Chin. Geogra. Sci. 2019 Vol. 29 No. 4 pp. 652–666
https://doi.org/10.1007/s11769-019-1048-9

Springer Science Press
www.springerlink.com/content/1002-0063

Directional Analysis of Urban Expansion Based on Sub-pixel and Regional Scale: A Case Study of Main Districts in Guangzhou, China

ZHAO Yi[1, 2, 3, 4, 5], ZHONG Kaiwen[2, 4, 5], XU Jianhui[2, 4, 5], SUN Caige[2, 4, 5], WANG Yunpeng[1, 3]

(1. *Guangzhou Institute of Geochemistry*, *Chinese Academy of Sciences*, *Guangzhou* 510640, *China*; 2. *Guangzhou Institute of Geography*, *Guangzhou* 510070, *China*; 3. *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*; 4. *Key Laboratory of Guangdong for Utilization of Remote Sensing and Geographical Information System*, *Guangzhou* 510070, *China*; 5. *Guangdong Open Laboratory of Geospatial Information Technology and Application*, *Guangzhou* 510070, *China*)

Abstract: Multi-scale data have had a wide-ranging level of performance in the area of urban change monitoring. Herein we investigate the correlation between the impervious surface fraction (ISF) and the Defense Meteorological Satellite Program/Operational Linescan System (DMSP/OLS) nighttime stable light (NTL) data with respect to the urban expansion in the main districts of Guangzhou. Landsat 5 Thematic Mapper and Landsat 8 Operational Land Image (OLI) data from 1988 to 2015 were used to extract the ISF using the linear spectral mixture analysis model and normal difference build-up index at the sub-pixel scale. DMSP/OLS NTL data from 1992 to 2013 were calibrated to illustrate the urban nighttime light conditions at the regional scale. Urban expansion directions were identified by statistics and kernel density analysis for the ISF study area at the sub-pixel scale. In addition, the correlation between the ISF and DMSP/OLS NTL data were illustrated by linear regression analysis. Furthermore, Profile Graph in ArcGIS was employed to illustrate the urban expansion from the differences in correlation in different directions. The conclusions are as follows: 1) The impervious surface (IS) in the study area has expanded to the northeast and the east, starting with the old urban zones, and the high-density IS area has increased by 321.14 km^2. 2) The linear regression analysis reveals a positive correlation between the ISF and the DMSP/OLS NTL data. The multi-scale data changes are consistent with the actual urban planning of Guangzhou. 3) The DMSP/OLS NTL data overestimate the urban extent because of its saturation and blooming effects, causing its correlation with ISF to decrease. The pattern of urban expansion influences the saturation and blooming effects of the DMSP/OLS NTL data.

Keywords: impervious surface fraction; linear spectral mixture analysis; nighttime light; kernel density; profile graph

Citation: ZHAO Yi, ZHONG Kaiwen, XU Jianhui, SUN Caige, WANG Yunpeng, 2019. Directional Analysis of Urban Expansion Based on Sub-pixel and Regional Scale: A Case Study of Main Districts in Guangzhou, China. *Chinese Geographical Science*, 29(4): 652–666. https://doi.org/10.1007/s11769-019-1048-9

1 Introduction

In recent decades, rapid urbanization has occurred in many regions around the world (Zhang et al., 2015). Cities are hotspots of demographic and economic development, generating more than 90% of the global gross value added (Fragkias et al., 2013). Spatiotemporal typologies of urbanization and their dynamics of changes have been studied intensely by geographers, economists, and social scientists for many decades (Haase and Nuissl, 2010). China has attracted widespread international attention from academics as a result

Received date: 2018-06-11; accepted date: 2018-10-10

Foundation item: Under the auspices of the Special Project of Science and Technology Development (No. 2017GDASCX-0101), the Science and Technology Planning Project of Guangdong Province (No. 2017A020217005, 2018B020207002), Guangdong Innovative and Entrepreneurial Research Team Program (No. 2016ZT06D336)

Corresponding author: ZHONG Kaiwen. E-mail: zkw@gdas.ac.cn

of its present rapid urbanization (Fang et al., 2015).

The urban impervious surface (IS) has been recognized as a key indicator of urbanization and the urban environment (Lu and Weng, 2006; Xu, 2009). In fact, a number of studies (Brodley, 1995; Yuan et al., 2008; Zhu and Lv, 2010; Wang et al., 2016) regarding IS extraction using multiple datasets (Xu et al., 2017) at different scales have been conducted, using analyses such as decision tree classifier (DTC), artificial neural network, support vector machine, and linear spectral mixture analysis (LSMA) (Weng and Lu, 2008), and indexes such as the normalized difference impervious surface index (Xu, 2008; Xu, 2010) at the pixel scale and biophysical composition index (Deng and Wu, 2012), with more approaches proposed (Chen et al., 2006). LSMA computes the impervious surface fraction (ISF) from mixed pixels at the sub-pixel scale from images with different spatial resolutions, based on the V-I-S model (Ridd, 1995; Pan et al., 2010; Li, 2015). In remote sensing images, as mixed pixels usually exist, LSMA has been a popular method to extract ISF at the sub-pixel scale. Liu et al. (2012b) analyzed the spatio-temporal dynamics of IS in Shenzhen, using LSMA from the Landsat Thematic Mapper (TM) and Enhanced TM. Combined with different methods, LSMA is frequently modified to extract the urban IS. Lu and Weng (2009) extracted the urban IS from IKONOS images, using LSMA and DTC. Fan et al. (2015) enhanced urban IS estimation by conventional LSMA, using spectral indices, and the results illustrated the advantage of using the normal difference build-up index (NDBI) to mask the results of the low-albedo fraction, improving the accuracy of LSMA. Previous studies have shown that LSMA, combined with other methods, can be applied to extract urban ISF to analyze urbanization and the urban environment. Xu et al. (2016) used NDBI to mask the LSMA urban ISF results (high-albedo fraction and low-albedo fraction) to simulate surface runoff.

DMSP/OLS NTL data were also a significant source for urban socioeconomic parameters (Li, 2015). Time series DMSP/OLS NTL data were used to develop an IS layer in 2007 (Elvidge et al., 2007), and calibrated DMSP/OLS data have been used to analyze urban indicators (Xin et al., 2017), such as gross domestic product and population. Fu et al. (2017) proposed a nighttime light economy index to investigate urbanization and the relationship between economic indicators and urbanization. Xie and Weng (2016; 2017) updated urban extents using a threshold from DMSP/OLS data and assessed large-scale urban dynamics by enhancing time-series DMSP/OLS data to eliminate saturation and blooming effects. A series of studies have shown that there is a positive correlation between the value of the DMSP/OLS NTL data and urban extent.

ISF and DMSP/OLS NTL data were both impactful references for urbanization. There were many studies regarding urbanization at different scales. Generally, the ISF by LSMA from Landsat TM and Landsat Operational Land Image (OLI) were employed to analyze the urban dynamics at the sub-pixel scale, and the DMSP/OLS data were employed to illustrate the urban scope at regional scales. However, there were few studies regarding the correlation between ISF and DMSP/OLS NTL data to reveal the urban expansion direction corresponding with actual urban area changes.

The goal of this study is: 1) to illustrate the correlation between ISF and the DMSP/OLS NTL data by linear regression analysis of actual values of samples with respect to urban expansion direction; 2) to reveal and analyze the urban expansion direction at multi-scales based on the correlation between ISF and DMSP/OLS NTL data and the actual urban area changes. The process is as follows: 1) extract the urban ISF using LSMA and NDBI to ascertain the urbanization direction by kernel density analysis and statistics at the sub-pixel scale; 2) calibrate the DMSP/OLS NTL data to study the correlation between DMSP/OLS NTL data and ISF data using linear regression analysis of urbanization direction; 3) combine the actual urban area changes, illustrate the direction properties of the urban expansion by profile graphs using the ISF and DMSP/OLS NTL data, and discuss the performance of the ISF and DMSP/OLS NTL data in revealing urban change.

