Chin. Geogra. Sci. 2017 Vol. 27 No. 4 pp. 614–625
doi: 10.1007/s11769-017-0882-x

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

Integrating CART Algorithm and Multi-source Remote Sensing Data to Estimate Sub-pixel Impervious Surface Coverage: A Case Study from Beijing Municipality, China

HU Deyong[1], CHEN Shanshan[1], QIAO Kun[2], CAO Shisong[1]

(1. *College of Resource Environment and Tourism*, *Capital Normal University*, *Beijing* 100048, *China*; 2. *College of Resources Science and Technology*, *Beijing Normal University*, *Beijing* 100875, *China*)

Abstract: The sub-pixel impervious surface percentage (SPIS) is the fraction of impervious surface area in one pixel, and it is an important indicator of urbanization. Using remote sensing data, the spatial distribution of SPIS values over large areas can be extracted, and these data are significant for studies of urban climate, environment and hydrology. To develop a stabilized, multi-temporal SPIS estimation method suitable for typical temperate semi-arid climate zones with distinct seasons, an optimal model for estimating SPIS values within Beijing Municipality was built that is based on the classification and regression tree (CART) algorithm. First, models with different input variables for SPIS estimation were built by integrating multi-source remote sensing data with other auxiliary data. The optimal model was selected through the analysis and comparison of the assessed accuracy of these models. Subsequently, multi-temporal SPIS mapping was carried out based on the optimal model. The results are as follows: 1) multi-seasonal images and nighttime light (NTL) data are the optimal input variables for SPIS estimation within Beijing Municipality, where the intra-annual variability in vegetation is distinct. The different spectral characteristics in the cultivated land caused by the different farming characteristics and vegetation phenology can be detected by the multi-seasonal images effectively. NLT data can effectively reduce the misestimation caused by the spectral similarity between bare land and impervious surfaces. After testing, the SPIS modeling correlation coefficient (*r*) is approximately 0.86, the average error (*AE*) is approximately 12.8%, and the relative error (*RE*) is approximately 0.39. 2) The SPIS results have been divided into areas with high-density impervious cover (70%–100%), medium-density impervious cover (40%–70%), low-density impervious cover (10%–40%) and natural cover (0%–10%). The SPIS model performed better in estimating values for high-density urban areas than other categories. 3) Multi-temporal SPIS mapping (1991–2016) was conducted based on the optimized SPIS results for 2005. After testing, *AE* ranges from 12.7% to 15.2%, *RE* ranges from 0.39 to 0.46, and *r* ranges from 0.81 to 0.86. It is demonstrated that the proposed approach for estimating sub-pixel level impervious surface by integrating the CART algorithm and multi-source remote sensing data is feasible and suitable for multi-temporal SPIS mapping of areas with distinct intra-annual variability in vegetation.

Keywords: impervious surface; impervious surface percentage; classification and regression tree (CART); sub-pixel; sub-pixel impervious surface percentage (SPIS); time series

Citation: Hu Deyong, Chen Shanshan, Qiao Kun, Cao Shisong, 2017. Integrating CART algorithm and multi-source remote sensing data to estimate sub-pixel impervious surface coverage: a case study from Beijing Municipality, China. *Chinese Geographical Science*, 27(4): 614–625. doi: 10.1007/s11769-017-0882-x

1 Introduction

Impervious surfaces, in contrast with permeable vegetation and soil surfaces, are usually defined as the collection of anthropogenic features through which water can not infiltrate, typically including buildings, roads, park-

Received date: 2016-05-02; accepted date: 2016-09-01
Foundation item: Under the auspices of National Natural Science Foundation of China (No. 41671339)
Corresponding author: CHEN Shanshan. E-mail: amchenshanshan@163.com

ing lots, sidewalks and other built surfaces (Arnold and Gibbons, 1996). Regional environmental changes and urban landscape pattern changes are mainly dominated by the rapid increase in impervious surfaces produced by human-induced changes during rapid urbanization (Xiao *et al*., 2007; Imhoff *et al*., 2010; Li *et al*., 2011; Zhou *et al*., 2013). The area, coverage and spatial pattern of impervious surfaces can be identified as a key indicator of urbanization rates. Therefore, it is significant to assess impervious surfaces accurately.

Large-scale impervious surface mapping has been supported by the rapid development of earth observing technology. Landsat remote sensing data, which contains significant amounts of information, can be used for the long term observation of the land surface (Zhang and Weng, 2016). However, because of the limitations imposed by spatial resolution and the complexity of urban surfaces, the pixels of Landsat images in urban areas are often composed of several land-cover/land-use types, and the problem of mixed pixels is created (Weng, 2012). If the sub-pixel impervious surface percentage (SPIS) can be extracted accurately, the composition of heterogeneous landscapes (particularly urban landscapes), and the processes associated with them, can be better understood.

Many techniques have been applied to characterize and quantify impervious surfaces using either ground measurements or remotely sensed data (Weng, 2012). Currently, SPIS estimation methods are mainly based on statistical models, machine learning methods, *etc*., such as multiple regression methods (Yang, 2006; Jin, 2013), artificial neural network methods (Sanbum and Lathrop, 2006; Cao *et al.*, 2012; Patel and Mukherjee, 2014), decision tree methods (Yang *et al*., 2003; Xian and Crane, 2005), and spectral mixing models (Wu and Murray, 2003; Zhou *et al*., 2007; Wang *et al*., 2011). Among these methods, the decision tree method establishes SPIS prediction models based on a series of tree-structured decision rules, and has advantages in terms of its nonlinear studying capability, simple implementation and high operational efficiency in the regression of continuous variables. The classification and regression tree (CART) algorithm, a general decision tree method, can be applied different decision tree types, such as classification trees when the target variable is discrete and regression trees when the target variable is continuous (Breiman *et al*., 1984). CART inherits all the advantages of general decision trees, and not only can be used for classification but can also carry out the prediction and regression of continuous variables (Friedl *et al*., 1999; Lawrence *et al*., 2004). It has been successfully applied in the geological analysis and monitoring program (GAM) of the United States Geological Survey (USGS), and the data on impervious surfaces has been added to the US National Land Cover Database (Homer *et al*., 2015).

