Mapping Urban Areas in China Using Multisource Data With a Novel Ensemble SVM Method

Xin Huang, *Senior Member, IEEE*, Ting Hu, Jiayi Li, *Member, IEEE*, Qing Wang, and Jón Atli Benediktsson, *Fellow, IEEE*

***Abstract*— The mapping of urban areas at regional to global scales is a crucial task due to its value for environmental monitoring, habitat and biodiversity conservation, and decision-making. In most current applications, two techniques (i.e., supervised classification and data fusion) are widely applied in large-scale urban mapping. However, the costly training sample collection, inadequate data-source descriptions, and diverse urban characteristics (e.g., shape, size, socioeconomic status, and physical environment) are challenging problems for the urban mapping approaches. In this context, aiming at effectively deriving accurate urban areas at a large scale, we propose a novel ensemble support vector machine (SVM) method which consists of three steps: 1) the automatic generation of training data to reduce labor costs; 2) the construction of an ensemble SVM model to effectively combine the multisource data (including remote sensing and socioeconomic data); and 3) an adaptive patch-based thresholding technique to tackle the diverse urban characteristics. The proposed method is employed to map urban areas of China in 2005 and 2010, and the resulting maps are compared with the existing urban maps for 287 prefecture-level cities. It is found that our results present a satisfactory superiority, especially in challenging small cities, with a significant improvement in median Kappa (0.174 for 2005 and 0.203 for 2010). When incorporating moderate-resolution imaging spectroradiometer multispectral data as an additional source, the Kappa coefficient can be further raised by 0.028 for 2010. In general, the proposed method shows great potential for accurately mapping urban areas at regional, continental, or even global scales in a cost-effective manner.**

***Index Terms*— Data fusion, ensemble support vector machine (SVM), training samples, socioeconomic data, urban areas, urban mapping.**

Manuscript received October 23, 2017; revised December 31, 2017; accepted February 6, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 41522110 and Grant 41771360, in part by the Hubei Provincial Natural Science Foundation of China under Grant 2017CFA029, in part by the National Key Research and Development Program of China under Grant 2016YFB0501403 and Grant 2017YFB0504103, and in part by the Open Fund of Key Laboratory of Ministry of Education for Spatial Data Mining and Information Sharing under Grant 2016LSDMIS04. *(Corresponding author: Jiayi Li.)*

X. Huang is with the Department of Remote Sensing, School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China, and also with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: huang_whu@163.com; xhuang@whu.edu.cn).

T. Hu and J. Li are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China (e-mail: hutingrs@whu.edu.cn; zjjercia@whu.edu.cn).

Q. Wang is with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China.

J. A. Benediktsson is with the Department of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland (e-mail: benedikt@hi.is).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2018.2805829

I. Introduction

ALTHOUGH occupying only a small fraction (less than 1%) of the earth's surface, urban areas are hotspots of human activity, containing more than 50% of the world's population (POP) and 70%–90% of the global economy [1]. The rapid growth of urban areas has promoted economic growth [2] and housed the growing urban POP [3]. In turn, it has simultaneously aggravated environmental problems such as urban heat islands [4], air pollution, water contamination, habitat loss, and biodiversity decline [5], [6]. In this regard, the monitoring of urban areas, especially extracting the spatial information (location and extent), is of great importance.

Despite the efforts made in recent decades in the mapping and monitoring of urban areas, most of these studies have concentrated on a local scale (i.e., within a city or several cities) [7], [8]. However, to focus on the overview of national/continental urbanization, deriving the spatial information of urban areas at large-scale remains a matter of concern [7]. To date, a number of global urban maps have been produced (Table I) [1], [9]–[11] by the use of supervised/unsupervised and single/multisource data fusion [1]. In the literature, the supervised classification technique is a popular method that uses prior knowledge collected in advance to build rules for labeling the urban areas [e.g., MCD12Q1 and moderate-resolution imaging spectroradiometer (MODIS) 1 km in Table I]. In general, it is believed that both the quality of such prior knowledge (referred to as "training samples" in the supervised classification) and the classification algorithms are vital for the performance of the mapping [12]. Meanwhile, by integrating multisource data, the data fusion technique can often yield a better mapping performance than the counterpart single-source techniques [9], [11], [13] (e.g., LandScan, History Database of the Global Environment v3, global impervious surface, Global Rural-Urban Mapping Project, and Global Land Cover-SHARE of year 2014, as presented in Table I). With regard to the data fusion technique, both suitable data sources that portray the urban areas in a complementary manner and an appropriate fusion method that is concerned with the complex urban characteristics are the two main research questions. However, improving urban area extraction performance at a large scale in a more efficient way, on the basis of the current research, still calls for further attention.

First, alleviating the training sample collection cost can help to facilitate large-scale urban area mapping, and can even promote the potential for routine operation. In the conventional approaches, the training data are derived

TABLE I
11 MAJOR GLOBAL URBAN MAPS (ADOPTED AND UPDATED FROM [1])

Abbreviation	Map	Year	Resolution	Definition of urban area	Data sources	Method
VMAP	VMAP Map Level Zero [23]	1992	Scale 1:1000000	Populated places	Maps and navigational charts	Digitization
GLC2000	Global Land Cover 2000 [24]	2000	998 m	Artificial surfaces and associated areas	Nighttime lights data, SPOT-vegetation	Unsupervised classification
GlobCover	GlobCover v2 [25]	2000/05/10	309 m	Artificial surfaces and associated areas	MERIS	Unsupervised classification
LandScan	LandScan [26]	1998–2014	927 m	Ambient population distribution	Geocover maps, MODIS 1-km, Landsat, census data, etc.	Data fusion
HYDE	History Database of the Global Environment v3 [22]	1700–2000	9000 m	Built-up, cities	LandScan, UN census data, city gazetteers	Data fusion
IMSPA	Global Impervious Surface [27]	2000	927 m	Density of impervious surfaces	LandScan, nighttime lights data	Data fusion by linear regression
MCD12Q1	MODIS Land Cover Type Product 2012 [28][29]	2001–2012	500 m	Urban and built-up land	MODIS data	Supervised classification
MODIS 1-km	MODIS Urban Land Cover 1-km [11]	2000	1000 m	Urban and built-up areas	MODIS 1-km data, nighttime lights, population density	Supervised classification, data fusion
GRUMP	Global Rural-Urban Mapping Project [30]	1995	927 m	Urban extent	VMAP, census data, nighttime lights, etc.	Data fusion
Lights	DMSP/OLS Nighttime Lights [31]	1992–2015	927 m	Nighttime intensity	DMSP/OLS data	Data composition
GLC-SHARE	Global Land Cover-SHARE of year 2014 [9]	2012	1000 m	Artificial surface	Eleven major thematic land-cover layers	Data fusion

The maps circa 2005 (ESA 2005 and MCD12Q1 2005) were compared with the 2005 urban map obtained by our method, and the maps circa 2010 (ESA 2010, GLC-SHARE 2012, and MCD12Q1 2010) were compared with the 2010 urban map obtained by our method.

by manual interpretation from high-resolution satellite images [12]. However, with regard to the extremely high labor expense, manual interpretation at a large scale is not practical [14]. Automatic/semiautomatic training sample collection is, therefore, preferred, but most of the recent studies have addressed this issue in the local-scale mapping tasks [14], [15]. To deal with a large-scale task, a semiautomatic collection method was proposed [16], but the blooming effect caused by night-time light (NTL) data still hindered the effectiveness of the approach.

Second, the diversity of urban areas across geographic space, in terms of shape, size, socioeconomic status, and physical environment, aggravates the mapping challenges. To alleviate such effects, region-specific classification methods can be a practical solution [8], [17]. For example, in [8], the East Asia was divided into three regions according to their climate characteristics, and an independent classifier was built for each region. However, it should be mentioned that the urban areas of different cities and the different urban patches (defined as contiguous areas of urban pixels) are also diverse, and thus the variations need to be further considered. To tackle this problem, several researchers have designed object-based thresholding techniques to derive the optimal threshold of each urban patch [18], [19]. However, in these studies, the optimal thresholds of the urban patches were estimated by comparing the NTL data with the reference urban maps, and hence, the establishment of the optimal thresholds was strongly dependent on the reference data provided in advance.

Third, the comprehensive characterization of urban areas is somewhat difficult to define, and the definition of urban areas in global urban maps is inconsistent when different data types are used [1] (Table I). In practice, maps derived from socioeconomic data are more closely related to POP or economic distribution, while maps derived from physical data (e.g., remote sensing data) correspond to built-up areas [1], [7]. Hence, selecting the appropriate data source is of great importance. An urban area actually centralizes both socioeconomic activities and the built environment [20], and hence, it can be considered as a local maximum of a combination of multiple factors [e.g., POP, gross domestic product (GDP), or buildings] that may vary across time and space [21]. Therefore, in this paper, an urban area is defined as the combination of socioeconomic [POP, GDP, and physical factors (e.g., buildings, streets, and other infrastructure)] [20], [21]. In view of this, combining physical data and socioeconomic data for urban area mapping is a logical approach. To date, a number of attempts have been made to synthesize multiple data types for the delineation of urban areas. Several of the global products listed in Table I aim to fuse multisource data, but some of them mainly depend on fusion of existing urban products (e.g., GLC-SHARE [9]), and more intelligent data fusion methods to combine the multisource data are therefore required [11]. For instance, a Bayes' rule guided fusion method that integrates NTL data and gridded POP data with MODIS 1-km data (Table I) has been proposed [11]. Although the fusion algorithm is based on a hard-to-meet hypothesis (i.e., normal distribution of urban areas), it hints at the possibility of complementary data-source integration for urban area mapping.

Despite several studies [7], [11], [22] having focused on multisource data fusion for urban mapping, it is worth mentioning that the contributions and weights of data sources for

different urban areas (within a city or across different cities) should be different with regard to diverse urban characteristics. In this way, directly using a machine learning classifier, which views the multiple sources with a relatively equal importance, is less reasonable for such a complex mapping task. By fully considering the issues mentioned above (i.e., the costly training sample collection, inadequate data-source descriptions, and diverse urban characteristics), we propose a novel ensemble support vector machine (SVM) method to automatically derive accurate urban areas, by integrating gridded POP data, GDP data, NTL data, and temperature data. The contributions presented in this paper can be summarized as follows.

1) The training data are automatically generated to reduce the labor and time cost of manual interpretation.
2) A novel ensemble SVM model is constructed for assembling the multisource features.
3) An adaptive patch-based thresholding technique is designed to tackle the diverse urban characteristics (e.g., shape, size, socioeconomic status, and physical conditions). It should be noted that, compared to the current methods [18], [19], the proposed thresholding method can automatically derive the segmentation threshold of each urban patch without the assistance of reference data provided in advance.

The proposed method was tested in China, which is a challenging and representative study region, in terms of the diverse climate and unbalanced distributions of POP and economic status. Multitemporal (i.e., 2005 and 2010) urban maps of China were generated by the proposed ensemble SVM method. A set of mainstream global data sets was compared with the results of the proposed method. In this paper, we define the minimum mapping unit as contiguous patches larger than 1 km^2, according to [1]. This paper is organized as follows. The study area and data sets are described in Section II. The proposed ensemble SVM method is introduced in Section III. The experiment results are presented in Section IV. The proposed method and experiments are discussed in Section V. Final conclusions are summarized in Section VI.