2 Study Area and Data Sets

2.1 Study Area

Guangzhou, the capital of Guangdong Province, is the core city of the Pearl River Delta and is located in the southern China. The study area includes six county-level units in Guangzhou (Yuexiu, Liwan, Haizhu, Baiyun, Huangpu, and Tianhe; Fig. 1). The city boundary is denoted by the Guangzhou administrative map of 2015. The first three counties are old urban zones, and the

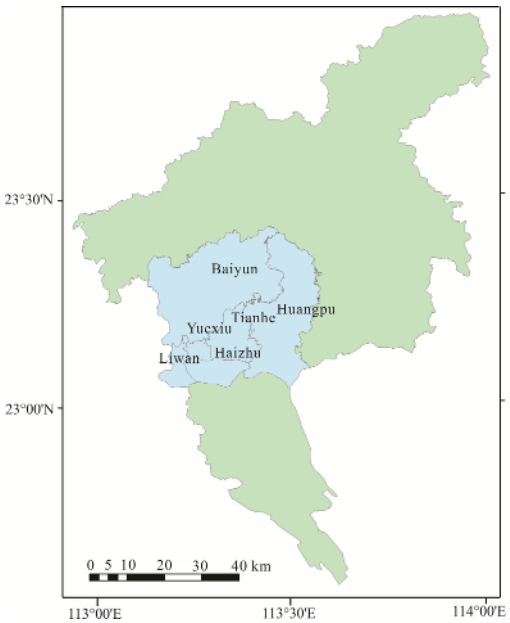

Fig. 1 The study area of main urban zones in Guangzhou (Baiyun, Huangpu, Yuexiu, Haizhu, Tianhe and Liwan) (source: https://zhfw.tianditu.gov.cn/)

latter three counties are new urban zones. There are typical IS areas, distributed in Yuexiu, Liwan, and Haizhu, as well as large areas of forest, main distributed in the north-central Huangpu and southeastern Baiyun. The study area includes land cover areas that have experienced significant and varied urbanization.

2.2 Data

In this study, Landsat 5 Thematic Mapper and Landsat 8 Operational Land Image (Path 122/Row 44) data were used to extract the urban ISF. The data were provided by the Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences (http://www.gscloud.cn). Some details for each image are shown in Table 1. Data from October to December were used to reduce the influence of vegetation and cloud coverage. Digital numbers (DNs) of the remote sensing images were converted to surface reflectance by preprocessing, including the atmospheric correction and geometric correction, which were provided by the United States Geological Survey (USGS), Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA) On Demand Interface. The surface reflectance products were rectified to the UTM-WGS84 projection system (zone 49). The modified normalized difference water index(MNDWI) threshold was applied for water elimination.

This study used DMSP/OLS NTL data from 1992 to 2013 in Version 4 DMSP/OLS Nighttime Lights Time Series (V4DNLTS) datasets, with a 6-bit data range and cell size of 1 km × 1 km. Data were obtained from the National Geophysical Data Center (NGDC) website (https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html) and included data acquired by six satellites: F10, F12, F14, F15, F16, and F18. The DMSP/OLS NTL data were extracted according to the Chinese administrative boundaries.

For the DMSP/OLS data used in this study, calibration included inter-calibration, intra-annual composition, and inter-annual calibration (Liu et al., 2012a) to solve the following problems: 1) discrepancies between DN values derived from different satellites for the same year; 2) abnormal fluctuations in DN values for different years derived from the same satellite; 3) discrepancies in the number of lit pixels between two satellites for the same year, and 4) abnormal decreases in the number of lit pixels derived from the same satellites between different years (Su, 2015).

For inter-calibration, Hong Kong was selected as the calibration area because of its relatively stable socio-economic development and high conformity of DN values over different years for the area. Satellite F16 in 2007 was used as the reference dataset because it had the highest cumulative DN value. The second-order regression model for each satellite is as follows (Su, 2015):

$$DN_{correct} = a \times DN^2 + b \times DN + c \quad (1)$$

where $DN_{correct}$ is the inter-calibrated NTL value; DN is the initial NTL value; and a, b, and c are the coefficients (Table 2). The DMSP/OLS NTL data were inter-calibrated for China for the period 1992–2013, using the regression model.

Table 1 Landsat dataset

Date	Sensor	Clouds (%)	Date	Sensor	Clouds (%)
1988-11-24	Landsat 5 TM	7.00	2004-10-19	Landsat 5 TM	0.00
1994-10-24	Landsat 5 TM	0.00	2009-11-02	Landsat 5 TM	0.00
2001-12-30	Landsat 5 TM	0.00	2015-10-18	Landsat 8 OLI	1.15

Table 2 Coefficients of the second-order regression models for inter-calibrating the DMSP/OLS NTL data

Satellite	Year	a	b	c	R^2	Satellite	Year	a	b	c	R^2
F10	1992	–0.006	1.328	0.596	0.856	F15	2001	0.005	0.631	1.542	0.980
	1993	–0.013	1.838	–2.941	0.913		2002	0.009	0.366	3.046	0.984
	1994	–0.011	1.668	–1.652	0.934		2003	–0.004	1.236	–0.025	0.978
F12	1994	–0.002	1.062	1.005	0.922		2004	–0.004	1.206	0.275	0.983
	1995	–0.004	1.210	–0.449	0.938		2005	–0.005	1.311	–0.279	0.985
	1996	–0.001	1.038	–0.137	0.942		2006	–0.003	1.153	0.121	0.989
	1997	0.001	0.864	0.616	0.949		2007	–0.006	1.428	–1.254	0.991
	1998	0.000	1.006	–1.132	0.968		2004	0.002	0.805	1.375	0.974
	1999	0.007	0.542	1.898	0.972	F16	2005	–0.009	1.544	–0.946	0.996
F14	1997	–0.006	1.343	–0.410	0.940		2006	–0.005	1.350	–1.663	0.991
	1998	–0.009	1.542	–1.571	0.977		2007	0	1	0	1
	1999	–0.003	1.176	–0.155	0.974		2008	–0.001	1.007	0.369	0.996
	2000	–0.003	1.141	0.797	0.973		2009	–0.005	1.308	0.163	0.982
	2001	–0.002	1.061	1.472	0.975		2010	0.005	0.646	1.094	0.952
	2002	–0.000	0.971	1.176	0.961	F18	2011	0.005	0.610	1.978	0.988
	2003	–0.001	0.999	1.340	0.981		2012	–0.001	1.063	–1.240	0.987
F15	2000	0.007	0.503	2.244	0.972		2013	0.001	0.892	–0.483	0.986

In Table 2, the minimum R^2 of the second-order regression models is greater than 0.85, illustrating a credible result. After calibration, the study area administrative map was used to extract the DMSP/OLS NTL data. Additionally, DMSP/OLS data were projected using WGS_84_UTM and resampled to a pixel size of 900 m × 900 m. To integrate with the multi-temporal images from the OLI and TM data, the DMSP/OLS NTL average values were calculated for each time period (1992–1995, 1996–2000, 2001–2005, 2006–2010, 2011–2013), expressed as NTL1992_1995, NTL1996_ 2000, NTL2001_2005, NTL2006_2010, and NTL2011_ 2013.

3 Methods

The flow diagram of the study is shown in Fig. 2. Landsat 5 Thematic Mapper (TM) and Landsat 8 Operational Land Image (OLI) data from 1988 to 2015 are used to extract the ISF using the linear spectral mixture analysis (LSMA) model and normal difference build-up index (NDBI) at the sub-pixel scale. DMSP/OLS NTL data from 1992 to 2013 are calibrated to illustrate the urban nighttime light conditions at the regional scale. Urban expansion directions are identified by statistics and kernel density analysis for the ISF study area at the sub-pixel scale, and the correlation between the ISF and

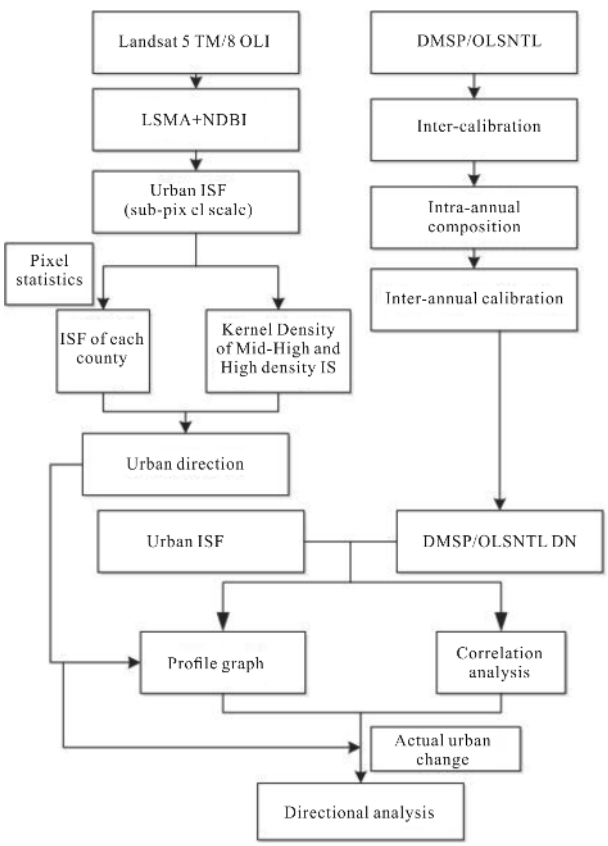

Fig. 2 Flow diagram of the study

DMSP/OLS NTL data are illustrated by linear regression analysis. Furthermore, Profile Graph in ArcGIS is employed to illustrate the urban expansion from the differences of correlation in different directions.

