

Article

Mapping Urban Areas with Integration of DMSP/OLS Nighttime Light and MODIS Data Using Machine Learning Techniques

Wenlong Jing^{1,2}, Yaping Yang^{1,3,*}, Xiafang Yue^{1,3} and Xiaodan Zhao^{1,3}

¹ State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; E-Mails: jingwl@lreis.ac.cn (W.J.); lexif@lreis.ac.cn (X.Y.); zhaoxd@lreis.ac.cn (X.Z.)

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China

* Author to whom correspondence should be addressed; E-Mail: yangyp@igsnr.ac.cn.

Academic Editors: Janet Nichol, Parth Sarathi Roy and Prasad S. Thenkabail

Received: 12 June 2015 / Accepted: 14 September 2015 / Published: 22 September 2015

Abstract: Mapping urban areas at global and regional scales is an urgent and crucial task for detecting urbanization and human activities throughout the world and is useful for discerning the influence of urban expansion upon the ecosystem and the surrounding environment. DMSP-OLS stable nighttime lights have provided an effective way to monitor human activities on a global scale. Threshold-based algorithms have been widely used for extracting urban areas and estimating urban expansion, but the accuracy can decrease because of the empirical and subjective selection of threshold values. This paper proposes an approach for extracting urban areas with the integration of DMSP-OLS stable nighttime lights and MODIS data utilizing training sample datasets selected from DMSP-OLS and MODIS NDVI based on several simple strategies. Four classification algorithms were implemented for comparison: the classification and regression tree (CART), k-nearest-neighbors (k-NN), support vector machine (SVM), and random forests (RF). A case study was carried out on the eastern part of China, covering 99 cities and 1,027,700 km². The classification results were validated using an independent land cover dataset, and then compared with an existing contextual classification method. The results showed that the new method can achieve results with comparable accuracies, and is easier to implement and less sensitive to the initial thresholds than the contextual method. Among

the four classifiers implemented, RF achieved the most stable results and the highest average Kappa. Meanwhile CART produced highly overestimated results compared to the other three classifiers. Although k-NN and SVM tended to produce similar accuracy, less-bright areas around the urban cores seemed to be ignored when using SVM, which led to the underestimation of urban areas. Furthermore, quantity assessment showed that the results produced by k-NN, SVM, and RFs exhibited better agreement in larger cities and low consistency in small cities.

Keywords: urban areas; DMSP-OLS; MODIS; SVM; random forests

1. Introduction

The urban areas on Earth's land surface have experienced rapid expansion rates through the last three decades [1]. Globally, urban expansion is one of the primary factors in habitat loss and species extinction and results in changes in land cover. Locally, urban areas and urbanization have great, irreversible impacts on their surrounding environments, further affecting local climate and hydrological systems through the modification of albedo and evapotranspiration [2–6]. The information on the magnitude, distribution, pattern, and scale of urban land use is urgently to be quantified at local and global scales, for understanding the spatial extents of urban areas, sustainable management of these areas, and evaluating the impacts of urbanization on environments [1,7].

Remote sensing-based techniques have provided an efficient approach for mapping urban areas at multiple scales. Urban areas or human settlements can be mapped at different scales utilizing remote sensing data with different spatial resolution. High (<10 m) (e.g., SPOT, IKONOS, QuickBird) and medium (10–100 m) (e.g., Landsat TM/ETM+, ASTER) spatial resolution remote sensing imagery have been applied worldwide in mapping urban areas or built-up areas for individual cities or city-regions. [8–11]; for mapping urban areas at regional and global scales, coarse spatial resolution (1–2 km) data are usually employed [12], and nighttime lights (NTL) images from the Defense Meteorological Satellite Program's Operational Line-scan System (DMSP-OLS) have provided effective and accessible data resources that can measure artificial illumination. NTL data have been increasingly used for mapping urban areas and urban expansions [13–15]. When using DMSP-OLS NTL images for mapping urban areas, the threshold technique is often implemented because of its simplicity [16,17]; however, the selection of threshold is almost always empirical or subjective, and high uncertainty can be found when compared across cities at different levels of development [7].

