathematical set that exhibits a repeating pattern which displays at every scale (Gouyet, 1996). Fractal geometry is a workable geometric middle ground between the excessive geometric order of Euclid and the geometric chaos of general mathematics. It is based on a form of symmetry that had previously been underused, namely invariance under contraction or dilation (Mandelbrot, 1983;

Mandelbrot & Blumen, 1989). Researches on spatial analysis have concluded that artificially planned and designed spatial objects such as urban forms and transportation networks can also be treated as fractals (Batty, 2008; Clauset, Shalizi, & Newman, 2009). The fractal structure of the urban form is more evident when the urbanized areas of cities, metropolis, or urban systems are seen as a whole. Looking at urban growth/expansion in a fractal manner is an alternative perspective compared with the conventional Euclidean way (Batty, 1991; Chen, 2008, 2015). Thus, the form of the urbanized areas in a city can be treated as a fractal and effective measurement for spatial analysis. Not only does the introduction of fractal geometry provide a powerful tool for the mathematical description of a city system, but it also makes it possible for a number of geographic fractal simulations to compensate for the lack of geographic experimental studies (Batty & Longley, 1986).

Urban expansion is a geographically uneven process (Wei & Ye, 2014a). For example, uneven urban expansion in developing countries like China is due to "uneven opportunities and dualism, including the household registration system, family planning, public resource allocation and land systems", "uneven process of globalization and economic

---

Module 1 Section 2.1, 4.1
Derivation

RS images
Road junctions
Point of Interests
Weibo records

Heavy tail detection
NC construction

Physical NCs
Morphological NCs
Functional NCs
Social NCs

Module 2 Section 2.2, 4.2
NC Evaluation and Evolution

Physical Evaluation → Physical Evolving
Morphological Evaluation → Morphological Evolving
Functional Evaluation → Functional Evolving
Social Evaluation → Social Evolving

t2-t1

Evaluation Dimension
Rank size   Average size
Ht-index   Density

Module 3 Section 2.3, 4.3
Expansion Evaluation

Overlay (t2-t1)   NC expansion on 4 dimensions

Nationwide
Magnitude evaluation
23 City regions

Nationwide
Pattern evaluation (quality)
23 City regions

OC expansion

Yearbooks (size)
Overlay (t2-t1)   RS images (size and pattern)
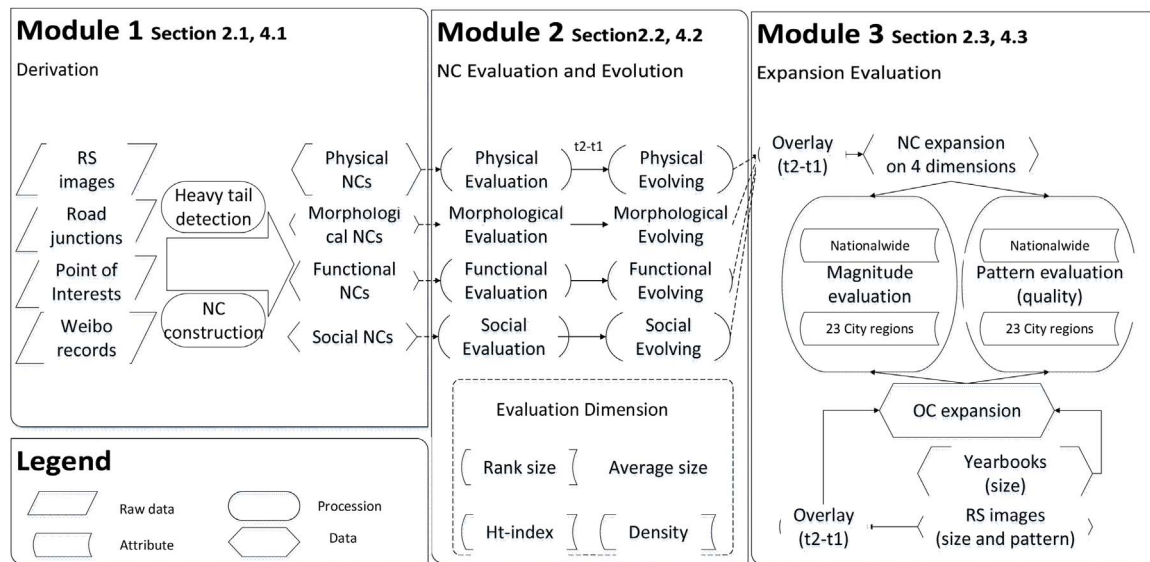
Legend
Raw data   Procession
Attribute   Data

Fig. 1. The overall framework for this study.
Note: t1 = 2009 and t2 = 2014 in this paper.

growth", and "unequal ecological carrying capacities ranging from overloading to richness" (Wei & Ye, 2014b, p759). Despite the importance of urban expansion research, the rate and magnitude have not been universally validated in various sources. For instance, the statistical yearbooks are the main data sources for describing urban expansion for Chinese official cities (OCs), each of which has the explicit administrative boundary. Yearbooks have long been the basic channel for understanding urban expansion in China (Anderson & Ge, 2004; Li, Wei, Liao, & Huang, 2015). Most of, if not all, existing urban studies for China adopt the administrative boundaries of OCs. The shortcomings of this dataset are summarized as the following. First, the existing OCs are not an objective way to observe the Chinese city system. Many towns have a large population, and they should be regarded as cities in the western context. However, they are not counted as the official city in China. In addition, some OCs are very rural. Second, the data quality of statistical yearbooks has been criticized for their resolution and informational richness. The geometrical properties and land use details of urban expansion are not recorded. Therefore, we cannot fully understand Chinese urban expansion using statistical yearbooks.

Besides the statistical yearbooks, remote sensing (RS) images are used to understand urbanization pattern and urban expansion. These RS images range from images with a resolution higher than 100 m by Landsat Thematic Mapper (TM), enhanced TM plus (ETM+) (Deng, Huang, Rozelle, & Uchida, 2010; Kuang, Liu, Zhang, Lu, & Xiang, 2013; Liu et al., 2005; Van de Voorde, Jacquet, & Canters, 2011) to the Operation Line-scan System (OLS) of the Defense Meteorological Satellite Program (DMSP) with 500 m resolution (He et al., 2006; Liu, He, Zhang, Huang, & Yang, 2012). Due to its labor and cost consuming nature in some cases (Liu, Liu, Zhuang, Zhang, & Deng, 2003; Liu et al., 2005, 2010), remote sensing images cannot be easily applied to study urban expansion for all the cities in a timely fashion. In addition, RS analysis emphasizes the physical perspective of urban expansion without taking into account other dimensions. Therefore, an alternative approach to redefine the city system and understand its urban expansion from various dimensions is urgently needed.

Open Data have provided opportunities for quantitative urban studies (Reichman, Jones, & Schildhauer, 2011). Open data is the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control (Auer, Bizer, & Kobilarov, 2007). Our definition of open urban data is broad and refers to urban data that are openly accessible to the public, mainly coming from three over-

lapping though different sources (Gurin, 2014): official data portals, big data initiatives, and the broader open data community. Ubiquitous open data from governments, commercial and social media websites are increasingly available to researchers (Shaw, Tsou, & Ye, 2016; Wang, Jiang, Liu, Ye, & Wang, 2016; Wang, Ye, & Tsou, 2016). Open urban data have been a viable and cost-efficient option for understanding the built environment from both physical and social aspects, which has been particularly useful for the developing countries where data infrastructure is not sufficient (Long & Liu, 2015). The applications of open data have demonstrated their importance and suitability for urban studies, planning practice and commercial consultancy (Liu, Derudder, & Wu, 2016; Long & Liu, 2015; Wang, Wang, Ye, Zhu, & Lee, 2015; Wang, Jiang et al., 2016; Wang, Ye et al. 2016).

In this study, we first redefine the city system by means of NCs using open data. We then derive urban expansion from various dimensions based on the derived NCs. In this paper, we initially distinguish urban expansion in two basic categories: size expansion and spatial expansion, corresponding to the conventional urban expansion using yearbooks and RS images, respectively. In this line, we are able to evaluate the identified size/pattern of urban expansion for various types of NCs for better understanding the process of urban growth. The identification and evaluation results for OCs can be benchmarked with those for NCs. Our proposed framework for understanding urban expansion with NCs is based on the fractal concept, within which the dynamics of internal urban contexts are reflected in a quasi-real-time manner. The expansion is apart from an absolute one in the traditional method.

This paper is structured as follows: Section 2 introduces the main methods of the NC generation as well as the identification and characteristics of urban expansion. Section 3 presents the study area of and data used in the research. Section 4 outlines the results of urban area extraction, urban expansion identification and evaluation, along with benchmarking with other existing data sources. Section 5 summarizes contributions and indicates future work. Finally, Section 6 draws a conclusion.

