Habitat International 59 (2017) 32–43

Contents lists available at ScienceDirect

Habitat International

journal homepage: www.elsevier.com/locate/habitatint

# Measuring urban agglomeration using a city-scale dasymetric population map: A study in the Pearl River Delta, China

Chunzhu Wei [a, *], Hannes Taubenböck [b], Thomas Blaschke [a]

[a] Department of Geoinformatics - Z_GIS, University of Salzburg, Schillerstrasse 30, 5020 Salzburg, Austria
[b] German Aerospace Center (DLR), German Remote Sensing Data Center, Oberpfaffenhofen, 82234 Weßling, Germany

ARTICLE INFO

*Article history:*
Received 15 July 2016
Received in revised form
23 September 2016
Accepted 11 November 2016

*Keywords:*
Dasymetric population map
City scale
Megacities
Limiting variable algorithm
Urban agglomeration
Connectivity metrics
Pearl River Delta

ABSTRACT

The rates of urbanization and increase in urban sprawl that have occurred in China over the past thirty years have been unprecedented. This article presents a new city-scale dasymetric modelling approach that incorporates historical census data for 28 cities in the Pearl River Delta area of southern China. It combines Landsat imagery (from 2000, 2005, 2010, and 2015) with a 'limiting variable' estimation algorithm to generate a gridded estimate of population density. These gridded population patches are organized as a city-network to reveal the influence of urban agglomeration on population spreading processes. We then combine population patches and graph-based connectivity metrics to describe the spatial-temporal evolution of each city within the urban agglomeration. Our population disaggregation results yield accuracy improvements of 40%–60% over three traditional population disaggregation methods, to reflect the population distribution characteristics more explicitly and in greater detail. The probability of connectivity metrics from dasymetric population maps in Pearl River Delta (1) outline the role of urban agglomeration in population spread, (2) simulate the evolution of 'polycentric' urban agglomeration, and (3) outline the individual components of the polycentric megaregion. Our outlined approach is a transferable and an improved means of producing city-scale dasymetric population maps. Our case study provides practical guidance on wide applications of the medium resolution remote sensing data in delineating, measuring, and quantifying the evolution of urban agglomeration across different jurisdictional boundaries and time periods.



## 1. Introduction

Urbanization has become a worldwide phenomenon over recent decades (Taubenböck, Wegmann, Roth, Mehl, and Dech, 2009). Economic activities and services, transportation development, and traffic flow all have profound implications for international networks of cities. Cities are often no longer isolated but increasingly concentrated and inextricably linked together in the evolutionary term-'megaregion', sharing infrastructure systems, environmental systems, economic linkages, land use patterns and culture (Robinson, 2006; Ross & Woo, 2011). This phenomenon is known as urban agglomeration (Yue, Zhang, & Liu, 2016; Zhou, Xu, Wang, & Lin, 2015). Urban agglomeration is generally characterized by the size of the territory associated with continuity between separate urbanized areas, contiguous economic and social relationships, and a population concentration (He et al., 2016; Lang, Chen, & Li, 2016; Listengurt, 1975). Nevertheless, urban agglomeration remains a diffuse and elusive concept and there is no general agreement on what agglomeration means, how it can be recognized, or how to delineate the spatially contiguous regions (Frankhauser, 1998; Glaeser, 2008).

Commonly used approaches to delineating urban agglomeration are mainly based on subjective perceptions of the growth rates for different forms of land use, on socio-economic aspects of specific areas (Poyil and Misra, 2015; Salvati, 2014), on quantifications of urban landscape configurations and estimates of the structure characteristic of each urban form through spatial metrics (Taubenböck and Wiesner, 2015), or on accessibility as defined by a variety of transportation models (Kim and Han, 2016). Studies of urban agglomeration also generally take into account population densities. Studying population distributions has been shown to be useful for urban demographic and geographic investigations, urban planning, and environmental protection, as well as for other applications (Brennan, 1999). Researchers have studied relationships

* Corresponding author.
*E-mail address:* weichunzhuzhu@gmail.com (C. Wei).

http://dx.doi.org/10.1016/j.habitatint.2016.11.007

between population distribution characteristics and the evolution of urban agglomerations (Schleicher, Biedermann, and Kleyer, 2011; Urban, Minor, Treml, and Schick, 2009; Plowright et al., 2011; Schumaker, 1996). However, the current understanding of the different effects of clustered populations versus dispersed populations is limited. Likewise, we know little about the effects of monocentric habitats versus polycentric habitats on urban agglomeration, as well as how urban agglomeration processes influence population spatial distributions along the gradient of decreasing population density from an urban center to its periphery (Arthur and McNicoll 1975; Wilson et al., 2001).

To overcome this deficiency it may be helpful to use real-world spatial and temporal dasymetric population models to disaggregate population distributions into multi-scale spatial population density patches (Foltête and Giraudoux 2012). These population density patches can provide a spatial framework within which to elucidate and spatially quantify the evolution of urban agglomeration through graph-based connectivity metrics (Foltête and Giraudoux 2012; Saura and Pascual-Hortal 2007). Moreover, a series of graph-based connectivity metrics, like the probability of connectivity metrics (Saura and Pascual-Hortal, 2007), the landscape coincidence probability metric (Pascual-Hortal and Saura, 2006), etc., have led to an increasing interest in considering connectivity for urban planning purpose (Nazara and Hewings 2003). With the advantages of measuring the connectivity, resilience and competition of landscape patches in the network, these well-applied graph-based connectivity metrics can also provide a valuable way of incorporating the spatial structure of spatial population density patches into an urban agglomeration analysis (Vaz, Zhao, and Cusimano, 2016). The challenge is therefore to establish a meaningful and useful spatial dasymetric population model with suitable scale and to add quantitative information that will help to identify the evolution of spatial population patches under the urban agglomeration.

With the development of Remote Sensing (RS) and Geo-informatics Science (GIS), the acquisition ability of population dasymetric maps derived by the integration of multi-disciplinary data, namely global remote sensing, human settlement and socio-economic has greatly improved (Wu, Qiu, and Wang 2005; Langford and Unwin, 2013). Previously well-cited coarse-scale (1 km–100 m) population maps include, for example, the Gridded Population of the World (GPW) method (Deichmann, Balk, and Yetman, 2001), the Landscan method (Dobson, Bright, Coleman, Durfee, and Worley, 2000), WorldPop (Stevens, Gaughan, Linard, and Tatem, 2005), amongst others. These methods establish a co-relationship between mean population densities and RS/GIS-based population distribution information (including land use types, DEMs, transportation, night-time images, various landmarks, slope) to disaggregate the population from province-scale or national-scale administrative unit into each cell of the (satellite) Geodata (L.Imhoff, Lawrence, Stutzer, & Elvidge, 1997; Zeng, Zhou, Wang, Yan, & Zhao, 2011). They are able to accurately express the inner cities' divergence within each country. However, most of these models are constrained by the coarse resolution of remotes sensing data, making the generation of city-scale (taking the city administrate boundary but not the province/national administrate boundary as the specific areal unit) dasymetric population maps a challenging research topic.

The Limiting Variable (LV) method, therefore, has been proposed by Martin (1996) and Gallego, Batista, Rocha, and Mubareka (2011) to integrate the pycnophylactic method into the population disaggregation. This method starts with a homogeneous population density in each initial zone, which is modified by applying upper limits to the less populated land-cover classes and redistributing the excess population to the more populated classes. This means that the LV method, when combined with reliable census data and finer resolution of satellite imagery, can evaluate the population density at a variety of regional scales (Mennis, 2003; Gallego et al., 2011). Therefore, the LV method makes it become possible to integrate medium-resolution of imageries with the city-scale administrate boundary into disaggregating the population density.

In summary, the sheer magnitude of population growth is an important factor affecting the evolution of urban agglomeration. It has a direct effect on the spatial concentration of urban agglomeration, as well as other causes of environmental stress (Tan et al. 2008). We therefore propose that city-scale dasymetric population maps are one crucial approach suited for the identification and delineation of spatial population characteristics, and for tracing its spatial evolution. For our investigations we chose an area covering the Pearl River Delta (PRD) megaregion in southern China and adopted a combined form of Landsat data (in 2000, 2005, 2010 and 2015) with the LV algorithm to delineate the city-scale dasymetric population maps. Then we integrate the population density patches with a series of graph-based connectivity metrics, to address the following question:

- How can city-scale dasymetric population maps delineate the city-network of the PRD megaregion and the spatial and temporal evolution of its urban agglomeration?