II. Study Area and Data

A. Study Area

China, which covers a total land area of approximately 9.6 million km^2, is one of the most populous countries in the world. The climate of China is diverse, ranging from temperate continental in the arid north to subtropical monsoon in the wet south. The POP distribution is also skewed. More than 96% of the POP lives in the eastern part of China, which comprises only 36% of the national land area [32]. Moreover, the level of development between cities is also unbalanced. For example, the 16 megacities of mainland China account for less than 5% of the total number of cities in China, but contribute more than 25% of GDP, according to the China Statistical Yearbook 2010 [33]. Thereby, China is a challenging and valuable study area for urban extent mapping, because of the complicated characteristics mentioned earlier.

TABLE II
Criteria for Dividing City Ranks

Type	City rank	Population (million)	The number of cities
Prefecture-level	Megacity	>5	16
	Large city	1–5	120
	Medium city	0.5–1	104
	Small city	<0.5	47
Other	---	---	---

In addition, China has a large number of cities of various sizes. In terms of the municipal level, these cities can be classified into two types: prefecture-level cities and other administrative units (i.e., prefectures, leagues, and autonomous prefectures). In this paper, to analyze the performance of the proposed ensemble SVM method across a range of city sizes (Fig. 1), 287 prefecture-level cities are further stratified into four ranks, by referring to the 2010 urban resident POP (Table II) [33], [34].

B. Data

First, two data sets were exploited to acquire training samples: the municipal-level administrative boundary of China in GIS format downloaded from Global Administrative Areas (http://www.gadm.org/) and the MODIS Land-Cover Type Yearly Global 500-m product (MCD12Q1) for 2001–2010 obtained from the website of Maryland University (http://www.landcover.org/) [28], [29]. Second, the feature sources used in the ensemble SVM method comprised physical data [i.e., land surface temperature (LST) and vegetation adjusted NTL urban index (VANUI)] [35] and gridded socioeconomic data (i.e., POP [36] and GDP [37]). The gridded GDP and POP data can characterize the socioeconomic properties of urban areas, while remote sensing LST and VANUI data can characterize the physical properties of urban areas.

Because of the thermal difference between urban and nonurban areas, LST was employed to capture urban extent [7]. LST data obtained by remote sensing satellites record the radiation energy emitted from the ground surface in the thermal infrared band, and usually show high values in urban areas and low values in nonurban areas [38], [39]. Compared with daytime LST, night-time LST has a stronger ability to distinguish urban and nonurban areas [40]. In this paper, the annual average remote sensing LST data were generated using the MODIS 8-day 1-km night-time LST product (MOD11A2) downloaded from the MODIS website (http://ladsweb.nascom.nasa.gov/data/search.html).

The VANUI data were calculated by integrating NTL data and normalized difference vegetation index (NDVI) data [35]. NTL data usually reflect the intensity of human activity, which is closely related to built-up areas [18], [41]. In this paper, the utilized NTL data were the version 4 DMSP/OLS stable NTL data with a 1-km spatial resolution, derived from the National Geophysical Data Center (NGDC) website (http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html). These data exclude ephemeral lights and retained persistent

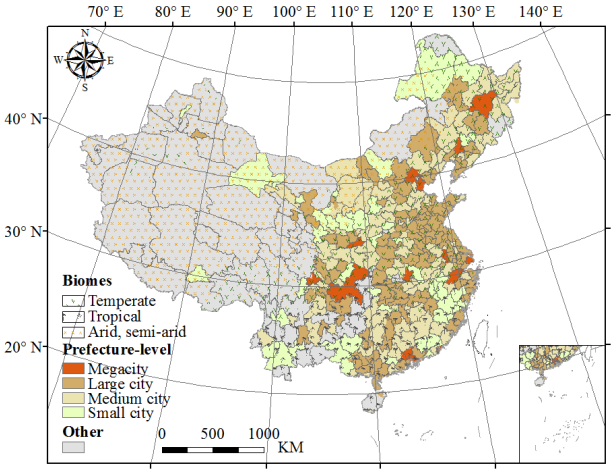


Fig. 1. Map of the study area and the prefecture-level cities in the three biomes [8].

lights from cities, towns, and other sites (e.g., gas flare areas) [17], [42]. In this paper, to tackle the inherent incompatibility of the DMSP/OLS NTL time series data set, a series of data correction steps, including intercalibration [42] and intra-annual composition [17], was adopted. The pseudolight pixels in the areas of water or gas flares were then removed by utilizing a gas flare mask and water mask [19]. The gas flare mask was obtained from the NGDC [42], and the water mask was derived from the MODIS water mask product (MOD44W). In addition, when using NTL data to describe urban areas, some endogenous obstacles, such as the saturation effect and the blooming effect [35], [43], [44], need to be addressed. In view of this, the NDVI can be used to alleviate such problems [35] and to distinguish the urban and nonurban areas, thereby enhancing the discriminative ability of NTL [45]. In this paper, the MODIS 16-day 500-m vegetation index product (MOD13A1) was used to generate annual mean NDVI. By combining the NTL and NDVI data, the derived VANUI data can alleviate the saturation effect of NTL data and overcome the confusion of NDVI data between urban areas and bare land [35], [45].

Gridded GDP data reflect the intensity of economic activities, and gridded POP data denote the density of POP. Compared with census data of GDP and POP, gridded GDP and POP data can provide detailed spatial distribution information, and thus they are more suitable for depicting the location and extent of urban areas from the perspective of socioeconomics. In this paper, the gridded GDP and POP data were obtained from the Resources and Environmental Scientific Data Center, Chinese Academy of Sciences (http://www.resdc.cn/), with a 1-km spatial resolution. Note that all the data sources mentioned earlier (i.e., VANUI, LST, GDP, and POP) were resampled to a pixel size of 500 m, the finest spatial solution among the data sources.

III. Methodology

In this section, the proposed ensemble SVM method is described. The proposed method (Fig. 2) consists of three steps: 1) generation of automated training data; 2) construction of an ensemble SVM model; and 3) implementation of an adaptive patch-based thresholding technique.

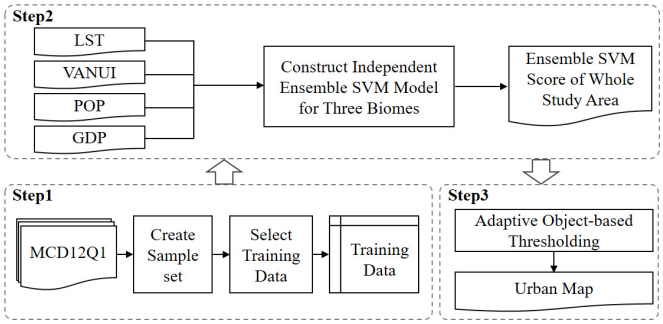


Fig. 2. Flowchart of the proposed ensemble SVM method.

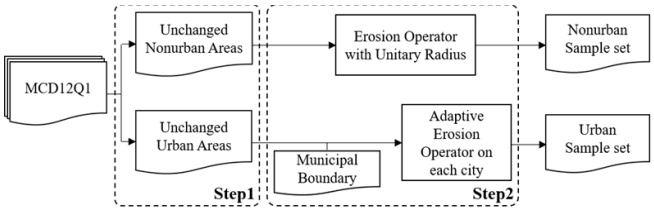


Fig. 3. Flowchart of sample set generation.

A. Automated Generation of Training Data

High-quality (i.e., reliable, diverse, and with an even spatial distribution) training data are necessary to guarantee a satisfactory classification result. In this section, an automated approach is presented to derive high-quality training data. The proposed approach consists of two parts: 1) creation of reliable and various sample sets and 2) selection of evenly distributed training data from the sample sets.

In the process of sample set generation (Fig. 3), considering that the MCD12Q1 product contains the spatial information of urban and nonurban classes, it is feasible to use this product as the training sample pool. In this way, pixels in the pool labeled as 1/0 are assigned as urban/nonurban pixels, respectively. With regard to the uncertainty of this product [46], a purification strategy is adopted not only to guarantee the reliability of the samples but also to meet the physical criterion of the comprehension definition of urban areas. To remove the samples with wrong labels in MCD12Q1 data, the unchanged areas of the multitemporal (i.e., 2001–2010) images are viewed as candidate urban/nonurban samples by an intersection operation.

Considering the complicated urban characteristics (e.g., the various and small sizes, as well as the diverse shapes), the urban samples generated by previous studies have often been skewed to big cities [47]. Therefore, an adaptive radius (R) of the erosion operator is employed to take into account the characteristics of different urban patches within each city. Here, the adaptive radius (R) for each city is obtained by a piecewise linear function (1). The piecewise linear function consists of two geometric parameters, i.e., the area (Area) and minor axis (L_{minor}) of the dominant urban patch within each city. The dominant urban patch is selected due to its high reliability. The minor axis, instead of the major axis, is considered to prevent the elongated urban patches being entirely filtered out. The coefficients of L_{minor}in (1) are determined by our trials and experiences. The parameter sensitivity is analyzed

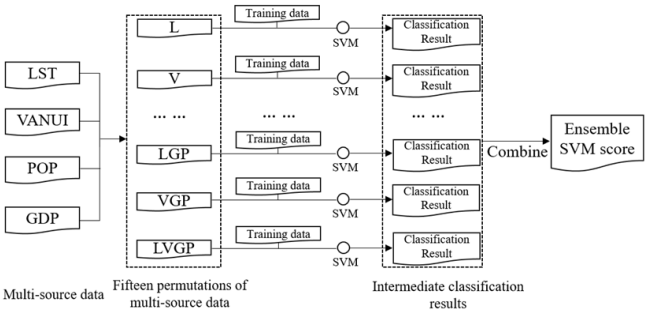


Fig. 4. Flowchart of the proposed ensemble SVM model implemented for each biome.

in Section V-B1.

$$R = \begin{cases} 0.2 \times L_{\text{minor}} & 0 < \text{Area} < 100 \\ 0.25 \times L_{\text{minor}} & 100 \leq \text{Area} < 500 \\ 0.2 \times L_{\text{minor}} & 500 \leq \text{Area} < 2000 \\ 0.15 \times L_{\text{minor}} & \text{Area} \geq 2000. \end{cases} \quad (1)$$

In this paper, to further select evenly distributed training data across the geographic space, the sample set of China was divided into 10×10 grids. In each grid, 10% of the urban samples were randomly picked as urban training data, and the nonurban training samples with the same number were randomly generated from the nonurban sample set. To reduce the possible bias induced by the random sampling, 10 groups of training data were individually created.

B. Multisource Ensemble SVM Model

In this section, the variations of urban areas between the three biomes are preliminarily considered [8], and an independent ensemble SVM model for each biome is constructed by synthesizing the multisource data. In this paper, the ensemble SVM model is a collection of several individual SVM classifiers (Fig. 4), aiming to derive the possibility of each pixel belonging to the urban class. In the ensemble model, each SVM classifier is trained by training samples with the input feature (i.e., the permutation of multisource data shown in Table III), and the unlabeled pixel is identified by majority voting since the 10 groups of training data are used to build each SVM. In Table III, for simplification, LST+VANUI (LV) means combined input of LST and VANUI data into the SVM classifier. SVM is an efficient machine learning classifier [14], [48], [49]. Its basic principle is to discriminate two classes by seeking an optimal linear separating hyperplane [50]. In addition, when dealing with nonlinear separability, a kernel trick that involves mapping the original feature into an implicit high-dimensional space can be exploited. In this paper, the widely used radial basis function is employed as the kernel function.