2. Methodology

This paper adopts the analytical framework consisting of three modules depicted in Fig. 1. The first module focuses on the identification of NCs. The second concentrates on the evaluation and the shift in the evolution of NCs on physical, morphological, functional and social dimensions for understanding urban expansion in size. The third module illustrates urban expansion in space and evaluates its quality
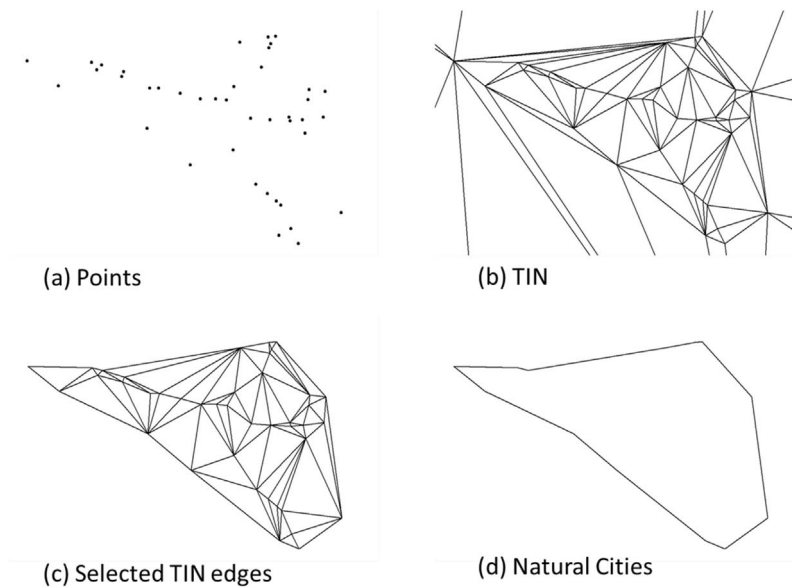
(a) Points

(b) TIN

(c) Selected TIN edges

(d) Natural Cities

Fig. 2. The four procedures for deriving natural cities from points.

in terms of magnitude and pattern.

2.1. Deriving natural cities from various dimensions (module 1)

A city is a relatively large human settlement according to the conventional definition. There is consistently a minimum population requirement. In this paper, we refer to Jiang and Liu (2012)'s approach to extract NCs.[1] It is based on the head/tail division rule stating given a variable *x*, if its values follow a heavy tailed distribution, the mean value of the values can divide all the values into two parts: a high percentage in the tail, and a low percentage in the head. The heavy tailed distribution refers to the statistical distributions that are right-skewed, for example, power law, lognormal, and exponential. In the previous studies, the densities of street nodes, the sizes of street blocks, and the nighttime imagery pixel values (DN) demonstrated a heavy tailed distribution (Jiang, 2013, 2015a, 2015b, 2015c). In this paper, the standard NC deriving procedures include four steps: 1) create a TIN with points (like road junctions); 2) convert TIN into TIN edges; 3) choose the edges less than the mean; 4) dissolve all the shorter edges into polygons to be individual NCs. This process is exemplified in Fig. 2. Unlike the OCs defined by the administrative boundaries, NCs refer to human settlements or human activities in general on earth's surface that are objectively or naturally defined and delineated from massive geographic information. NC method is closer to the inherent structure of spatial entities, thus, providing us an alternative and valuable way to revisit the complex city system.

2.2. Evaluating evolving natural cities and their expansion in size (module 2)

Four types of NCs can be derived from multiple data sources, namely the physical NCs from remote sensing images, morphological NCs from road junctions, functional NCs from points of interest (POIs) and social NCs from social media records. The physical NCs extracted from the remote sensing images represent the skeleton of urban built-up areas. The road junctions are the elements of space connections in the built environment, thus making the corresponding NCs indicate the morphology of urban areas. The POIs-generated NCs characterize the urban function distribution. The NCs from the social media records describe social significance of human activities. The physical-morphological-func-

tional-social sequence is a path from physical developments to social developments, representing a full process of urbanization.

The evolution of NCs can be detected through deriving NCs during different periods. In doing so, five indicators are utilized to evaluate the NCs' evolution in size: rank size distribution, ht-index, average city size, density and expansion size. When we arrange all the cities according to their sizes, we get the rank-size plot. Considering the probability distribution of city sizes of all the derived NCs in China, we employ the head/tail division rule, a novel classification scheme for data with a heavy tailed distribution (Jiang, 2013, 2015a; Jiang & Yin, 2014). The ht-index measures the reoccurrence of the possible head/tail division, capturing the fractal properties of the NCs to describe the development degree (Jiang, 2015b, 2015c). The head/tail division rule is superior to conventional classification methods in terms of capturing the inherent scaling pattern and the ht-index of geographic feature h exists if there are far more small geographic features than larger one at different scale when recursive times are h-1 (Jiang, 2013; Jiang & Yin, 2014). As for average size, that is literally. Density is the number of points (road junctions, POIs and social media records) per km^2 within some specific urban areas. A higher density represents more road junctions, POIs or social media records in the specific region, indicating more development or prosperity. Last but not least, we define the expansion in size of NCs as the total urban area variation of NCs in a time period. For benchmarking with the expansion in size of the four types of NCs, we also derive the expansion in size for OCs using yearbooks or remote sensing images. Based on the comparison between OCs and NCs, we could gain knowledge from the definition of two different viewpoints of the urban area.

2.3. Evaluating urban expansion in space and its quality (module 3)

In addition to evaluating evolving NCs and their expansion in size, we can further explore the spatially expanded urban areas both in NCs and OCs, comparing the expanded areas with the originally existing urban areas to gain more knowledge on the quality of urban expansion in space. Different from the urban evolution in module 2, this module mainly concentrates on the urban shape and quality in different times. This can be achieved by using the overlay function in almost all GIS software, fed with urban areas from OCs or NCs in the aforementioned four dimensions.

Two indicators are proposed for evaluating urban expansion in space, magnitude and relative maturity. The magnitude of urban expansion in space is defined by the geometrical size of expanded urban areas. The ratio of magnitude and the urban area size in t_1

---

[1] We admit that some NCs may be greatly smaller than a city in common sense. These NCs as natural agglomerations may correspond to towns or streets in the real world. In this paper, we use this term following Jiang's definition (Jiang, 2015a).

indicates the spatial urban expansion speed. The relative maturity indicator is used for measuring the quality of urban expansion in space in contrast to the originally existing urban areas, and is calculated as the ratio between the density of the newly expanded urban areas and that of the original urban areas in t_1. A higher quality represents a better development level of expansion, showing that there are a rapidly expansion increasing speed on some specific aspect. We suggest the ratio be calculated using road junctions for road junctions-derived NCs, POIs for POIs-derived NCs and social media records for social media records-derived NCs. Road junctions, POIs and social media records are applied to the remote sensing images derived OCs.

3. Data

3.1. The official cities in China and the observed expansion from yearbooks in 2009 and 2014

The Chinese city system has a long history of being defined from the administrative view. Most of the statistical data correspond to administrative Chinese cities (official cities, OCs), rather than concepts like metropolitan statistical area (MSA) in the United States or functional urban area (FUA) in European Union countries. There are totally 653 OCs in 2014 in China (Fig. 3), with the sum of the area of jurisdiction 782,330.9 km². On the basis of the Chinese administrative system (Ministry of Housing and Urban Development, 2013; Marceau, 2005), there are mainly five categories of cities: municipalities directly led by the nation (MD, 4 cities), sub-provincial cities (SPC, 15 cities), other provincial capital cities (OPCC, 17 cities), prefecture-level cities (PLC, 250 cities), and county-level cities (CLC, 367). The study area of this paper covers the whole Chinese territory, not limited to administrative areas of OCs in Fig. 3. However, the administrative boundaries are used to extract our analysis results in the whole country for benchmarking with the conventional indicators available in yearbooks.

China has been experiencing a rapid process of urbanization, with the urban population percentage growing from 17.9% in 1978 to 49.7% in 2010 (Nation Bureau of Statistics, 1982, 2011). As a result, China has become the world's largest and most rapidly transforming urban society (Bai, Shi, & Liu, 2014; Wei & Ye, 2014a). Besides, the uneven urban expansion rates are particularly salient in China, where annual rates of urban land expansion (1970–2000) vary from 13.3% in coastal areas to 3.9% in western regions (Seto, Fragkias, Güneralp, & Reilly, 2011). Due to the dramatic urbanization, large amounts of rural lands have been converted to built-up areas, and the irreversible and human-dominated land use transformation has led to both environmental and socio-economic problems (Qi, Ye, Zhang, & Yu, 2014; Qiao, Lee, & Ye, 2016; Wan et al., 2015; Zhang et al., 2013). Urbanization in China has therefore attracted wide attentions from researchers, policy makers, investors, and the public (Yue, Ye, Xu, Xu, & Lee, 2014). As Wei and Ye (2014b) state, "an understanding of urbanization and land use change in China is required for appropriate strategies and policies to facilitate future sustainable development" (p757). For instance, the recently released National New-Type Urbanization Plan (2014–2020) has marked a new consensus among the country's ruling elites to facilitate sustainable and human-centered urbanization patterns (Bai et al., 2014; Wallace, 2014).

For gaining knowledge on the urban expansion of OCs, we extract urban areas of Chinese OCs in 2009 and 2014 from the yearbook of Ministry of Housing and Urban Development (2013). In 2014, the total urban area (UA) for the 653 Chinese cities is 49904.1 km². The urban expansion from 2009 to 2014 is plotted in Fig. 4.