## 2. Methods

### 2.1. Case study

The PRD is one of the most densely urbanized regions in the world, and one of the most populous, rapidly commercialized and urbanized economic regions in China (shown in Fig. 1). The average annual precipitation in PRD is over 1500 mm, with an average annual temperature of 23 °C. The humid subtropical climate, fertile alluvial soils, and a water system good for year-round irrigation and transportation in PRD have supported more than 56 million people. According to the World Bank Group (2015), the PRD have become

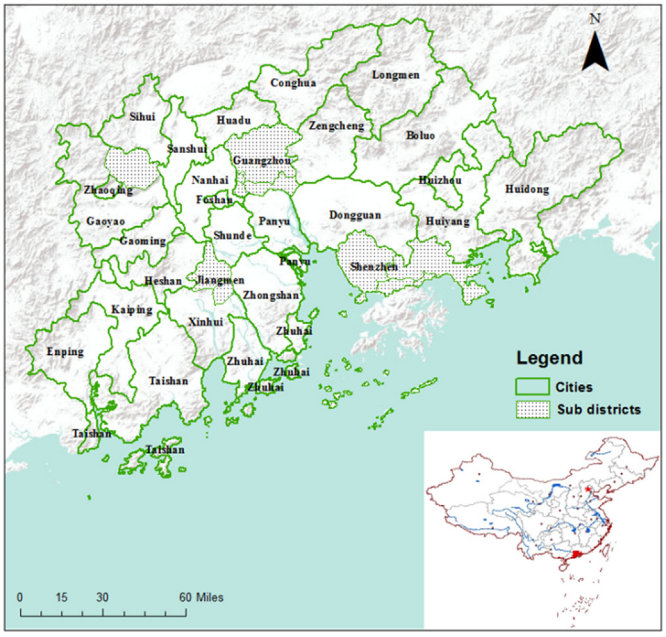


**Fig. 1.** Study area-the Pearl River Delta megaregion area.

the largest megaregion in the world in terms of both surface area and population. Long-term monitoring by Taubenböck, Wegmann, Roth, Mehl, and Dech (2014) of the spatial growth rates of settlement areas revealed that its spatial extent in 2011 was 13.14 times greater than in 1975, making this megaregion among the most dynamic areas in the world, even outperforming the two-dimensional spatial growth rates of China's other megacities such as Shanghai (ca. 6 times) or Beijing (7.5 times) (Taubenböck and Wiesner, 2015). Our area of investigation extends from 21°N to 24°N, and from 112°E to 115°E.

### 2.2. Data collection

The imagery used for the land-use and land-cover (LULC) classification was obtained from Landsat Surface Reflection products (L4-5 TM: 2000, 2005, 2010 and L8 OLI/TIRS: 2015). The PRD covers eight Landsat images, and the Landsat data used in our study were mainly collected during July and October (as shown in Table 1). We determined this time period through analysis of images free from cloud containment. Atmospheric and geometric corrections were applied to the Landsat Surface Reflection products to facilitate the comparison between Landsat imageries over space and time to support land surface change studies. Although some variations due to satellite drift/changeover, as well as incomplete corrections for calibration loss and atmospheric effects (clouds, aerosols, etc.) may still exist, here we assume that such influences are smaller than those caused by environmental drivers and are negligible.

Since PRD has subtropical climate and the typical vegetation includes subtropical evergreen broad-leaved forest, broadleaf shrubs and woodlands (Piao et al., 2003), these vegetation areas do not have dramatic changes during the period of July to October (Tewari, Guenther and Wiedinmyer 2009), and there are less seasonal variations NDVI in this area.

Coarse resolution LULC images (1 km) in 2000, 2005 and 2010 provided by Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC) (http://www.resdc.cn) were used as ancillary data for the fine resolution LULC classification. High resolution images data (<10 m) which were obtained online from http://www.terraserver.com/, and urban footprint data with a spatial resolution of 30 m provided by the German Aerospace Center (from Taubenböck & Wiesner, 2015), were both used as reference data to evaluate the LULC accuracy for the PRD. 30 m SRTM digital elevation data were used to integrate with the LULC data to estimate the accessibility of different cities in PRD.

The original population census data are from the Guangdong Statistical Yearbooks for 2000, 2005, 2010, and 2014 (http://www.gdstats.gov.cn/tjsj/gdtjnj/). We use census data from these four years, which are available for 28 original city-scale administrative units, to estimate the population density. A further 10 sub-regional units of population census data for 2000 and 2010, which were provided by Population Statistical center of Guangzhou Province, are used for validation purposes. The reference population density information for population disaggregation was provided by Zeng et al. (2011).

### 2.3. Models

#### 2.3.1. Framework

We propose a hierarchical framework (Fig. 2) to first establish a city-scale dasymetric population model, based on a combination of Landsat TM/8 data and population census data. Dasymetric population maps for the four reference years are used to analyze the urban agglomeration network and its individual components using graph-based connectivity metrics.

#### 2.3.2. Land use and land cover classification

The limiting variable method, as one of the most popular population downscaling methods, has been tested in Europe (Gallego et al., 2011) and Australia (Li, Pullar, Corcoran, & Stimson, 2007). It yielded better results when compared with the binary method (Langford & Unwin, 2013) and the three-class method (Eicher and Brewer 2001). Because the limiting variable method requires the reference population density coefficients for different land use classes as input, it has rarely been tested at city-scale of Chinese administrative units. In our work, we used as reference input data sets the coarse resolution (1 km) population density coefficients for three land use classes (urban areas, rural areas and vegetation areas (excluding forest)) in China provided by Zeng et al. (2011). We therefore integrated the limiting variable model with the above mentioned population density coefficients to assign different population densities to the three land use classes in PRD, to further validate the transferability of reference input data sets to our study area in population disaggregation analysis with the limiting variable model.

We first classified the Landsat images into LULC maps using an Object Based Image Analysis (OBIA) method and then disaggregate census data from the whole PRD megaregion into a system of smaller city-scale zones with the help of the LULC maps. A review of the advantages of OBIA approaches, such as (1) image-objects

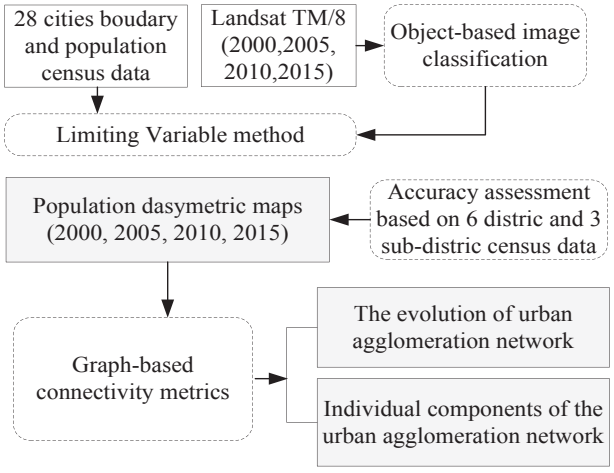


**Fig. 2.** Flowchart of urban agglomeration analyse is based on dasymetric population map. White boxes represent the data collection, white boxes with dashed borders represent the specific methods and gray boxes represent the results.

**Table 1**
Data collection of Landsat 4–5/8 from 2000 to 2015.

| Year | (Path, Row) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (121,44) | (121,45) | (122,43) | (122,44) | (122,45) | (123,43) | (123,44) | (123,45) |
| 2000 | Sep 15 | Sep 15 | Aug 21 | Aug 21 | Aug 21 | Oct 08 | Oct 08 | Spe13 |
| 2005 | Aug 12 | Aug 12 | Jul 18 | Jul 18 | Jul 18 | Sep 11 | Sep 11 | Sep 11 |
| 2010 | Oct 29 | Oct 29 | Sep 18 | Sep 18 | Sep 18 | Jul 07 | Jul 07 | Jul 07 |
| 2015 | Aug 08 | Aug 08 | Oct 18 | Oct 18 | Oct 18 | Aug 06 | Aug 06 | Aug 22 |

exhibit useful features (e.g. shape, texture), and (2) image-objects keep good correspondence with the phenomenon, can be found in Blaschke (2010). Several studies have already referred to object-based methods for classification of Landsat TM images as an alternative to pixel-based methods (Dorren, Maier, and Seijmonsbergen, 2003; Zhu and Woodcock 2012). We used both multi-level image segmentation and fuzzy rules in our OBIA process. One of the main advantages of multi-level segmentation is its ability to combine spatial, textual and contextual information between the values of proximate pixels to divide the satellite image into spatially continuous and homogeneous objects. This segmentation process provides the possibility for patch-based population disaggregation by reducing the local spectral variation of objects and avoids the pixel-based population disaggregation which may be influenced by the "salt and pepper" effect based on traditional pixel-based methods. For a fuzzy rule-based classification, it takes into account the uncertainty in sensor measurements, and consists of an n-dimensional tuple of memberships to describe the degree of class assignments. Hence, fuzzy classification is well suited to handle most sources of vagueness in remote sensing information extraction. Many previous studies have also proved that fuzzy classification is not only a time efficient process, but also an approach with high classification accuracy and reliability (Benz, Hofmann, Willhauck, Lingenfelder, & Heynen, 2004; Hofman, Blaschke, and Strobl 2011; Baraldi, Puzzolo, Blonda, Bruzzone, & Tarantino, 2006).