Once all the single SVM classifiers are built, their outputs (i.e., the intermediate classification results in Fig. 4) need to be combined. A satisfactory combination strategy should pick up the desirable complementary information of the multisource data. With regard to the comprehensive definition of "urban area," in advance there is no significant preference between physical properties and socioeconomic factors. All the

TABLE III
PERMUTATIONS OF STACKING THE MULTISOURCE DATA

One-source permutations	Two-source permutations	Three-source permutations	Four-source permutation
LST (L)	LST+VANUI (LV)	LST+VANUI+GDP (LVG)	LST+VANUI+GDP+POP (LVGP)
VANUI (V)	LST+GDP (LG)	LST+VANUI+POP (LVP)	
GDP (G)	LST+POP (LP)	LST+GDP+POP (LGP)	
POP (P)	VANUI+GDP (VG)	VANUI+GDP+POP (VGP)	
	VANUI+POP (VP)		
	GDP+POP (GP)		

For simplification, "LST+VANUI" (stacking these two data sources) is shortened to LV.

N intermediate classification results are summed, and pixels with scores less than $N/2$ (i.e., half of N) are reset as zero. The combination method to derive the ensemble SVM score can be formulated as

$$\text{DN}_{\text{en}} = \begin{cases} \sum_{i=1}^{N} \text{DN}_{\text{tempi}} & \text{if } \sum_{i=1}^{N} \text{DN}_{\text{tempi}} \geq \frac{N}{2} \\ 0 & \text{else} \end{cases} \quad (2)$$

where DN_{en} is the digital number (DN) of the ensemble SVM score, DN_{tempi} is the DN of the ith intermediate classification result (labeled as 1/0), $i \in [1, N]$, $N = C_m^1 + C_m^2 + \ldots C_m^m$ is the number of intermediate classification results, and m represents the number of data sources. For instance, N is equal to 15 when four data sources are considered.

In theory, the computational cost mainly consists of the SVM training and testing. For m data sets, we suppose that the size of the training sample is l. With the linear convergence of the machine learning algorithm (i.e., SVM solved in [51]), for a permutation with $p(p \leq m)$ data sources, the computational complexity for the SVM training should be $O(pql)$ (see [51, Sec. 5.7]), where q is the number of iterations. For the testing, we suppose that the size of the test sample is M, and the complexity of the testing should be $O(Mp)$. By enumerating all permutations, the total computational cost is

$$\begin{aligned} O\left(\left(\Sigma pC_m^p\right)(ql + M)\right) &< O\left(\left(\Sigma C_m^p\right)m(ql + M)\right) \\ &= O(2^m m(ql + M)). \end{aligned} \quad (3)$$

Although roughly incorporating the data sources would aggravate the cost, for the urban mapping task, it is believed that m should not be larger than 10 in reality. In view of this, when m is less than 10, it can be said that the computational complexity is linear with the size of test sample, as ql is much less than M (for instance, in the case study of China, q was set to 200 by default, l was roughly equal to 1.5×10^3 per biome, and M was roughly equal to 3.8×10^7).

C. Adaptive Patch-Based Thresholding

Considering that different urban patches exhibit diverse characteristics, in terms of their socioeconomic status and physical environment, a patch-based method for urban mapping is necessary and essential. Several studies [18], [19] have

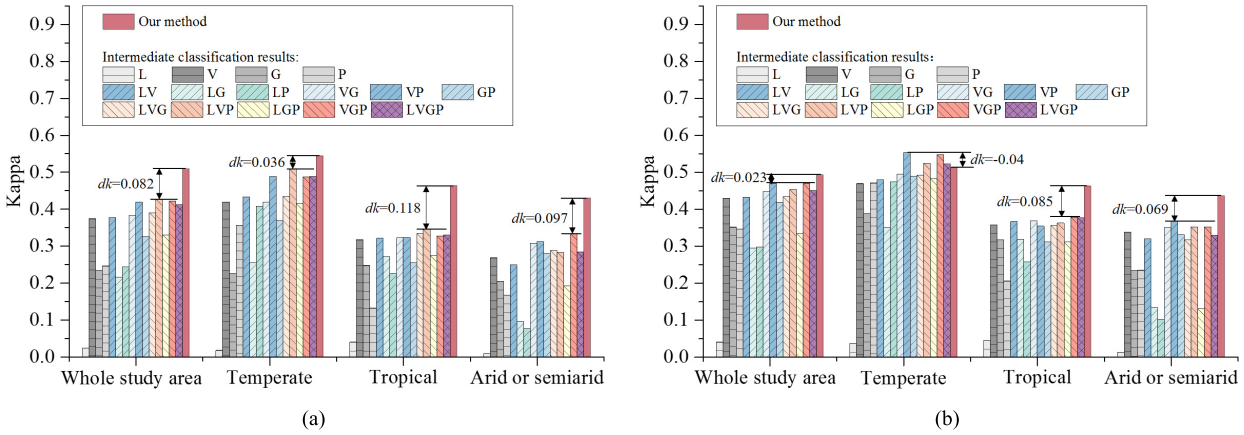


Fig. 5. Kappa coefficients of the mapping results for (a) 2005 and (b) 2010 in different regions, including the whole study area and the three biomes (temperate, tropical, and arid or semiarid in Table II). The accuracies obtained by our method and the intermediate classification results are compared. Here, *dk* is the difference between the Kappa of our method and the highest Kappa of all the intermediate classification results.

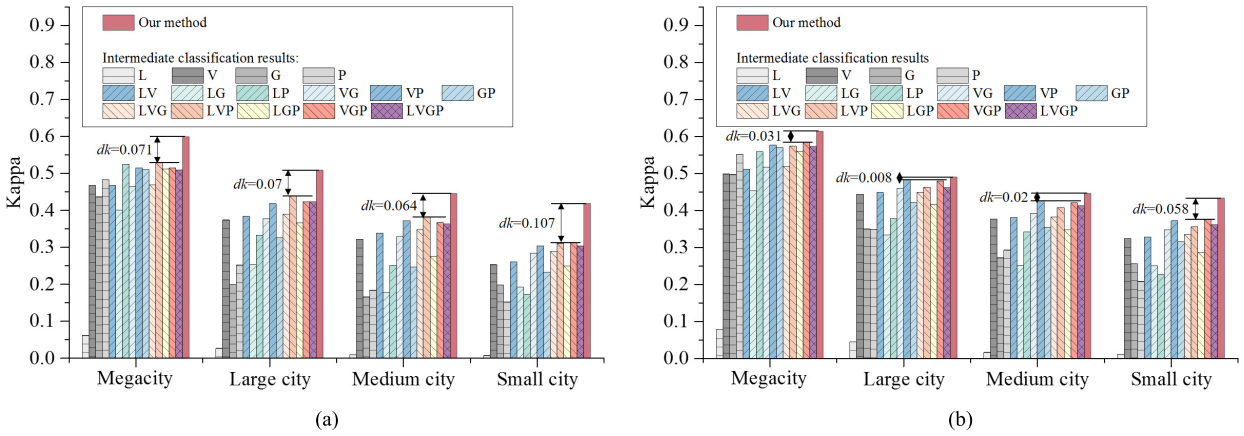


Fig. 6. Kappa coefficients of the mapping results for (a) 2005 and (b) 2010 for different city ranks (megacity, large city, medium city, and small city in Table II). The accuracies obtained by our method and the intermediate classification results are compared. Here, *dk* is the difference between the Kappa of our method and the highest Kappa of all the intermediate classification results.

employed an object-based thresholding method to tackle the diverse urban characteristics and obtain urban areas from NTL data. However, it is noted that these existing techniques require additional reference from high-resolution land-cover data to estimate the threshold. In this paper, based on the ensemble SVM score, an adaptive patch-based thresholding approach, without the assistance of high-resolution land-cover data, is proposed to derive the final urban extent. This technique can automatically derive the segmentation threshold of each urban patch. In this adaptive patch-based thresholding technique, the pixels with zero scores are filtered out and only the pixels whose ensemble SVM scores are nonzeros are aggregated as candidate urban patches. Second, Otsu's thresholding method is employed on each potential urban patch to capture the final urban map. Therein, Otsu's thresholding method [52] is an adaptive thresholding technique that can automatically capture the optimum threshold separating the two classes.

IV. Results

To validate the effectiveness of our method, the national land-use/land-cover data sets of China (CLUD) at a 30-m spatial resolution (for 2005 and 2010) were utilized as reference data. The CLUD data sets were produced by expert interpretation of Landsat Thematic Mapper and HJ-1A/B satellite images [53] and were derived from the Earth System Scientific Data Sharing Network at the Chinese Academy of Sciences. The overall accuracy of the CLUD data is larger than 90%. In this paper, to match the spatial resolution of our results, the spatial resolution of CLUD was aggregated to 500 m using a simple majority rule, and all the pixels of CLUD were used as test samples. In this section, the intermediate classification results by the individual SVM classifiers (shown in Fig. 4) and the map produced by our ensemble SVM method are validated (Figs. 5–7). Here, the intermediate classification results are named by the corresponding abbreviations of the data-source permutations (as shown in Table III), e.g., the VANUI+POP (VP) map represents the intermediate result generated by the use of the VANUI and POP data. In addition, the Kappa coefficient (k) [54] based on the confusion matrix is utilized to evaluate the quantitative accuracy [55].

The Kappa coefficients of all the intermediate classification results and our method for different regions and city ranks are presented in Figs. 5 and 6. Moreover, the urban extent is visually depicted in Fig. 7. The general conclu-