3.2. Road networks in ordnance surveys in 2009 and 2014-morphological data

We amass the Chinese equivalent of Ordnance Survey data (hence the use of 'ORDNANCE' as the abbreviation) from China Siwei
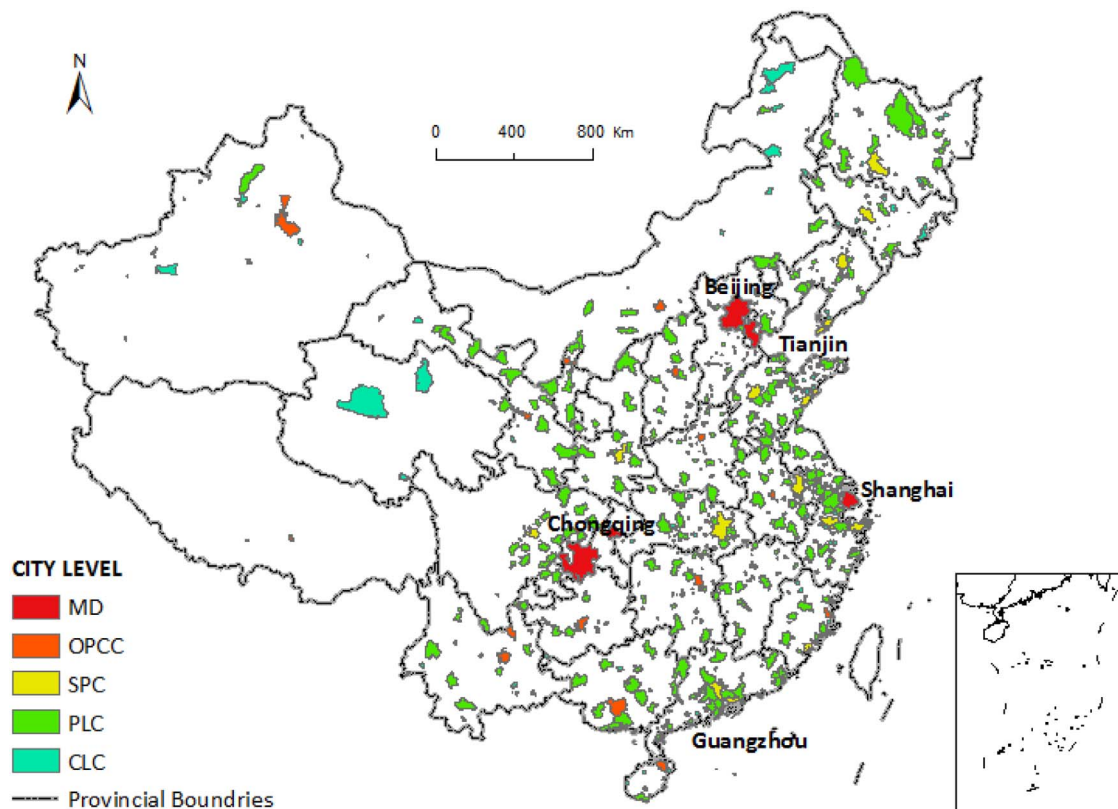

Fig. 3. Administrative areas of all Chinese official cities.
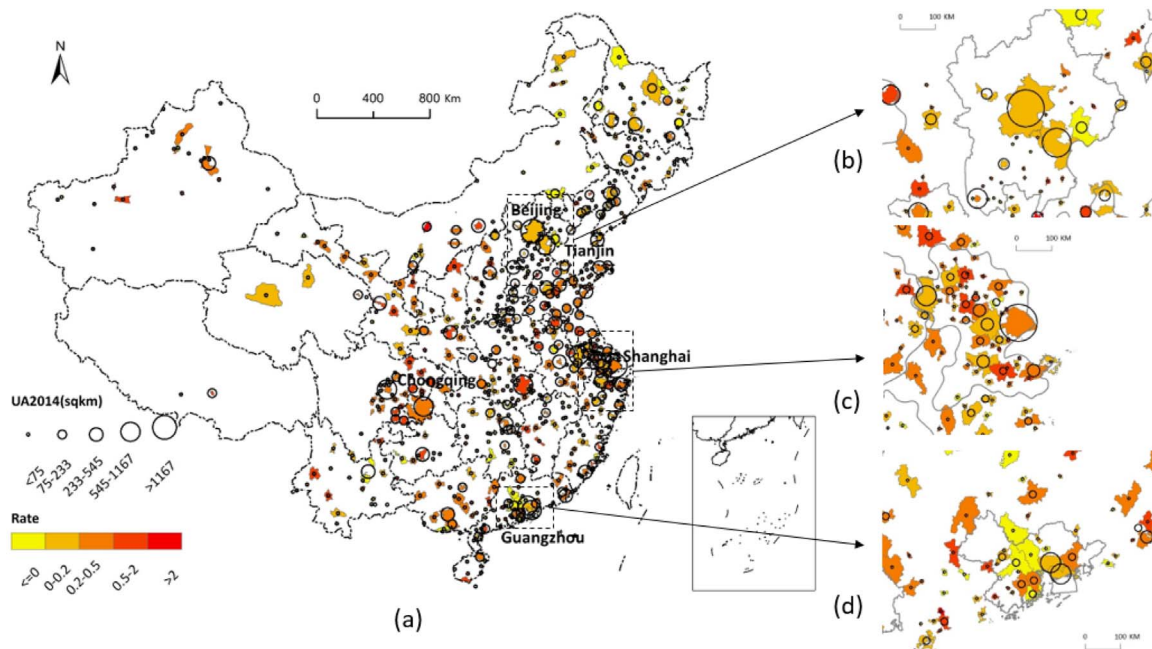Note: Missing data in Taiwan

Fig. 4. Urban areas in 2014 at the city level. (a) China, (b) Beijing-Tianjin-Hebei (BTH), (c) Yangtze River Delta (YRD), (d) Pearl River Delta (PRD).
Note: The split thresholds in the process of head/tail division rules are chosen as the criteria for the legend classification. The rate denotes (UA2014-UA2009)/UA2009 for each city. Missing data in Taiwan.

Surveying and Mapping Technology Co. Ltd., one of the largest navigation companies in China. Detailed road networks are available in the data. The total amount of all levels of road segments in 2009 and 2014 are 3,342,322 and 9,532,984, respectively. Nearby but topologically separated lines are connected and the road junctions are generated from the road networks by using Network Analyst Tools in ESRI ArcGIS. There are 2,832,045 road junctions in 2009 and 9,054,366 road junctions in 2014, as shown in Fig. 5(a, b).

3.3. Points of interest (POIs) in 2009 and 2014-functional data

POIs in 2009 and 2014 are gathered and geo-coded from business cataloguing websites (www.baidu.com). The initial twenty POI types are clustered into eight more general assemblies, among which commercial sites account for most POIs, followed by business establishments, transportation facilities, and government buildings. The data quality is secured through manually checking randomly selected POIs. There are 1,563,624 POIs in 2009 and 10,589,322 POIs in 2014, as shown in Fig. 5(c, d).

3.4. Urban areas interpreted from night light images in 2009 and 2014-physical data

To derive NCs from the physical dimension, we gather the urban areas presented in the 1 km resolution night light images retrieved from DMSP/OLS in 2009 (He, Liu, Tian, & Ma, 2014) and in the 500 m resolution night light images retrieved from VIIRS/DNB (the updated version of DMSP/OLS) in 2014. The total urban areas in 2009 are 52,251.0 km^2 with 2,166 patches (the mean patch size is 24.1 km^2) in China. We find that 26.4% of urban areas in 2009 is not urban in 2014, due to the different approaches adopted for image interpretation and the distinct spatial resolution of raw images. There are 12,752 urban patches totaling 59,530.5 km^2 in 2014, and the mean patch size is 4.7 km^2.

3.5. Weibo records in 2011 and 2014-social data

Sina Weibo, the Chinese version of Twitter, is a leading online social

media (microblogging) in China, which allows users to publish, share and discuss short messages on the website (http://www.weibo.com). Weibo is used by over 30% of Internet users in China, with a market penetration similar to Twitter. It was launched by SINA Corporation on 14 August 2009, and at the end of the fourth quarter of 2014, the number of monthly active users has reached 176 million, 80% of which come from mobile terminals. We collected the Weibo check-in records with geographical information in 2011 and 2014 (about 1% Weibos have geo tags). There are 54,486 Weibo record points in 2011 and 874,542 Weibo records in 2014, along with the sharp increase in its user quantity. Due to incompleteness of data in Jiangxi and Hunan, our analysis on Weibo records does not cover the two provinces. In our research, we set the Weibo records in 2011 as the benchmark instead of 2009 in view of data availability.

4. Results

4.1. Derived natural cities from various dimensions (module 1)

Following the method mentioned above, we establish NCs based on remote sensing images, road junctions, POIs and Weibo records in the year 2009 and 2014, respectively (Weibo records in 2011 and 2014). The results are shown in Fig. 6. For better visualizing these NCs, we use the centroid to present an NC and apply the head/tail division rule for categorizing city sizes in area, and the threshold for the head/tail division is set at 40% in this study.