However, as a weak learning algorithm, CART also has limitations; it can be adversely affected by the learning ability of unbalanced input variables. The results of the CART algorithm are greatly influenced by the input variables. Several studies have been carried out. For example, Yang *et al*. (2003) used different input variables of different scale regions for SPIS modeling: 1) leaf-on data (Landsat 5 TM images from the growing season) are selected as the optimal input variable for local-scale areas; 2) bands 1, 4, 5, 6, 7 of the leaf-on data are selected for meso-scale areas; 3) band 6 of the leaf-on and leaf-off data (Landsat5 TM images from the non-growing season) and brightness, greenness and humidity as extracted using the tasseled cap transformation from the leaf-on and leaf-off data are selected for macro-scale area. The average error of the three prediction models at different scales was approximately 9.6%, 9.1% and 9.3%, respectively. Furthermore, four bands of Systeme Probatoire d'Observation de la Terre (SPOT) 5 high-resolution geometric (HRG) imagery and three parameters derived from European Remote Sensing (ERS)-2 single look complex (SLC) SAR images were used as input variables for the CART algorithm for SPIS estimation; the average error was approximately 12.9% and the correlation coefficient was approximately 0.77 (Yang *et al*., 2009). Gao *et al*. (2010) built an SPIS estimation model using six bands of Landsat 5 TM images, excluding the thermal-infrared band, and the average error was approximately 15.1%, the correlation coefficient was approximately 0.71, and the relative error was approximately 0.64. Zhang *et al*. (2010) showed that the near-infrared band (NIR band) had the highest contribution rate to the SPIS estimation when three kinds of mid-resolution remote sensing images (SPOT 5 HRG, CBERS-02, Landsat 7 ETM+) were used to build a CART model. Li *et al*. (2013) proposed an ensemble learning method based on the CART algorithm to estimate SPIS values. The correlation coefficient of the results was approximately 0.89 and the average error was

approximately 9.5%; however, this method requires the support of a large number of indexes.

According to the above studies, the selection of input variables is the main direction of SPIS estimation because of its effect on model accuracy. Meanwhile, it is necessary to consider the geographical characteristics of the study area. Massive uncertainty has been contributed by image features related to the effects of phenology on different land-cover types in modeling. Therefore, further studies of SPIS extraction in different geographical regions are necessary. This study attempts to integrate the CART algorithm with medium- and high-resolution remote sensing images to estimate SPIS values within Beijing Municipality of China and to analysis the applicability of input variables selected in CART algorithm. More specifically, the goals of this study were to develop a SPIS estimation approach suitable for typical temperate semi-arid climate zone with distinct intra-annual variability in vegetation. In addition, to further verify the effectiveness of the approach, it intends to explore a multi-temporal SPIS mapping scheme which is needed to monitor regional environmental changes and urban landscape pattern changes.

2 Materials and Methods

2.1 Study area

Beijing Municipality, which includes urban and suburban areas, is our study area. It is located in the northern part of the North China Plain at approximately 39.4°N–41.6°N, 115.7°E–117.4°E, and it experiences a typical continental monsoon climate with four distinct seasons characterized by hot and rainy summers and cold and dry winters. Moreover, there is also considerable phenological variability because the deciduous vegetation in the area loses its leaves seasonally. Therefore, there is a mixed land-use pattern in the study area that includes buildings, roads, vegetated land, bare land, cultivated land, water (river, reservoir), *etc*. The study area is shown in Fig. 1.

2.2 Data and processing

Landsat 5 TM images covering the study area were selected as the main data source for SPIS estimation. All of the images were acquired in cloud-free conditions. Landsat 5 TM images (http://glovis.usgs.gov/) acquired on May 6, 2005 and November 14, 2005 were used to represent leaf-on data (during the growing season) and leaf-off data (during the non-growing season), respectively. Two path/rows (p123r32 and p123r33) were obtained to contain the full extent of Beijing Municipality.

A multi-spectral QuickBird satellite image acquired on July 5, 2005 was used to develop the reference data sets for SPIS prediction and performance assessment. The main districts of the study area were covered by the high spatial resolution image, and the reference data sets are shown in Fig. 1b.

Furthermore, some auxiliary data that covered the whole study area were selected, including DMSP/OLS nighttime light (NLT) data (1-km resolution, http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html), the GDEM digital elevation model (DEM, 30-m resolution) and slope data extracted from GDEM. Data preprocessing was carried out in order to facilitate the subsequent processing. First, Landsat 5 TM and QuickBird images were geometrically corrected and re-projected into the UTM/WGS 84 projection with a pixel size of 30 m by 30 m for all bands, and then the NLT data were re-sampled as well. In addition, the information extracted from the TM images is further employed as an input variable of the SPIS model: 1) the brightness, greenness and humidity values were extracted by applying the tasseled cap transformation to the Landsat 5 TM images; 2) normalized differential vegetation index (NDVI) and normalized water index (NDWI) values were extracted from the Landsat images.

2.3 Methods

2.3.1 CART algorithm

The general classification and regression tree (CART) algorithm is a method for spatial data mining that was put forward by Breiman in 1984 (Breiman *et al*., 1984). The CART algorithm carries out a binary recursive partitioning process. Each parent node is split into two child nodes, and each child node is treated as a potential parent node in the process. The regression tree algorithm produces rules based on training data. Each rule set defines the conditions under which a multivariate linear regression model is established. Furthermore, the continuous variables can be estimated based on the rules.

Regression tree models can account for nonlinear relationships between input and target variables and allow both continuous and discrete variables as input variables. Therefore, the CART algorithm has been widely