3.1 ISF Extraction

3.1.1 Linear Spectral Mixture Analysis (LSMA)

LSMA has a good theoretical basis and computational framework for treating mixed pixels in actual remote sensing images (Xu and Wang, 2016). It is frequently applied to mid-resolution images, with the assumption that the reflectance of a mixed pixel is a linear combination of the reflectance of its endmembers (Zhang and Zhao, 1999; Wu, 2004).

$$\begin{cases} R_i = \sum_{k\text{-}1}^{n} f_k R_{ik} + ER_i \\ \sum_{k=1}^{n} f_k = 1 \\ f_k \geq 0 \end{cases} \tag{2}$$

where i = 1, 2, …, M, M is the band number, and n is the number of endmembers. R is the reflectance of band I, f is the fraction of endmember K, and ER_i is the residual error of band i (Wang et al., 2009).

Before LSMA, the modified normalized difference water index (MNDWI) (Xu, 2005) was applied for water elimination:

$$MNDWI = \frac{B_{\text{GREEN}} - B_{\text{MIR}}}{B_{\text{GREEN}} + B_{\text{MIR}}} \tag{3}$$

where B_{GREEN} is green band, and B_{MIR} is middle infrared band.

RMS is applied to check the accuracy of LSMA (Zhou and Xu, 2007), whereby smaller RMS corresponds with higher LSMA precision:

$$RMS = \left(\frac{\sum_{i=1}^{M} ER_i^2}{M}\right)^{\frac{1}{2}} \tag{4}$$

where ER_i is the residual and M is the number of the spectral bands.

It is assumed that a mixed pixel consists of vegetation, soil, and the high-albedo and low-albedo in the LSMA unmixed results. Urban IS is assumed to be a combination of the high albedo and low albedo (Wu and Murray, 2003). Within the end-member selection, in which mixed pixels and the interrelationships of different spectra exist, the ISF in some areas can be overrated by LSMA where the ISF is lower than 20% (Zhu et al., 2013).

3.1.2 Normalized Difference Building Index (NDBI)

NDBI is usually used to distinguish buildings in urban land cover:

$$NDBI = \frac{B_{\text{MIR}} - B_{\text{NIR}}}{B_{\text{MIR}} + B_{\text{NIR}}} \tag{5}$$

where the B_{MIR} is middle infrared band, and B_{NIR} is near infrared band.

In this study, the NDBI threshold to remove the abnormal area at the pixel scale masked the combination of the high-albedo fraction and low-albedo fraction. Theoretically, pixels with an NDBI value greater than 0 correspond to buildings. However, the interrelationships of spectra in mixed pixels differ with respect to differences in land cover in each image, and the NDBI and LSMA are based on different scales. Therefore, the NDBI thresholds for masking the ISF results of LSMA for each remote sensing image in this study were not the same.

3.2 Urban Expansion Analysis

3.2.1 Correlation analysis

Linear regression is a statistical analysis method that uses regression analysis in mathematical statistics to determine the quantitative relationship between two or more variables. The linear regression is expressed as:

$$y = ax + \text{b} \tag{6}$$

In regression analysis, only one independent variable x and one dependent variable y are included, and the relationship between them can be approximately represented by scatter diagram and the parameters a and b. In this study, the ISF was taken as the independent variable and the NTL DN was taken as the dependent variable (Xu, 2009).

3.2.2 ISF classification

To reveal the dynamics of urbanization in different areas, all area units were classified into four groups, according to the ISF distribution in the study area, empirical thresholds, and studies showing the dominant components of the four categories, as shown in Table 3 (Gao et al., 2010):

Table 3 Impervious surface fraction (ISF) classification

Categories of land use	Low density ISF	Mid-low density ISF	Mid-high density ISF	High density ISF
ISF (%)	0–30	30–50	50–80	80–100
Dominant feature components	Agricultural land and urban green space	Residential land and a spot of roads	Roads and commercial land	Commercial land and industrial storage sites

In this study, the variation tendency of urban IS was analyzed by counting the different grades of urban ISF for each time phrase.

3.2.3 Kernel Density Analysis

Kernel density analysis was implemented in the ArcGIS platform to calculate the density of mid-high density IS and high density IS points around each output pixel. In theory, there is a smooth surface above each point, and the value at the location of that point is higher than that at distances farther away from the point. The value at the distance equal to the search radius is zero. Kernel density (Wei et al., 2015) is expressed as:

$$P(x)=\frac{1}{nr}\sum_{i=1}^{n}\left\{K\left[\frac{d(x,x_i)}{r}\right]\right\} \tag{7}$$

where n is the number of IS points within a circular radius of r, K is the kernel density function, and d is the Euclidean distance between two points.

In this study, the spatial resolution of original Landsat images (OLI and TM) is 30 m × 30 m, and 300 m × 300 m fishnets were taken to count the mid-high density IS and high density IS in each net to simplify the calculation. By format conversion, ISF images were taken for kernel density analysis, the output pixel size is set to 174, and the circular radius r is 1450 m by default.

3.2.4 Profile Graph

Profile graphs from ArcGIS consist of a continuous horizontal distance on the x-axis and ISF and NTL DN values on the y-axis. The horizontal distance refers to the distance from the beginning of each 3D line when projected onto an imaginary horizontal plane, measured along the line. The profile graphs of ISF and the NTL DN illustrate the directions of the urban expansion.

4 Results and Discussion

4.1 Urban ISF Extraction

The results of ISF extraction are shown in Fig. 3. Statistically, the IS areas from 1988 to 2015 are 70.327, 157.698, 260.506, 50.363, 472.004, and 580.819 km^2. The Root Mean Square (RMS) of LSMA and the NDBI thresholds for multi-temporal images are shown in Table 4. The values of RMS are no more than 0.020, and the accuracy meets the research precision requirements. In addition, high spatial resolution remote sensing images (1 m × 1 m; 2015; 2009) were used to test the combination of LSMA and NDBI using root mean square error (RMSE), systematic error (SE), mean absolute error (MAE), and correlation coefficient (R) using 100 random samples (480 m × 480 m). After vectorization and statistics, the results show that the method of LSMA combined with NDBI could extract ISF with high precision and meet the research needs (Fig. 4, Table 5).

4.2 Directional Analysis of Urban Expansion

4.2.1 ISF statistics and kernel density analysis

Due to differences in initial states, the 2015 administrative map was used to count the IS area of each county in the multi-temporal images. From Fig. 5a, it is evident that the IS area of Yuexiu increased most slowly, followed by Liwan and Huangpu, as these three counties are examples of old urban zones with rapid development and higher economic saturation occurring in earlier years. The IS area in the new urban zones increased much faster than that in the old urban zones. With the strategic layout of Guangzhou, Tianhe has developed as the strongest economic zone in Guangzhou over the past thirty years, with a number of business centers forming after the National Games in 1987. In Huangpu, which merged with Luogang in 2014, many business zones appeared with the opening of the subway. Baiyun, which previously consisted of large forests and rural areas, was an important developing area in 1994–2001. The IS area of Baiyun sharply increased in 2001, 2009, and 2015, in relation to the urban and rural planning of Guangzhou. The results of ISF classification and statistical analysis (Fig. 5b) show that from 1994 to 2015, the mid-low density IS area was comparatively stable, only increasing by 57.010 km^2. The low density impervious surface area decreased by 668.610 km^2, the high density IS area increased by 321.140 km^2, and the mid-high density IS area increased by 290.460 km^2.