Our multi-level image segmentation and fuzzy rules prototype were designed on the basis of the eCognition Developer 9.1 software for Landsat TM+/8 data and distinguishes five land-cover classes: water, urban areas, rural areas, vegetation areas (excluding forest), and forest. Among various land use types, the water and forest were assumed to have no residents. While using the land use data, these two land cover types were just treated as background and set to zero in the imageries.

According to the Land Use Coding Standard provided by RESDC (http://www.resdc.cn), urban area is defined as the large, medium and small cities and towns over the construction lands. Rural area is defined as the rural settlements independent of the towns and cities. In our study, we defined all the buildings, road, all constructions of infrastructure and other artificially sealed or paved areas in the cities and towns as the urban areas. Meanwhile, we defined the scattered agricultural habitation (agricultural buildings and shelters which are far away from the urban areas) or scattered main and secondary residences in natural or agricultural areas as rural areas, and the vegetation and bare soil were excluded. The coarse resolution (1 km) LULC data provided by RESDC has helped us to get the primary visual interpretation of the urban and rural areas in PRD. Through combining the population statistical data (from Guangdong Statistical Yearbooks), the urban footprint maps (from Taubenböck and Wiesner, 2015) and high resolution images data (<10 m) (from http://www.terraserver.com/and Google earth), we treated the built-up areas which had merged into the urban corridor (from Georg, Blaschke and Taubenböck 2016) into urban areas. The scattered and small patches, which were far away from the megaregion, were treated as rural areas.

We first chose the multi-level image segmentation to do the automatic spatial pattern recognition of Landsat images. The segmentation scale (scale = 15) was determined by the local variance in ESP2 tool (Drăguţ, Csillik, Eisank, and Tiede, 2014). Then we divided our fuzzy rule-based land use classification procedure into three top-down strategies which gradually divide land use classes into more comprehensive classes. The SRTM 30 m digital elevation data are added as an extra band to provide the DEM information.

i. We simply used the ancillary data provided by RESDC to help generate coarse resolution (1 km) rural masks and urban masks, which served as the base for further classification.
ii. We classified the built-up areas using the parameters of the normalized difference vegetation index (NDVI, the fuzzy membership ranged from 0.24 to 0.38), the normalized difference water index (NDWI, the fuzzy membership ranged from 0.04 to 0.08), the specific leaf area vegetation index (SLAVI, the fuzzy membership ranged from 0.53 to 0.58), the land and water masks index (LWMI, the fuzzy membership ranged 105 to 200), and the normalized difference moisture index (NDMI, the fuzzy membership ranged from 0.03 to 0.05), where the membership functions for each property are combined by the fuzzy-logic operator 'fuzzy-and'. The integration areas between the built-up areas and the rural masks were assigned as the class of rural areas. We then merged the built-up areas and assigned the built-up areas with *'total areas >500'* as the temporal class of *urban corridor*. The built-up areas with *'distance to urban corridor 30'* pixel and *'total areas <300 pixels'* were assigned as the rural areas. The rest built-up areas and the urban corridor were all classified into urban areas.
iii. We classified the water class with the parameters of LWMI (the fuzzy membership ranged 45 to 120) and NDVI (the fuzzy membership ranged from 0 to 0.15). The forest class was classified by the parameters of DEM (the fuzzy membership ranged from 25 to 100) and NDVI (the fuzzy membership ranged from 0.45 to 0.72). The rest non-built up areas were categorized into the class of vegetation (*excluding forest and water*).

Our experiment shows that the fuzzy rules are transferable to the surface reflectance products of Landsat TM/8 from 2000 to 2015. Finally, we used the high resolution images (2005, 2010, and 2015) from google earth to modify our classification through visual interpretations. Given the resolution of satellite image and the characteristics of our LULC classification results, the minimum mapping size of rural areas is 0.36 ha (4 pixels), and the minimum mapping size of urban areas is 18 ha (20 pixels).

### 2.3.3. City-scale dasymetric population model

With reference to previous research in China using dasymetric population models based on remote sensing data and night time imagery (Zeng et al., 2011), this study also derives population density estimates for only three land-cover types, these being urban areas, rural areas, and vegetation areas (exclude the forest areas).

Population density ratios between the different land use classes are calculated in order to identify internal variations of population distributions within the PRD megaregion. To generate city-scale (taken the city administrate boundary as the specific areal unit) dasymetric population maps, we take 28 cities boundaries as the specific areal unit and apply the LV method (Eicher and Brewer, 2001; Gallego et al., 2011) to divide the census data for each city between the three LULC classes. The LV method, which enables the disaggregation of population density from an initial zone to multi-scale subzones with heterogeneous attributes, has been shown by Gallego et al. (2011) to yield better results than other fixed-ratio models in the European countries. It is therefore chosen as the first try to transfer population census data from the whole PRD megaregion to a smaller scale targeting individual cities, in order to investigate the heterogeneity of their population densities. The steps of the LV method has clearly described by Gallego et al. (2011), involving: (1) Attribution of a uniform population density to all the land−cover classes in each city, (2) Computation of the upper limits

of population density for each land cover class in each city, (3) Modification of the population density for each land cover class in each city using the uniform population density and the upper limits of population density, (4) Redistribution of any exceed population to all classes in proportion. As reference population density information, we use the data from Zeng et al. (2011), who provided 28 different population density categories for Chinese cities in their study.

After producing the dasymetric population maps, we use 6 district-based census data and 3 sub-district based census data to evaluate the population estimation accuracies by the Pearson coefficient (R), Root-Mean-Square Error (RMSE), and the Mean Absolute Error (MAE). Our city-scale population density results will also be compared with those obtained using the GPW (Deichmann, Balk, and Yetman, 2001), Cnpopulation from Data Center of Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC) (http://www.resdc.cn)), and WorldPop (Stevens, Gaughan, Linard, and Tatem, 2015) methodologies. Meanwhile, due to the insufficient information about the input datasets and modelling methods of LandScan methodology, we exclude it from the direct comparison.

#### 2.3.4. Probability of connectivity (PC) metric to investigate population spread processes within the urban agglomeration network

The evolution of urban agglomeration in developing countries has required considerable effort in devising strategies that will foster the growth of secondary cities and promote regional development (Brennan, 1999). In the absence of any empirical information on economic connections between the 28 cities such as transportation volumes, financial transactions, and so on, we decided to rely on regional density functions to describe how the population density varies within the cluster networks of the PRD megaregion. The Probability of Connectivity (PC) metric allows to provide quantitative descriptions of the speed and intensity of population spread for two types of analyses: (1) *Connectivity analysis*, to compute the spatial and temporal links between population density patches using spatial distance correlations within the city networks of the PRD megaregion, and (2) *Individual components analysis*, using the correlations between population data and the probability of connectivity metrics to quantify and describe the individual components of the whole PRD network structure.

*2.3.4.1. Connectivity.* All the inhabited patches in our research are defined as two-dimensional entities. Depending on how the links between these entities are defined, the first basic parameter of the links (the Probability of Connectivity (PC) metric) is used as the spatial metric for measuring the distance between patches. The PC metric, which was developed by Saura and Pascual-Hortal (2007), is defined as 'the probability that two animals randomly placed within the landscape fall into habitat areas that are reachable from each other'. This index provides a possible spatial framework within which to elucidate the spatial and temporal changes in accessibility of multiple cities within the PRD. It was calculated as follows:

$$PC = \frac{\sum_{i=1}^{n}\sum_{j}^{n} a_i a_j p_{ij}}{A^2}$$

Where n is the total number of patches, $a_i$ and $a_j$ are the cities i and j, respectively, $p_{ij} = e^{-kd_{ij}}$ is the accessibility between i and j are connected, and A is the total area of the PRD. Since the least-cost distance can well demonstrate the path of least resistance from one patch to another (Porter, Dueser, and Moncrief, 2015), we adopt the least-cost path as $d_{ij}$ to express the accessibility of a city within the PRD to an inhabitant from another. Setting the resistance value of each cell in the satellite data on the basis of the attributed land cover types is then crucial for calculating the connectivity. Setting these resistance values is a difficult process which makes use of expert judgement and the data available in published literature (see Walker and Craighead, 1997; Adriaensen et al., 2003). We therefore establish our resistance value rule on the basis of previous empirical least-cost distance modelling (Stucky et al., 1998; Singleton and Lehmkuhl, 2001; Adriaensen et al., 2003; Driezen, Adriaensen, Rondinini, Patrick Doncaster, and Matthysen, 2007; Damoiseaux and Greicius, 2009). In order to describe the heterogeneous landscape characteristics and the population capacities of each landscape element, we range the values from 1 (weak resistance, easily passable) to 10 (high resistance, passable only with difficulty) according to the information available on the different land-cover types (including urban areas, rural areas, farmland, forest, and water), also taking 'slope' into account. The slopes within the PRD area are derived from a DEM and reclassified from 1 to 10 on an evaluation scale. The slope accounted for 60% of the influence in the least-cost distance calculation (Stucky et al., 1998; Damoiseaux and Greicius, 2009), with the six land-cover classes accounted for the remaining 40% of the influence. A resistance value of 1 was assigned to urban areas, while landscape elements that are presumably only moderately favorable to movement (such as farmland) and rural areas were assigned resistance values of 3 and 5, respectively. Forest and water areas are considered difficult to traverse and were therefore assigned a resistance value of 10 (Singleton and Lehmkuhl, 2001; Adriaensen et al., 2003; Driezen et al., 2007).