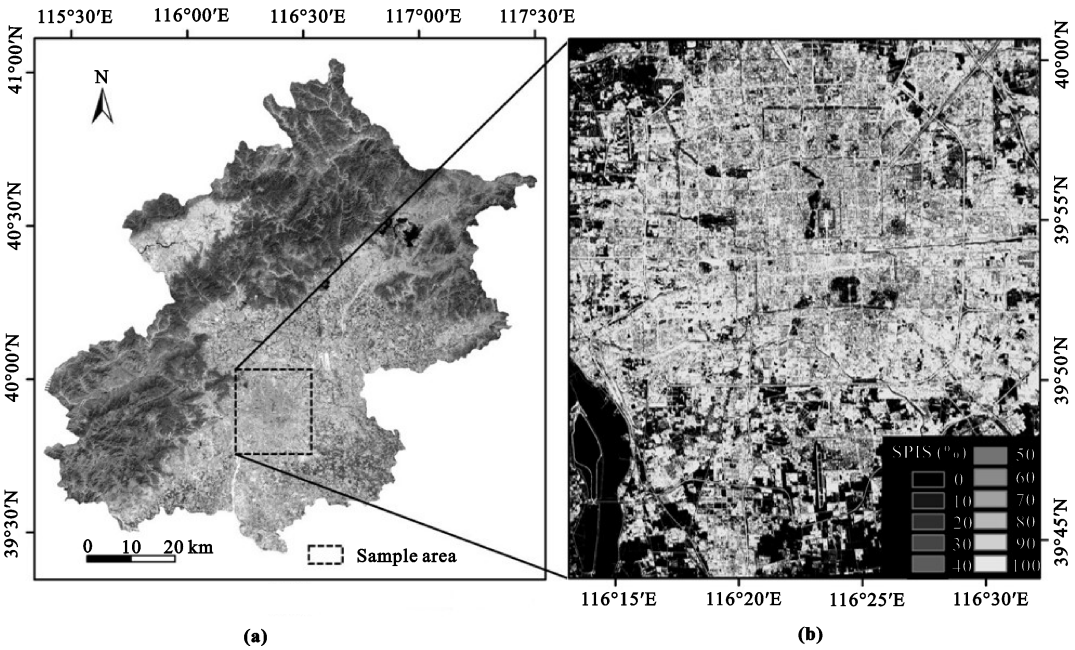

Fig. 1 Images showing the study area: (a) Landsat 5 TM, resolution 30 m; (b) Sub-pixel impervious surface percentage (SPIS) in a sample area extracted from a Quickbird image, resolution 30 m

used in data mining. It has been shown that accuracy and predictability of regression tree models are often better than those of simple linear regression models. (Huang and Townshend, 2003; Liu *et al*., 2005). The initial predictive models were developed based on CART algorithm by using the Cubist software package, which is produced by Rulequest (http://www.rulequest.com/), for data mining.

2.3.2 Training-test data collection

The training data used for modeling is composed of input variables and target variables. Here, the input variable is the spectral characteristics of the impervious surface, and the target variable is the SPIS values of the corresponding pixels. Sufficient training samples need to be collected to represent the impervious surface spectral features of different prediction units and the SPIS values of the corresponding pixels.

Many researchers have carried out a large number of studies on the selection of the input variables (Yang *et al*., 2003; Jiang *et al*., 2008; Li *et al*., 2010; Zhang *et al*., 2010). To determine the characteristic variables for the study area, the Landsat TM visible light, near infrared, and thermal infrared bands; the first three components of the tasseled cap transformation based on Landsat TM; NDVI; night light data (NLT); DEM data; and slope data were collected for modeling. These characteristic variables were available for use as input variables.

The target variables were extracted from high- resolution remote sensing data. First, high-resolution remote sensing images were classified into a variety of surface types, such as impervious surfaces, water, vegetation, bare land, *etc*., and then the shadows caused by high-rise buildings were classified into these surface types manually. The final results were recoded into two categories, including impervious surfaces (mainly composed of artificial buildings, roads and cement floors) and permeable surfaces (water, vegetation, and bare soil). The impervious surface pixels were counted, and then the percentages of 30 m by 30 m pixels that were covered by impervious surfaces were calculated for all bands. In addition, the calculation results at 30 m resolution were regarded as training data for the CART algorithm and the target variable of the prediction model.

Ideally, the training data for modeling and the test data for evaluation should be obtained from different sources, and they must be kept spatially independent. Due to the large scale dealt with by this study, the 30 m reference data derived from high-resolution images were used for both training and validation. Each high-resolution reference data set was divided into equal-sized blocks, with two thirds of the blocks (approximately 65 000 points) randomly selected as training

samples and the remaining blocks reserved as test samples. Splitting the reference points by pixel block rather than by pixel can reduce the spatial autocorrelation between training and test samples. The predictive accuracy can be estimated using *n*-fold cross-validation in Cubist. Therefore, the training data set can be divided into *n* blocks of roughly equal size. For each block in turn, a model is built from the data in the remaining blocks and tested using the holdout block (Michie *et al*., 1994).

2.3.3 CART modeling and validation

To compare the accuracy of SPIS estimation using different input variables, the collected characteristic variables were grouped differently in the models. All of these models were established based on the CART algorithm to estimate SPIS values. The input variables used in the different models are shown in Table 1.

The input variables in each model were used as the independent variables in Cubist. The SPIS values obtained from the classification and statistics of the QuickBird image were used as the target variables. A large number of samples were randomly selected from the independent and target variables as training data. The CART algorithm was used to study these training samples, and the initial prediction models to estimate SPIS values were established.

The final CART model was built using the most relevant input variables and target variables and was then applied to map SPIS values over a large area. The models were tested using the remaining test samples. Three evaluation indexes, the average error (*AE*) and relative error (*RE*), correlation coefficient (*r*) between actual and predicted values, were used in the study to evaluate model performance (Yang *et al*., 2003).

Table 1 Models and their input variables

Model	On b1–7	Off b1–7	On TC	Off TC	DEM	Light	Slope
1	√	√					
2	√	√	√	√	√	√	√
3	√	√	√	√		√	
4	√	√				√	
5	√					√	
6		√				√	

Notes: On b1–7 includes seven bands from the leaf-on data (visible, near-infrared, mid-infrared and thermal-infrared bands); Off b1–7 includes seven bands from the leaf-off data (visible, near-infrared, mid-infrared and thermal-infrared bands); On TC includes the brightness, greenness and humidity extracted using the tasseled cap transformation from the leaf-on data; Off TC includes the brightness, greenness and humidity extracted using the tasseled cap transformation from leaf-off data; DEM is the digital elevation model; Light is the NLT data; slope represents the gradient data extracted from the DEM

The method used to estimate sub-pixel level impervious surface consists of the following steps: 1) collection of training and testing data, including the input variables and the target variables; 2) estimation of models built using the CART algorithm and selection of the optimal model with the relevant input variables. An optimal combination of variables is selected for the study area by comparing the SPIS estimation accuracies of different models; 3) estimation and mapping of SPIS values over a large area in 2005, and the optimization and verification of the outputs; 4) Development of the multi-temporal SPIS mapping scheme: multi-temporal SPIS modeling and mapping are established based on the optimized SPIS results for 2005. An overview of the process is shown in Fig. 2.