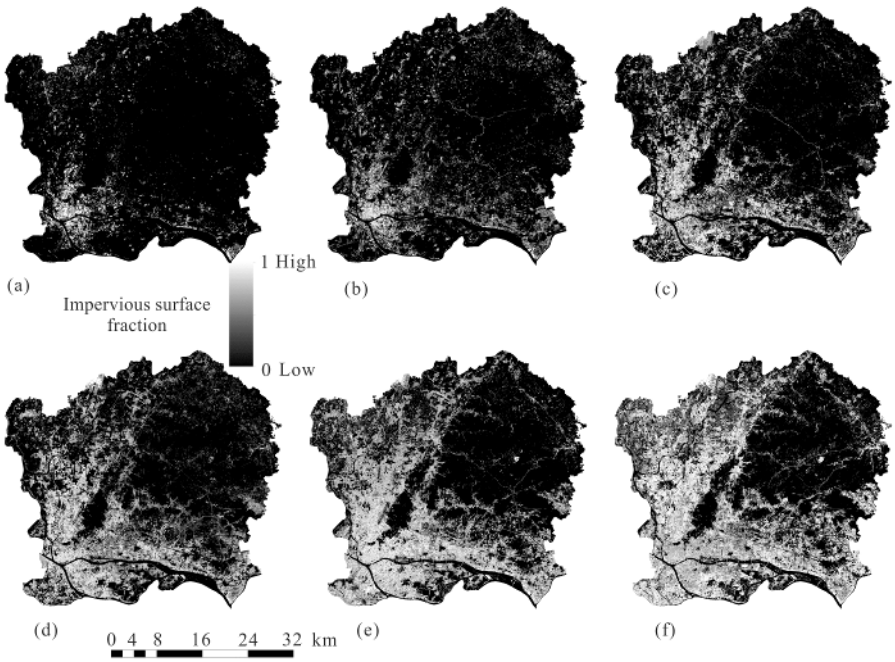

Fig. 3 Impervious surface fraction (ISF) by LSMA and NDBI: (a) 1988; (b) 1994; (c) 2001; (d) 2004; (e) 2009; (f) 2015

Table 4 Parameters of ISF extraction

Parameter	1988	1994	2001	2004	2009	2015
RMS	0.012	0.013	0.010	0.011	0.020	0.009
NDBI Threshold	–0.050	0	–0.100	–0.100	–0.150	–0.150

Many commercial areas (high density IS areas) have been formed with urbanization, as the kernel density analysis results show (Fig. 6). The high density IS areas have been mainly located in old urban zones such as Yuexiu in 1988. In the east, between the old urban zones and southeastern Huangpu, the value of the kernel density increases toward the center from both sides, while northeast of the old urban zones, the value of the kernel density increases towards the northeastern Baiyun. The number of ISF accumulated areas and polycentric urban structures has increased in the time series kernel density results. By 2001, the high-density IS had diffused to the east and northeast, with Yuexiu, northern Liwan, and western Haizhu, which correspond to the old urban zones, serving as the starting point. In addition, many roads appeared while urban development gradually shifted into a boom period, and the economy increased sharply in the old urban zones, approaching saturation.

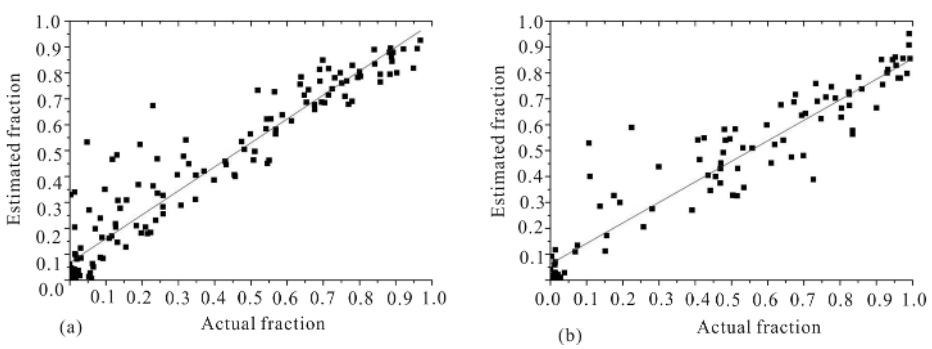

Fig. 4 Precision test result using 100 random areas: a) Charts of ISF from high resolution (*x*-axis) and ISF by LSMA and NDBI (*y*-axis) in 2009; b) Charts of ISF from high resolution (*x*-axis) and ISF by LSMA and NDBI (*y*-axis) in 2015

Table 5 Accuracy assessment of ISF extraction in 2009 and 2015

Year	RMSE	SE	MAE	R^2	R
2009	0.110	0.035	0.065	0.906	0.952
2015	0.120	–0.035	0.092	0.890	0.943

The directions of urbanization in the study area proceed toward the east and northeast from Yuexiu, northern Liwan, and western Haizhu. Based on Guangzhou urban planning (extend the city to the east, construct economical cores, such as Zhujiang New Town, and maintain ecological environment protection, such as ecotourism development in the north), the results of the ISF statistics and kernel density analysis are highly credible. In the east, the directional property of urban expansion is weaker than that in the northeast, because southern Huangpu developed in the early years, in the same period as the old urban zones.

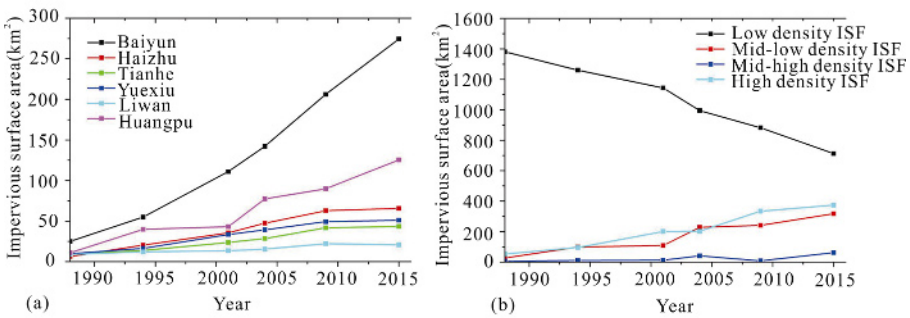

Fig. 5 Chart of the urban IS variation tendency in Guangzhou City: a) IS area statistics of six counties; b) IS density statistics in the study area

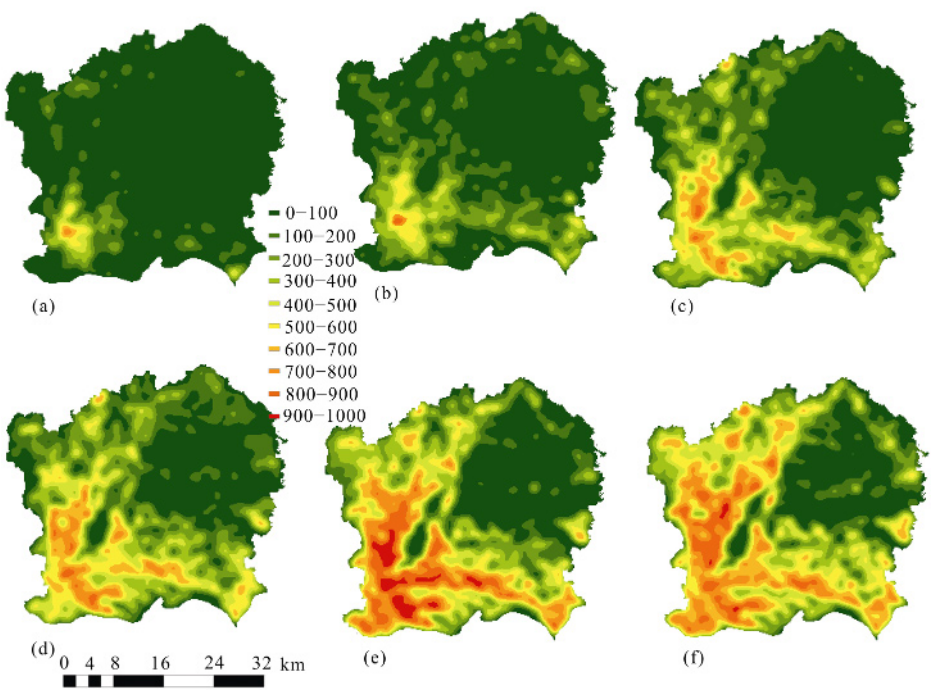

Fig. 6 Kernel density of mid-high density impervious surface (IS) and high density IS: a) 1988; b) 1994; c) 2001; d) 2004; e) 2009; f) 2015

4.2.2 Linear regression of ISF and DMSP/OLS NTL data

DMSP/OLS NTL data and ISF are significant references for urbanization. Linear regression analysis between DMSP/OLS NTL data and ISF were carried out in SPSS software.

As the pixel size of the DMSP/OLS NTL data were resampled to 900 m × 900 m, much lower than that of the ISF data (30 m × 30 m), 100 random areas (1800 m × 1800 m) were selected in each direction from the two datasets to count the mean values for regression analysis. Considering the saturation and blooming effects of the DMSP/OLS NTL data, the samples were chosen at the edge of the busiest downtown area (Fig. 7), based on experiments.

The results are shown in Tables 6 and 7 as well as Figs. 7 and 8. In the northeast, R^2 is greater than 0.43, and the correlation coefficient R is greater than 0.65 (Table 6), implying a large positive correlation. In the east, R^2 is greater than 0.33, so the correlation coefficient R is greater than 0.53 (Table 6), illustrating a weaker positive correlation than that in the northeast. The results are consistent with the above results of the ISF statistics and kernel density analysis.