In this paper, we present four machine learning methods, including classification and regression tree (CART), k-nearest neighbors (k-NN), random forests (RF), and support vector machine (SVM), to extract urban areas by using DMSP-OLS and MODIS NDVI data, attempting to develop a new approach to derive urban areas or human settlement from DMSP-OLS. Studies have found that the RF versions are likely to be the best classifiers, and SVM is the second best, without statistically significant differences [18]. CART and k-NN, as basic methods that commonly use machine learning classifiers, were employed for comparison in this study. A case study was carried out for eastern China cities, which have experienced the highest rate of urbanization in China.

2. Study Area and Data Resources

Eastern China cities were selected for the case study, because China has experienced rapid urbanization since the 1980s, and eastern China is the most developed area of China since the last three decades. The study area includes seven provinces and three municipalities, covering 99 cities, with an area of 1,027,700 km² (Figure 1). This accounts for only 10.65% of China's mainland area; however, the population is 553.14 million, accounting for 40.65% of China's total population, and the area accounts for 54.94% of GDP (China Statistical Yearbook, 2010, [19]).

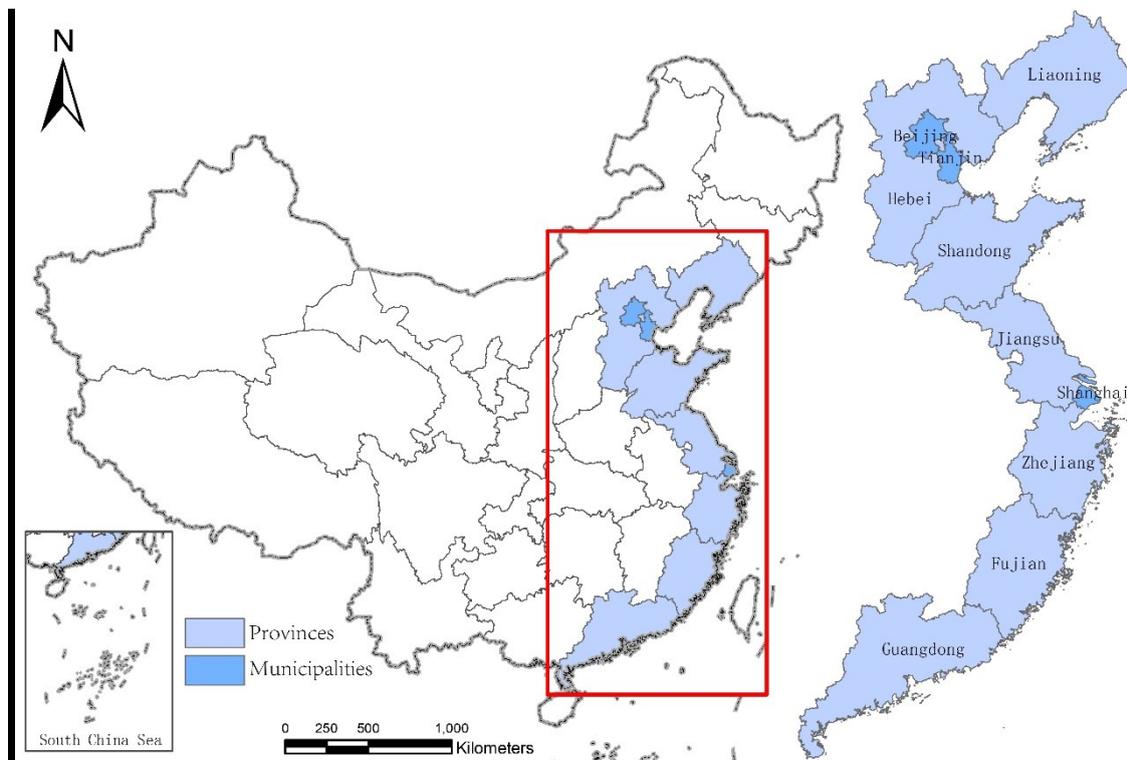


Figure 1. Case study area of eastern China includes seven provinces and three municipalities (Beijing, Tianjin, and Shanghai), which are of provincial-level.