We are able to benchmark the two types of OCs and the four types of derived NCs of China from various dimensions, as revealed in Table 1. The statistics of OCs from yearbooks and remote sensing images are similar partially due to the same city amount. There are much more NCs than OCs in each year, which is evident to confirm the arbitrariness of defining OCs. The total urban areas of the two types of physical and morphological NCs (RS images and road junctions) significantly increase from 2009 to 2014, while the functional NCs do not follow this trend. This indicates that the functional developments are far behind the physical and morphological developments of Chinese cities and might need more time to forester. The maximum NC size and standard deviation decline, indicating a tendency of homogenization
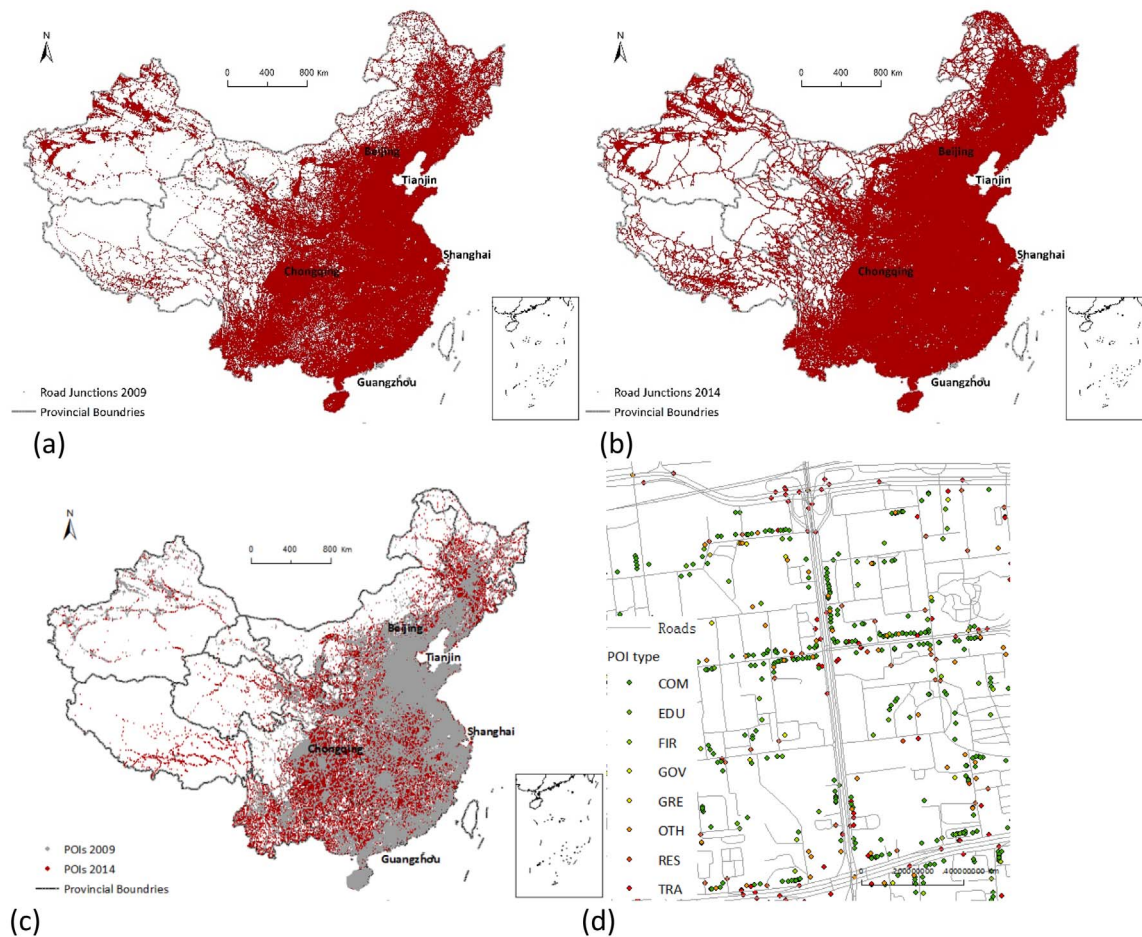
Fig. 5. Road junctions and POIs of China in 2009 and 2014 (a) road junctions in 2009, (b) road junctions in 2014, (c) POIs in China, (d) POIs in a part of Beijing.
Note: COM denotes commercial sites, EDU denotes education facilities, FIR denotes firms, GOV denotes government sites, GRE denotes green spaces, OTH denotes other types, RES denotes residence communities, and TRA denotes transport facilities. Miss data in Taiwan.

and fragmentation. The point data in 2014 are relatively dense, leading to a lower value of cut distance in 2014 than in 2009. The ht-index is computed for all the patches as a whole, representing the development level of the entire country. The ht-index computed by RS images and road junctions is unchanged, while that by POIs increases from 6 to 7, and that by Weibo records increases from 4 to 6.

4.2. Evolving natural cities and their expansion in size (module 2)

We order the NC patches according to their sizes, and the rank-size plots of the four kinds of NCs are indicated in Fig. 7. The sizes of NCs versus their ranks follow a heavy tailed distribution, suggesting that there are far more small patches than large ones. Two patterns can be observed from all distributions of NCs. The first is for the NCs from large-scale datasets (road junctions, POIs and Weibo records), and the coefficients are around −2. The second is for the NCs from RS images, and the coefficients are around −1 (Zip'f law featuring a coefficient of −1). The distributions at the tail part are also different for the two patterns, which may result from fewer city numbers in the RS images derived NCs, compared to large-scale data derived NCs.

We are able to gain knowledge on the urban expansion in size by comparing the features of the OCs and derived NCs. As indicated in Table 2, we find that the four types of NCs become compact in size from 2009 to 2014 from all four perspectives, as revealed by the average city size. The size expansion rate from 2009 to 2014 is 21.4% for the yearbook derived OCs and 3.0% for the RS images derived OCs, which may suggest that the official urban expansion speed has been underestimated by the yearbooks. However, the rate for NCs is 50.4% in the

morphological dimension, 22.0% in the physical dimension, and −4.8% in the functional dimension, 37.6% in the social dimension. In contrast to the official rate from the yearbooks, the transport infrastructure based urbanization is faster than the official rate, but the urban function developments are again confirmed to be behind the official rate and morphological rate.

We also evaluate NCs from the view of density, which is calculated as the number of points (road junctions, POIs and Weibo records) per km² within NCs. This indicator represents the development prosperity. The NCs in 2014 are associated with higher density than that in 2009, in both morphological and functional dimensions, indicating the morphological and functional developments are dramatically progressing within the NCs. In addition, the density of road junctions (64.7 km^{-2} and 62.4 km^{-2}), POIs (143.0 km^{-2} and 136.6 km^{-2}) and Weibo records (12.6 km^{-2} and 11.9 km^{-2}) of OCs by remote sensing images are also shown in Table 2 for comparison. The RS images derived OCs are with a higher density of road junctions and a lower density of POIs.

4.3. Urban expansion in space and its quality evaluation (module 3)

4.3.1. The magnitude evaluation

Urban expansion in space can be derived by employing the method in Section 2.3. The results from remote sensing images, road junctions, POIs and Weibo records shown in Table 3 reflect the physical, morphological, functional and social transformation. Different from urban expansion in size, the expanded urban areas generated from a spatial overlay operation are the areas that are non-urban in 2009 and
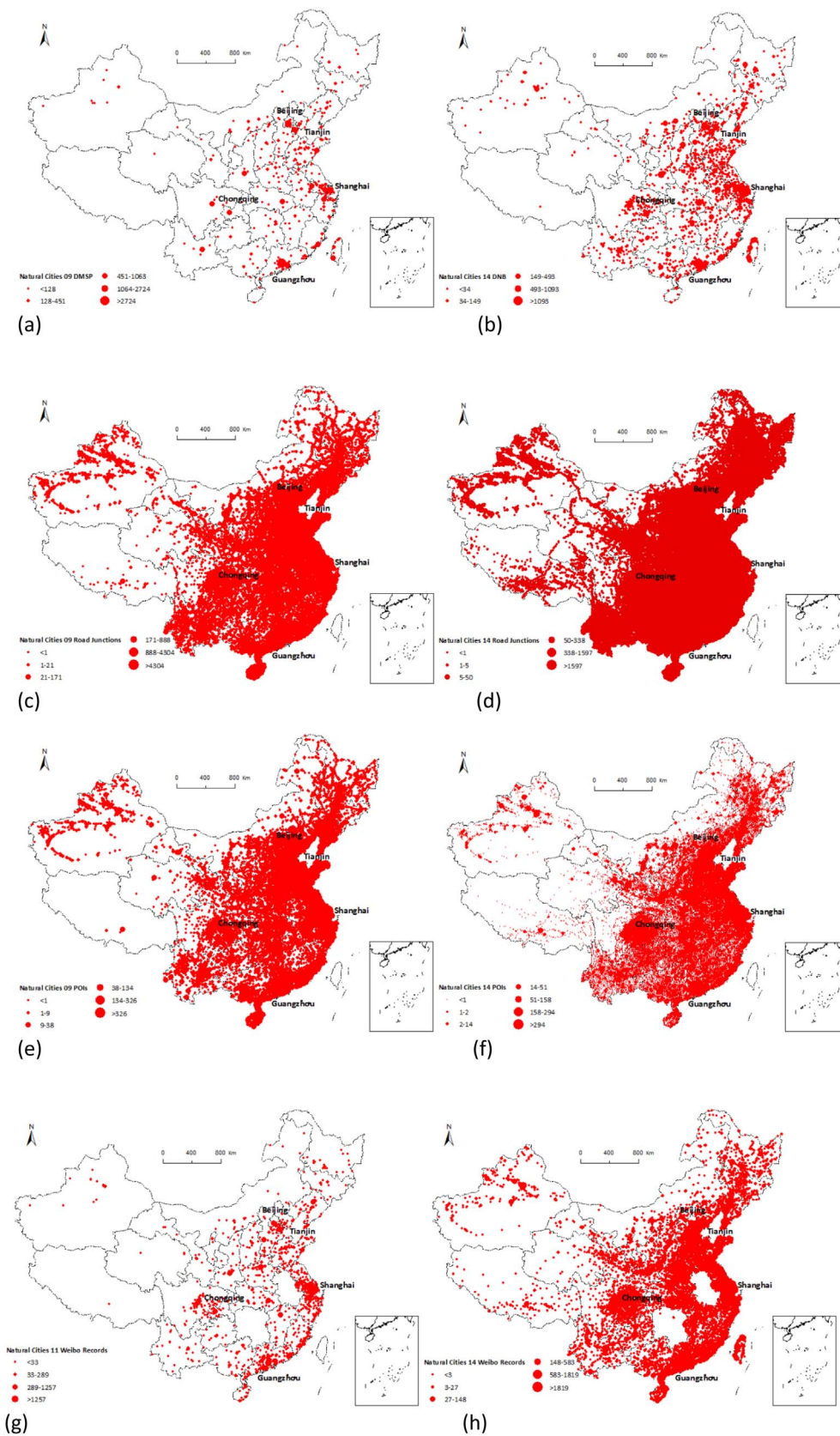
Fig. 6. Four Types of Natural cities.
Note: The NCs are derived from DMSP in 2009(a) and DNB in 2014(b), from road junctions in 2009(c) and 2014(d), from POIs in 2009(e) and 2014(f), and from Weibo records in 2011(g) and 2014(h)