As shown in Fig. 8, most samples with higher NTL DN have a smaller ISF value. From Figs. 8a and 8b, the linear regressions of NTL1992_1995/IS1994 and NTL1996_2000/IS2001 perform more like a conic fitting. Based on Fig. 8c, after 2004, the DMSP/OLS NTL data and ISF maintained a better positive correlation, and the samples with high values were clustered together. With the progress of urban expansion, the performance of the DMSP/OLS NTL data and ISF, with respect to urbanization, became analogous, illustrating that in the early years, the urban IS area was less prominent. However, because of the saturation and blooming effect, the NTL DN had a higher value, and with the increase in IS, the positive correlation became significant.

As seen in Fig. 9, the linear regression of NTL1992_1995/ISF1994 and NTL1996_2000/ISF2001 performs more like a conic fitting. Most deviated samples have a higher NTL DN. In the regression chart of NTL2011_2013 and IS2015 (Fig. 9e), most samples have an NTL DN over 60 and ISF values above 0.6. Compared with Fig. 7d and 7e, the NTL DN in Fig. 8d and 8e are higher, showing that the correlation of DMSP/OLS NTL data and ISF data are weaker (Table 7).

From Table 6 and Table 7, the DMSP/OLS NTL data are consistent with ISF data (Tables 6 and 7). From Fig. 8 and Fig. 9, the distribution of most samples from the two-scale data have similar tendencies. These results show that the DMSP/OLS NTL data and ISF data have a certain positive correlation for urban expansion. However, because of the different development and data properties, the correlations in the two directions are different.

4.2.3 Profile Graph

In this study, Profile Graph on the ArcGIS platform was employed in two directions to illustrate the differences in the correlation between ISF and DMSP/OLS NTL data. To simplify the calculation and make it fit with the DMSP/OLS NTL data, the spatial resolution of the ISF was resampled to 900 m × 900 m by calculating the average pixel value. The profile graphs show the values

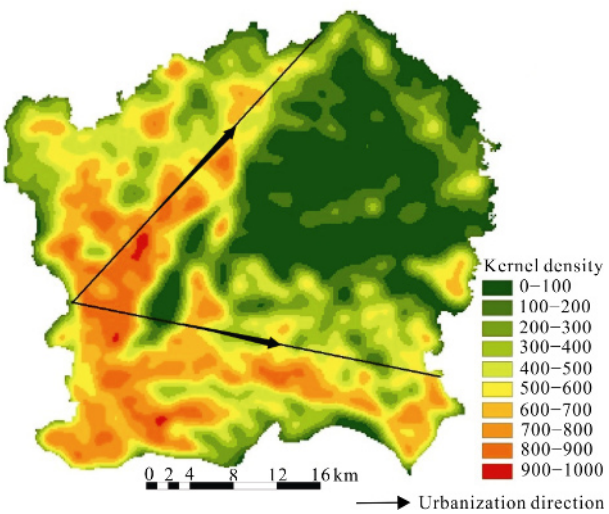

Fig. 7 The diagram of urbanization directions

Table 6 Linear regression coefficients of DMSP/OLS NTL data and ISF result in the northeast NI

DMSP/OLS NTL data/ISF result	a	b	R^2	R
NTL1992_1995/ISF1994	97.239	23.820	0.434	0.659
NTL1996_2000/ISF2001	69.290	25.511	0.542	0.736
NTL2001_2005/ISF2004	66.598	24.348	0.552	0.743
NTL2006_2010/ISF2009	46.215	28.704	0.564	0.751
NTL2011_2013/ISF2015	39.632	30.980	0.513	0.716

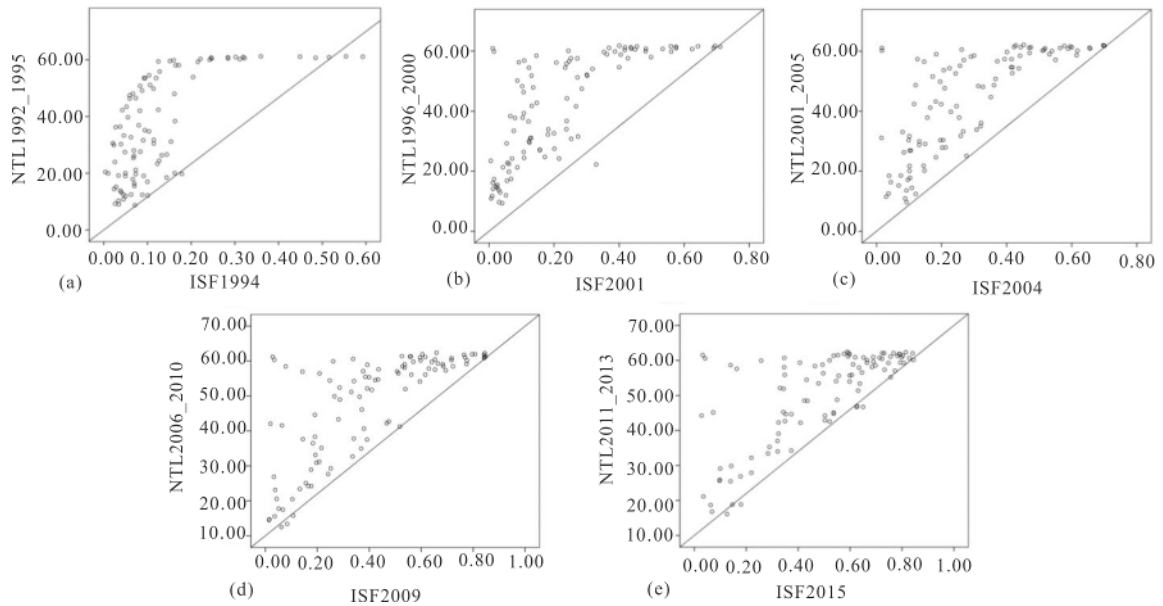

Fig. 8 Regression chart of DMSP/OLS NTL data and ISF by OLI and TM with respect to 100 random areas in the northeast from the starting area: a) NTL1992_1995and IS 1994; b) NTL1996_2000 and IS 2001; c) NTL2001_2005 and IS2004; d) NTL2006_2010 and IS 2009; e) NTL2011_2013 and IS 2015

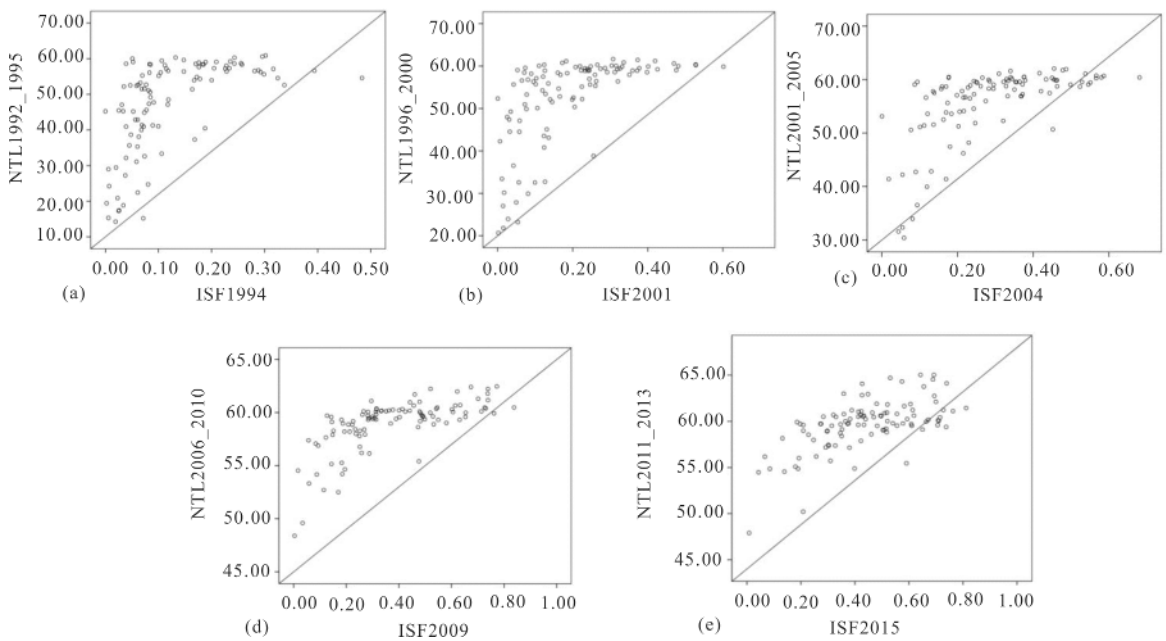

Fig. 9 Regression scatter plot of DMSP/OLS NTL data and ISF using OLI and TM data for 100 random areas in the east from the starting area: a):NTL1992_1995 and ISF 1994; b): NTL1996_2000 and ISF 2001; c): NTL2001_2005 and ISF2004; d): NTL2006_2010 and ISF 2009; e): NTL2011_2013 and ISF 2015

of the ISF and DMSP/OLS NTL data in the two directions (Fig. 7), from the urbanization starting position (the old urban zones) (Figs. 9–12). In the early years, the values of ISF and NTL DN in the old urban zones increased quickly. As a result of economic development and urban planning, the values of ISF and DMSP/OLS NTL data in the new urban zones have increased faster than those in the old urban zones. Meanwhile, the values

Table 7 Linear regression coefficients of DMSP/OLS NTL data and ISF data in the east

DMSP/OLS NTL data/ISF	a	b	R^2	R
NTL1992_1995/ISF1994	75.942	37.724	0.332	0.576
NTL1996_2000/ISF2001	45.311	43.552	0.355	0.596
NTL2001_2005/ISF2004	28.962	47.209	0.382	0.618
NTL2006_2010/ISF2009	8.122	55.849	0.444	0.667
NTL2011_2013/ISF2015	4.760	57.726	0.393	0.627

of ISF and NTL DN in the old urban zones have increased slowly until becoming stable, illustrating a saturation state. As a whole, the NTL DN curves (Fig. 12) trends are analogous to the curve trends of the ISF profiles. However, as the spatial resolution of the DMSP/OLS NTL data are much lower than that of the ISF, the curves of the NTL DN profiles are relatively smooth with less fluctuation, compared with the curves of the ISF profiles.