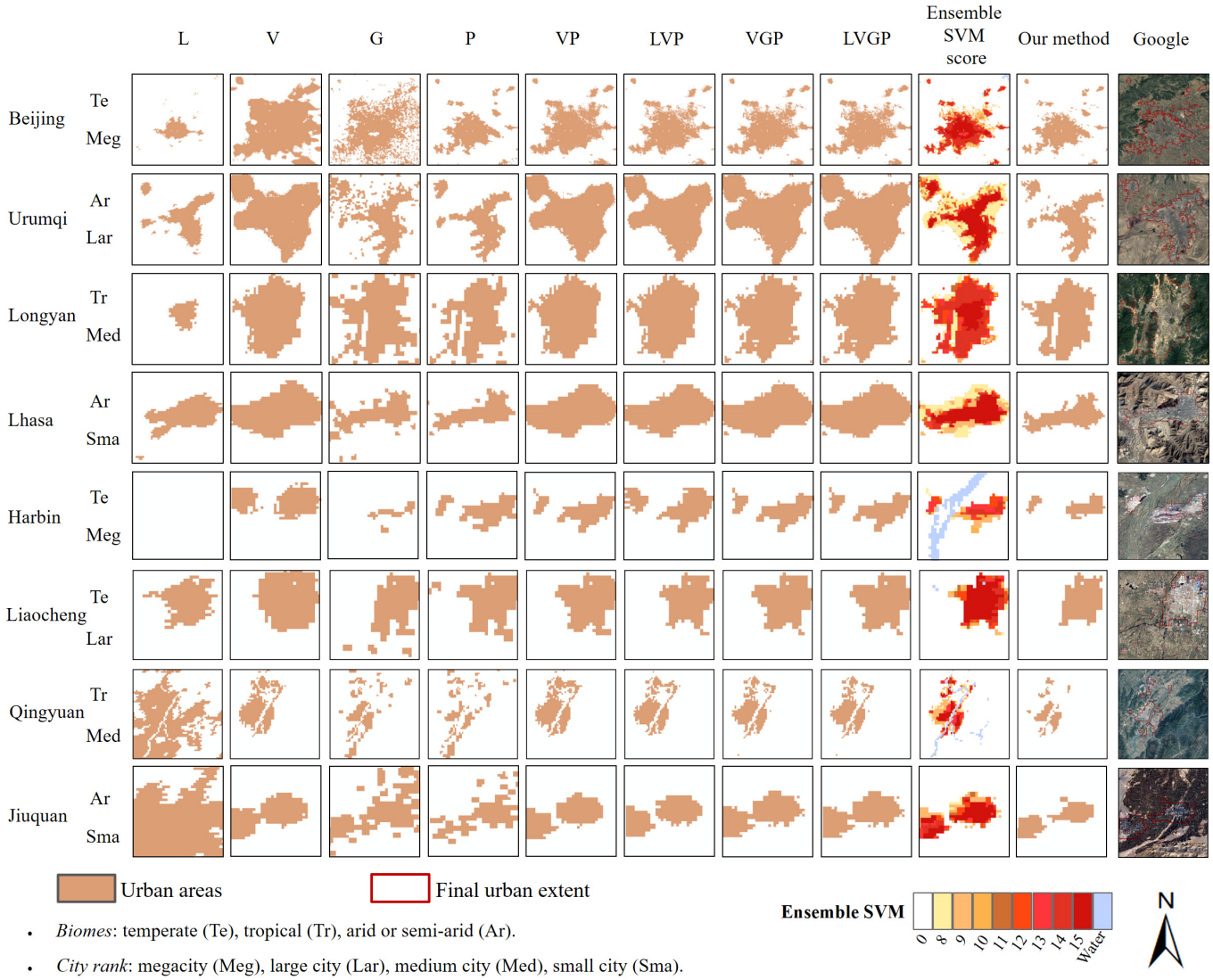


Fig. 7. Comparison among the 2010 urban mapping results for eight cities. (Left–Right) Maps include the intermediate classification results, the ensemble SVM score, the map obtained by our method, and the 2010 Google map overlaid by the final urban extent (highlighted in the red lines). The intermediate classification results include L, V, G, P, and VP (the two source permutation results with the highest Kappa), LVP and VANUI+GDP+POP (the three source permutation results with the highest Kappa), as well as LVGP. Note that the eight cities (superscript: the biome; subscript: the city rank) were chosen from different biomes and city ranks.

sions regarding the results in Figs. 5–7 are summarized as follows.

1) The Kappa values generally become higher and more stable with the addition of data sources, and the map obtained by our method presents the best performance, in both visual inspection (Fig. 7) and quantitative assessment (Fig. 5). From Fig. 5, it can be seen that, for mapping the whole study area, our method holds a higher Kappa ($k = 0.509$ in 2005 and $k = 0.493$ in 2010) than all the intermediate classification results ($k = 0.024$–0.427 in 2005 and $k = 0.040$–0.470 in 2010).
2) When focusing on the single source permutations, it is demonstrated that the V map of the whole study area (Fig. 5) obtains the highest k ($k = 0.374$ in 2005 and $k = 0.430$ in 2010), but presents the largest false-positive errors (Fig. 7), which is probably due to the blooming effect partially remaining in the VANUI data. On the other hand, the L map of the whole study area (Fig. 5) obtains the worst results ($k = 0.024$ in 2005 and $k = 0.040$ in 2010). It is deemed that the use of the LST difference for distinguishing urban/nonurban areas can only play an auxiliary role, owing to the complex temperature variation across large geographic and climatic scales [7].
3) Due to the blooming effect remaining in the VANUI data, the intermediate classification results aligned with this data can only alleviate this problem to some degree [e.g., Longyan of LST+VANUI+GDP+POP (LVGP) in Fig. 7]. However, the ensemble SVM score (as shown in Fig. 7) suggests that the proposed approach has the potential to tackle this issue and presents robust results in extracting urban areas, as shown in the urban map obtained by our method, which can be attributed to the fusion of the multiple properties.

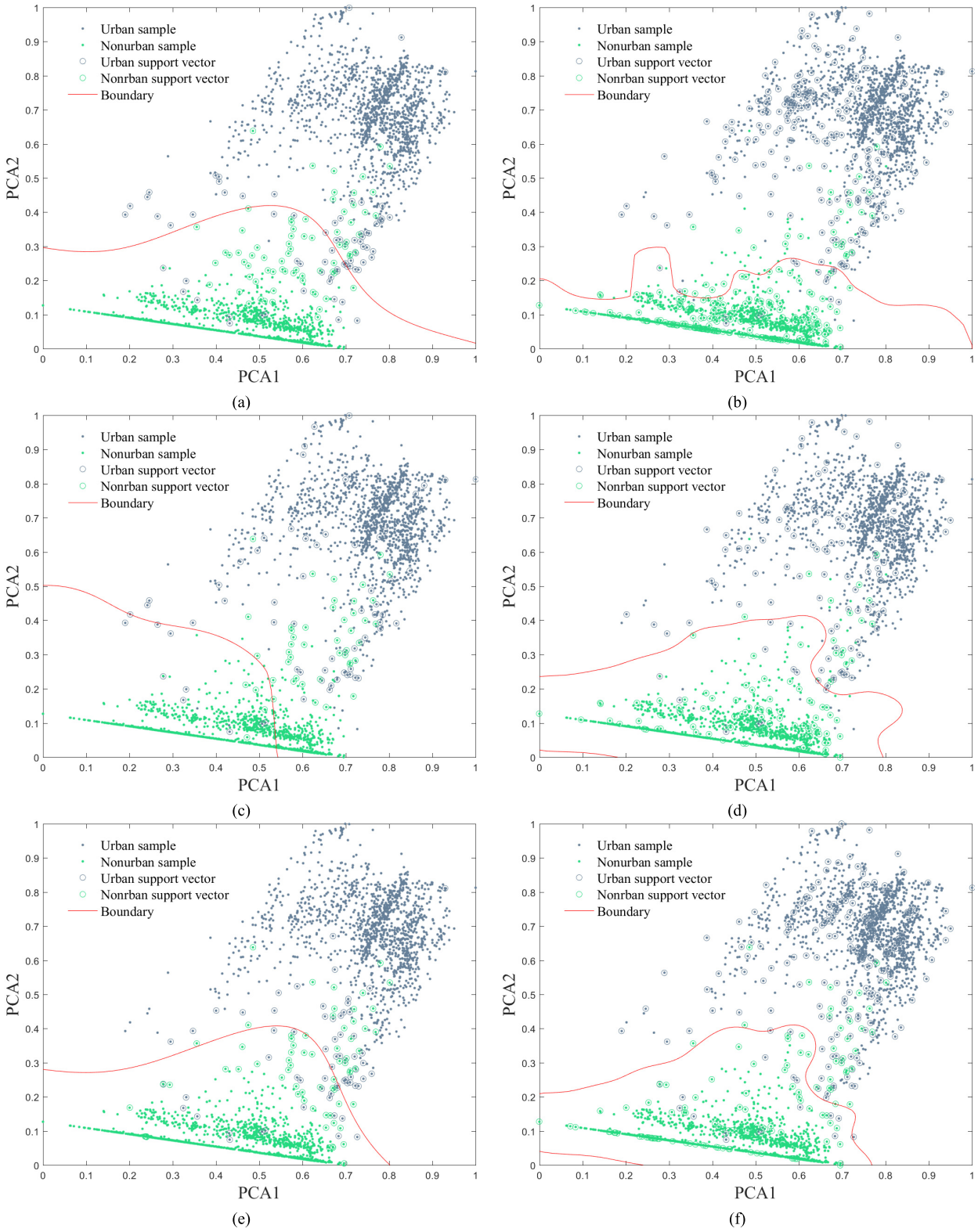


Fig. 8. Support vectors and decision boundaries of the individual classifiers, with different feature combinations. (a) LV. (b) LG. (c) VP. (d) LVG. (e) LVP. (f) LVGP.

4) The results in the three biomes (Fig. 5) and four city ranks (Fig. 6) are similar to those in the whole study area, while some subtle disparities still exist between different biomes and different city ranks. The G map shows a higher accuracy for the tropical biome and small cities, while the P map is superior for the temperate biome and other city ranks. In general, the Kappa coefficients tend to decline from the temperate biome to the arid or semiarid biome (i.e., from left to right in Fig. 5), similar to the tendency presented in the

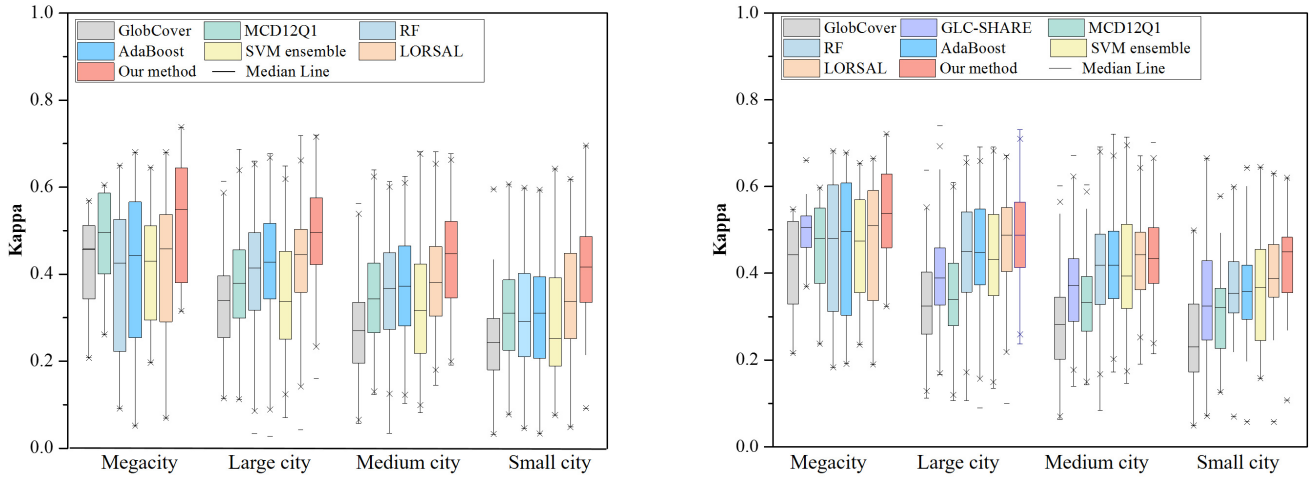


Fig. 9. Box plots that compare the Kappa coefficients among our method, the mainstream global urban maps (Table I), and the maps obtained by four state-of-the-art classifiers (RF, AdaBoost, SVM ensemble, and LORSAL) for (a) 2005 and (b) 2010 for four city ranks (Table II), utilizing the CLUD data as a reference. The mainstream data sets circa 2005 include GlobCover 2005 and MCD12Q1 2005, and the mainstream data sets circa 2010 are GlobCover 2010, GLC-SHARE 2012, and MCD12Q1 2010.

four city ranks (i.e., from large cities to small cities in Fig. 6). This indicates that mapping the urban extent in the latter two biomes [1] and the small cities [18] is a challenging task. In the urban map obtained by our method, such difficulties can be alleviated, in view of the accuracy increments (referred to as dk in Figs. 5 and 6) in the tropical ($dk = 0.118$ in 2005 and $dk = 0.085$ in 2010) and arid or semiarid biomes ($dk = 0.097$ in 2005 and $dk = 0.069$ in 2010) and small city ranks ($dk = 0.107$ in 2005 and $dk = 0.058$ in 2010).