*2.3.4.2. Evolution of urban agglomeration networks.* The incorporation of surrounding counties and statutory towns into core municipal metropolitan regions has increased in coastal China, and especially within the PRD region. From this perspective, identifying the core cities, their regions of influence, and how their roles gradually change within the megaregion becomes crucial to governance. We therefore adopt the percentage of the variation in PC (*dPC*) to further describe the prioritization and ranking of each city by their contribution to the whole megaregion.

$$dPC = 100 \times \frac{PC - PC_i}{PC}$$

The *dPC* value can be further partitioned into three separate fractions (*dPCintra, dPCflux and dPCconnector*) to quantify the different ways in which individual cities can contribute to the evolution of urban agglomeration and availability in the landscape (Saura and Rubio, 2010):

$$dPC = dPCintra + dPCflux + dPCconnector$$

where *dPCintra* represents the contribution of each city to city-network within a megaregion, *dPCflux* describes how well each city is linked to each of the other cities, and *dPCconnector* shows the contribution made by a city as a connecting element that allows the connectivity between cities to be maintained.

The pixel-based population density information is treated as the multi-scale patches to explore which cities in the PRD make significant contributions to the metropolitan network. We thus select the 28 cities in which the population density exceeded 10 inhabitants per 900 $m^2$ as this is the size of the patches used to analyze connectivity within the urban agglomeration process.

## 3. Results

### 3.1. Accuracy assessment of LULC classification

Fig. 3 reveals very high spatial urban growth dynamics at mega-region level in PRD. The most rapid urban growth happened in the cities of Guangzhou, Foshan, Dongguan, and Shenzhen among 2000 to 2010. We adopt two methods for assessing the accuracy of the LULC classification. In the first method we randomly generate 2000 reference samples for each image in order to assess the classification result. We verify the attributes of these samples through visual interpretation and by consulting many historical high resolution (0.5 m) images dating from 2009 that were obtained online from http://www.terraserver.com/. The overall accuracies of the four images indicated by the assessment are 75.64%, 78.51%, 79.36%, and 80.22%. These classification accuracies are considered to be generally satisfactory. In the second method we use two urban footprint maps (Taubenböck et al., 2014), one for the year 2000 and the other for the year 2010, for the verification of past urban and rural areas. The overlaps between the urban and rural areas in the urban footprints maps and those in our own research were 90.3% in 2000 and 86.5% in 2010 revealing high consistency.

### 3.2. Comparison and accuracy assessment of different population disaggregation models

For validation purposes, we categorized 28 original city-scale administrative areas (as shown in Fig. 1) 6 district-scale administrative areas and 3 sub-district scale areas into four sets according to their surface areas and population, in order to evaluate the estimation accuracy based on different regional scales:

- Set 1 - consisting of all 37 areas together, with an average area for each area of 1534.45 km$^2$ and a total of 23.89 million inhabitants in 2000, and 27.20 million in 2010.
- Set 2 - consisting of 3 district-scale administrative areas (covering Baiyun, Yuexiu and Liwan, Baoan and Nanshan, and Longgang, Futian, and Luohu), with an average area for each area of 939 km$^2$ and a total of 3.33 million inhabitants in 2000 and 4.41 million in 2010.
- Set 3 - consisting of 3 district-scale administrative areas (covering the cities of Zhaoqing, Jiangmen, and Zhuhai) with an average area for each area of 543.25 km$^2$ and a total of 1.32 million inhabitants in 2000 and 1.57 million inhabitants in 2010.
- Set 4 - consisting of 3 sub districts-scale administrative areas (covering Tianhe and Huangpu, Fangcun and Haizhu, and Yantian), with an average area for each area of 155.33 km$^2$ and a total of 3.17 million inhabitants in 2000 and 4.12 million in 2010.

We use three parameters to evaluate the accuracy of the different models: the Pearson coefficient (R), the root-mean-square error (RMSE), and the mean absolute error (MAE). Results from the different validations are shown in Table 2. In each of these four sets of accuracy assessment the best results were obtained using a modified version of the LV method, yielding R = 0.998–0.999 and RMSE = 74,091–85,795. For Set 1, an extremely strong correlation is evident between the LV and WorldPop methods, with R greater than 0.99, but the RMSE for the WorldPop method is about 30 times

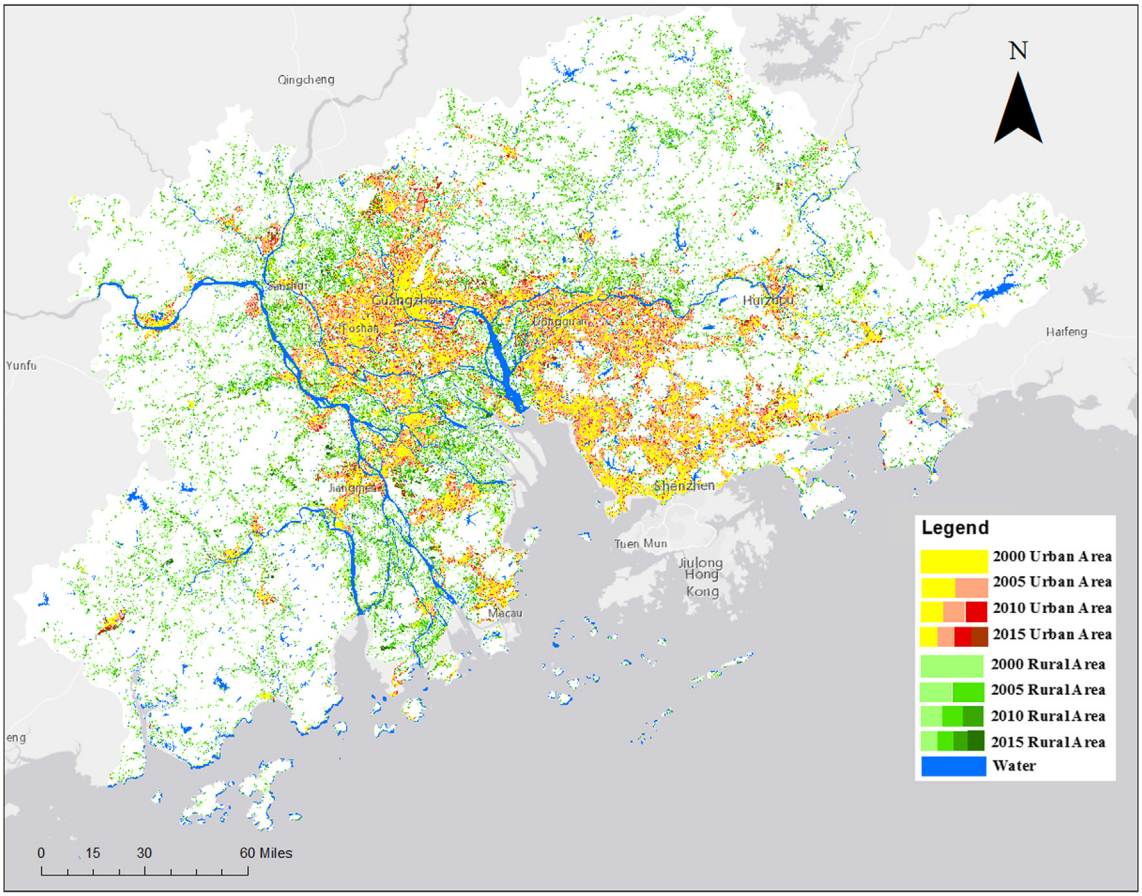


**Fig. 3.** Urban growth of PRD from multi-temporal Landsat data, 2000 to 2015.

higher than that for the LV method. For the smaller validation units in Set 2 to 4, there are smaller differences between the RMSE and the MAE when using the LV method and those when using the other methods, but the RMSE and MAE values for the Cnpopulation, GPW, and WorldPop methods are all more than double those for the LV method.

The overall accuracy of the LV-based disaggregated maps approach was 40–60% higher than that of the Cnpopulation, GPW, and WorldPop approaches. The variances in R, RMSE, and MAE reveal that the LV algorithm yields less extreme estimation errors within the different sizes of administrative units. The differences between the errors are smaller in both, the large and small census units, indicating that the use of city-scale administrative units in the LV algorithm also increased the reliability of the population density estimates at smaller scales, and especially at the scale of small administrative areas (<0.01 $km^2$).