In the northeast, the ISF values in 1988 were low (no more than 0.1) in general (Fig. 10a). From the starting position, the urban IS area significantly increased (maximum increased to greater than 0.3), as the land cover changed from 1988 to 1994 (Fig. 10b). Numerous business centers and more high density IS areas appeared in the old urban zones (Yuexiu, Liwan, and Haizhu) because the urbanization process accelerated, yielding high ISF values over a short distance (maximum increased to 0.6 and most pixel values were greater than 0.1) (Fig. 10c). As seen in Fig. 10, the first crest before 5 km indicates that the old urban zones developed early and fast. In the earlier years, there were large areas of agricultural production. The troughs within distances of 10, 17, and 28 km on the *x*-axis represent the farming regions where the ISF values are lower and the range is shrunk, with the development of buildings and roads. Near 13 and 15 km, there was a crest, representing the business circles where the IS areas increased sharply from the northward extension of subway Line 3 in Guangzhou in 2010 and where different grades of residential areas were shaped. The urbanization process was significant along the profile, and the IS coverage increased annually with distance from the starting position. All the waves of the profile curves are consistent with the actual conditions of Guangzhou. Performance continuously increased significantly, exemplified by the crest at a distance of 30 km, which increased from less than 0.1 in 1988 to 0.7 in 2015, and the trough at a distance of 20 km, which increased from less than 0.1 before 2009 to greater than 0.5 in 2015. Compared with the ISF of the old urban zones, the ISF of the new urban zones maintained an upward trend at long distances soon after 2001. Meanwhile, the development of the old urban zones was gentle, and the ISF rates of increase were slow.

As shown in Fig. 11, the DN value of NTL DN data in the old urban zones are relatively higher than those in the new zones, and the curve trends are similar to those of the ISF data. In Fig. 11, the curves of (a) and (b) at 10, 17, and 28 km exhibit troughs, and the curves between 13 and 15 km exhibit a crest. In Fig. 11(a), (b), and (c), there is a significant trough at 20 km and a crest at 30 km. However, from the profile graph of NTL2006_2010, the curves become smooth, and the values are greater than 40. On account of the saturation and blooming effects of the DMSP/OLS NTL data, the locations of the trough and crest on the *x*-axis are relatively farther, compared with the curves of the ISF, even to the point of disappearing. In the early years, the NTL DN in the old urban zones was relatively higher, different from the curves of the ISF because the old urban zones were under construction, and the population was large. As the IS increased, the curves of the NTL DN and the ISF became similar. Nevertheless, urbanization is accelerating, and the correlation between the two kinds of data has weakened, because of the saturation and blooming effects, causing the troughs and the crest of the NTL DN curves to disappear.

From Fig. 12, the increase in the trend of the ISF values in the old urban zones in the east is similar to that in the northeast. The ISF profile graphs show that in the new urban zones, the urban area is expanding, similar to the conditions in the northeast. In multi-temporal ISF profiles graphs, there is an apparent trough at a distance of 10 km (Fig. 12). The trough represents Baiyun Mountain, the famous beautiful spot in the city, where vegetation coverage is high, and the ecological environment is protected well. Between 10 and 20 km, there is an apparent crest, which represents Tianhe, a new county. In Tianhe, the economy has developed rapidly, and the IS area has increased rapidly. From Fig. 12b,

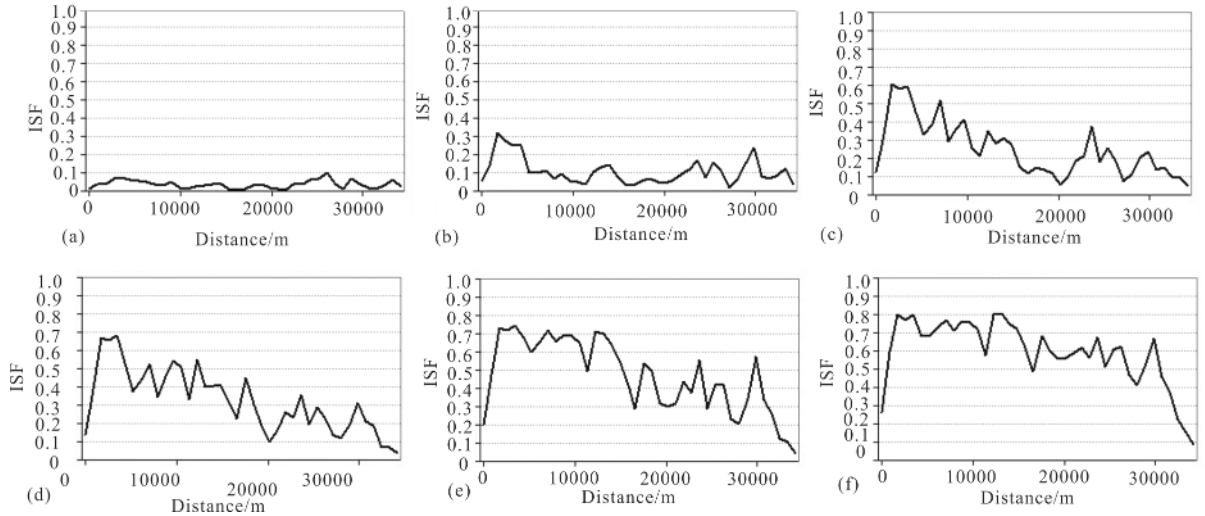

Fig. 10 Time series profile of impervious surface fraction in the northeast: a) 1988, b) 1994, c) 2001, d) 2004, e) 2009, f) 2015

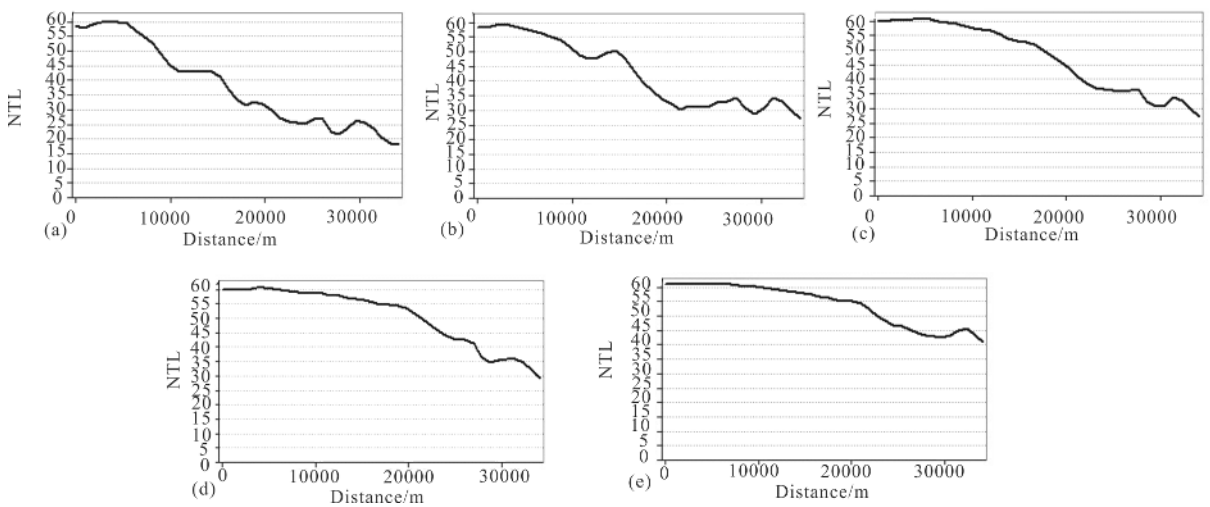

Fig. 11 Time series profile of DMSP/OLS NTL data in the northeast: a) NTL1992_1995, b) NTL1996_2000, c) NTL2001_2005, d) NTL2006_2010, e) NTL2011_2013

there is a crest after 30 km, which represents southern Huangpu. With the development of the new urban zones, the ISF in the southern Huangpu has increased, and the overall trend is low in the middle and high on both sides, illustrating that in the east, the directivity of urban expansion is weaker than that in the northeast.