3 Results and Analyses

3.1 CART modeling and accuracy assessment

All of the models shown in Table 1 were established

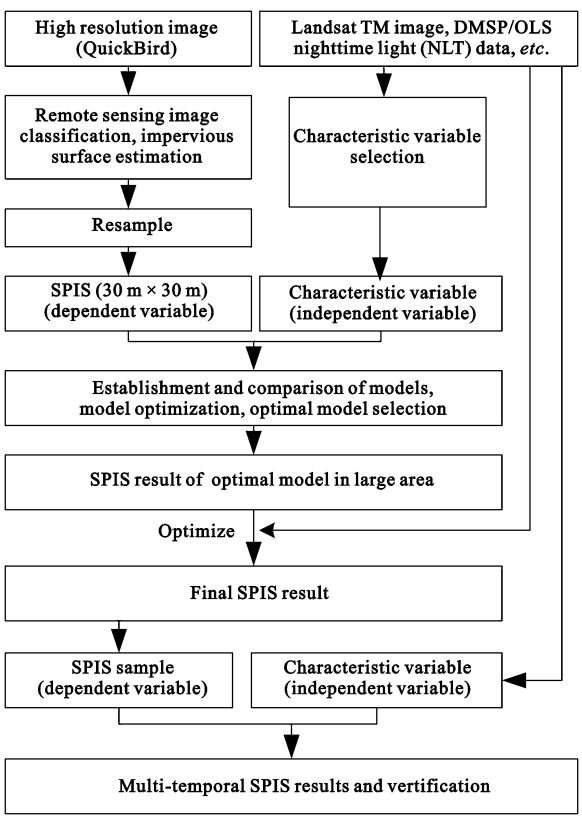

Fig. 2 Procedure for SPIS mapping using the classification and regression tree (CART) algorithm and multi-source remote sensing data

using different input variables to estimate SPIS values. Table 2 shows the results of accuracy assessment for these models.

As shown in Table 2, the differences in accuracy among the models were not noticeable over the sample area. *AE* ranges from 12.7% to 15.2%, *RE* ranges from 0.39 to 0.46, and *r* ranges from 0.81 to 0.86. The accuracies of models 3 and 4 are better than the other models. It can be seen that all these models can be used to estimate SPIS values in the sample area, and then the predictions of these models for the entire study area should be investigated.

3.2 Regional extension of model and input variable selection

3.2.1 Overall assessment of regional modeling

These models were used to estimate SPIS values over the whole study area, and then an analysis and comparison of the spatial pattern of SPIS mapping results was made to address the key problems in this research.

The SPIS results of different models were analyzed from two aspects. The first one is the accuracy of the spatial distribution of SPIS, i.e., the corresponding analysis of Landsat TM and SPIS values. The other one is the accuracy of the SPIS values. According to the empirical analysis, there SPIS values should be higher (>50%) in the city center, while relatively low SPIS values (20%–40%) should be observed in the suburban areas empirically. The comparison of the spatial pattern of SPIS mapping results produced by these models and the selection of the optimal input variables were based on the two aspects mentioned above.

Table 2 Accuracy assessment for each model

Evaluation index	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
AE (%)	13.5	13.6	12.7	12.8	13.6	15.2
RE	0.41	0.42	0.39	0.39	0.41	0.46
r	0.84	0.84	0.86	0.86	0.84	0.81

Notes: *AE* is the average error, *RE* is the relative error, and *r* is the correlation coefficient

The SPIS mapping results from models 2, 3, 4, and 6 are shown in Fig. 3. The results of model 1 and 5 were not noticeable on the whole scale, thus the analysis was performed at local scale later. The SPIS values in large mountain areas coved by vegetation were supposed to be 0%, but they were overestimated in models 2, 3, and 6. It does not accord with the actual surface features. However, the SPIS results of model 4 showed good performance. After comprehensive analysis of the SPIS mapping results and performing an accuracy assessment for these models, it can be seen that model 4 performed best when estimating both impervious surfaces and permeable surfaces (e.g., vegetation, bare land, cultivated land and water).

3.2.2 Analysis of modeling with NLT data

The results of model 1 and model 4 were compared to analyze the effects of employing NLT data in CART modeling.

The SPIS results from model 4, which includes the NLT data, and model 1, which does not include the NLT data, are shown in Figs. 4c and 4d, respectively. It is seen that the bare land in suburban areas was misestimated

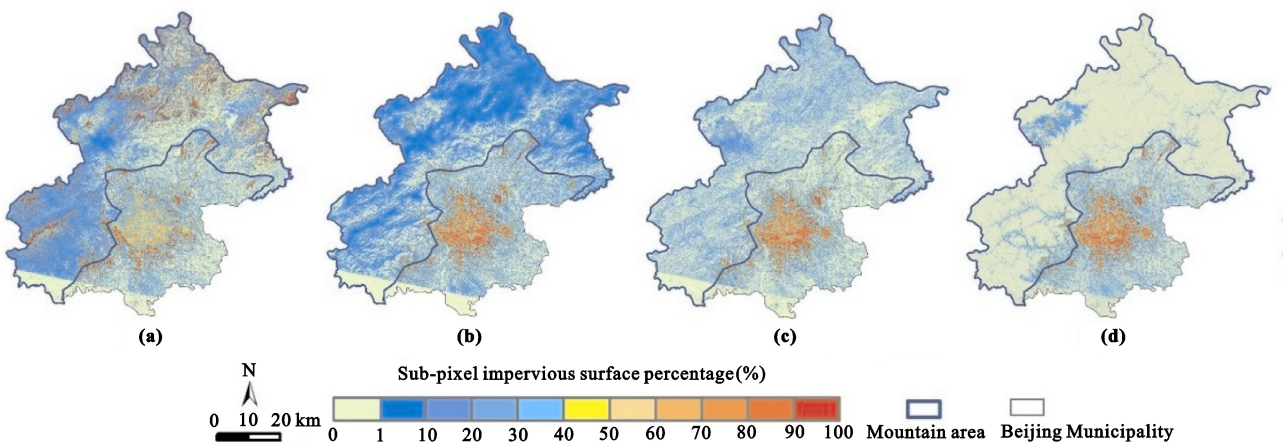

Fig. 3 Examples showing overestimation of sub-pixel impervious surface percentage (SPIS) in mountain area: (a) Model 2 (On b1–7, Off b1–7, On TC, Off TC, DEM, Light, Slope); (b) Model 3 (On b1–7, Off b1–7, On TC, Off TC, Light); (c) Model 6 (Off b1–7, Light); (d) Model 4 (On b1–7, Off b1–7, Light)