In order to show the diversity of individual classifiers, the support vectors and decision boundaries of the SVMs constructed by the training samples in the temperate biomes in 2010 are illustrated in Fig. 8. The first and second principal components (PCA1 and PCA2) of all the four features, i.e., L, V, G, and P, are utilized to represent the feature space. Fig. 8 shows the support vectors as well as the decision boundaries of the six individual classifiers [with LV, LST+GDP (LG), VP, LST+VANUI+GDP (LVG), LST+VANUI+POP (LVP), and LVGP as input features]. It can be clearly observed that both the distributions of the support vectors and the shapes of decision boundaries are distinct for different classifiers. Thus, it can be said that the individual SVMs are relatively diverse, and the proposed ensemble method has the potential to improve the mapping performance.

To assess the computational cost, without loss of generality, taking the urban mapping of China for 2010 as an example, the running time (including training and testing) of the proposed method with different numbers of multisource inputs is listed in Table IV. The experiments were implemented in the LibSVM package [51] and carried out using MATLAB R2015b on a PC with a single i7-6700K 4.00-GHz processer and 64 GB of RAM. Although the processing time becomes longer with the addition of data sources (i.e., m), it is acceptable since the study area covers the mainland of China, and it seems reasonable to consider more inputs due to the accuracy increments.

Given four sources in this paper, the running time for m sources is the average of all the combinations. In the study case of China, q was set to 200 by default, l was roughly equal to 1.5×10^3 per biome, and M was roughly equal to 3.8×10^7).

TABLE IV
AVERAGE PROCESSING TIME OF THE PROPOSED METHOD WITH DIFFERENT NUMBERS OF MULTISOURCE INPUTS

Input data (m)	**Training time** (minutes)	**Testing time** (minutes)
One-source ($m = 1$)	6.99	11.93
Two-source ($m = 2$)	15.28	29.44
Three-source ($m = 3$)	25.85	53.18
Four-source ($m = 4$)	40.00	90.98

V. DISCUSSION

A. Comparison With Existing Products and Methods

A set of mainstream global urban data sets (Table I) was chosen as comparisons with the urban map obtained by our method. In addition, the urban maps derived by the state-of-the-art classifiers, including random forest (RF) [56], AdaBoost [57], SVM ensemble [58], and variable splitting and augmented Lagrangian (LORSAL) classifiers [59], were also compared. The input data for the four classifiers and the training sample set were the same as those utilized in our method. The Kappa coefficients of the mainstream global urban maps (Table I), the RF map (obtained by the RF classifier), AdaBoost map, SVM ensemble map, LORSAL map and the map obtained by our method for 2005 and 2010 for the four city ranks (Table II) are presented in Fig. 9, using the CLUD data as reference.

Compared with the mainstream global urban data sets, our method always provides the highest median k (median $k = 0.417$–0.549 for 2005 and median $k = 0.434$–0.538 for 2010) for all the city ranks. With regard to the maps generated by the four state-of-the-art classifiers (i.e., RF, AdaBoost, SVM ensemble, and LORSAL), the urban map obtained by our method achieves the best performance in nearly all the cases in terms of the Kappa coefficients, showing that the proposed method is able to exploit the multisource features effectively to produce an accurate urban map. Moreover,

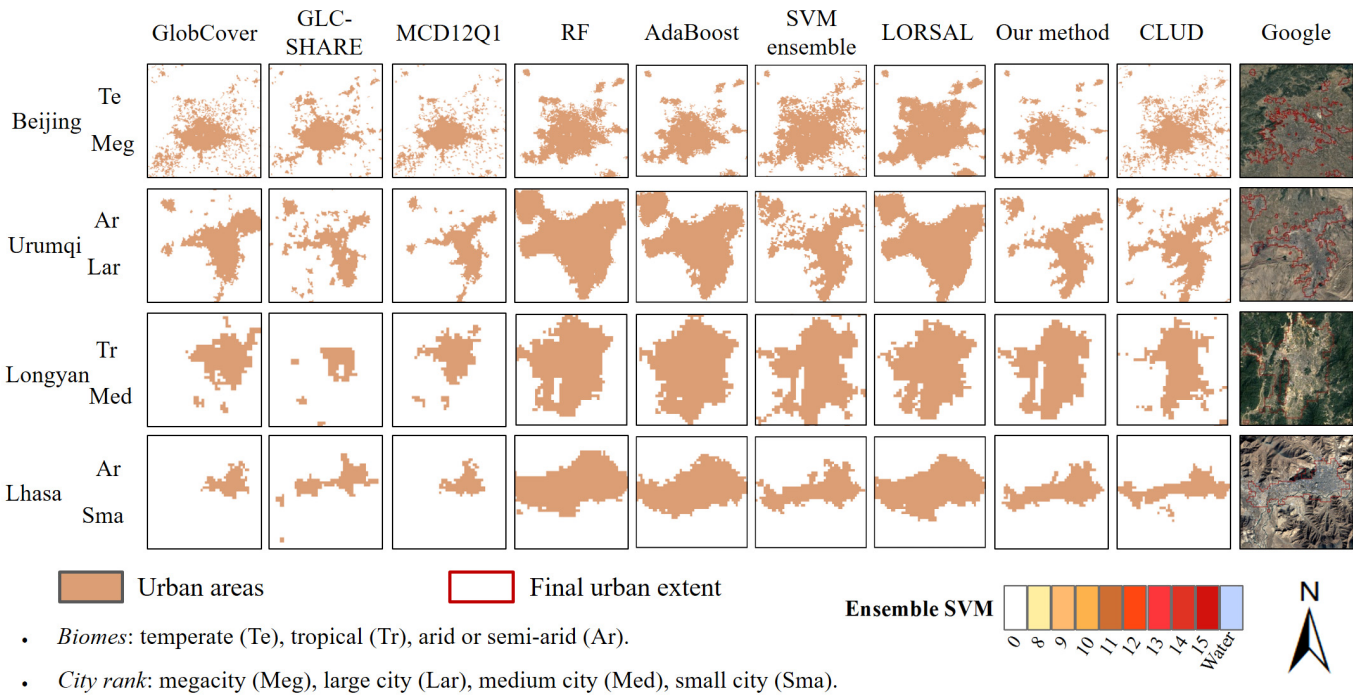


Fig. 10. (Left–Right) Comparison among the 2010 urban extent maps of GlobCover, GLC-SHARE, MCD12Q1, RF, AdaBoost, SVM ensemble, LORSAL, our method, CLUD (as a reference), and the 2010 Google map overlaid by the final urban extent (highlighted in red lines) in four typical cities (Top–Bottom): Beijing, Urumqi, Longyan, and Lhasa. Note that the four cities (superscript: the biome; subscript: the city rank) were chosen from different biomes and city ranks.

TABLE V

KAPPA COEFFICIENTS OF THE WHOLE STUDY AREA, DIFFERENT BIOMES, AND DIFFERENT CITY RANKS WITH DIFFERENT EROSION PARAMETERS

Parameter	Whole	Biomes			City ranks				# Samples
		Te	Tr	Ar	Meg	Lar	Med	Sma	
0.10	0.497	0.539	0.444	0.398	0.608	0.502	0.454	0.417	70417
0.15	0.498	0.532	0.455	0.427	0.614	0.501	0.450	0.419	41661
0.20	0.500	0.517	0.472	0.434	0.613	0.493	0.447	0.431	24153
0.25	0.498	0.518	0.488	0.433	0.618	0.501	0.443	0.437	13865
0.30	0.490	0.500	0.491	0.431	0.620	0.490	0.446	0.423	6954

Biomes: temperate (Te), tropical (Tr), arid or semi-arid (Ar);
City ranks: megacity (Meg), large city (Lar), medium city (Med), small city (Sma);
Samples: the number of training samples for the class of urban.

a visual comparison of the 2010 urban extent maps in the four typical cities (Fig. 10) shows that the urban map produced by our proposed method is in the best agreement with both the Google map and the CLUD reference data, particularly in the medium (e.g., Longyan) and small city ranks (e.g., Lhasa), while underestimation can be found in the mainstream global urban data sets and some overestimation can be observed in the RF map, AdaBoost map, SVM ensemble map, and LORSAL map for all the four city ranks (Fig. 10).

B. Considerations for Accurate Urban Mapping

To accurately map urban extent at a large scale in an efficient way, three challenges need to be addressed: 1) the costly acquisition of sample sets [14]–[16]; 2) the inadequate description of the multisource data [11]; and 3) the diverse characteristics of urban areas [18], [19]. Hence, Sections V-B1–V-B3 focus on these aspects and detail the countermeasures proposed in the ensemble SVM approach.

1) Automated Sample Set Generation: In this section, we investigated the influence of the adaptive erosion radius to the accuracy of the urban mapping. The four coefficients in (1) are set the same and the results with different coefficients are listed in Table V. It can be seen that its value does not significantly affect the mapping results when ranging from 0.1 to 0.3 and the peak of the overall accuracy is reached when the parameter is set to 0.2. However, the number of available training samples decreases sharply when the erosion parameter increases. Therefore, in this paper, an appropriate parameter range is around 0.2 and further fine-tuned in the experiments, considering the satisfactory mapping accuracy and the sufficient number of training samples available.

Subsequently, the CLUD data were utilized as a reference to validate the reliability of the automatically generated sample set. The validation result (Table VI) confirms the reliability of the derived sample sets (mean Kappa $>$ 0.8 and standard error $\sim$0.004). Meanwhile, with regard to the complicated

TABLE VI

ACCURACY ASSESSMENT OF THE 10 RANDOM SAMPLE GROUPS, UTILIZING 2005 AND 2010 CLUD DATA AS THE REFERENCE

Year	Mean Kappa	Standard error of Kappa
2005	0.808	0.004
2010	0.841	0.004

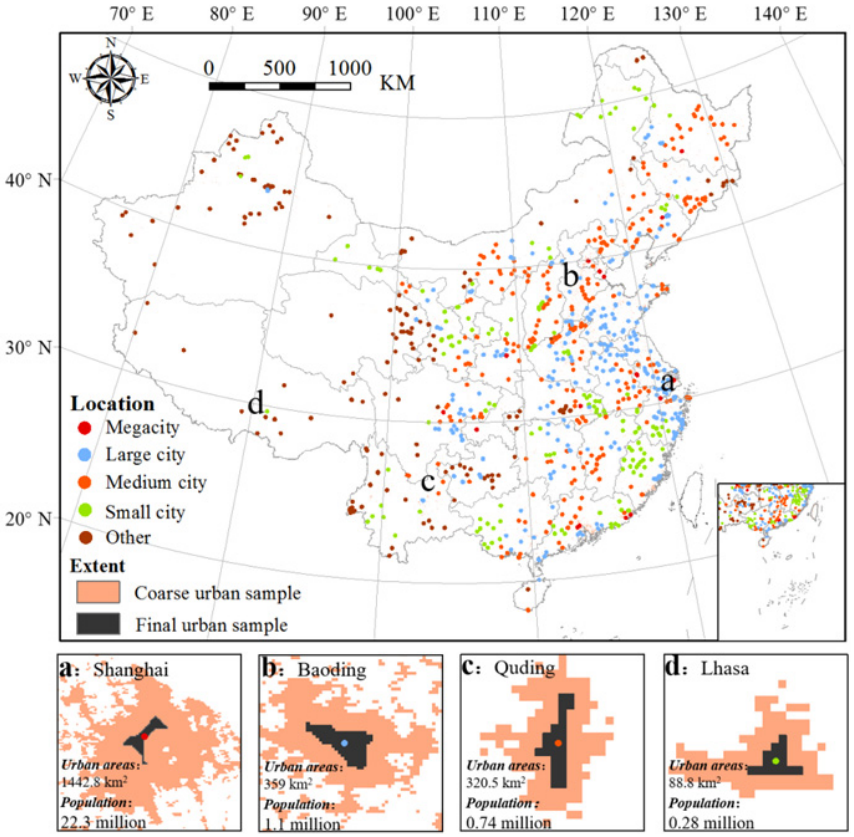


Fig. 11. Geographic distribution of the urban sample set.