In the east, the NTL DN increases suddenly and is inconsistent with the actual conditions (Fig. 13). From Fig. 13a, there is a small trough representing Baiyun Mountain, where the urban IS area is smaller, aligning with Fig. 12. However, the NTL DN is higher. There is a smaller crest between 10 and 20 km that disappears in Fig. 13b, which does not correspond with Fig. 12. The trough between 20 and 30 km is roughly in line with Fig. 12, indicating that the intermediate region develops later. From Fig. 13b, there is higher NTL DN, after 30 km, indicating the development of southern Huangpu, as shown in Fig. 12. Over time, the NTL DN profile curves become smooth, and the crests and troughs gradually disappear.

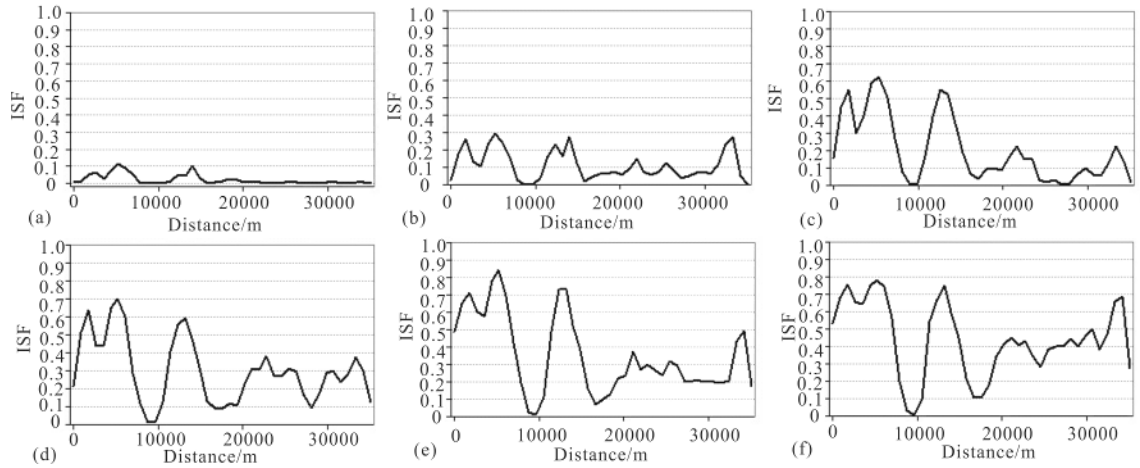

Fig. 12 Time series profile of ISF in the east: (a) 1988, (b) 1994, (c) 2001, (d) 2004, (e) 2009, (f) 2015

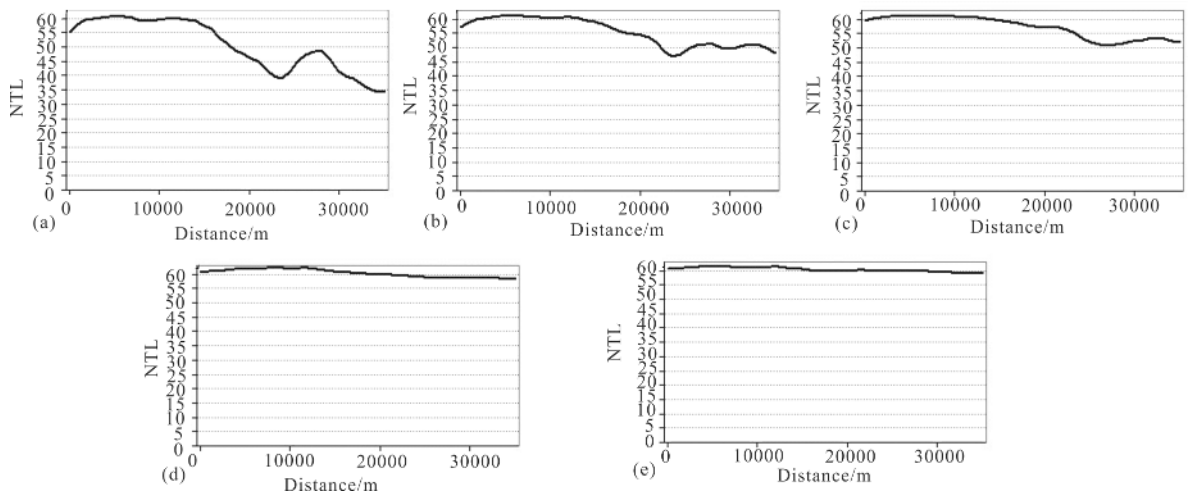

Fig. 13 Time series profile of DMSP/OLS NTL data in the east: (a) NTL1992_1995, (b) NTL1996_2000, (c) NTL2001_2005, (d) NTL2006_2010, (e) NTL2011_2013

Pixels generally have higher values for NTL DN, because of the saturation and blooming effects, just as in Fig. 13. With respect to the urban differences in the two directions, there is a weaker correlation in the east because of the saturation and blooming effects in the DMSP/OLS NTL data and the different development patterns of urbanization. In the northeast, the directional property of urbanization is stronger than that in the east. In the east, the urbanization is more dispersed, and southern Huangpu developed early, causing the urbanization to progress gradually from both sides toward the center. This type of urbanization pattern makes the saturation and blooming effects of the DMSP/OLS NTL data more serious.

5 Conclusions

In this study, linear regression from ArcGIS was employed to indicate the positive correlation between the DMSP/OLS NTL data and ISF data in the directions of urban expansion, obtained by kernel density analysis of ISF on the sub-pixel scale. Using Profile Graph, the two-scale data performs differently as a result of the data properties and different urban expansion patterns. Through the analysis of these differences and the typical directional urban development of the study area, it is identified that the ISF and the DMSP/OLS NTL data have similar functions in describing urban development. The main findings are as follows:

First, more than 668 km^2 of the low-density urban IS in the study area has been replaced by higher density IS areas, particularly by high-density urban IS areas. The rates of increase of the IS area in Yuexiu, Liwan, and Haizhu are lower than those of Tianhe, Baiyun, and Huangpu. Kernel density analysis shows that urbanization proceeds toward the east and northeast from the old urban zones, illustrating a polycentric urban structure.

Next, based on linear regression analysis, there is a positive correlation between the ISF and DMSP/OLS NTL data. In addition, the Profile Graph results show varied performance for the ISF and DMSP/OLS NTL data in describing urbanization. Based on the profile graphs, the overall trends of the two datasets are similar, but the details show that the DMSP/OLS NTL data overestimates the urban scope. There is a deviation in the correlation between the ISF and DMSP/OLS NTL data in different directions because of the saturation and blooming effects of the DMSP/OLS NTL data and the patterns of urbanization.

In addition, the ISF and DMSP/OLS NTL data are urban parameters at different scales. DMSP/OLS NTL data are usually used for large areas. According to this study, DMSP/OLS NTL data may be used as an important reference dataset for small-scale regions that develop directionally and belong to developing areas.

Based on the above analysis, urban ISF and the DMSP/OLS NTL data can be combined for different-scale urbanization in which the DMSP/OLS NTL data could play a more constructive role. Despite a great deal of pre-processing and corrections performed for DMSP Version 4 images, a known saturation effect remains at higher levels in the stable light band (Letu et al., 2010).

There has been much work conducted to enhance the DMSO/OLS data, and parameters such as ISF may be applied as references to correct this data. DMSP/OLS NTL data may be used for urban analysis and social parameters on a small scale in future studies.