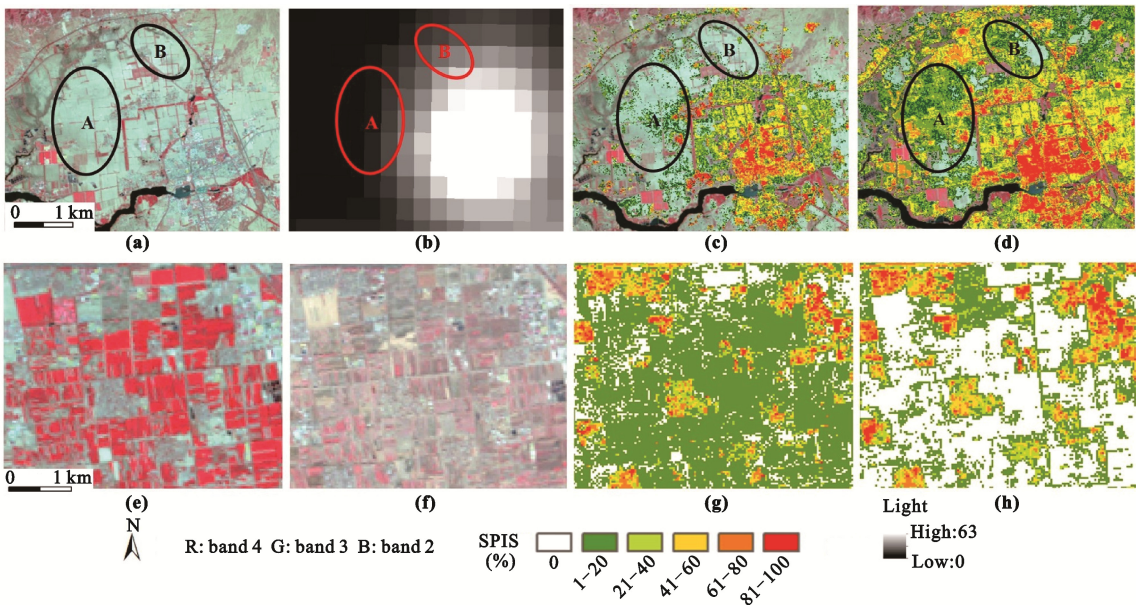

Fig. 4 Effects of including NLT (nighttime light data) data on estimation of SPIS values (example: A and B) and suburban cultivated land misestimation: (a) Standard false color Landsat 5 TM image in the northwestern part of the area; (b) NLT data; (c) results from model 4 (On b1–7, Off b1–7, Light); (d) results from model 1 (On b1–7, Off b1–7); (e) standard false color leaf-on image in the southeastern part of the area; (f) standard false color leaf-off image; (g) results from model 5 (On b1–7, Light); (h) results from model 4 (On b1–7, Off b1–7, Light)

as impervious surface in model 1, because of the spectral similarity between bare land and impervious surface. However, the SPIS results of model 4 perform well over the bare land because of the NLT data. For example, it can be seen that the SPIS estimation in bare land of model 4 performs better than model 1, as is seen in locations A and B. In addition, the accuracy of the model that includes NLT data is more precise.

From the above analysis, it can be seen that the NLT data can effectively reduce the misestimation caused by the spectral similarity between bare land and impervious surfaces. Meanwhile, NDVI and NDWI extracted from Landsat TM images were used to solve the problems caused by the saturation of NLT data in urban area and to optimize the SPIS results.

3.2.3 Analysis of modeling using multi-seasonal images

Comparing the SPIS results from model 4 and model 5 to explore the effects of multi-seasonal images, the difference between them is whether both the leaf-on data and leaf-off data were used as input variables or not.

As shown in Fig. 4, we can see that the SPIS values of the cultivated land were misestimated in model 5, which only used single season images as input variables. The leaf-on and leaf-off data are shown in Fig. 4e and 4f, respectively. It can be seen that the cultivated land shows different spectral characteristics in different seasons. As shown in Figs. 4g and 4h, the SPIS values of large areas of cultivated land were misestimated by model 5, which only used leaf-on data as an input variable, while this misestimation is reduced in model 4, which uses multi-season images. Moreover, the accuracy of model 4 is higher than that of model 5.

There is considerable variability in the spectral characteristics of cultivated land caused by differences in farming practices and vegetation phenology. However, these differences can be detected in the multi-seasonal image effectively. Therefore, the multi-seasonal images can be used to optimize the modeling and improve the precision of the SPIS results effectively.

3.3 Final CART model and SPIS results

Analysis and comparison of the SPIS mapping results shows that model 4 has the highest accuracy, and the SPIS results from model 4 are most consistent with actual surface conditions. Hence, the leaf-on data, leaf-off data and NLT data were used as the optimal input variables.

To improve the accuracy of SPIS mapping, it is necessary to further edit the output of the CART model.

Areas of vegetation and the small rivers were estimated as having low SPIS values because of the saturation inherently associated with NTL data in urban areas. The NTL data suffer from the problem of saturation, which makes the SPIS inaccurate within urban areas. Based on previous research, the status of vegetation and water can be reflected by the NDVI and NDWI, respectively, as extracted from Landsat5 TM images. The recognition of vegetation and water can be conducted by setting appropriate thresholds for the NDVI and NDWI (Ma *et al*., 2014; Su *et al*., 2015).

Finally, after accuracy validation and error editing, the SPIS values of Beijing Municipality in 2005 were mapped as shown in Fig. 5.

3.4 Multi-temporal SPIS mapping

To further validate the effectiveness and generalization of the optimal model, multi-temporal SPIS values were estimated based on the optimized SPIS results from 2005. SPIS values covering a period including several years before and after 2005 were estimated based on the optimal model. According to trends in urbanization, the SPIS values mostly increase with urban sprawl. Therefore, the sample data were recollected from the optimized SPIS results for 2005 in highly urbanized areas.

An accuracy assessment of the multi-temporal SPIS results was performed based on the sample data using cross-validation, including the *AE*, *RE*, and *r* of the SPIS results in these years. The accuracy assessment is shown in Table 3.

As shown in Table 3, the range of *AE* is approximately 7.3–14.9, *RE* lies in the range between 0.34 and 0.44, and *r* lies in the range between 0.73 and 0.86. In addition, r has gradually decreased with time relative to 2005, and the degree of decrease is approximately 10% roughly every 10 years. The method of multi-temporal SPIS estimation proposed in this paper can be used to address multiple time series, and the prediction accuracy can be guaranteed. Therefore, it is shown that this approach for estimating sub-pixel level impervious surfaces by integrating the CART algorithm and multi-source remote sensing data is feasible and suitable for the multi-temporal SPIS mapping of areas with distinct intra-annual variability in vegetation.