### 3.3. The variations in population density from 2000 to 2015

The LV method is applied for the PRD in the years 2000, 2005, 2010, and 2015. The resulting dasymetric population maps (Fig. 4a and d) suggest that the PRD experienced significant population growth.

All 28 cities in the PRD have gradually become contiguous and form a continuous urban area. The total megaregion population increased from 39 million in 2000 to 54 million in 2015, indicating an average annual growth rate of approximately 1.7%. The built-up areas within the metropolitan area increased at an annual rate of 3.35% over the past 15 years. All of the high population density areas (>10 inhabitants/900 $m^2$) are located within the built-up areas; the average population density in urban areas decreased from 13.56 inhabitants/900 $m^2$ in 2000 to 10.70 in 2015, after having reached a maximum of 17.78 in 2005. The average population density also decreased at an annual rate of 1% in the rural areas, but fluctuated within a steady range (from 8.12 to 9.59 inhabitants/900 $m^2$) in farmland areas. Forested areas and water bodies remained relatively unchanged.

### 3.4. The variations in population connectivity from 2000 to 2015

Least-cost distances firstly are used to provide an insight into the strong influence that urban expansion has on the spread of populations within the megaregion. All of the studied cities appear to have experienced an improvement in connectivity to their neighborhood over the past 15 years (Fig. 5). Some cities (Guangzhou, Nanhai, Dongguan, Panyu, Foshan, Jiangmen, Zhongshan, and Shunde) are closer connected than others, which suggests that these cities have already become focal centers within the larger metropolitan areas, consisting of large, continuous, urbanized land areas. Furthermore, the cities of Boluo, Sihui, Huiyang, and Huidong experienced high rates of increase in connectivity (>12%/year) during this period, indicating that urban agglomeration is proceeding more rapidly in these cities than in the other 24 cities.

In the PRD different cities may fulfil different roles in urban agglomeration, depending on their topological position and intrinsic population characteristics. In order to understand the functions of individual population patches, we rank them by population density and only took into consideration those patches with high population densities. High population density (Fig. 6a) was defined as a population density that was higher than the average population density (9.49 inhabitants/900 $m^2$) in 28 different types of cities in China (Zeng et al., 2011).

Fig. 7a illustrates that in the year 2000 the high population density areas (>9.49 inhabitants/900 $m^2$) were distributed within 14 cities of the PRD, most of which are in the south-eastern part of the area (including Guangzhou, Dongguan, Shenzhen, and Huidong). As a result of expansion of built-up areas, the high population density in Dongguan and Huizhou then gradually decreased to less than 10 inhabitants/900 $m^2$ and most of the high population density areas were subsequently concentrated in the cities of Guangzhou, Shenzhen, and Zhongshan. Nevertheless, since the total population in Dongguan remained in the top three cities of the PRD over the past 15 years, we still included Dongguan in the city-network analysis. High density of dasymetric population maps (Fig. 6b) provide a clear representation of the influence that the coalescing multi-nucleus urban landscape had on the spatial characteristics of the population distribution. Over the past 15 years, the coalescing processes have resulted in a relative predominance of large urban areas, especially within the nuclei of the Guangzhou and Shenzhen.

The dPC variable for all 14 high population density areas increased over the past 15 years (Fig. 7a). The valleys of connectivity curves reveal the prioritization of each city within the city-network, with the cities of Guangzhou and Shenzhen becoming the nuclei of the high population density network.

In addition, there are close relationships between the dPC, dPCintra, and dPCflux indexes and the population capacity (the proportion of total population, termed pop%) of each city over the past 15 years. Taking the year 2015 as an example, logarithmic processing of all the variables reveal a weak positive correlation (>0.3) between these three connectivity indexes and the population capacity (Fig. 7b); the dPCconnector and population density trends are similar and show a negative correlation with the other four indicators. The dPCintra index, which is relevant to the habitat availability in each city, is used to represent the fact that Guangzhou and Shenzhen made the most significant contribution to the population distribution in terms of the connectivity within the megaregion. The metrics of dPCflux to measure the maximum ability of connecting elements also suggest that Guangzhou and Shenzhen are important to form connections between other cities in the metropolitan network. The dPCconnector metric represents the accessibility that people can travel to a particular neighborhood; it indicates that as the population density in the connecting

**Table 2**
Accuracy assessment results for the LV, cnpopulation, GPW, and Worldpop modelling methods.

| Year | Method | Set 1 | | Set 2 | | Set 3 | | Set 4 | |
|---|---|---|---|---|---|---|---|---|---|
| | | R | RMES | RMES | MAE | RMES | MAE | RMES | MAE |
| 2000 | LV | 0.998 | 74,091 | 85,795 | 58,927 | 85,143 | 83,592 | 85,221 | 84,338 |
| | Cnpopulation | 0.585 | 1,154,321 | 275,861 | 258,927 | 220,832 | 173,683 | 2,133,865 | 1,713,187 |
| | GPW | 0.862 | 1,633,236 | 1,228,516 | 1,050,247 | 458,313 | 386,158 | 2,799,510 | 2,530,698 |
| 2010 | LV | 0.998 | 55,351 | 65,482 | 64,697 | 38,172 | 35,147 | 98,539 | 86,478 |
| | Cnpopulation | 0.685 | 1,631,538 | 977,562 | 789,502 | 450,558 | 402,645 | 3,162,284 | 2,719,879 |
| | GPW | 0.853 | 2,211,306 | 1,273,668 | 1,032,518 | 629,800 | 555,927 | 4,284,186 | 3,847,088 |
| | WorldPop | 0.994 | 2,354,352 | 1,621,574 | 1,337,305 | 623,056 | 552,085 | 4,346,886 | 3,842,041 |