Multiple datasets were utilized in this study, including remote sensing data and land cover products derived from different remote sensing resources. A brief description of data resources is listed in Table 1.

The version 4 DMSP-OLS stable nighttime lights annual image composites for the year 2010 was downloaded from the NOAA National Geophysical Data Center; the original spatial resolution of the products was 30 arcsecond, and the data were re-projected to Albers Conical Equal Area projection and resampled to 1-km resolution using the nearest neighbor resampling algorithm during the re-projection. The DN values of DMSP-OLS NTL range from 0 to 63; classifiers such as k-NN and SVM are not scale invariant, so we scaled the DN values to the range of 0 to 1.0 by using the following normalized algorithm:

$$VA_i = \frac{V_i - V_{min}}{V_{max} - V_{min}} \quad (1)$$

where VA_i is the normalized value of the i -th pixel, V_i is the original value of the i -th pixel, and V_{min} and V_{max} are the minimum and maximum values of all pixels.

Two MODIS products, monthly NDVI (MOD13A3) and land water mask (MOD44W), were downloaded from the NASA Land Processes Distributed Active Archive Center (LP DAAC). These two products having a sinusoidal projection were re-projected to the Albers Conical Equal Area projection, and the nearest neighbor resampling algorithm was used to resample MODIS NDVI images to maintain the pixel size 1 km by 1 km; the majority resampling algorithm was used to resample the land water mask of 250 m resolution to 1 km resolution for maintaining correspondence with NTL and NDVI data. The annual average NDVI was calculated using the average algorithm from January to December. We used the average NDVI instead of the annual maximum NDVI because of the former's stability and reduced sensitivity to seasonal and inter-annual fluctuations [14]. NDVI values were constrained to the range of non-negative values between 0 and 1.0 by using the same normalized algorithm as Equation (1).

The Finer Resolution Observation and Monitoring-Global Land Cover-Hierarchy (FROM-GLC-Hierarchy) dataset was utilized as reference data to validate the extracted results [20]. FROM-GLC-Hierarchy is a global land cover dataset produced using the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) data, and has been aggregated with multi-resolution (*i.e.*, 30 m, 250 m, 500 m, 1 km, 5 km, 10 km, 25 km, 50 km, and 100 km) to meet the requirements for different resolutions from different applications. The 30-m base map of FROM-GLC-Hierarchy has been improved from FROM-GLC-agg [20], and the other multi-resolution data were produced using additional coarse resolution datasets to reduce land cover type confusion [21,22].

We selected the land cover data at a resolution of 1 km for validation in this study to match the spatial scale of NTL and MODIS NDVI. The data were reclassified into two classes, urban and not-urban.

Table 1. A brief description of data resources.

Abbreviation	Data Description	Spatial Resolution	Time
NTL	DMSP-OLS stable nighttime light	1 km	2010
MOD13A3	MODIS monthly NDVI	1 km	2010
MOD44W	MODIS land water mask	250 m	NA
FROM-GLC	Global land cover	1 km	2010