Table 1
Official cities (OCs) and derived natural cities (NCs).

	Average city size in km^2	Maximum size	Total urban area in km^2	Standard deviation in km^2	# cities	Cut Distance in meters	Ht-index
2009 OCs (yearbooks)	63.3	2429.1	41,094.6	140.2	649	N/A	N/A
2014 OCs (yearbooks)	75.7	2915.6	49,904.1	168.0	659	N/A	N/A
2009 OCs (DMSP)	65.9	2090.3	43,401.5	183.4	649	N/A	N/A
2014 OCs (DNB)	67.8	2235.5	44,707.4	178.7	659	N/A	N/A
2009 NCs (DMSP)	128.8	4725.4	41,191.6	320.7	326	N/A	5
2014 NCs (DNB)	34.5	3995.2	50,245.5	142.1	1458	N/A	5
2009 NCs (road junctions)	1.6	8439.1	102,017.1	57.0	65,707	880.8	6
2014 NCs (road junctions)	0.4	11,396.3	153,441.5	29.5	347,886	587.0	6
2009 NCs (POIs)	1.0	1263.6	30,817.2	13.2	29,674	963.0	6
2014 NCs (POIs)	0.2	724.4	29,332.8	4.3	138,470	395.3	7
2011 NCs (Weibo records)	33.6	9712.8	34,637.8	327.9	1031	4937.0	4
2014 NCs (Weibo records)	3.2	6743.8	47,676.2	62.6	14,741	1274.1	6

Note: The cut distance is the length threshold used to select edges for constructing NC. For example, the cut distance of the 8,496,103 edges of NCs from road junctions in 2009 is 880.8 m, which splits all the edges into two unbalanced parts: 1,679,315 in the head (19.8%) and 6,816,788 in the tail (80.2%). The edges in the head part are used to build NCs. Because the RS images are not vector files, the data processing step is slightly different the other types of data sources. From the original RS images, we extract different patches according to the intensity of each pixel. Next, we select the patches which are larger than the average patch size as the natural cities. In this paper, 40% is chosen as the breakdown threshold. We also conduct an experiment with 50% as threshold, and acquire similar results of Ht-index: 5 5 6 6 6 7 5 7.

urban in 2014 due to the adjustment of administrative boundaries. The spatial expansion rate is the ratio between expanded urban areas and urban areas of OCs or NCs in 2009. It is worth mentioning that there are some regions in NCs which existed in the year 2009 but disappeared in 2014. These areas are then not accounted for in urban expansion in space, but accounted for the size expansion calculation. We find that the spatial expansion rate of OCs has been largely underestimated when compared with the results of NCs. The four urban expansion rates in China are uneven, among which the one derived by road junctions is the most salient one, following by the one derived by Weibo records, much higher than that of functional NCs. This is consistent with the findings for urban expansion in size.

For further investigation upon the urban expansion in space, we detail our analysis for the 23 urban agglomerations in China (Fang & Yu, 2017; Liu et al., 2015; Long, Wu, & Mao, 2014). These 23 urban agglomerations account for 21.3% of the total land resources, 55.7% of the total population, 64.2% of the total non-agriculture population, and 79.7% of the gross domestic product (GDP) in China in 2010 (Fig. 8). The urban areas in 2009 and 2014, as well as the expanded urban areas of the 23 urban agglomerations are plotted in

Fig. 9, in which the urban agglomerations along the x axis follow the sequence 1–23 in Fig. 8. As revealed by Fig. 9, the urban expansion in space is a significant process across China in all the dimensions, both OCs and NCs. The rank of urban areas in 2014 in all dimensions is similar, indicating that the cities with larger urban areas in one dimension are often coming with larger areas in the other dimensions. In each dimension, there are great differences among the 23 urban agglomerations, no matter regarding the urban areas in 2009 and 2014 or the urban expansion area. What is more, the variations among different dimensions within each urban agglomeration are still obvious: the morphological and social NCs form a sharp increasing from 2009 to 2014, while the functional NCs are not as remarkable.

We pay more attention to the three most dominating urban agglomerations of China, Beijing-Tianjin-Hebei (BTH), Yangtze River Delta (YRD) and Pearl River Delta (PRD). The results are available in Fig. 10. We notice significant pattern distinctions among all four types of NCs in 2009, as well as obvious differences among the four spatial expansion rates, revealing the mismatches among urban developments in various aspects. This provides a new lens for understanding urbanization, especially urban expansion, in China, in addition to the
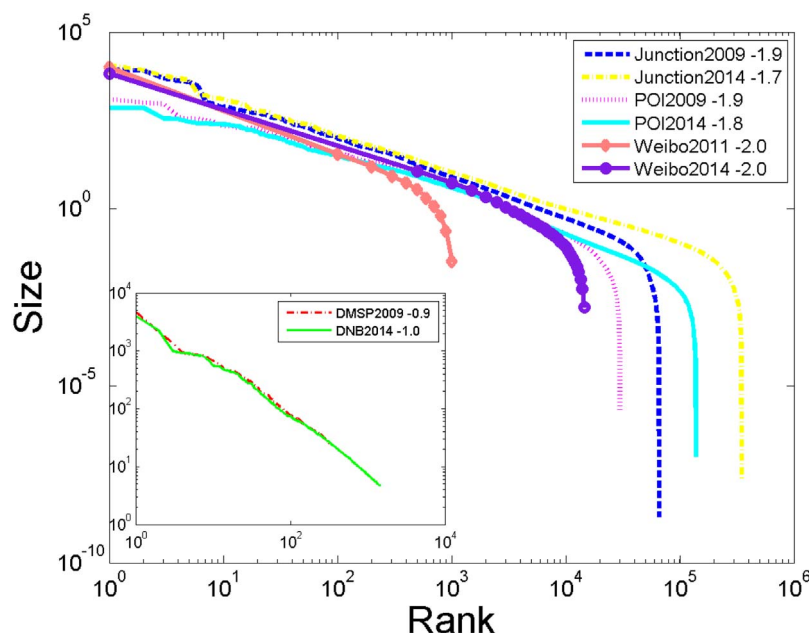

Fig. 7. Rank-size distributions of four types of natural cities.
Note: The coefficients of regressions between log(size) and log(rank) are indicated at the end of the legends.

Table 2
Descriptive information on evolving official cities (OCs) and natural cities (NCs).

	Official Cities		Natural Cities			
	Yearbooks	RS images	RS images	Road junctions	POIs	Weibo records
Average city size in 2009(km^2)	63.3	65.9	128.8	1.6	1.0	33.6
Average city size in 2014(km^2)	75.7	67.8	34.5	0.4	0.2	3.2
Urban expansion in size (km^2)	8809.5	1305.9	9,048.9	51,424.4	−1484.3	13,038.4
Size expansion rate (%)	21.4	3.0	22.0	50.4	−4.8	37.6
Density in 2009 (km^{-2})	N/A	64.7, 143.0, 12.6	N/A	44.3	245.8	15.5
Density in 2014 (km^{-2})	N/A	62.4, 136.6,11.9	N/A	49.3	330.4	17.1

Note: Urban expansion in size shows the difference between urban area sizes in 2014 and 2009. Size expansion rate is the ratio between increased urban areas and the urban areas in 2009. For Weibo records, the data are in 2011 and 2014.