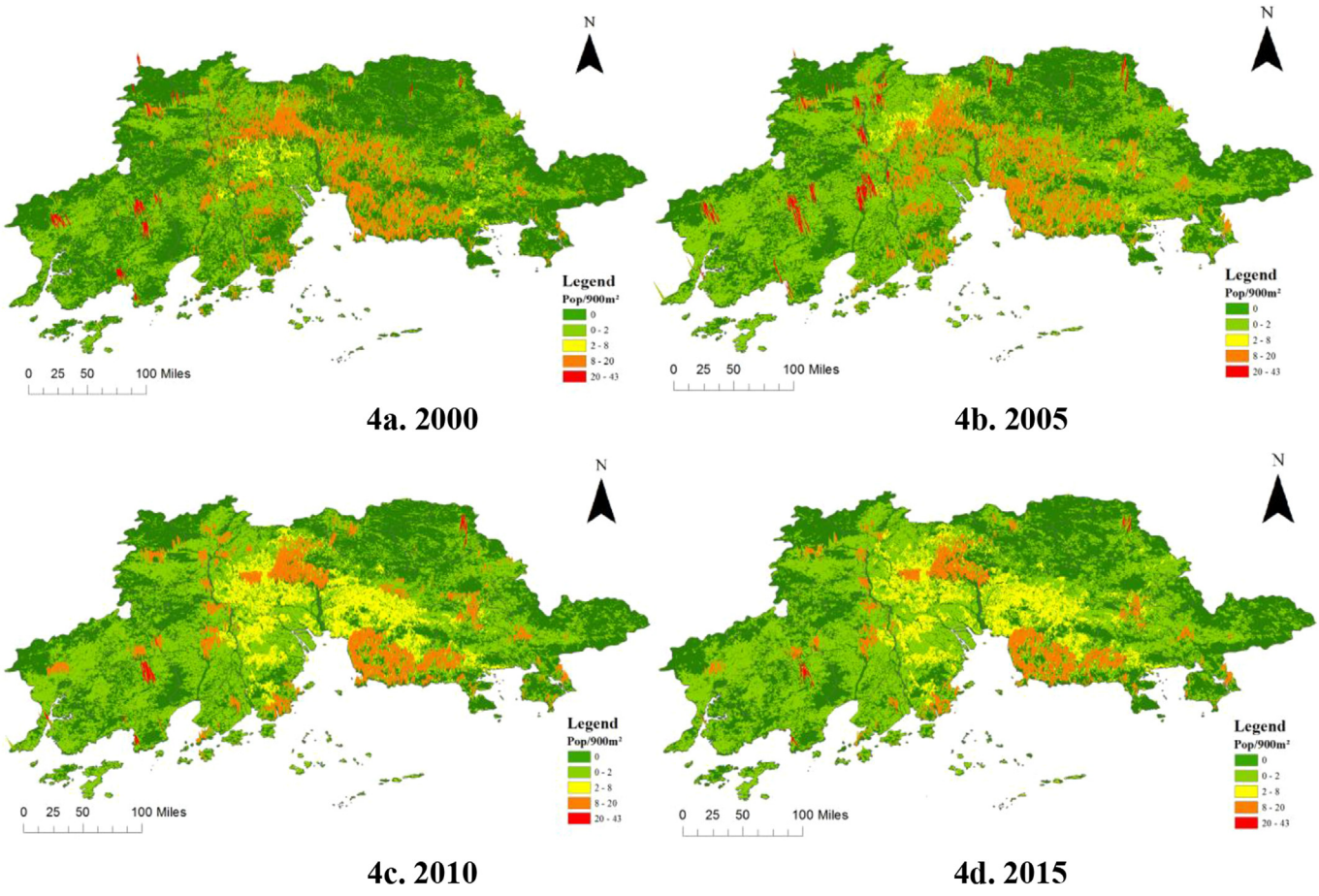


**Fig. 4.** Population density change detection for the Pearl River Delta, 2000–2015.

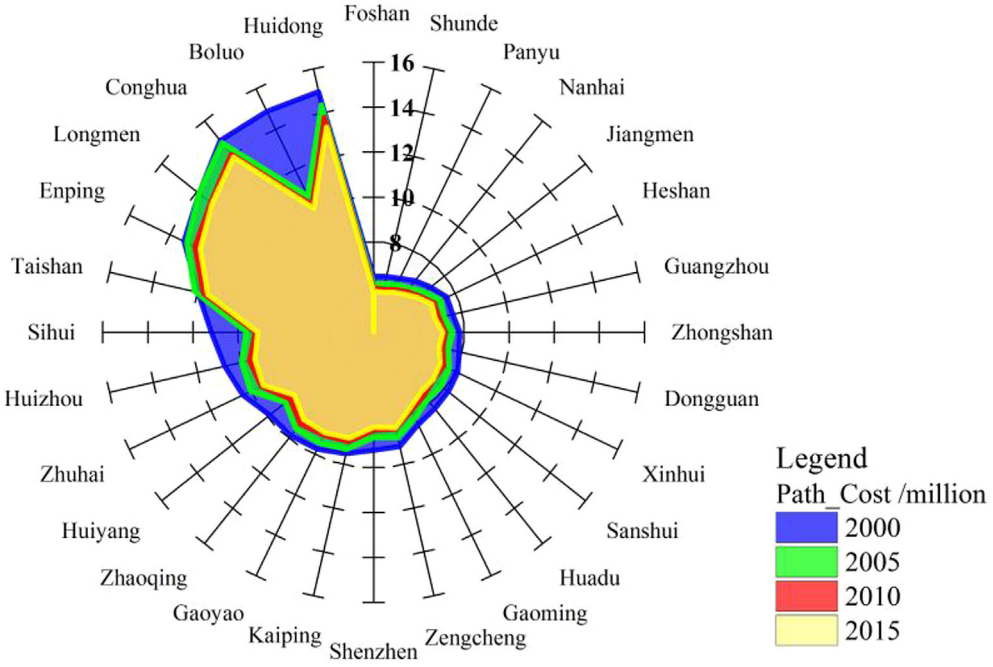


**Fig. 5.** The trends of least-cost distance of each city to its neighborhood, 2000–2015.

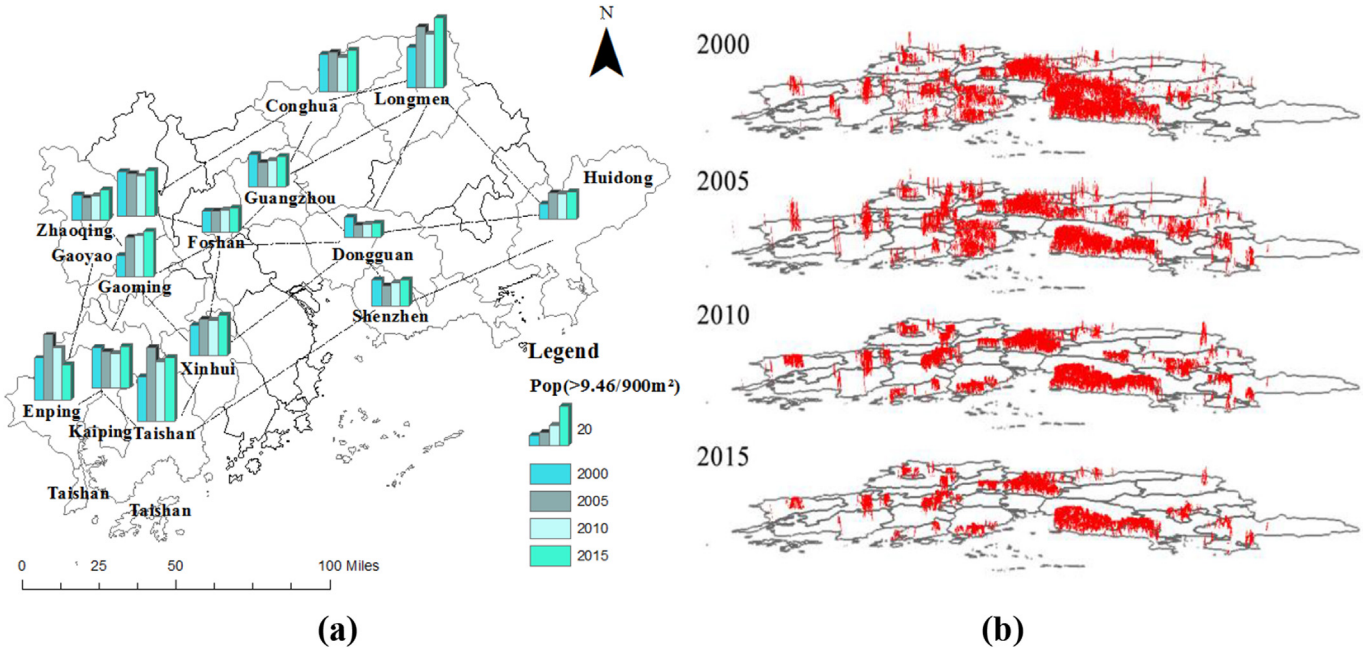


**Fig. 6.** The trends of high population density for the Pearl River Delta, 2000–2015.

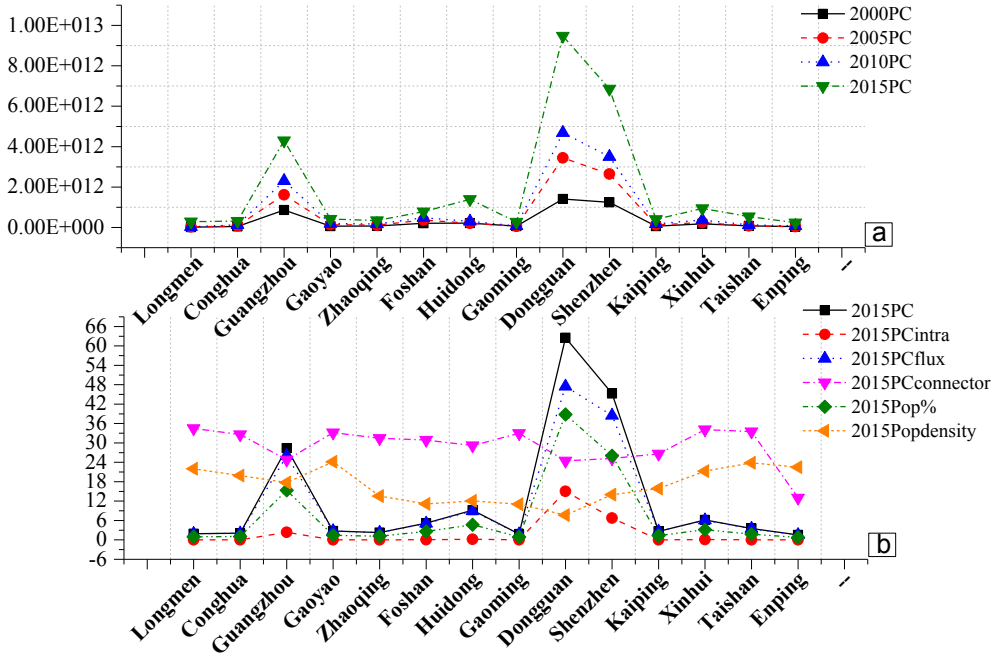


**Fig. 7.** The interregional interaction between economic factors and population density: (a) Trend in dPC from 2000 to 2015. Note that PCconnector is shown at a log10 scale; (b) Correlations between the connectivity indexes (dPC, dPCintra, dPCflux and dPCconnector) and population capacity (the percentage of population (Pop%) and population density) in 2015. Note that correlations are shown at different scale.

areas decreases, the accessibility of the residential areas increases, leading to an increasingly extensive of the PRD city-network.