TABLE VII

URBAN SAMPLE RATE FOR THE FOUR CITY RANKS

City rank	Urban area (pixels) (2005,2010)	Urban sample (pixels)	Urban sample rate (2005,2010)
Megacity	71188, 92321	4328	6.1%, 4.7%
Large city	186728, 284211	9076	4.9%, 3.2%
Medium city	81616, 112457	4557	5.6%, 4.1%
Small city	23707, 31075	1718	7.3%, 5.5%

The urban area was calculated with the CLUD data.

characteristics of urban areas (e.g., various sizes and shapes), maintaining diversity of the urban sample set is of vital importance. To quantitatively measure the diversity, we define it as the difference between the urban sample rates (i.e., the ratio between the urban samples and the total urban area) of the four city ranks. A relatively small difference indicates that each city rank provides approximate sample proportions, which ensures diversity of the generated urban samples. In Table VII, the urban sample rates for each city rank are listed, which reveal the satisfactory diversity of the generated urban samples, with regard to the small difference of the sample rate [i.e., 2.4% (7.3% − 4.9%) for 2005 and 2.3% (5.5% − 3.2%) for 2010]. In addition, Fig. 11 shows that the geographic distribution of the urban samples is comparatively even. To sum up, it can be stated that the automated sample sets generated in a cost-effective manner are reliable, diverse, and evenly distributed.

2) Multisource Data Integration: The four sample sites in Fig. 12 are presented to illustrate the performance of integrating physical and socioeconomic sources and the effectiveness of the ensemble SVM method, by using the high spatial resolution Google map as a reference. The regions of interest that are highlighted by red rectangles in Fig. 12 reflect the discrepancies between the Google map (located in the center, with the magnified subregion circled in yellow) and the various classification maps (located on both sides). Here, the urban map obtained by our method and the LVGP map are chosen to describe the significance of the multisource integration. As a benchmark, two of the intermediate classification results in Fig. 3 are picked to represent the urban maps produced by

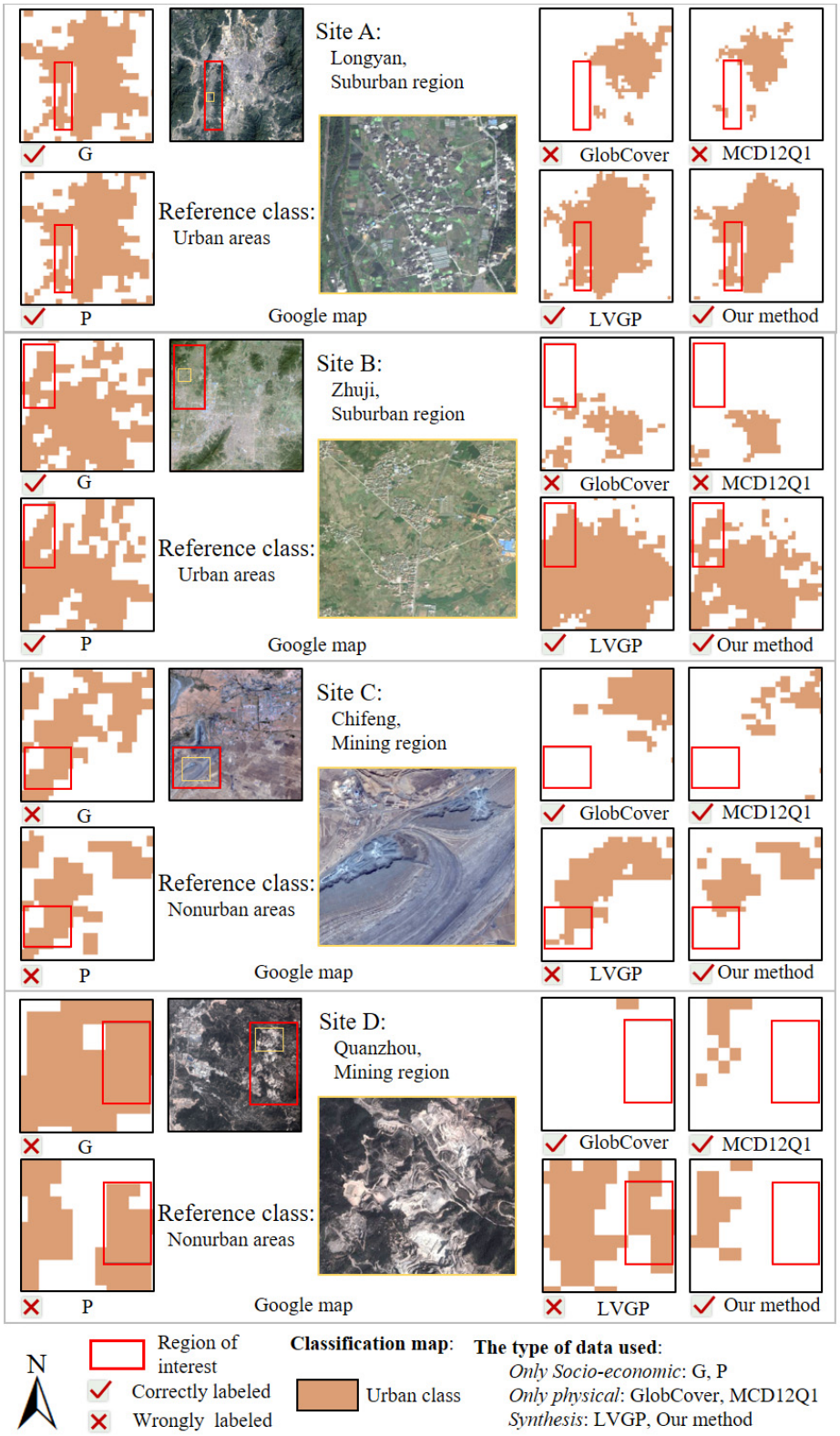
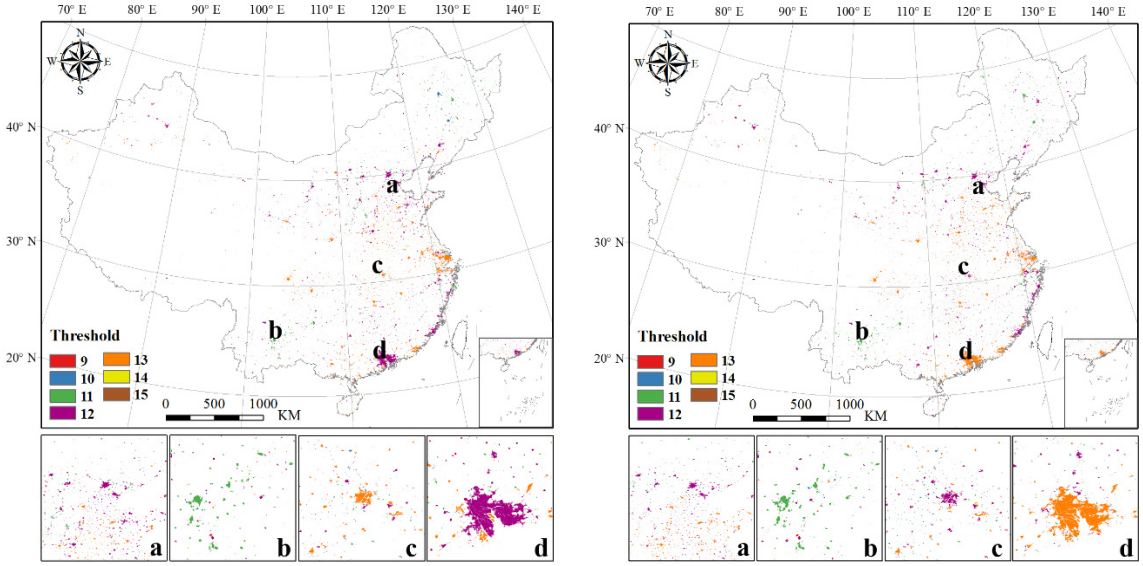


Fig. 12. Four sample sites illustrating the significance of combining both socioeconomic data and physical data. All the maps are circa 2010. The Google images are provided as ground reference. Readers can focus on the red rectangles for observing the discrepancies between different maps, and the areas circled in yellow are magnified. In addition, the right/wrong tick under these urban maps indicates that the map correctly/wrongly labels the region of interest.

the socioeconomic factors, and two existing urban maps (i.e., GlobCover and MCD12Q1) are selected to represent the urban maps produced by the physical factors [1]. In this manner, by observing the regions of interest highlighted by the red rectangles, the following three conclusions can be drawn.

1) Some suburban areas omitted by physical data can successfully be recognized by socioeconomic data. For example, in sites A and B, correctly identify the



(I) (II)

Fig. 13. Distribution of the thresholds derived by the proposed adaptive thresholding algorithm.

TABLE VIII

KAPPA COEFFICIENTS OF THE WHOLE STUDY AREA, DIFFERENT BIOMES, AND DIFFERENT CITY RANKS WHEN MODIS MULTISPECTRAL BANDS ARE ADDED IN THE PROPOSED ENSEMBLE FRAMEWORK

Features	Whole	Biomes			City ranks				Time (minutes)
		Te	Tr	Ar	Meg	Lar	Med	Sma	
LVGP	0.493	0.513	0.463	0.436	0.614	0.490	0.445	0.433	130.98
LVGPM	0.520	0.544	0.503	0.448	0.644	0.515	0.461	0.469	325.14
dk	0.028	0.031	0.040	0.012	0.030	0.025	0.016	0.036	–

Biomes: temperate (Te), tropical (Tr), arid or semi-arid (Ar);
City ranks: megacity (Meg), large city (Lar), medium city (Med), small city (Sma);
'LVGPM' represents the incorporation of MODIS data; '*dk*' is increment of the Kappa when considering the MODIS multispectral bands in the proposed mapping framework.

suburban areas, while GlobCover and MCD12Q1 mislabel these areas. This phenomenon is probably due to the fact that suburban areas show moderate economic and POP activity, but low-density buildings.

2) Similarly, consideration of physical data can reduce the areas mislabeled by socioeconomic data. For instance, for the mining areas with high socioeconomic values but scarce buildings (as shown in sites C and D), they are correctly classified as nonurban areas in GlobCover and MCD12Q1, while they are mislabeled as urban areas in the G and P maps.
3) To integrate the virtues of the multisource data, the urban map produced by the proposed ensemble method is superior to the LVGP map, which was obtained by directly stacking and classifying all the sources. As shown in Fig. 12, the regions of interest for all sites are correctly labeled in the urban map obtained by our method, while LVGP correctly labels these regions only in sites A and B. This superiority partly comes from the flexible combinations of the data sources in our method, i.e., a comprehensive fusion of all the individual and intermediate classifiers.