Table 3
Urban expansion in space for OCs and NCs.

	Official Cities		Natural Cities			
	Yearbooks	RS images	RS images	Road junctions	POIs	Weibo records
Expanded urban areas (km^2)	8809.5	11,869.5	18,392.2	71,363.1	9745.2	21,459.0
Spatial expansion rate (%)	21.4	27.4	43.8	70.0	31.6	62.0

Note: We borrow the results of increased total urban areas and increased rate in Table 1 for the yearbooks derived OCs.
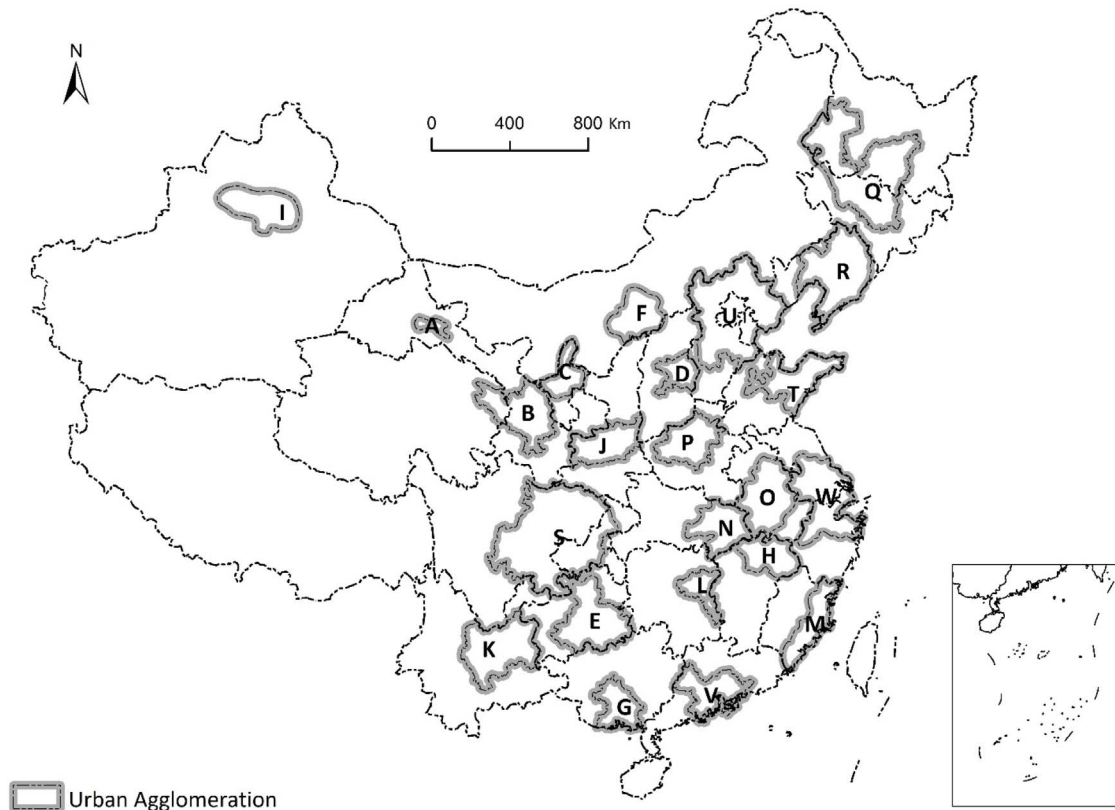

Fig. 8. The 23 urban agglomerations in China.
Note: The urban agglomerations are 1. Jiuquan-Jiayuguan-Yumen, 2. Lanzhou–Xining, 3. Northern Ningxia, 4. Central Shanxi, 5. Central Guizhou, 6. Central Inner Mongolia, 7. Southern Guangxi, 8. Northern Jiangxi, 9. Tianshan Mountains, 10. Guanzhong Plain, 11. Central Guizhou, 12. Eastern Hunan, 13. Eastern Fujian, 14. Eastern Hubei, 15. Central Anhui, 16. Central Plain, 17. Harbin–Changchun, 18. Liaodong Peninsula, 19. Chengdu–Chongqing, 20. Shandong Peninsula, 21. Beijing-Tianjin-Hebei, 22. Pearl River Delta, 23. Yangtze River Delta.

conventional aspect from yearbooks. There are also significant characteristics among the three typical urban agglomerations. In BTH, it expands greatly on the morphological and social dimension, while its functional development does not match. The land use policies and the rigid demand to working and living in Beijing primarily contribute to this irreversible expanding progress. Though Beijing and Tianjin have achieved a comparatively high urbanization level within a large scale,

the expansion tendency is evidently intensive. Driven by the economic benefit, the local officials raise the revenue by transferring land use rights to developers, bringing about the substantially high housing prices and considerable expansion in Beijing. What is more, accounting for the transportation cost (monetary cost and time cost), the residents would prefer the land near the center as expected, and the center outwards pattern is derived. In YRD, the functional expansion rate is the
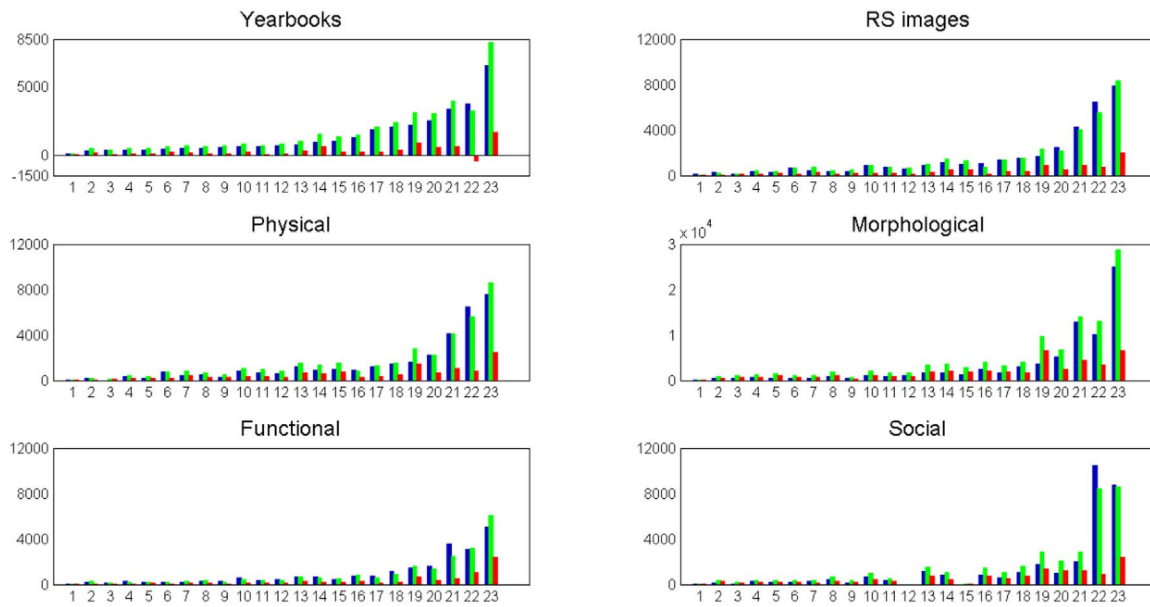
Fig. 9. Magnitude evaluation of expanded urban areas of 23 urban agglomerations for OCs (the first two subfigures in the first row) and NCs (the other three subfigures).
Note: The bars in blue, green and red represent the urban area sizes in 2009(2011), the urban area sizes in 2014 and the expanded urban area sizes. The y-axis stands for the urban areas in km^2. The expanded urban areas in the first subfigure is calculated as the increased urban areas in yearbooks from 2009 to 2014. Initially we used the same classes/ranges on the y axis on Fig. 9; while the 6 subfigures differ greatly and the universal classes/ranges make it awkward. Therefore, we do not use the same class/ranges on the y axis.

largest percentage, indicating a considerable human activity evolvement. As to the PRD, the NC from remote sensing images has the lowest expansion level, while the morphological and functional expansion form a coordinated development.

4.3.2. The quality of urban expansion

In addition to the magnitude evaluation on urban expansion in space, we measure the quality of expansion with the relative maturity indicator, which is calculated as the ratio between the density of the expanded urban areas and that of the urban areas in 2009, as shown in

Table 4. The spatial expansion density of NCs is 42.5 km^{-2} and 39.1 km^{-2} for road junctions and POIs. The results for the remote sensing images derived OCs are calculated as well, and the density is much lower than that of NCs. In this regard, the quality of urban expansion from the view of NCs is good and approaching the quality of existing urban areas in 2009. However, the quality examination on urban expansion from the view of OCs indicates a very low level. This gap arises from the definition of a city e.g. a NC is generated based on the density of road junctions or POIs and those low density areas are excluded from the NC at the very derivation phrase. Nevertheless, the

Fig. 10. Urban expansion patterns of NCs in the three typical urban agglomerations.
Note: The red denotes expanded urban areas from 2009 to 2014 and the blue denotes existing urban areas in 2009. The percentage under each subfigure represents the spatial expansion rate in a urban agglomeration. For Weibo records, the data are in 2011 and 2014.