## 4. Discussion

### 4.1. Methodology evaluation

The LV model results in a fine-scale (city-scale) dasymetric population weighting scheme using remote sensing data in PRD. This model has some conceptual difference with worldwide products GPW, WorldPop models, LandScanTM, etc., in our case the area covered is smaller but the spatial resolution is finer. We compare this approach to other population disaggregation models (the Cnpopulation). Both GPW and Cnpopulation are area-weighted population disaggregation models calculated from a class-based combination of 1 km resolution land cover classes. These models divide up the population into different types of land-use and land-cover information across province-based census units. Their

accuracies are limited by the large scale of their population redistribution weighting schemes which are unable to reflect the fine-scale heterogeneity of the population distribution, leading to a 30-fold greater inaccuracy in population density estimation at a city-scale than the LV model. The WorldPop model incorporates multiple ancillary data sources (such as land-cover, climate zone, etc.) based on the Random Forest algorithm. This model offers a potential solution to the problem of coarse resolution in dasymetric population mapping, yielding a good fit in cross-scale validation ($R > 0.8$ in this study). However, the RMSE and MAE for this model are reported to be at least 20 times higher than for the LV model. Furthermore, LV model that has been tested only in the European LULC system is also suitable for the city-scale Chinese population estimation through combining with the province-based reference data of population density in each LULC type. The LV model does not require vast quantities of spatial data and is based solely on land-use and land-cover information derived from Landsat TM/ETM+/8 imageries. Machine learning algorithms have advantages but are associated with workload, efficiency, and uncertainty problems when dealing with large data sets. As Langford and Unwin (2013) points out, complex methods to produce dasymetric population maps are a major obstacle for many users. It means the limit-based model has fewer variables than other ratio-based models, making it more flexible and preferable for semi-automated mapping of population distributions on a fine scale.

There are also some limitations to the city-scale dasymetric population model. This model can ensure the accuracy of the aggregation process within each census block but it cannot accurately reflect the consistent population redistribution along the area boundaries, especially when two cities have almost merged with each other and their built-up densities remain basically the same. Meanwhile, Street by street census data, or high resolution of parcel data will be required in the future in order to solve this problem, so that population density estimation based on different types of building areas would be able to disaggregate population densities with high levels of accuracy. Since the LULC information was constrained by the spatial resolution of Landsat satellite data, we only used three types of land cover classes to disaggregate the population density. The medium geometric resolution of the Landsat data as well as the related problem of mixed pixels make it necessary to take higher resolution (0.5–5 m) images into consideration for the finer-scale population disaggregation analysis in the future.

### 4.2. The evolution of population density in urban agglomeration measurement

A clear advantage of using fine resolution remote sensing data is the ability to produce consistent mapping and carry out periodic monitoring of large agglomerations (such as megaregions) on a fine scale, and at a small fraction of the cost that would be required for field surveys and censuses. With regard to population density, when we dissect the city population growth into each type of land-use there are 14 cities that have experienced a simultaneous increase in both population density and urban areas. With regard to urban expansion and the degree of population clusters, the cities of Shenzhen, Dongguan, and Guangzhou rank highest having the largest built-up areas and the largest populations. Huizhou, Panyu, and Shenzhen rank highest in terms of the annual rate of growth in population density, with growth rates of 3.34%, 3.26%, and 2.27%, respectively. Gaoming, Huidong, and Longmen rank highest in terms of the annual rate of growth in population density, with growth rates of 8.02%, 5.22%, and 4.63%, respectively.

The dasymetric population maps generally provide a realistic portrayal of population densities within specific areas in the PRD, where new urban agglomeration has taken place during the past 15 years. The population density results have confirmed a polycentric urban agglomeration process in the PRD. Foshan, Jiangmen, Zhongshan, and Zhuhai combined over the last 15 years to form a sub-center within the primary center formed by Guangzhou, Dongguan, and Shenzhen. The urban area extends far beyond the individual cities and such extensive built-up areas are facing heavy population pressure.

### 4.3. Connectivity

The least-cost distances of the investigated cities in the PRD all decreased over the past 15 years, which is to be expected under a pervasive expansion scenario. The areas that experienced the greatest reductions in the least-cost distance were those furthest from the established multi-nuclei of the mega-cities (Guangzhou, Nanhai, Jiangmen, Foshan, and Panyu). This demonstrates the trend in urban development towards increasing agglomeration, with growth occurring in low density areas rather than on the edge of, or adjacent to, the megaregion cores, and with all cities merging into each other. Since the year 2000, change in the probability of connectivity index (*dPC*) has shown a significant reduction, providing evidence of a coalescent process towards multi-nuclei mega-cities in the PRD. The *dPCconnector*, *dPCflux*, and *dPCintra* indexes of the 14 cities indicate the important role that Guangzhou and Shenzhen have played in the development of other cities. These two cities fulfilled important roles in their topological position, enhancing the connectivity and spatial cohesion of the agglomeration network, facilitating the dispersal of inhabitants from other cities, and maintaining the overall population of the community. The constant decrease in all four of these indexes indicates the densification process in all cities in the PRD, which is accompanied by problems of adaptation faced by inhabitants at broad temporal and spatial scales. However, we are aware that our analysis is based on two-dimensional urban expansion. Thus, implicitly due to data availability we neglect the growth of cities in the third dimension, which definitively has also immense impact.

The graph-based method used, nevertheless, may be restricted by the underlying assumption that each land-use class in each city represents an inhabited patch. As an approximation we just used the center of each patch in the connectivity analysis, how the dimension and the shape of each patch can influence the connectivity results will be the subject of future investigations. Meanwhile, we are also aware that the thresholds for the connectivity analysis are at risk to be subjective. The use of thresholds in our study is consistent with the ordinal nature of much subjective well-being data, as it requires no assumptions about the cardinality of scale responses. Nevertheless, how to identify meaningful thresholds that have real-world validity remains an essential and difficulty question.

## 5. Conclusion

The main objective of the proposed framework was to create spatially disaggregated population maps for a number of time steps, covering the PRD in southern China. Using dasymetric population maps obtained from medium to high resolution remote sensing data, we have been able to characterize the population distribution at three different scales with comparatively high accuracies. The evolution of the urban agglomerations was transferred onto a landscape network. The graphic structure allowed the delineation and analysis of 28 cities and their relationships within the PRD megaregion network.

We are confident that the methodology can be transferred across the globe with similar input data. In areas with uncomplicated access to census data it may be relatively straightforward to

replicate this approach. For the PRD area the methodology reveals new insights into the evolution of the 28 cities and their multi-scale agglomeration pattern. It will be interesting from both methodological and sociological perspectives to re-apply the method to urban agglomeration analyses in other megaregions. This framework is transferable and is able to provide a spatially explicit and tractable representation of complex coalescent megaregions; it allows investigators to draw a broad picture of delineating the spatial agglomeration in megacity regions, and to assist in urban planning and sustainable development efforts.

## Acknowledgements

Work performed by the editorial office and the anonymous referees is greatly appreciated, and their comments and suggestions have significantly improved this manuscript. We also wish to thank the GeoData Institute of the University of Southampton for providing the WorldPop data, the Earth Institute of Columbia University for providing the GPW data, the Data Center for Resources and Environmental Sciences of the Chinese Academy of Sciences (RESDC) for providing the 1 km LULC data and Cnpopulation data, and Dr. Futao Wang and Cuiqing Zeng for providing the population reference data and population survey data. This research has also partially been funded by the China Scholarship Council Scholarship (contract No.CSC 201306070014), and the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W 1237-N23).

## References

Adriaensen, F., Chardon, J. P., De Blust, G., Swinnen, E., Villalba, S., Gulinck, H., et al. (2003). The application of 'least-Cost' modelling as a functional landscape model. *Landscape and Urban Planning, 64*(4), 233–247.

Arthur, W. B., & McNicoll, G. (1975). Large-scale simulation models in population and Development: What use to planners? *Population and Development Review, 1*(2), 251–265.

Baraldi, A., Puzzolo, V., Blonda, P., Bruzzone, L., & Tarantino, C. (2006). Automatic spectral rule-based preliminary mapping of calibrated Landsat TM and ETM+ images. *IEEE Transactions on geoscience and remote sensing, 44*(9), 2563–2586.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of photogrammetry and remote sensing, 58*(3), 239–258.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*(1), 2–16.

Brennan, E. M. (1999). Population, urbanization, environment, and security: A summary of the issues. *Environmental Change and Security Project Report, 5*, 4–14.

Damoiseaux, J. S., & Greicius, M. D. (2009). Greater than the sum of its parts: A review of studies combining structural connectivity and resting-state functional connectivity. *Brain Structure and Function, 213*(6), 525–533.

Deichmann, U., Balk, D., & Yetman, G. (2001). *Transforming population data for interdisciplinary usages: From census to grid*.

Dobson, J. E., Bright, E. A., Coleman, P. R., Durfee, R. C., & Worley, B. A. (2000). LandScan: A global population database for estimating populations at risk. *Photogrammetric Engineering and Remote Sensing, 66*(7), 849–857.

Dorren, L. K. A., Maier, B., & Seijmonsbergen, A. C. (2003). Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. *Forest Ecology and Management, 183*(1), 31–46.