3) Adaptive Patch-Based Thresholding Technique: In this paper, we have proposed an adaptive patch-based thresholding method, in response to the variations of urban areas. In Fig. 13, the distribution of the thresholds calculated by the proposed adaptive patch-based thresholding algorithm is presented, showing a similar threshold distribution pattern for 2005 and 2010, in spite of some local inconsistencies (e.g., sites C and D). Moreover, for each year, the cities with different urbanization levels and different geographical locations present various thresholds, e.g., sites A–D in Fig. 13. To further illustrate the heterogeneity of the thresholds of patches within a city, Wuhan (I) and Lanzhou (II) are selected as sample cities in Fig. 14. It is demonstrated that the urban extent obtained by our method matches well with the actual areas (i.e., Google map), and the proposed adaptive thresholding algorithm presents a satisfactory capability in detecting challenging small villages (e.g., site A in Wuhan and Lanzhou). In addition, the adaptive thresholds for each site in Wuhan (site A: 11, site B: 13, and site C: 12) and Lanzhou (site A: 13, site B: 11, and site C: 13) are different, showing that the preferred threshold of each urban patch should be adapted to its own characteristics, rather than a uniform criterion within a city. For example, for Wuhan, if the threshold of site B was applied to site A, the urban areas of site A would be overwhelmed, as the threshold of site B (i.e., 13) is larger than the maximum DN of site A (i.e., 11). Therefore, it is meaningful to employ adaptive thresholds to derive more reasonable and detailed urban areas.

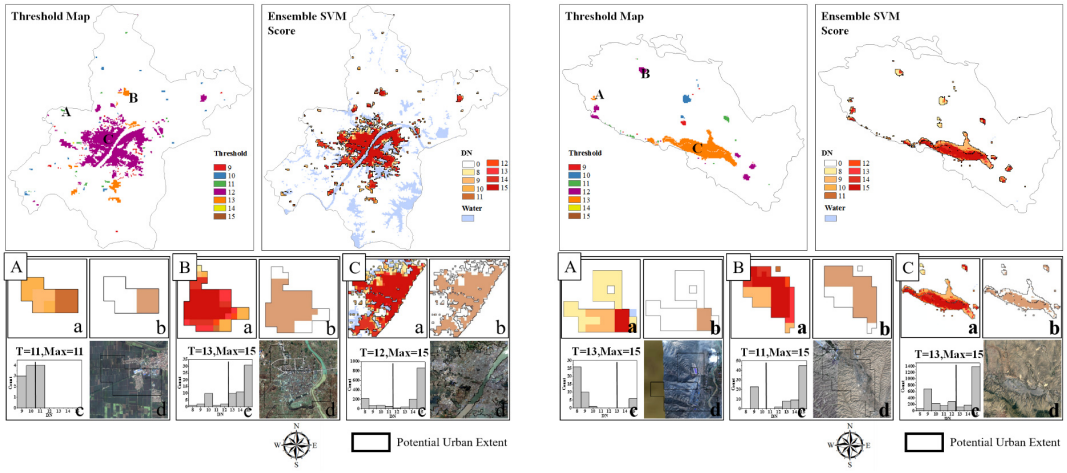


Fig. 14. Two sample cities illustrating the performance of the adaptive thresholding technique. For each city, three sample sites are shown: sites A–C. Site A shows the urban patches of village and town, site B shows the urban patches of county and district, and site C shows the urban patches of the city core. For each site, four subimages are provided. (a) Potential urban patches, which represent the nonzero pixels of the ensemble SVM score. (b) Urban map obtained by our method. (c) Histogram of potential urban patches, where T is the threshold, and Max is the maximum scores of potential urban patches. (d) 2010 Google map overlaid by potential urban patches (marked in the black lines).

C. Incorporation of MODIS Multispectral Data

The MODIS multispectral bands have been considered and incorporated as an additional feature into the proposed ensemble framework. Specifically, the eight-day composite MODIS Surface Reflectance Product (MOD09A1) at 500-m resolution is taken into account, covering the spectrum of visible and near-infrared: bands 1 (red: 620–670 nm), 2 (near-infrared 1: 841–876 nm), 3 (blue: 459–479 nm), 4 (green: 545–565 nm), 5 (near-infrared 2: 1230–1250 nm), 6 (shortwave infrared 1: 1628–1652 nm), and 7 (shortwave infrared 2: 2105–2155 nm). Similar to [60], we use the percentile images extracted from the image collection of Year 2010 to exclude clouds and cloud shadows.

The experimental results with MODIS spectral bands incorporated are listed in Table VIII. Compared to the results obtained by the four features (LST, VANUI, GDP, and POP), mapping accuracy increments (referred to as *dk* in Table VIII) can be obtained for all the biomes and city ranks ($dk = 0.028$ for the whole study area) when considering MODIS multispectral data as an additional feature. On the other hand, however, the computational time (including both training and testing) increases. Therefore, it can be stated that the proposed mapping framework can be further enhanced by taking more effective input features into account.

VI. Conclusion

In this paper, we have proposed a novel ensemble SVM method for deriving accurate urban areas by combining socioeconomic and physical data. The proposed method contains three contributions: 1) reliable and diverse samples are generated by an automatic procedure; 2) the information in the multisource data is effectively utilized by constructing an ensemble SVM model; and 3) an accurate and reasonable urban extent can be captured by employing an adaptive patch-based thresholding method. Finally, the urban map produced in this paper was validated with the CLUD data set and compared with the mainstream global data sets. The results confirmed the superiority of the proposed method, especially in challenging areas (such as tropical and arid or semiarid biomes [1], as well as small cities [18]). In addition, the mapping accuracy can be further increased when MODIS multispectral data are incorporated as an additional source to the proposed ensemble method.

In the future, it is possible to include more data sources, due to the high flexibility of the proposed method. Moreover, the proposed method has the potential to map urban areas and monitor urbanization at global scale, due to its cost-effectiveness and high accuracy. Moreover, the proposed method can be useful for studies of the effect of urbanization on the environment, e.g., urban heat islands [4], air pollution, water contamination, habitat loss, and biodiversity decline [5], [6]. In view of this, further exploring the potential of the proposed method and generating a global urban map are in our future agenda.

Acknowledgment

The authors would like to thank anonymous reviewers for the insightful and constructive suggestions, which significantly improved the quality of this paper.

References

[1] A. Schneider, M. A. Friedl, and D. Potere, "Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions,"' *Remote Sens. Environ.*, vol. 114, no. 8, pp. 1733–1746, Aug. 2010.

[2] X. Bai, J. Chen, and P. Shi, "Landscape urbanization and economic growth in China: Positive feedbacks and sustainability dilemmas," *Environ. Sci. Technol.*, vol. 46, no. 1, pp. 132–139, Nov. 2012.

[3] P. Longley, "Global mapping of human settlement: Experiences, datasets, and prospects," *Photogramm. Rec.*, vol. 25, no. 130, pp. 205–207, 2010.

[4] D. X. Tran, F. Pla, P. Latorre-Carmona, S. W. Myint, M. Caetano, and H. V. Kieu, "Characterizing the relationship between land use land cover change and land surface temperature," *ISPRS J. Photogramm. Remote Sens.*, vol. 124, pp. 119–132, Feb. 2017.

[5] J. A. Foley *et al.*, "Global consequences of land use," *Science*, vol. 309, no. 5734, pp. 570–574, 2005.

[6] M. L. Imhoff, C. J. Tucker, W. T. Lawrence, and D. C. Stutzer, "The use of multisource satellite and geospatial data to study the effect of urbanization on primary productivity in the United States," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 6, pp. 2549–2556, Nov. 2000.

[7] X. Huang, A. Schneider, and M. A. Friedl, "Mapping sub-pixel urban expansion in China using MODIS and DMSP/OLS nighttime lights," *Remote Sens. Environ.*, vol. 175, pp. 92–108, Mar. 2016.

[8] C. M. Mertes, A. Schneider, D. Sulla-Menashe, A. J. Tatem, and B. Tan, "Detecting change in urban areas at continental scales with MODIS data," *Remote Sens. Environ.*, vol. 158, pp. 331–347, Mar. 2015.

[9] J. Latham, R. Cumani, I. Rosati, and M. Bloise, "Global land cover share (GLC-SHARE) database beta-release version 1.0-2014," FAO, Rome, Italy, 2014.

[10] D. Potere, A. Schneider, S. Angel, and D. L. Civco, "Mapping urban areas on a global scale: Which of the eight maps now available is more accurate?" *Int. J. Remote Sens.*, vol. 30, no. 24, pp. 6531–6558, Dec. 2009.

[11] A. Schneider, M. A. Friedl, D. K. McIver, and C. E. Woodcock, "Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data," *Photogramm. Eng. Remote Sens.*, vol. 69, pp. 1377–1386, Dec. 2003.

[12] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis*, vol. 10, no. 2. Berlin, Germany: Springer, 2006.

[13] T. Zhang, X. Huang, D. Wen, and J. Li, "Urban building density estimation from high-resolution imagery using multiple features and support vector regression," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 7, pp. 3265–3280, Jul. 2017.

[14] C. Huang *et al.*, "Use of a dark object concept and support vector machines to automate forest cover change analysis," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 970–985, 2008.

[15] X. Huang, C. Weng, Q. Lu, T. Feng, and L. Zhang, "Automatic labelling and selection of training samples for high-resolution remote sensing image classification over urban areas," *Remote Sens.*, vol. 7, no. 12, pp. 16024–16044, Dec. 2015.

[16] X. Cao, J. Chen, H. Imura, and O. Higashi, "A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data," *Remote Sens. Environ.*, vol. 113, no. 10, pp. 2205–2209, Oct. 2009.

[17] Z. Liu, C. He, Q. Zhang, Q. Huang, and Y. Yang, "Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008," *Landscape Urban Planning*, vol. 106, no. 1, pp. 62–72, May 2012.

[18] Y. Xie and Q. Weng, "Updating urban extents with nighttime light imagery by using an object-based thresholding method," *Remote Sens. Environ.*, vol. 187, pp. 1–13, Dec. 2016.

[19] Y. Zhou, S. J. Smith, C. D. Elvidge, K. Zhao, A. Thomson, and M. Imhoff, "A cluster-based method to map urban area from DMSP/OLS nightlights," *Remote Sens. Environ.*, vol. 147, pp. 173–185, May 2014.

[20] Q. Zhang and K. C. Seto, "Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data," *Remote Sens. Environ.*, vol. 115, no. 9, pp. 2320–2329, Sep. 2011.

[21] C. S. Small, "Mapping urban growth and development as continuous fields in space and time," *Rev. Dept. Geografia*, vol. 1, pp. 155–179, Aug. 2014.

[22] K. K. Goldewijk, "Three centuries of global population growth: A spatial referenced population (density) database for 1700–2000," *Population Environ.*, vol. 26, no. 4, pp. 343–367, Mar. 2005.

[23] D. M. Danko, "The digital chart of the world project," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 8, pp. 1125–1128, 1992.

[24] E. Bartholomé and A. S. Belward, "GLC2000: A new approach to global land cover mapping from Earth observation data," *Int. J. Remote Sens.*, vol. 26, no. 9, pp. 1959–1977, Feb. 2007.