3. Methods

The DMSP-OLS NTL and MODIS NDVI datasets were obtained to extract the urban areas. The key steps of the method are as follows: First stage, the water mask was used to remove the pixels dominated by water bodies. Second, a spectral index, the Vegetation Adjusted NTL Urban Index (VANUI), which combines MODIS NDVI and NTL [14], was calculated as an independent variable of the classification algorithms. Training samples were selected on the basis of thresholds drawn from MODIS NDVI and NTL images. Potential urban pixels were defined with NTL DN greater than 40 and NDVI less than 0.4, and potential non-urban pixels were defined with NDVI greater than 0.4 [7,23]. To ensure that the training pixels were uniformly distributed in the study area, a stratified random sampling was carried out on the basis of the administrative boundary layers. First, a random value between 0.0 and 1.0 was generated for each potential pixel. Then, the probability of selection (P) was calculated:

$$P = N_s/N \quad (2)$$

where N_S is the number of pixels to be drawn for each stratum, and N is the total number of potential pixels. The pixel was selected for the sample if its random value was $<P$. N_S was set to 15% of potential pixels for each stratum. The training set was input into the classifiers for training; after training, the classifiers were applied to the NTL and NDVI data to classify the unknown pixels. A flowchart of the classification procedure is shown in Figure 2.

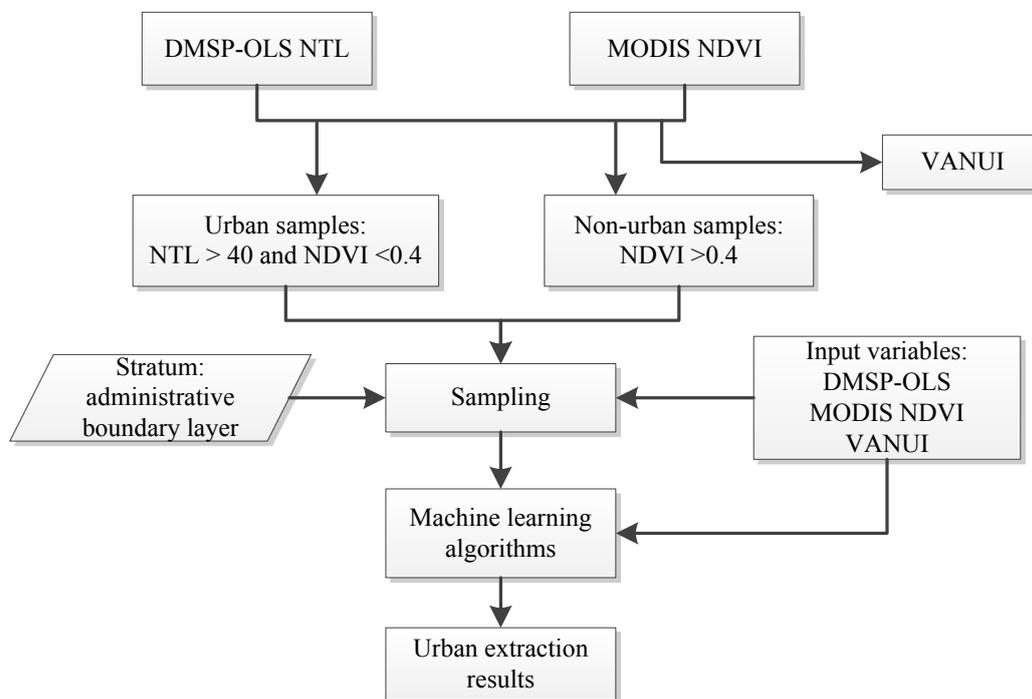


Figure 2. Flowchart of urban area extraction method.

A contextual classification method proposed by Cao *et al.* [7] was also used to compare the performance with that of the new method. This contextual classification algorithm did not classify all unknown pixels at one time, but instead used the classifier in an iterative procedure to classify the pixels within a 3×3 window of each seed pixel step-by-step. After each iteration, newly classified urban pixels were assigned as a new set for the new training procedure. This iterative procedure was carried out until the number of newly identified urban pixels was zero. Non-urban training samples were defined with the same condition aforementioned, and because this method is based on the high degree of spatial clustering of DMSP-OLS nighttime lights, urban samples were selected as the pixels with maximum OLS DN values in each 9×9 window with OLS DN greater than 40 to ensure the inclusion of all the potential urban patches instead of selecting urban samples using a random process based on probability selection.