References

Brodley C E, 1995. Recursive automatic bias selection for classifier construction. *Machine Learning*, 20(1–2): 63–94. doi: 10.1007/bf00993475

Chen Yunhao, Feng Tong, Shi Peijun et al., 2006. Classification of remot sensing image based on object oriented and class rules. *Geomatics and Information Science of Wuhan University*, 31(4): 316–320. (in Chinese)

Deng C B, Wu C S, 2012. BCI: a biophysical composition index for remote sensing of urban environments. *Remote Sensing of Environment*, 127: 247–259. doi: 10.1016/j.rse.2012.09.009

Elvidge C D, Tuttle B T, Sutton P S et al., 2007. Global distribution and density of constructed impervious surfaces. *Sensors*, 7(9): 1962–1979. doi: 10.3390/s7091962

Fan F L, Fan W, Weng Q H, 2015. Improving urban impervious surface mapping by linear spectral mixture analysis and using spectral indices. *Canadian Journal of Remote Sensing: Journal Canadien de Télédétection*, 41(6): 577–586. doi: 10.1080/07038992.2015.1112730

Fang C L, Ma H T, Wang J, 2015. A regional categorization for 'new-type urbanization' in China. *PLoS One*, 10(8): e0134253. doi: 10.1371/journal.pone.0134253

Fragkias M, Güneralp B, Seto K C et al., 2013. A synthesis of global urbanization projections. In: Assessment G A, (ed.) *Urbanization*, *Biodiversity and Ecosystem Services: Challenges and Opportunities*, Dordrecht: Springer, 409–435. doi: 10.1007/978-94-007-7088-1_21

Fu H Y, Shao Z F, Fu P et al., 2017. The dynamic analysis between urban nighttime economy and urbanization using the DMSP/OLS nighttime light data in China from 1992 to 2012. *Remote Sensing*, 9(5): 416. doi: 10.3390/rs9050416

Gao Zhihong, Zhang Lu, Li Xinyan et al., 2010. Detection and analysis of urban land use changes through multi-temporal impervious surface mapping. *Journal of Remote Sensing*, 14(3): 593–606. (in Chinese)

Haase D, Nuissl H, 2010. The urban-to-rural gradient of land use change and impervious cover: a long-term trajectory for the city of Leipzig. *Journal of Land Use Science*, 5(2): 123–141. doi: 10.1080/1747423X.2010.481079

Letu H, Hara M, Yagi H et al., 2010. Estimating energy consumption from night-time DMPS/OLS imagery after correcting for saturation effects. *International Journal of Remote Sensing*, 31(16): 4443–4458. doi: 10.1080/01431160903277464

Li Deren, 2015. An overview on data mining of nighttime light remote sensing. *Acta Geodaetica et Cartographica Sinica*, 44(6): 591–601. (in Chinese)

Li Xinyu, 2015. *On the Urban Growth of Jiangsu Province from 1985–2014 Based on Impervious Surface Information from Remote Sensing Imagery*. Nanjing: Nanjing University. (in Chinese)

Liu Z F, He C Y, Zhang Q F et al., 2012a. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landscape and Urban Planning*, 106(1): 62–72. doi: 10.1016/j.landurbplan.2012.02.013

Liu Zhenhuan, Wang Yanglin, Peng Jian, et al., 2012b. Quatifying spatiotemporal patterns dynamics of impervious surface in Shenzhen. *Geographical Research*, 31(8): 1535–1545. (in Chinese)

Lu D S, Weng Q H, 2006. Use of impervious surface in urban land-use classification. *Remote Sensing of Environment*, 102(1–2): 146–160. doi: 10.1016/j.rse.2006.02.010

Lu D S, Weng Q H, 2009. Extraction of urban impervious surfaces from an IKONOS image. *International Journal of Remote Sensing*, 30(5): 1297–1311. doi: 10.1080/ 0143116080

2508985

Pan Jinghu, Li Xiaoxue, Feng Zhaodong et al., 2010. Analysis of Spatial and Temporal Patterns of Impervious Surfaces and Vegetation Covers in Lanzhou Based on the V-I-AP Model. *Resources Science*, 32(3): 520–527. (in Chinese)

Ridd M K, 1995. Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *International Journal of Remote Sensing*, 16(12): 2165–2185. doi: 10.1080/01431169508954549

Su Yongxian, 2015. *Study on the Carbon Emissions from Energy Consumption in China Using DMSP/OLS Night Light Imageries*. Guangzhou: Guangzhou Institute of Geochemistry, Chinese Academy of Sciences. (in Chinese)

Wang Qinjun, Lin Qizhong, Li Mingxiao et al., 2009. Comparison of two spectral mixture analysis models. *Spectroscopy and Spectral Analysis*, 29(10): 2602–2605. (in Chinese)

Wang W, Yao X F, Ji M, 2016. Integrating seasonal optical and thermal infrared spectra to characterize urban impervious surfaces with extreme spectral complexity: a Shanghai case study. *Journal of Applied Remote Sensing*, 10(1): 016018. doi: 10.1117/1.jrs.10.016018

Wei Haiyang, Jing Changfeng, Du Mingyi, 2015. Kernel density analysis of different scales for distributiontrend of urban. *Urban Geotechnical Investigation & Surveying*, (1): 18–20. (in Chinese)

Weng Q H, Lu D S, 2008. A sub-pixel analysis of urbanization effect on land surface temperature and its interplay with impervious surface and vegetation coverage in Indianapolis, United States. *International Journal of Applied Earth Observation and Geoinformation*, 10(1): 68–83. doi: 10.1016/j.jag.2007.05.002

Wu C S, Murray A T, 2003. Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment*, 84(4): 493–505. doi: 10.1016/S0034-4257(02)00136-0

Wu C S, 2004. Normalized spectral mixture analysis for monitoring urban composition using ETM + imagery. *Remote Sensing of Environment,* 93(4): 480–492. doi: 10.1016/j.rse.2004.08. 003

Xie Y H, Weng Q H, 2016. Updating urban extents with nighttime light imagery by using an object-based thresholding method. *Remote Sensing of Environment,* 187: 1–13. doi: 10.1016/j.rse.2016.10.002

Xie Y H, Weng Q H, 2017. Spatiotemporally enhancing time- series DMSP/OLS nighttime light imagery for assessing large-scale urban dynamics. *Isprs Journal of Photogrammetry and Remote Sensing*, 128: 1–15. doi: 10.1016/j.isprsjprs.2017. 03.003

Xin X, Liu B, Di K et al, 2017. Monitoring urban expansion using time series of night-time light data: a case study in Wuhan, China. *International Journal of Remote Sensing, 38*(21): 6110–6128. doi: 10.1080/01431161.2017.1312623

Xu Hanqiu, 2005. A study on information extraction of water body with the modified normalized difference water index (MNDWI). *Journal of Remote Sensing*, 9(5): 589–595. (in Chinese)

Xu Hanqiu, 2008. A new remote sensing index for fastly extracting impervious surface information. *Geomatics and Information Science of Wuhan University*, 33(11): 1150–1153. (in Chinese)

Xu Hanqiu, 2009. Quantitative analysis on the relationship of urban impervious surface with other components of the urban ecosystem. *Acta Ecologica Sinica*, 29(5): 2456–2462. (in Chinese)

Xu H Q, 2010. Analysis of impervious surface and its impact on urban heat environment using the normalized difference impervious surface index (NDISI). *Photogrammetric Engineering & Remote Sensing*, 76(5): 557–565. doi: 10.14358/pers.76.5.557

Xu Hanqiu, Wang Meiya, 2016. Remote sensing-based retrieval of ground impervious surfaces. *Journal of Remote Sensing*, 20(5): 1270–1289. (in Chinese)

Xu J H, Zhao Y, Zhong K W et al., 2016. Coupling modified linear spectral mixture analysis and soil conservation service curve number (SCS-CN) models to simulate surface runoff: application to the main urban area of Guangzhou, China. *Water*, 8(12): 550. doi: 10.3390/w8120550

Xu R, Zhang H S, Lin H, 2017. Urban impervious surfaces estimation from optical and sar imagery: a comprehensive comparison. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(9): 4010–4021. doi: 10.1109/JSTARS.2017.2706747

Yuan Linshan, Du peijun, Zhang Huapeng et al., 2008. CBERS imagery classification based on decision tree and derformance analysis. *Remote Sensing for Land & Resources*, (2): 92–98. (in Chinese)

Zhang Xichuan, Zhao Yingshi, 1999. Application of line spectral mixture to rapid assessment of land degradation in semiarid area. *Journal of Graduate School*, *Academia Sinica*, 16(2): 169–176. (in Chinese)

Zhang H S, Lin H, Zhang Y et al., 2015. *Remote Sensing of Impervious Surfaces in Tropical and Subtropical Areas.* Boca Raton: CRC Press, 14–20.

Zhou Cunlin, Xu Hanqiu, 2007. A spectral mixture analysis and mapping of impervious surfaces in built-up land of Fuzhou city. *Journal of Image and Graphics*, 12(5): 875–881. (in Chinese)

Zhu Aili, Lv Chengwen, 2010. Advances in the methods of extracting urban impervious surface based on remote sensing. *Journal of Anhui Normal University* (*Natural Science*), 33(5): 485–489 (in Chinese)

Zhu H L, Ying L, Fu B L, 2013. Estimating impervious surfaces by linear spectral mixture analysis under semi-constrained condition. In: *RSETE 2013.* Atlantis Press, 357–360. doi: 10.2991/rsete.2013.87