[25] S. Bontemps, P. Defourny, E. Van Bogaert, O. Arino, V. Kalogirou, and J. R. Perez, "GLOBCOVER 2009—Products description and validation report," ESA, Univ. Catholique Louvain, Louvain-la-Neuve, Belgium, Tech. Rep., 2011.

[26] B. Bhaduri, E. Bright, P. Coleman, and J. Dobson, "LandScan: Locating people is what matters," *Geoinformatics*, vol. 5, no. 2, pp. 34–37, Apr. 2002.

[27] C. D. Elvidge, B. T. Tuttle, and P. C. Sutton, "Global distribution and density of constructed impervious surfaces," *Sensors*, vol. 7, no. 9, pp. 1962–1979, 2007.

[28] S. Channan, K. Collins, and W. R. Emanuel, "Global mosaics of the standard MODIS land cover type data," Pacific Northwest Nat. Lab., Univ. Maryland, College Park, College Park, MD, USA, Tech. Rep., 2014, vol. 30.

[29] M. A. Friedl *et al.*, "MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets," *Remote Sens. Environ.*, vol. 114, no. 1, pp. 168–182, Jan. 2010.

[30] CIESIN. (2004). *Global Rural-Urban Mapping Project (GRUMP), Alpha Version: Urban Extents*. Accessed: Aug. 1, 2009. [Online]. Available: http//sedac.ciesin.columbia.edu/gpwlast

[31] C. D. Elvidge *et al.*, "Night-time lights of the world: 1994–1995," *ISPRS J. Photogramm. Remote Sens.*, vol. 56, no. 2, pp. 81–99, Dec. 2001.

[32] X. Bai, P. Shi, and Y. Liu, "Society: Realizing China's urban dream," *Nature*, vol. 509, no. 7499, pp. 158–160, May 2014.

[33] Department of Urban Surveys of National Bureau of Statistics of the People's Republic of China, *China City Statistical Yearbook*, (in Chinese). Beijing, China: China Statistics Press, 2011.

[34] B. Gao, Q. Huang, C. He, Z. Sun, and D. Zhang, "How does sprawl differ across cities in China? A multi-scale investigation using nighttime light and census data," *Landscape Urban Planning*, vol. 148, pp. 89–98, Apr. 2016.

[35] Q. Zhang, C. Schaaf, and K. C. Seto, "The vegetation adjusted NTL urban index: A new approach to reduce saturation and increase variation in nighttime luminosity," *Remote Sens. Environ.*, vol. 129, pp. 32–41, Feb. 2013.

[36] J. Fu, D. Jiang, and Y. Huang, "1 km grid population data of China (2005, 2010)," (in Chinese), Global Change Research Data Publishing & Repository, 2014, doi: 10.3947/geodb.2014.01.06.V1.

[37] Y. Huang, D. Jiang, and J. Fu, "1 km grid GDP data of China (2005, 2010)," (in Chinese), Global Change Research Data Publishing & Repository, 2014, doi: 10.3947/geodb.2014.01.07.V1.

[38] D. Lu and Q. Weng, "Use of impervious surface in urban land-use classification," *Remote Sens. Environ.*, vol. 102, nos. 1–2, pp. 146–160, May 2006.

[39] J. Peng, P. Xie, Y. Liu, and J. Ma, "Urban thermal environment dynamics and associated landscape pattern factors: A case study in the Beijing metropolitan region," *Remote Sens. Environ.*, vol. 173, pp. 145–155, Feb. 2016.

[40] I. Keramitsoglou, C. T. Kiranoudis, G. Ceriola, Q. Weng, and U. Rajasekar, "Identification and analysis of urban surface temperature patterns in greater Athens, Greece, using MODIS imagery," *Remote Sens. Environ.*, vol. 115, no. 12, pp. 3080–3090, Dec. 2011.

[41] Z. Chen *et al.*, "A new approach for detecting urban centers and their spatial structure with nighttime light remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6305–6319, Nov. 2017.

[42] C. D. Elvidge *et al.*, "A fifteen year record of global natural gas flaring derived from satellite data," *Energies*, vol. 2, no. 3, pp. 595–622, Aug. 2009.

[43] H. Letu, M. Hara, G. Tana, and F. Nishio, "A saturated light correction method for DMSP/OLS nighttime satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 389–396, Feb. 2012.

[44] C. Small, F. Pozzi, and C. D. Elvidge, "Spatial analysis of global urban extent from DMSP-OLS night lights," *Remote Sens. Environ.*, vol. 96, nos. 3–4, pp. 277–291, Jun. 2005.

[45] R. C. Sharma, R. Tateishi, K. Hara, S. Gharechelou, and K. Iizuka, "Global mapping of urban built-up areas of year 2014 by combining MODIS multispectral data with VIIRS nighttime light data," *Int. J. Digit. Earth*, vol. 9, no. 10, pp. 1004–1020, Apr. 2016.

[46] A. Schneider, M. A. Friedl, and D. Potere, "A new map of global urban extent from MODIS satellite data," *Environ. Res. Lett.*, vol. 4, no. 4, p. 044003, Oct. 2009.

[47] Y. Yang, C. He, Q. Zhang, L. Han, and S. Du, "Timely and accurate national-scale mapping of urban land in China using defense meteorological satellite program's operational linescan system nighttime stable light data," *J. Appl. Remote Sens.*, vol. 7, no. 1, p. 073535, 2013.

[48] M. Pal and G. M. Foody, "Feature selection for classification of hyperspectral data by SVM," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2297–2307, May 2010.

[49] B. Waske and J. A. Benediktsson, "Fusion of support vector machines for classification of multisensor data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3858–3866, Dec. 2007.

[50] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[51] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 2, 2011, Art. no. 27. [Online]. Available: http://www.scis.ntu.edu.tw/~cjlin/libsvm

[52] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.

[53] J. Liu *et al.*, "Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s," *J. Geograph. Sci.*, vol. 24, no. 2, pp. 195–210, 2014.

[54] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.

[55] O. Allouche, A. Tsoar, and R. Kadmon, "Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS)," *J. Appl. Ecol.*, vol. 43, no. 6, pp. 1223–1232, Sep. 2006.
[56] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[57] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
[58] X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 257–272, Jan. 2013.
[59] J. Li *et al.*, "Multiple feature learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1592–1606, Mar. 2015.
[60] G. Donchyts, J. Schellekens, H. Winsemius, E. Eisemann, and N. Giesen, "A 30 m resolution surface water mask including estimation of positional and thematic differences using Landsat 8, SRTM and OpenStreetMap: A case study in the Murray-Darling basin, Australia," *Remote Sens.*, vol. 8, no. 5, p. 386, 2016.

Xin Huang (M'13–SM'14) received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2009.

He was with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. He is currently a Luojia Distinguished Professor with Wuhan University, where he teaches remote sensing, photogrammetry, and image interpretation. He is also the Founder and the Director of the Institute of Remote Sensing Information Processing, School of Remote Sensing and Information Engineering, Wuhan University. He has been supported by the National Program for Support of Top-notch Young Professionals (2017), the China National Science Fund for Excellent Young Scholars (2015), and the New Century Excellent Talents in University from the Ministry of Education of China (2011). He has authored more than 110 peer-reviewed articles (SCI papers) in the international journals. His research interests include remote sensing image processing methods and applications.

Dr. Huang was a recipient of the Boeing Award for the Best Paper in Image Analysis and Interpretation from the American Society for Photogrammetry and Remote Sensing (ASPRS) in 2010, the Second Place for the John I. Davidson President's Award from ASPRS in 2018, the National Excellent Doctoral Dissertation Award of China in 2012, and the winner of the IEEE Geoscience and Remote Sensing Society (GRSS) 2014 Data Fusion Contest. In 2011, he was recognized by the IEEE GRSS as the best reviewer for IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He was the Lead Guest Editor of the special issue on information extraction from high-spatial-resolution optical remotely sensed imagery for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (vol. 8, no. 5, 2015), and the special issue on Sparsity-Driven High Dimensional Remote Sensing Image Processing and Analysis for the *Journal of Applied Remote Sensing* (vol. 10, no. 4, 2016). Since 2016, he has been an Associate Editor of the *Photogrammetric Engineering and Remote Sensing*. Since 2014, he has been an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

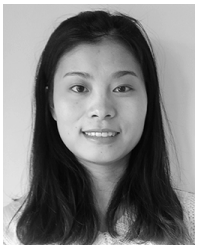
Ting Hu received the B.S. and M.S. degrees from Wuhan University, Wuhan, China, in 2012 and 2015, respectively, where she is currently pursuing the Ph.D. degree with the School of Remote Sensing and Information Engineering.

Her research interests include machine learning, remote sensing mapping, and remote sensing applications.

Jiayi Li (S'12–M'15) received the B.S. degree from Central South University, Changsha, China, in 2011, and the Ph.D. degree from Wuhan University, Wuhan, China, in 2016.

She is currently an Assistant Professor with the School of Remote Sensing and Information Engineering, Wuhan University. Her research interests include hyperspectral imagery, sparse representation, computation vision and pattern recognition, and remote sensing images.

Dr. Li is a reviewer for more than 10 international journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, the IEEE SIGNAL PROCESSING LETTERS, and the *International Journal of Remote Sensing*.



Qing Wang received the B.S. degree from Wuhan University, Wuhan, China, in 2015, where she is currently pursuing the M.S. degree with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing.

Her research interests include machine learning, data fusion, and remote sensing applications.

Jón Atli Benediktsson (S'84–M'90–SM'99–F'04) received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, Iceland, in 1984, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1987 and 1990, respectively.

From 2009 to 2015, he was the Pro Rector of Science and Academic Affairs and a Professor of electrical and computer engineering with the University of Iceland, where he was the President and the Rector in 2015. He is a Co-Founder of the biomedical startup company Oxymap (www.oxymap.com). His research interests include remote sensing, biomedical analysis of signals, pattern recognition, image processing, and signal processing. He has authored extensively in those fields.

Dr. Benediktsson was a member of the 2014 IEEE Fellow Committee. He is a member of the Association of Chartered Engineers in Iceland, Societas Scinetiarum Islandica, and Tau Beta Pi. He is a fellow of the Society of Photo-Optical Instrumentation Engineers. He was a recipient of the Stevan J. Kristof Award from Purdue University in 1991, as Outstanding Graduate Student in Remote Sensing, the Icelandic Research Council's Outstanding Young Researcher Award in 2000, the IEEE Third Millennium Medal in 2004, the Yearly Research Award from the Engineering Research Institute, University of Iceland, in 2007, the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society (GRSS), the OECE Award from the School of ECE, Purdue University, in 2016, and the IEEE/VFI Electrical Engineer of the Year Award in 2013. He was a co-recipient of the University of Iceland's Technology Innovation Award in 2006, the 2012 IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Paper Award, the IEEE GRSS Highest Impact Paper Award in 2013, and the International Journal of Image and Data Fusion Best Paper Award in 2014. He was the Chairman of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING from 2007 to 2010. He was the 2011–2012 President of IEEE GRSS and has been on the GRSS AdCom since 2000. Since 1999, he has been an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS). He was the Editor-in-Chief of IEEE TGRS from 2003 to 2008. He has been an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS since 2003 and the IEEE ACCESS since 2013. He serves on the Editorial Board of the PROCEEDINGS OF THE IEEE, the International Editorial Board of the *International Journal of Image and Data Fusion*, and the Editorial Board of *Remote Sensing*.