Mean overall accuracy was calculated to assess the performance of these aforementioned approaches. Overall accuracy (OA) and Kappa coefficient were implemented for each city for further validation.

3.1. VANUI

VANUI is a spectral index proposed by Zhang *et al.* [14] that combines MODIS NDVI and NTL, to reduce the effects of NTL saturation and increase variation of the NTL signal, especially within urban areas. Additionally, the index is intuitive, simple to implement, and was found to correspond to urban characteristics and the percent of imperviousness [14,24]. VANUI was calculated as

$$\text{VANUI} = \text{NTL} \times (1 - \text{NDVI}) \quad (3)$$

where NTL is the normalized DMSP-OLS stable nighttime lights, and NDVI is normalized annual average MODIS NDVI.

3.2. Machine Learning Methods

We experimented with four types of machine learning methods: CART (an improved version of C4.5), k-NN, SVM, and RF.

CART is a commonly used method [25], in which the basic idea is to construct a tree-like graph or model of decisions and their possible consequences by generating relative homogeneous subgroups by recursively partitioning the training dataset to the maximum variance between groups of independent variables and dependent variables [26]. The problem in this study is binary (two-class) classification and involves only three variables; the maximum depth of the tree was set to 50 to avoid creating over-complex trees generated from the training data [27].

The k-NN method is a non-parametric method used for classification and regression [28]. For k-NN classification, an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors; Euclidean distance as a common distance metric for continuous variables is used to define the neighbors. Typically, k (number of neighbors) is a positive integer less than 20 [29]; too-small values of k increase the effect of noise on classification [30], so we set k equal to 10 in this study.

The SVM classifier has been widely used and reported as an outstanding classifier [31]. The basic idea of SVM is to classify the input vectors into two classes using a hyperplane with maximal margin. The maximal margin is derived by solving the constrained quadratic problem:

$$\text{Maximize } W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (4)$$

$$\text{Subject to } \left\{ \sum_{i=1}^n \alpha_i y_i = 0 \text{ and } 0 \leq \alpha_i \leq T \text{ for } i = 1, 2, \dots, n \right\} \quad (5)$$

where $x_i \in R_d$ are the training sample vectors, $y_i \in \{-1, +1\}$ is the corresponding class label, and $K(u, v)$ is the kernel function. We used the radial basis function as the kernel function [8,21,29,32] and default parameters in the implementation.

The RF method is an ensemble learning method for classification and regression; it is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forests. Comparatively, RF does not overfit because of the law of large numbers [33]. Its performance has been reported to be the best among 179 classifiers arising from 17 families (discriminant analysis, Bayesian, neural networks,

SVMs, decision trees, rule-based classifiers, boosting, bagging, stacking, RF and other ensembles, generalized linear models, nearest neighbors, partial least squares and principal component regression, logistic and multinomial regression, multiple adaptive regression splines, and other methods) [18]. Although RF does not overfit and one can run as many trees as desired according to Breiman, who first introduced RF [33], we set the number of trees to 25, which was sufficient depending on the size and nature of the training set (three variables, two classes) and can guarantee efficiency.

The classifiers used in this study were implemented in scikit-learn, which is a Python module integrating a wide range of state-of-the-art machine learning algorithms for medium-scale supervised and unsupervised problems [32]. The parameter set for each algorithm is listed in Table 2.

Table 2. Classification methods and parameters.

Classifiers	Abbreviation	Parameters	Remarks
Classification and Regression Tree	CART	Maximum depth: 50	
k-Nearest Neighbors	k-NN	Number of neighbors: 10	
Support Vectors Machine	SVM	Kernel : RBF C(cost):1.0 gamma: 0.1	Data were scaled to [0,1] before training and classification
Random Forests	RF	Probability estimates: false Number of trees: 25	

4. Results and Discussion

4.1. Performance of Machine Learning Methods

To quantify the performance of the four methods, for every region, 85% of the sample dataset was selected as the input training dataset of the four methods. The overall accuracy (OA) of each method was calculated for both training and testing datasets. The training and testing processes were carried out region by region.