Table 4
Quality evaluation of urban expansion in space.

City Type	Official Cities			Natural Cities		
	Road junctions	POIs	Weibo records	Road junctions	POIs	Weibo records
Density of 2009 (km^{-2})	64.7	143.0	12.6	44.3	245.8	15.5
Spatial expansion density (km^{-2})	29.7	39.1	2.9	42.5	225.7	13.4
Relative maturity (%)	45.9	27.3	23.2	95.9	91.8	86.4

Note: For Weibo records, the data are in 2011 and 2014.

findings on the quality evaluation of Chinese urban expansion have provided a picture of understanding urban expansion from a new perspective.

We further examine the relative maturity of three typical urban agglomerations BTH, YRD and PRD for NCs and OCs, as indicated in Table 5. The relative maturity of BTH is 77.1% for morphological NCs and 140.1% for functional NCs. The conditions are different in the other two urban agglomerations. The morphological relative maturity indicators in YRD and PRD are 47.2% and 23.9%, much lower than the national average. The functional relative maturity indicators in YRD and PRD are 65.1% and 63.5%, also lower than the national average. The social relative maturities are consistent with the functional one, and for three regions they are 53.2%, 75.0% and 140.8%. The low relative maturities on the three dimensions indicate two facts: Primarily, there are obvious distinctions between the highly developed central regions and the undeveloped marginal regions, especially within the large agglomerations. Secondly, the development of these urban agglomerations is approaching saturated; in other words, there seems to be a limitation for the urban expansion. The relative maturity examination for OCs is also scrutinized for benchmarking. The morphological relative maturity indicators of BTH, YRD and PRD are 39.2%, 46.1% and 62.6%; while in functional dimension, they are 17.9%, 25.5% and 44.4% and 11.4%, 14.4% and 44.8% in social dimension. This finding from the OC view indicates that the quality of urban expansion in BTH is the worst, in contrast to PRD and YRD. The findings are not supported by the examination on the quality of urban expansion from the NC view, which again arises from the definition of a city. That is, some new developed areas of YRD are not regarded as urban areas in the derived NCs, due to low densities of both road junctions and POIs.

5. Discussion

5.1. Contributions

In this paper, we conduct an empirical study adopting NCs in China. This novel method comes from a completely fractal viewpoint and helps acquire the timely information concerning the dynamics, effects of uneven urban expansion in China, and is also of the essence for policy making. The study contributes to existing urban expansion studies from the following aspects in detail. **First**, we apply the ideas of NCs to the case of urban expansion in China. The lightweight and straightforward methodology can facilitate timely urban expansion monitoring and evaluation for a large geographical area in the light of ubiquitous open data. The objectives and outcomes accurately reveal the differentiation of development, and prevent the influence of political and economic factors. The number of NCs from different open data varies, but it seems to have no obvious relation to the number of input data or the chosen cut distance. For example, the number of POIs in 2014 is almost ten times as large as that in 2009 and the cut distance shrinks to 40%, while the total functional NC urban area decreases slightly, because of the urban functional agglomeration development mode. **Second**, we re-evaluate the urban system and five-year evolution of NCs which is well reflected in this study. As the fractal perspective is a different way from the traditional Euclidean method, the findings come from different viewpoints. The rank size, ht-index, average size, and density are considered as measurements to evaluate the complexity and integrity of all cities in China, thus providing an overall review of the redefined Chinese city system. The obtained NCs suggest that there is a quantity of regions with high-level development but not defined as a city in the bureau. The increase of ht-index for NCs from POIs suggest that functional expansion might need the support of its complexity, so the POIs are getting more and more clustered and complex with a relatively lower degree of expansion. This is a total different expansion process. **Third,** we investigate the urban expansion magnitude and pattern in China. Road junctions represent the city's comprehensive transportation hub. Therefore, a large expansion rate of road junctions demonstrates a high degree of expansion on the morphological dimension. The POIs represent the urban functional developments, which obviously lag behind the morphological expansion in both space and quality. The social development is also obvious in both space and quality, which may arise from the phenomenon that Sina Weibo was lately created and the users proliferated. The results from NC method have proven the uneven urban expansions in China in different dimensions. The phenomena exist nationwide and locally. The OC is a top-down definition from the government, which sometimes has policy guidance functions and usually fall lag the timely urban development. Therefore, the OC is supposed to be a biased and incomplete way of defining urban areas. In contrast, NC is a bottom-up method which is constructed on the basis of open data, reflecting the accurate urban city shape and urbanization progresses, and it is able to suggest regions which could not be identified by OC methods. Considering the requirements of cartography and also for simplicity, the OCs tend to omit some small patches of urban areas and aggregate the urban areas to be an integral entity, while the NCs artlessly demonstrate the natural urban forms; therefore, NCs outnumber the OCs in each year. The distinctions between NCs and OCs on the magnitude and quality evaluation with respect to the Chinese urban expansion may arise from the definition of a city, demonstrating that the proposed bottom-up NC method contributes to our comprehensive understanding of China's urban expansion from a new perspective.

5.2. Limitations and future research

This study focuses on the uneven urban expansion in China from an

Table 5
Relative maturity of three urban agglomerations.

Urban agglomeration	Official Cities			Natural Cities		
	morphological	functional	social	morphological	functional	social
BTH	39.2%	17.9%	11.4%	77.1%	140.1%	53.2%
YRD	46.1%	25.5%	14.4%	47.2%	65.1%	75.0%
PRD	62.6%	44.4%	44.8%	23.9%	63.5%	140.8%

alternative perspective, and the adoption of emerging open data for natural city construction contributes to the study. Nevertheless, several limitations exist in the current study which should be highlighted. **First**, NCs in 2014 do not entirely cover NCs of 2009, so there are several regions which are considered as natural cities in 2009 but not in 2014. This phenomenon is incompatible with the urbanization process in China. **Second**, the complicated urban expansion is supposed to be more than four aspects aforementioned in our research, and more open data sources are demanded to feed our proposed framework. **Third**, the inconsistency and incomparability of different open data sources may give rise to the disagreement or incongruity of the application of this proposed framework which is hard to avoid so far. **Last but not least,** unrepresentativeness and occasional inaccessibility of the open data may contribute to the bias of the conclusion. More comprehensive data acquisition methods are supposed to reduce the inadequacy and impartiality to the utmost.

6. Conclusions

Research on China's urbanization has unfolded a complex landscape of uneven urban development (Ye & Xie, 2012). There has been a renewed interest in spatial inequality of urbanization, fueled by the concern over the effects of globalization and liberalization, and facilitated by theoretical and methodological developments in geography, economics, and computational social science. As Wei (2015) argues, "the popularity of social media provides new tools to study the nature of human behavior and spatial interaction. Social media has also produced new space of inequality, which has yet to be fully understood" (p7). This paper demonstrates such an implementation from open data perspectives including social media data.

In this paper, we examine urban expansion in China based on the theory of natural cities (NCs), by using remote sensing images (physical dimension), road junctions (morphological dimension), POIs (functional dimension) and Weibo records (social dimension) over time, facilitating the understanding of uneven urban expansion in China. From the view of NCs, we explore urban expansion in size and in space. The findings on four dimensions of NCs are also benchmarked with yearbooks and remote sensing images derived from official cities (OCs) to demonstrate the applicability of our proposed framework. The research on Chinese cities heavily relied on statistical yearbooks and reports, prior to the emergence of open data. The aggregated information corresponds to administrative cities, rather at the intra-urban level. As we have indicated in the introduction section of the paper, the administrative cities in China are not perfect for understanding the Chinese city system, since some cities are not included in administrative cities and some large administrative cities are not real cities from the spatial entity dimension. The accuracy of the official information is also not always reliable. The open data used in this study do help in facilitating us understanding a more comprehensive and objective Chinese city system from various dimensions like morphology, function and activity. The open data make it possible to derive and evaluate NCs, and characterize the magnitude and pattern of their temporal evolution. It can facilitate the next-step research on identifying distinct models of urban development and the significant role of driving forces (Li, Ye, Lee, Gong, & Qin, 2016).

The findings of this paper are as follows. First, The NCs outnumber the OCs in each year, further confirming the setting problems of OCs. The total urban areas of the four types of NCs derived significantly increase from 2009 to 2014, especially for the morphological and social ones. Second, within each type of NC, we could observe a tendency of homogenization and fragmentation. The heavy tailed distribution for the sizes of NCs versus their ranks has been confirmed, indicating that there are far more small NCs than large ones. Third, while defining size expansion rate as the ratio between increased urban areas and the urban areas in 2009, we find that in contrast to the official rate from yearbooks, the transport infrastructure based urbanization in terms of

road junction development is faster than the official rate, but the urban function developments are again confirmed to be behind the official rate and morphological rate. Fourth, given density for points (road junctions, POIs and Weibo records), the NCs form an increasing trend while the OCs decrease from 2009 to 2014. Fifth, the spatial expansion rate is set as the ratio between expanded urban areas and urban areas of OCs or NCs in 2009 for magnitude evaluation. Among the four expansion rates of NCs, the one derived by road junctions is the most significant, which is consistent with the urban expansion in size findings. The expansion derived by Weibo records is also noticeable in size and quality which may result from the sharp increase of Sina Weibo users from 2011 to 2014. Last, we measure the quality of expansion with the relative maturity indicator, which is calculated as the ratio between the density of expanded urban areas and that of the urban areas in 2009. We find a large gap between the relative maturity of OCs and NCs which may arise from the city definition.