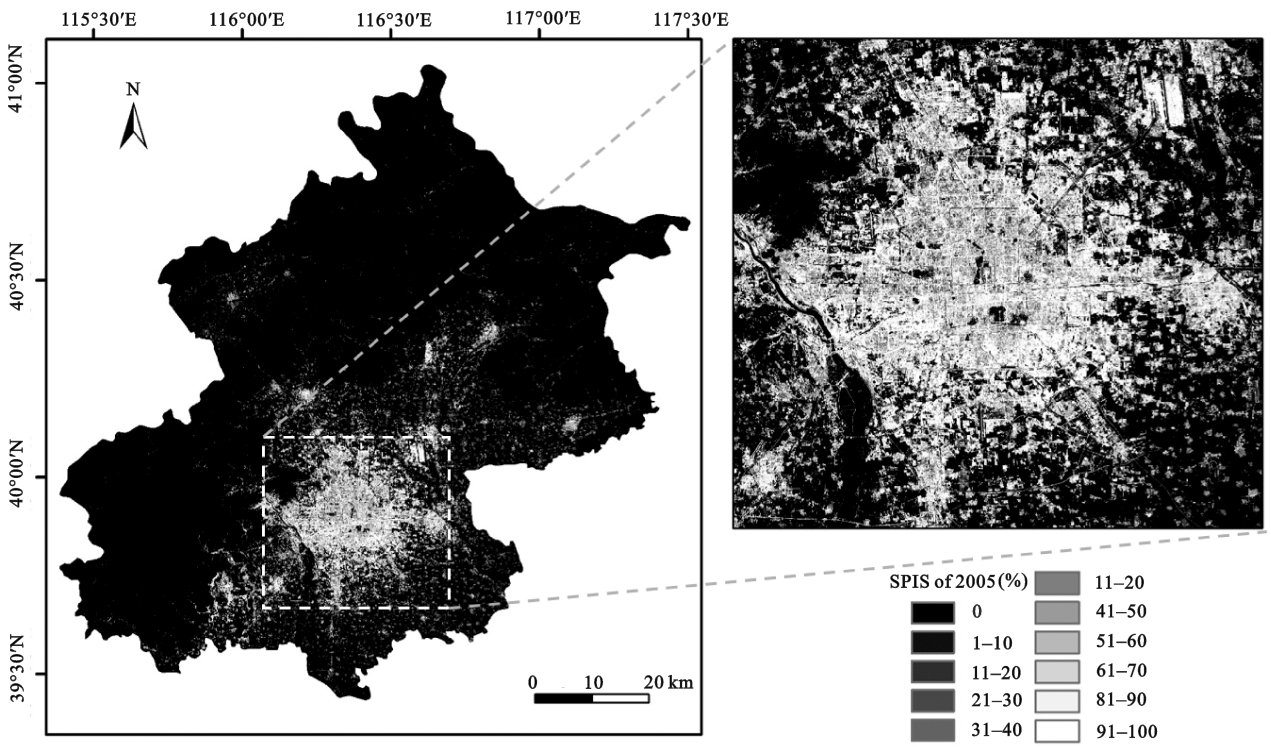

Fig. 5 Distribution of SPIS values within Beijing Municipality in 2005

Table 3 Accuracy assessment for multi-temporal SPIS results

Evaluation index	1991	1995	2001	2005	2007	2009	2011	2016
AE (%)	10.4	9.7	8.9	12.8	7.3	8.6	8.6	14.9
RE	0.43	0.41	0.36	0.39	0.35	0.34	0.44	0.43
r	0.73	0.75	0.8	0.86	0.84	0.8	0.76	0.76

4 Discussion

4.1 Effect of bands selection on the SPIS estimates

To avoid redundancy within the input variables, the number of bands used as the input variables is reduced to improve algorithm efficiency.

When all seven bands of Landsat 5 TM were used as the independent variables of estimation models, the model contribution rate (Contr) and the conditional percentage (Conds) of each band, as output from Cubist, are shown in Table 4. It can be seen that blue band (B1), near-infrared band (B4), mid-infrared band (B5) and thermal-infrared band (B6) are the most relevant bands.

Bands 1, 4, and 5 of Landsat 5 TM were combined as the independent variables of some models (model 7 and 8) to reduce the numbers of bands. Meanwhile, bands 1, 4, 5 and 6 were combined as the independent variables of other models (model 9, 10, 13 and 14). The corresponding NDVI and NLT data were added to set up the models (model 11 and 12). The specific combinations of input variables used the models are shown in Table 5.

Table 4 Contribution rate and conditional percentage of each band in the CART model

Temporal	Band	Contr (%)	Cond (%)
Leaf on	Band4	93	64
	Band1	99	61
	Band5	96	55
	Band6	71	49
	Band3	90	46
	Band2	72	0
	Band7	91	1
Leaf off	Band1	98	43
	Band6	81	32
	Band4	89	10
	Band5	97	8
	Band7	90	2
	Band2	73	1
	Band3	92	0

Notes: Temporal: multi-temporal Landsat data; Leaf on: Landsat 5 TM images from the growing season, Leaf off: Landsat 5 TM images from the non-growing season; Contr: model contribution rate, Conds: conditional percentage of each band

Table 5 Models and their input variables

Model	On b145	Off b145	On b1456	Off b1456	On NDVI	Off NDVI	Light
7	√	√					√
8	√						√
9			√	√			
10			√				
11			√	√	√	√	√
12			√		√		√
13			√	√			√
14			√				√

Notes: On b145 includes bands 1, 4, and 5 from the leaf-on data (blue, near-infrared, and mid-infrared bands); Off b145 includes bands 1, 4, and 5 from the leaf-off data (blue, near-infrared, and mid-infrared bands); On b1456 includes bands 1, 4, 5, and 6 from the leaf-on data (blue, near-infrared, mid-infrared and thermal-infrared bands); Off b1456 includes bands 1, 4, 5, and 6 from the leaf-off data (blue, near-infrared, mid-infrared and thermal-infrared bands); On NDVI is the normalized difference vegetation index extracted from the leaf-on data; Off NDVI is the normalized difference vegetation index extracted from the leaf-off data; Light is the NLT data

An accuracy assessment of these models was performed based on the sample data using cross-validation, including *AE*, *RE*, and *r*. The accuracy assessment of the different models is shown in Table 6. It can be seen that these models (model 7–14) can achieve good accuracy while also reducing the number of bands used.

Models 1 and 9, 4 and 13, and 5 and 14 were compared to determine whether or not bands 1–7 can be replaced by bands 1, 4, 5, and 6. The accuracy of model 9 with bands 1, 4, 5, and 6 is better than model 1 with bands 1–7, while the accuracies of models 4 and 5 with bands 1–7 are better than those of models 13 and 14 with bands 1, 4, 5, and 6, as shown in Table 2 and Table 6. There are no significant differences between bands 1–7 and bands 1, 4, 5, and 6 after comparing the SPIS results of these models. The misestimated pixel ratios of these models were counted roughly as shown in Fig. 6. It shows that there is no significant effect of the number of bands used on the prediction results when both leaf-on and leaf-off data were used as the input variables. However, the model using bands 1, 4, 5, and 6 as the input variables was more accurate when only leaf-on data were used.