Driezen, K., Adriaensen, F., Rondinini, C., Patrick Doncaster, C., & Matthysen, E. (2007). Evaluating least-cost model predictions with empirical dispersal data: A case-study using radiotracking data of hedgehogs (Erinaceus Europaeus). *Ecological Modelling, 209*(2–4), 314–322.

Drăguţ, L., Csillik, O., Eisank, C., & Tiede, D. (2014). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing, 88*(February), 119–127.

Eicher, C. L., & Brewer, C. A. (2001). Dasymetric mapping and areal Interpolation: Implementation and evaluation. *Cartography and Geographic Information Science, 28*(2), 125–138.

Foltête, J.-C., & Giraudoux, P. (2012). A graph-based approach to investigating the influence of the landscape on population spread processes. *Ecological Indicators, 18*(July), 684–692. http://dx.doi.org/10.1016/j.ecolind.2012.01.011.

Frankhauser, P. (1998). The fractal approach. A new tool for the spatial analysis of urban agglomerations. *Population, 10*(1), 205–240.

Gallego, F. J., Batista, F., Rocha, C., & Mubareka, S. (2011). Disaggregating population density of the European union with CORINE land cover. *International Journal of Geographical Information Science, 25*(12), 2051–2069.

Georg, I., Blaschke, T., & Taubenböck, H. (2016). New spatial dimensions of global cityscapes: From reviewing existing concepts to a conceptual spatial approach. *Journal of Geographical Sciences, 26*(3), 355–380.

Glaeser, E. L. (2008). *Cities, agglomeration, and spatial equilibrium. OUP catalogue*. Oxford University Press.

He, C., Chen, T., Mao, X., & Zhou, Y. (2016). Economic transition, urbanization and population redistribution in China. *Habitat International, 51*(February), 39–47.

Hofmann, P., Blaschke, T., & Strobl, J. (2011). Quantifying the robustness of fuzzy rule sets in object-based image analysis. *International Journal of Remote Sensing, 32*(22), 7359–7381.

Kim, J. Y., & Han, J. H. (2016). Straw effects of new highway construction on local population and employment growth. *Habitat International, 53*(April), 123–132.

L.Imhoff, M., Lawrence, W. T., Stutzer, D. C., & Elvidge, C. D. (1997). A technique for using composite DMSP/OLS 'city lights' satellite data to map urban area. *Remote Sensing of Environment, 61*(3), 361–370.

Lang, W., Chen, T., & Li, X. (2016). A new style of urbanization in China: Transformation of urban rural communities. *Habitat International, 55*(July), 1–9.

Langford, M., & Unwin, D. J. (2013). Generating and mapping population density surfaces within a geographical information system. *The Cartographic Journal, 18*(July), 21–26.

Li, T., Pullar, D., Corcoran, J., & Stimson, R. (2007). A comparison of spatial disaggregation techniques as applied to population estimation for South East Queensland (SEQ), Australia. *Applied GIS, 3*(9), 1–16.

Listengurt, F. M. (1975). Criteria for delineating large urban agglomerations in the USSR. *Soviet Geography, 16*(9), 559–568.

Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *The Professional Geographer, 55*(1), 31–42.

Martin, David (1996). An assessment of surface and zonal models of population. *International Journal of Geographical Information Systems, 10*(8), 973–989.

Nazara, S., & Hewings, G. J. D. (2003). Towards regional growth decomposition with Neighbor's effect: A new perspective on shift-share analysis. *Technical Series, 21*(July), 2–18.

Pascual-Hortal, L., & Saura, S. (2006). Comparison and development of new graph-based landscape connectivity Indices: Towards the priorization of habitat patches and corridors for conservation. *Landscape Ecology, 21*(7), 959–967.

Piao, S., Fang, J., Zhou, L., Guo, Q., Henderson, M., Ji, W., et al. (2003). Interannual variations of monthly and seasonal normalized difference vegetation index (NDVI) in China from 1982 to 1999. *Journal of Geophysical Research: Atmospheres, 108*(D14).

Plowright, R. K., Foley, P., Field, H. E., Dobson, A. P., Foley, J. E., Eby, P., et al. (2011). Urban Habituation, Ecological Connectivity and Epidemic Dampening: The Emergence of Hendra Virus from Flying Foxes (Pteropus Spp.). *Proceedings of the Royal Society of London B: Biological Sciences, 278*(1725), 3703–3712.

Porter, J. H., Raymond, D. Dueser, & Moncrief, N. D. (2015). Cost-distance analysis of mesopredators as a tool for avian habitat restoration on a naturally fragmented landscape. *The Journal of Wildlife Management, 79*(2), 220–234.

Poyil, R. P., & Misra, A. K. (2015). Urban agglomeration impact analysis using remote sensing and GIS techniques in Malegaon city, India. *International Journal of Sustainable Built Environment, 4*(1), 136–144.

Robinson, J. (2006). *Ordinary Cities: Between modernity and development*. Psychology Press.

Ross, C. L., & Woo, M. (2011). Megaregions and mobility. *Bridge, 41*(1), 24–34.

Salvati, L. (2014). The 'sprawl Divide': Comparing models of urban dispersion in mono-centric and polycentric Mediterranean cities. *European Urban and Regional Studies, 9*(January), 1–17.

Saura, S., & Pascual-Hortal, L. (2007). A new habitat availability index to integrate connectivity in landscape conservation Planning: Comparison with existing indices and application to a case study. *Landscape and Urban Planning, 83*(2–3), 91–103.

Saura, S., & Rubio, L. (2010). A common currency for the different ways in which patches and links can contribute to habitat availability and connectivity in the landscape. *Ecography, 33*(3), 523–537.

Schleicher, A., Biedermann, R., & Kleyer, M. (2011). Dispersal traits determine plant response to habitat connectivity in an urban landscape. *Landscape Ecology, 26*(4), 529–540.

Schumaker, N. H. (1996). Using landscape indices to predict habitat connectivity. *Ecology, 77*(4), 1210–1225.

Singleton, P. H., & Lehmkuhl, J. F. (2001). Using weighted distance and least-cost corridor analysis to evaluate regional-scale large carnivore habitat connectivity in Washington. *Road Ecology Center* (September).

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PLoS ONE, 10*(2), e0107042.

Stucky, Jay Lee, & Dan. (1998). On applying viewshed analysis for determining least-cost paths on digital elevation models. *International Journal of Geographical Information Science, 12*(8), 891–905.

Tan, M., Li, X., Lu, C., Luo, W., Kong, X., & Ma, S. (2008). Urban population densities and their policy implications in China. *Habitat International, 32*(4), 471–484.

Taubenböck, H., Wegmann, M., Roth, A., Mehl, H., & Dech, S. (2009). Urbanization in India – spatiotemporal analysis using remote sensing data. *Computers, Environment and Urban Systems, 33*(3), 179–188.

Taubenböck, H., & Wiesner, M. (2015). The spatial network of megaregions–Types of connectivity between cities based on settlement patterns derived from EO-

data. *Computers, Environment & Urban Systems, 54*, 165–180.
Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., & Dech, S. (2014). New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. *Applied Geography, 47*, 137–153.
Tewari, M., Guenther, A., & Wiedinmyer, C. (2009). Impacts of weather conditions modified by urban expansion on surface ozone: Comparison between the Pearl River Delta and Yangtze River Delta regions. *Advances in Atmospheric Sciences, 26*(5), 962–972.
Urban, D. L., Minor, E. S., Treml, E. A., & Schick, R. S. (2009). Graph models of habitat mosaics. *Ecology Letters, 12*(3), 260–273.
Vaz, E., Zhao, Y., & Cusimano, M. (2016). Urban habitats and the injury landscape. *Habitat International, 56*(August), 52–62.
Walker, R., and L. Craighead. , 1997. "Analyzing wildlife movement corridors in Montana using GIS." In Proceedings of the 1997 ESRI user conference, Redlands, USA.
Wilson, J. F., Weale, M. E., Smith, A. C., Gratrix, F., Benjamin Fletcher, M. G., Thomas, N. B., et al. (2001). Population genetic structure of variable drug response. *Nature Genetics, 29*(3), 265–269.
Wu, S. S., Qiu, X., & Wang, L. (2005). Population estimation methods in GIS and remote sensing: A review. *GIScience & Remote Sensing, 42*(1), 80–96.
Yue, W., Zhang, L., & Liu, Y. (2016). Measuring sprawl in large Chinese cities along the Yangtze River via combined single and multidimensional metrics. *Habitat International, 57*(October), 43–52.
Zeng, C., Zhou, Y., Wang, S., Yan, F., & Zhao, Q. (2011). Population spatialization in China based on night-time imagery and land use data. *International Journal of Remote Sensing, 32*(24), 9599–9620.
Zhou, D., Xu, J., Wang, L., & Lin, Z. (2015). Assessing urbanization quality using structure and function analyses: A case study of the urban agglomeration around Hangzhou Bay (UAHB), China. *Habitat International, 49*(October), 165–176.
Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment, 118*, 83–94.