References

Anderson, G., & Ge, Y. (2004). Do economic reforms accelerate urban growth? The case of China. *Urban Studies, 41,* 2197–2210.

Auer, S., Bizer, C., Kobilarov, G., et al. (2007). *Dbpedia: A nucleus for a web of open data [M]//the semantic web.* Berlin, Heidelberg: Springer, 722–735.

Bai, X., Shi, P., & Liu, Y. (2014). Society: Realizing China's urban dream. *Nature, 509,* 158.

Batty, M., & Longley, P. A. (1986). The fractal simulation of urban structure. *Environment and Planning A, 18*(9), 1143–1179.

Batty, M. (1991). *Cities as fractals: Simulating growth and form. Fractals and chaos,* . Springer43–69.

Batty, M. (2008). The size, scale, and shape of cities. *Science, 319,* 769–771.

Chen, Y. G. (2008). *Fractal urban systems: Scaling, symmetry, and spatial complexity.* Beijing, China: Scientific Press.

Chen, Y. (2015). Power-law distributions based on exponential distributions: latent scaling, spurious Zipf's law, and fractal rabbits. *Fractals-Complex Geometry Patterns and Scaling In Nature And Society,* 1550009.

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *Siam Review, 51,* 661–703.

Deng, X., Huang, J., Rozelle, S., & Uchida, E. (2010). Economic growth and the expansion of urban land in China. *Urban Studies, 47,* 813–843.

Fang, C., & Yu, D. (2017). Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and Urban Planning, 162,* 126–136.

Gouyet, J.-F. (1996). *Physics and fractal structures.* Paris/New York: Masson Springer.

Gurin, J. (2014). *Open data now: The secret to hot startups, smart investing, savvy marketing, and fast innovation.* New York, NY: McGraw-Hill.

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

He, C., Liu, Z., Tian, J., & Ma, Q. (2014). Urban expansion dynamics and natural habitat loss in China: A multiscale landscape perspective. *Global Change Biology, 20,* 2886–2902.

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,* 215–229.

Jiang, B., & Miao, Y. (2015). The evolution of natural cities from the perspective of location-based social media. *The Professional Geographer, 67,* 295–306.

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

Jiang, B. (2013). Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution. *The Professional Geographer, 65,* 482–494.

Jiang, B. (2015a). Head/tail breaks for visualization of city structure and dynamics. *Cities, 43,* 69–77.

Jiang, B. (2015b). Wholeness as a hierarchical graph to capture the nature of space. *International Journal of Geographical Information Science, 29*(9), 1632–1648.

Jiang, B. (2015c). A city is a complex network. In M. W. Mehaffy, & C. Alexander (Eds.), *A city is not a tree: 50th anniversary edition.* Portland, OR: Sustasis Press.

Kuang, W., Liu, J., Zhang, Z., Lu, D., & Xiang, B. (2013). Spatiotemporal dynamics of impervious surface areas across China during the early 21st century. *Chinese Science Bulletin, 58,* 1691–1701.

Li, H., Wei, Y., Liao, F., & Huang, Z. (2015). Administrative hierarchy and urban land

expansion in transitional China. *Applied Geography, 56*, 177–186.

Li, S., Ye, X., Lee, J., Gong, J., & Qin, C. (2016). Spatiotemporal analysis of housing prices in China: A big data perspective. *Applied Spatial Analysis and Policy*, 1–13.

Liu, J., Liu, M., Zhuang, D., Zhang, Z., & Deng, X. (2003). Study on spatial pattern of land-use change in China during 1995–2000. *Science in China Series D: Earth Sciences, 46*, 373–384.

Liu, J., Liu, M., Tian, H., Zhuang, D., Zhang, Z., Zhang, W., et al. (2005). Spatial and temporal patterns of China's cropland during 1990-2000: An analysis based on Landsat TM data. *Remote Sensing of Environment, 98*, 442–456.

Liu, J., Zhang, Z., Xu, X., Kuang, W., Zhou, W., Zhang, S., et al. (2010). Spatial patterns and driving forces of land use change in China during the early 21st century. *Journal of Geographical Sciences, 20*, 483–494.

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

Liu, X., Song, Y., Wu, K., Wang, J., Li, D., & Long, Y. (2015). Understanding urban China with open data. *Cities*.

Liu, X., Derudder, B., & Wu, K. (2016). Measuring polycentric urban development in China: An intercity transportation network perspective. *Regional Studies, 50*(8), 1302–1315.

Long, Y., & Liu, L. (2015). *Big/open data in Chinese urban studies and planning: A review*. *ISOCARP Review 11*.

Long, Y., Wu, K., & Mao, Q. (2014). Simulating urban expansion in the parcel level for all Chinese cities, arXiv preprint arXiv, 1402.3718.

Long, Y. (2016). Redefining Chinese city system with emerging new data. *Applied Geography, 75*, 36–48.

Mandelbrot, B. B., & Blumen, A. (1989). Fractal geometry: What is it, and what does it do? *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences The Royal Society, 423*(1864), 3–16.

Mandelbrot, B. B. (1983). *The fractal geometry of nature*. Macmillan.

Marceau, D. J. (2005). Exploration of spatial scale sensitivity in geographic cellular automata. *Environment and Planning B: Planning and Design, 32*, 693–714.

Ministry of Housing and Urban-rural Development of the People's Republic of China (2013). *Chinese city construction statistics yearbook 2012*. Beijing: China Planning Press.

Nation Bureau of Statistics, P (1982). *China's main demographic indicators from the 3rd national population census*.

Nation Bureau of Statistics, P (2011). *China's main demographic indicators from the 6th national population census*.

Qi, Z. F., Ye, X. Y., Zhang, H., & Yu, Z. L. (2014). Land fragmentation and variation of ecosystem services in the context of rapid urbanization: the case of Taizhou city,

China. *Stochastic Environmental Research and Risk Assessment, 28*(4), 843–855.

Qiao, J., Lee, J., & Ye, X. (2016). Spatiotemporal evolution of specialized villages and rural development: A case study of Henan province, China. *Annals of the American Association of Geographers, 106*(1), 57–75.

Reichman, O. J., Jones, M. B., & Schildhauer, M. P. (2011). Challenges and opportunities of open data in ecology. *Science, 331*.

Seto, K. C., Fragkias, M., Güneralp, B., & Reilly, M. K. (2011). A meta-analysis of global urban land expansion. *PloS One, 6*, e23777.

Shaw, S. L., Tsou, M. H., & Ye, X. (2016). Editorial: human dynamics in the mobile and big data era. *International Journal of Geographical Information Science, 30*(9), 1687–1693.

Van de Voorde, T., Jacquet, W., & Canters, F. (2011). Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. *Landscape and Urban Planning, 102*, 143–155.

Wallace, J. L. (2014). *Cities and stability: Urbanization, redistribution, and regime survival in China*. Oxford University Press.

Wan, L., Ye, X., Lee, J., Lu, X., Zheng, L., & Wu, K. (2015). Effects of urbanization on ecosystem service values in a mineral resource-based city. *Habitat International, 46*, 54–63.

Wang, Y., Wang, T., Ye, X., Zhu, J., & Lee, J. (2015). Using social media for emergency response and urban sustainability: A case study of the 2012 Beijing rainstorm. *Sustainability, 8*(1), 25.

Wang, Y., Jiang, W., Liu, S., Ye, X., & Wang, T. (2016). Evaluating trade areas using social media data with a Calibrated Huff Model. *ISPRS International Journal of Geo-Information, 5*(7), 112.

Wang, Z., Ye, X., & Tsou, M. H. (2016). Spatial, temporal, and content analysis of Twitter for wildfire hazards. *Natural Hazards, 83*(1), 523–540.

Wei, Y. D., & Ye, X. (2014a). Urbanization, land use, and sustainable development in China. *Stochastic Environmental Research and Risk Assessment, 28*(4), 755.

Wei, Y. D., & Ye, X. (2014b). Urbanization, urban land expansion and environmental change in China. *Stochastic Environmental Research and Risk Assessment, 28*(4), 757–765.

Wei, Y. D. (2015). Spatiality of regional inequality. *Applied Geography, 61*, 1–10.

Ye, X., & Xie, Y. (2012). Re-examination of Zipf's law and urban dynamic in China: A regional approach. *The Annals of Regional Science, 49*(1), 135–156.

Yue, W., Ye, X., Xu, J., Xu, L., & Lee, J. (2014). A brightness–darkness–greenness model for monitoring urban landscape evolution in a developing country? A case study of Shanghai. *Landscape and Urban Planning, 127*, 13–17.

Zhang, H., Qi, Z. F., Ye, X. Y., Cai, Y. B., Ma, W. C., & Chen, M. N. (2013). Analysis of land use/land cover change, population shift, and their effects on spatiotemporal patterns of urban heat islands in metropolitan Shanghai, China. *Applied Geography, 44*, 121–133.