Table 3. Overall accuracy of training and testing for the four methods.

Region	Training OA				Testing OA			
	CART	k-NN	SVM	RF	CART	k-NN	SVM	RF
Beijing	0.998	--	0.95	0.987	0.909	0.97	0.96	0.939
Tianjin	0.997	--	0.932	0.997	0.859	0.915	0.915	0.915
Shanghai	1	--	0.874	0.986	0.769	0.821	0.821	0.744
Hebei	0.996	--	0.988	0.993	0.982	0.986	0.983	0.986
Liaoning	0.992	--	0.983	0.988	0.979	0.983	0.987	0.985
Shandong	0.986	--	0.972	0.982	0.96	0.96	0.966	0.963
Jiangsu	0.988	--	0.973	0.981	0.957	0.961	0.965	0.964
Zhejiang	0.989	--	0.97	0.981	0.966	0.958	0.971	0.966
Fujian	0.999	--	0.99	0.996	0.988	0.99	0.988	0.986
Average	0.994	--	0.959	0.988	0.93	0.949	0.951	0.939

Table 3 shows the training OA and testing OA of each method. The k-NN is an instance-based learning algorithm, or lazy-learning algorithm; all computation is deferred until classification is performed [31], so

there was no training processing in the k-NN algorithm. On average, the best-performing method in training was CART (training OA = 0.994), and SVM was the worst, relatively (training OA = 0.959). The test results showed the method with the highest testing OA was SVM (testing OA = 0.951), and CART was the worst (testing OA = 0.930). Even though CART performs almost perfectly compared to the other three classifiers in training processing, its testing OA ranks last among the four classifiers; a possible explanation might be that overfitting problems can occur during training processes [31]. The testing OA for the four classifiers seems to result in close values; average differences are less than 0.01%, thus further quantity accuracy needs to be assessed to the extracted urban areas in the following sections.

4.2. Mapping Urban Areas

We extracted four urban area results using CART, k-NN, SVM, and RF from MODIS NDVI and DMSP-OLS NTL and VANUI of 2010 using the new method in this study. Figure 3 shows the extracted urban areas of seven cities (Beijing, Tianjin, Qingdao, Shenyang, Shanghai, Cangzhou, and Guangzhou), as well as FROM-GLC reclassification images (1 km); the DSMP-OLS NTL and VANUI are also displayed. Evident differences can be observed through a simple visual comparison between DMSP-OLS NTL and VANUI. VANUI can greatly decrease the saturation effect of DMSP-OLS NTL and enhance intra-urban variability, especially in large cities such as Shanghai, with most of the areas associated with red color. Therefore, it is appropriate to utilize VANUI as an indicator for extracting urban areas, and it is clear that results of the extracted urban areas can provide much finer details within the urban areas because of the effect of VANUI.

For different machine learning algorithms, urban areas extracted by CART were visually much larger than those extracted by the other three algorithms and the FROM-GLC urban areas. The SVM method tended to identify continuous urban extent in high illumination areas but was less likely to detect subtle pixels around the NTL saturation areas than RF and k-NN. Although the results extracted by using RF and k-NN were similar when compared visually, the results extracted by using k-NN had some noise around the urban areas.

Quantity accuracy assessments were performed on extracted urban area results. OA and Kappa coefficients were calculated on the basis of the extracted urban areas and the FROM-GLC images of 2010. Urban pixel counts of the extracted results and FROM-GLC of each province/municipality were derived on the basis of administrative boundaries. The accuracy assessment results for the four classifiers are listed in Table 4. Figure 4 shows the scatter plots between the extracted urban pixels and FROM-GLC urban pixel counts of each city.