Table 6 Accuracy assessment for each model

Evaluation index	4	7	8	9	10	11	12	13	14
AE (%)	12.8	13.4	14.2	13.4	14.6	12.9	13.7	13.2	14.0
RE	0.39	0.41	0.43	0.41	0.45	0.40	0.42	0.40	0.43
r	0.86	0.85	0.83	0.85	0.81	0.86	0.84	0.85	0.83

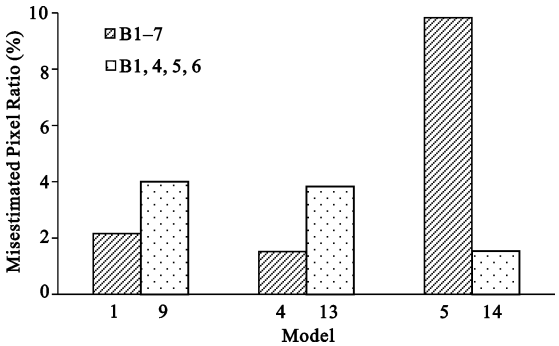

Fig. 6 Misestimated pixel ratio yielded by different models

4.2 Accuracy assessment in different categories of SPIS

According to previous research (Zhang *et al*., 2010), one challenging problem for impervious surfaces estimation is the overestimation in less developed areas and underestimation in developed areas. Hence, considering the spatial pattern of SPIS values and former studies comprehensively (Xian and Crane, 2005), the results have been divided into areas with high-density impervious cover (70%–100%), medium-density impervious cover (40%–70%), low-density impervious cover (10%–40%) and natural cover (0%–10%) to assess this problem.

Table 7 shows the accuracy assessment for different categories of impervious cover. Range refers to the range of the accuracies of all 14 models previously described. Generally speaking, the accuracy of the optimal model (model 4) performs well in terms of the three statistical indicators.

Fig. 7 shows the distribution of the accuracy for the different SPIS categories. It can be seen that: 1) the CART algorithm performed better in estimating SPIS values for high-density urban areas than other categories, with the *AE* of all models reaching the overall average error and the *RE* of all models better than the overall *RE*. 2) The overestimation rate of low-density SPIS value is far greater than the underestimated rate of high-density SPIS value. The CART algorithm for estimating low-density SPIS values did not appear to be as good as that for high-density SPIS values. The *AE* and *RE* of low-density urban areas were higher than the overall accuracy of the optimal model. The overestimation of low-density SPIS values may be due to the spectral similarity between bare land and impervious surfaces. 3) Values of *r* for all categories were lower than the overall correlation coefficient (0.86), but the *r* values of high-density and low-density urban areas were better than those of medium-density urban and natural cover.

5 Conclusions

An approach for quantifying impervious surfaces as a continuous variable, which is suitable for the typical temperate semi-arid climate zone, was developed and applied in Beijing Municipality. Our results demonstrate that multi-seasonal images and NLT data are the optimal input variables for SPIS estimation based on the CART algorithm. The results also show that the SPIS prediction model has high precision, in which the values of *r*, *AE* and *RE* are approximately 0.86, 12.8%, and 0.39, respectively and the SPIS model performed better in estimating values for high-density urban areas than other categories. Moreover, the multi-temporal (1991–2001, 2007–2016) SPIS values were estimated using the optimal model trained from the SPIS of 2005. They can be used as the basic data for the analyses of regional environmental changes and urban landscape pattern changes. Technically, the effectiveness and generalizability of the SPIS model have been validated.

It is shown that the proposed approach to estimate sub-pixel level impervious surfaces by integrating the CART algorithm and multi-source remote sensing data is feasible and suitable to the multi-temporal SPIS mapping of areas with distinct intra-annual variability in vegetation. However, the accuracy of SPIS estimation diminished with time relative to the year in which the optimal model was trained (2005). Therefore, further studies on the stability of SPIS estimation accuracy for long time series is needed in the future.

Table 7 Accuracy assessment for different categories of impervious cover

Land cover	*AE* (%)		*RE*		*r*	
	Range	Model 4	Range	Model 4	Range	Model 4
High-density urban	12.25–26.95	12.47	0.15–0.43	0.15	0.33–0.40	0.40
Medium-density urban	17.10–19.52	17.10	0.30–0.42	0.30	0.15–0.22	0.22
Low-density urban	13.96–22.21	17.76	0.45–0.55	0.48	0.36–0.45	0.45
Natural cover	8.92–12.33	9.43	0.93–0.95	0.94	0.16–0.27	0.23

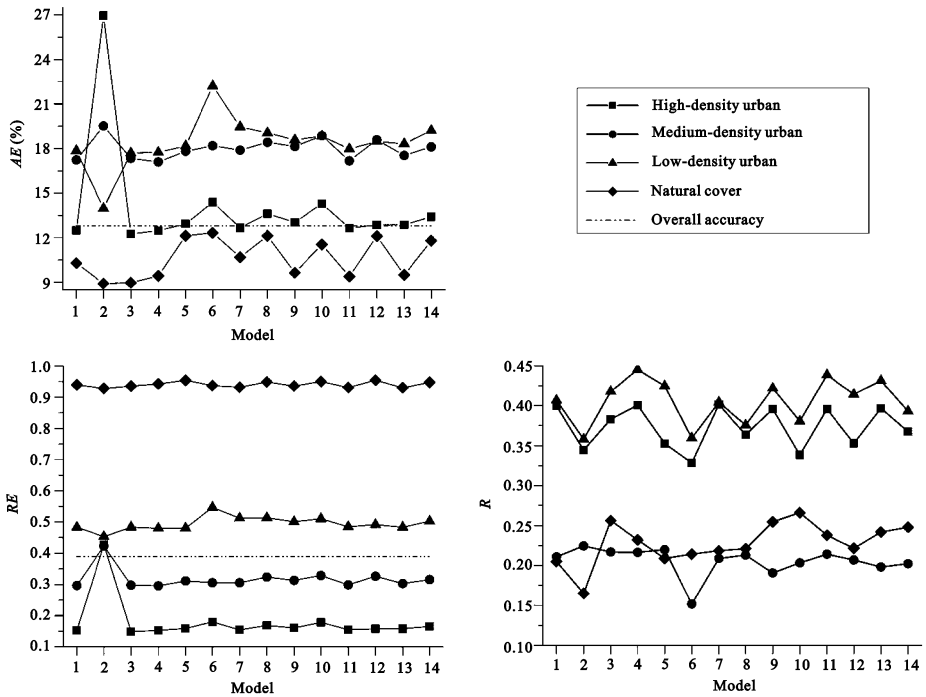

Fig. 7 Comparison of accuracy assessment in different categories of impervious cover

References

Arnold Jr C L, Gibbons C J, 1996. Impervious surface coverage: the emergence of a key environmental indicator. *Journal of the American Planning Association*, 62(2): 243–258. doi: 10.1080/01944369608975688

Breiman L, Friedman J, Olshen R, 1984. *Classification and Regression Tree*. New York: Chapman and Hall.