Note that the proportion of the non-urban background pixels can account for the majority and largely increase OA values. Kappa coefficient was considered the primary factor for the comparisons of the four classifiers. Comparing the average OA and Kappa and the scatter plot, the RF had the best coherency with FROM-GLC (OA = 0.964, Kappa = 0.598, $R^2 = 0.972$); k-NN and SVM had similar accuracy and also exhibited good agreement with the FROM-GLC data (k-NN: OA = 0.961, Kappa = 0.574, $R^2 = 0.972$; SVM: OA = 0.967, Kappa = 0.568, $R^2 = 0.962$), whereas CART had the worst results (OA = 0.943, Kappa = 0.525, $R^2 = 0.973$). From Table 4, we can see that the average urban pixels of the 99 cities extracted by using CART is 614.9, much more than the average FROM-GLC urban pixels (374.1), and in Figure 4 the scatter plots also show that the Beta of the fitted regression equation between

CART and FROM-GLC is 1.46, much higher than that of the other three classifiers, indicating that the CART tended to be highest at overestimating among the four classifiers. On average, RF and k-NN also tended to overestimate the urban areas and SVM tended to underestimate, but the RF was the closest to FROM-GLC; this conclusion was consistent with the previous visual comparison.

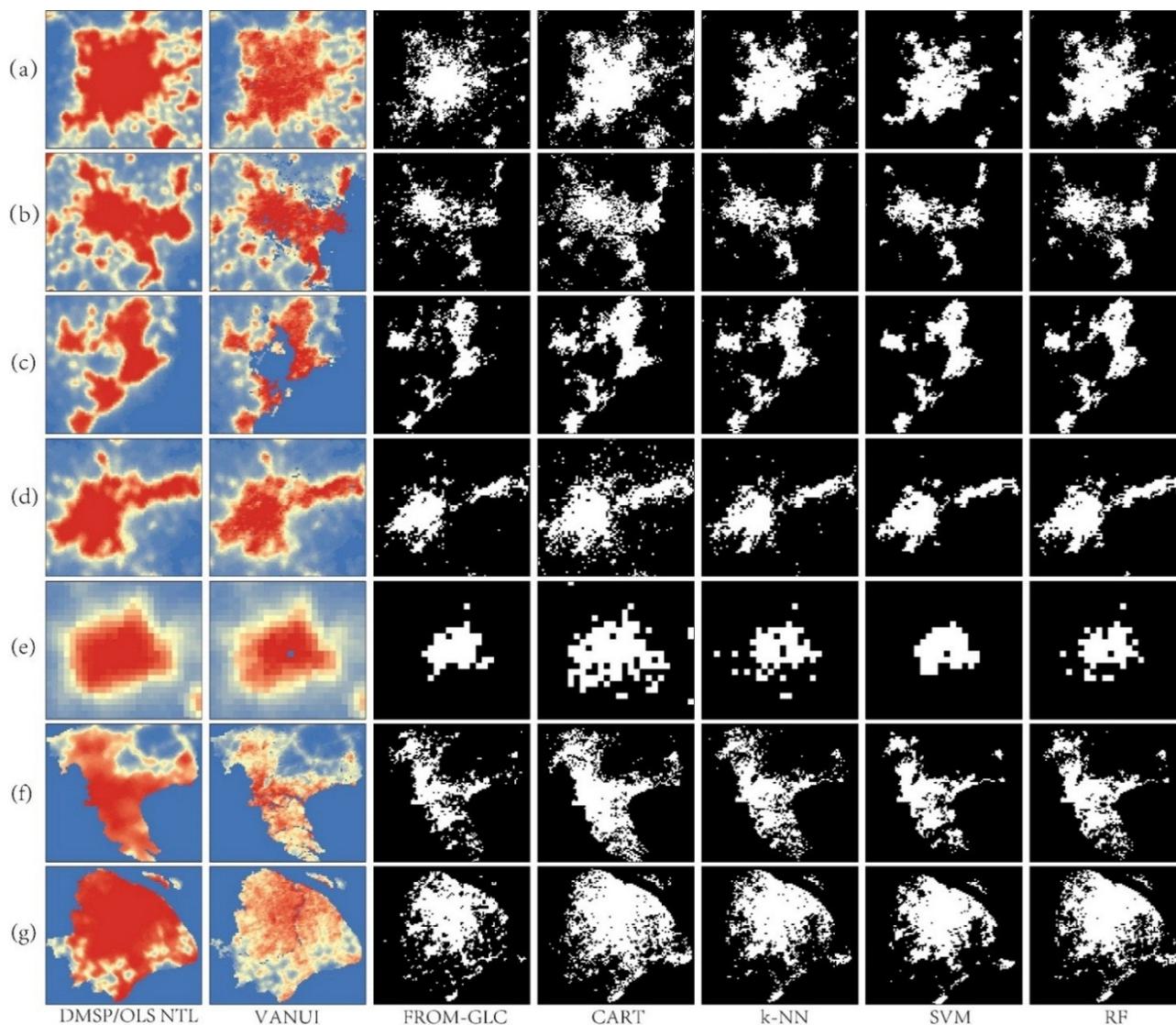


Figure 3. Urban area extraction results: (a) Beijing, (b) Tianjin, (c) Qingdao, (d) Shenyang, (e) Cangzhou, (f) Guangzhou, and (g) Shanghai.