Cao Liqin, Li Pingxiang, Zhang Liangpei *et al*., 2012. Estimating impervious surfaces using the fuzzy ARTMAP. *Geomatics and Information Science of Wuhan University*, 37(10): 1236–1239. (in Chinese)

Friedl M A, Brodley C E, Strahler A H, 1999. Maximizing land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Trans on Geoscience and Remote Sensing*, 37(2): 969–977. doi: 10.1109/36.752215

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)

Homer C, Dewitz J, Yang L *et al*., 2015. Completion of the 2011 national land cover database for the conterminous United States representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, 81(5): 346–354. doi: 10.14358/PERS.81.5.345

Huang C, Townshend J R G, 2003. A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover. *International Journal of Remote Sensing*, 24(1): 75–90. doi: 10.1080/01431160110115032

Imhoff M L, Zhang Ping, Wolfe R E *et al*., 2010. Remote sensing of the urban heat island effect across biomes in the continental USA. *Remote Sensing of Environment*, 114(3): 504–513. doi: 10.1016/j.rse.2009.10.008

Jiang Liming, Liao Mingsheng, Lin Hui *et al*., 2008. Estimating urban impervious surface percentage with ERS-1/2 InSAR data. *Journal of Remote Sensing*, 12(1): 176–185. (in Chinese)

Jin H, Mountrakis G, 2013. Integration of urban growth modelling products with image-based urban change analysis. *International Journal of Remote Sensing*, 34(15): 5468–5486. doi: 10.1080/01431161.2013.791760

Lawrence R, Bunn A, Powell S, 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment*, 90(3): 331–336. doi: 10.1016/j.rse.2004.01.007

Li J X, Song C H, Cao L *et al*., 2011. Impacts of landscape structure on surface urban heat islands: a case study of Shanghai, China. *Remote Sensing of Environment*, 115(12): 3249–3263. doi: 10.1016/j.rse.2011.07.008

Li Qian, Li Caili, Rui Hanyi *et al*., 2010. Estimate of impervious surface percent based on different brightness of cart method with remote sensing images. *Water Resources and Power*, (12): 45–48. (in Chinese)

Li Xiaoning, Zhang Youjing, She Yuanjian *et al*., 2013. Estimation of impervious surface percentage of river network regions using an ensemble leaning of CART analysis. *Remote Sensing for Land and Resources*, 25(4): 174–179. (in Chinese)

Liu Y H, Niu Z, Wang C Y, 2005. Research and application of the decision tree classification using MODIS data. *Journal of*

Remote Sensing, 9(4): 405–411. doi: 10.11834/jrs.20050459

Ma Q, He C, Wu J *et al*., 2014. Quantifying spatiotemporal patterns of urban impervious surfaces in China: an improved assessment using nighttime light data. *Landscape and Urban Planning*, 130(4): 36–49. doi: 10.1016/j.landurbplan.2014.06.009

Michie D, Spiegelhalter D J, Taylor C C, 1994. *Machine Learning*, *Neural and Statistical Classification*. New York: Ellis Horwood.

Patel N, Mukherjee R, 2014. Extraction of impervious features from spectral indices using artificial neural network. *Arabian Journal of Geosciences*, 8(6): 3729–3741. doi: 10.1007/s12517-014-1492-x

Sanbum L, Lathrop R G, 2006. Subpixel analysis of Landsat ETM/sup+using self-organizing map (SOM) neural networks for urban land cover characterization. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6): 1642–1654. doi: 10.1109/TGRS.2006.869984

Su Y, Chen X, Wang C, *et al*., 2015. A new method for extracting built-up urban areas using DMSP-OLS nighttime stable lights: a case study in the Pearl River Delta, southern China. *Giscience and Remote Sensing*, 52(2): 218–238. doi: 10.1080/15481603.2015.1007778

Wang H, Wu B F, Li X S, 2011. Extraction of impervious surface in Hai Basin using remote sensing. *Journal of Remote Sensing*, 15(2): 388–400. doi: 10.11834/jrs.20110288

Weng Q, 2012. Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends. *Remote Sensing of Environment*, 117(2): 34–49. doi: 10.1016/j.rse.2011.02.030

Wu C, Murray A T, 2003. Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment*, 84(23): 493–505. doi: 10.1016/S0034-4257(02)00136-0

Xian G, Crane M, 2005. Assessments of urban growth in the Tampa Bay watershed using remote sensing data. *Remote Sensing of Environment*, 97(22): 203–215. doi: 10.1016/j.rse.2005.04.017

Xiao Rongbo, Ouyang Zhiyun, Cai Yunnan, 2007. Urban landscape pattern study based on sub-pixel estimation of impervious surface. *Acta Ecologica Sinica*, 27(8): 3189–3197. (in Chinese)

Yang L M, Jiang L M, Lin H *et al*., 2009. Quantifying sub-pixel urban impervious surface through fusion of optical and InSAR imagery. *Giscience and Remote Sensing*, 46(2): 161–171. doi: 10.2747/1548-1603.46.2.161

Yang L, Huang C, Homer C G *et al*., 2003. An approach for mapping large-area impervious surfaces: synergistic use of Landsat-7 ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing*, 29(2): 230–240. doi: 10.5589/m02-098

Yang X, 2006. Estimating landscape imperviousness index from satellite imagery. *IEEE Geosience and Remote Sensing Letters*, 3(1): 6–9. doi: 10.1109/LGRS.2005.853929

Zhang L, Weng Q, 2016. Annual dynamics of impervious surface in the Pearl River Delta, China, from 1988 to 2013, using time series Landsat imagery. Isprs *Journal of Photogrammetry and Remote Sensing*, 113(3): 86–96. doi: 10.1016/j.isprsjprs.2016.01.003

Zhang Lu, Gao Zhihong, Liao Mingsheng *et al*., 2010. Estimating urban impervious surface percentage with multi-source remote sensing data. *Geomatics and Information Science of Wuhan University*, 35(10): 1212–1216. (in Chinese)

Zhou J, Chen Y H, Zhang X *et al*., 2013. Modelling the diurnal variations of urban heat islands with multi-source satellite data. *International Journal of Remote Sensing*, 34(21): 7568–7588. doi: 10.1080/01431161.2013.821576

Zhou Ji, Chen Yunhao, Zhang Jinshui *et al*., 2007. Urban impervious surface abundance estimation in Beijing based on remote sensing. *Remote Sensing for Land and Resources*, 19(3): 13–17. (in Chinese)