For individual provinces and cities, although SVM had higher Kappa values than the other three classifiers for Beijing, Tianjin, and Shanghai, RF was found to have the highest Kappa values in the other seven provinces on average. For Beijing, Qingdao, Huizhou, and Chaozhou, the latter three classifiers (SVM, k-NN, and RF) all produced high-accuracy predicted results ($Kappa > 0.7$). Good agreement ($Kappa > 0.6$) can be found in almost 60% of cities (59) among the 99 cities in the study area for RF, whereas they were 23%, 42%, and 47% for CART, k-NN, and SVM. Cities with fair agreement ($0.2 < Kappa < 0.4$) for CART, k-NN, SVM, and RF account for 11%, 4%, 11%, and 4% of the total quantity, respectively.

Table 4. Accuracy assessments of the four algorithms for extracting urban areas by province/municipality.

Province/ Municipality	FROM- GLC Urban Pixels	CART			k-NN			SVM			RF		
		OA	Kappa	Urban Pixels									
Beijing	2107	0.943	0.666	2892	0.961	0.727	2159	0.964	0.743	2000	0.96	0.733	2336
Tianjin	1323	0.907	0.54	2130	0.947	0.652	1293	0.956	0.696	1174	0.948	0.666	1407
Shanghai	2137	0.763	0.528	3222	0.836	0.649	2452	0.845	0.66	2257	0.832	0.642	2593
Hebei	296.4	0.964	0.527	518.9	0.978	0.569	258.6	0.981	0.588	194.1	0.979	0.593	242.8
Shandong	244.2	0.963	0.605	371.0	0.972	0.632	276.4	0.976	0.652	216.6	0.975	0.663	265.2
Liaoning	249.0	0.959	0.439	465.7	0.976	0.533	236.3	0.978	0.536	183.7	0.978	0.562	234.7
Jiangsu	302.1	0.941	0.509	513.9	0.959	0.53	312.9	0.967	0.548	215.5	0.962	0.571	313.0
Zhejiang	399.3	0.947	0.541	664.5	0.966	0.588	419.2	0.971	0.523	284.4	0.969	0.604	395.5
Fujian	199.8	0.964	0.517	365.5	0.978	0.6	221.1	0.981	0.555	160.0	0.979	0.611	231.3
Guangdong	408.2	0.903	0.522	643.0	0.929	0.557	484.1	0.939	0.535	408.2	0.933	0.571	459.5
Average	374.1	0.944	0.526	614.9	0.962	0.575	399.4	0.967	0.568	317.0	0.964	0.598	392.7

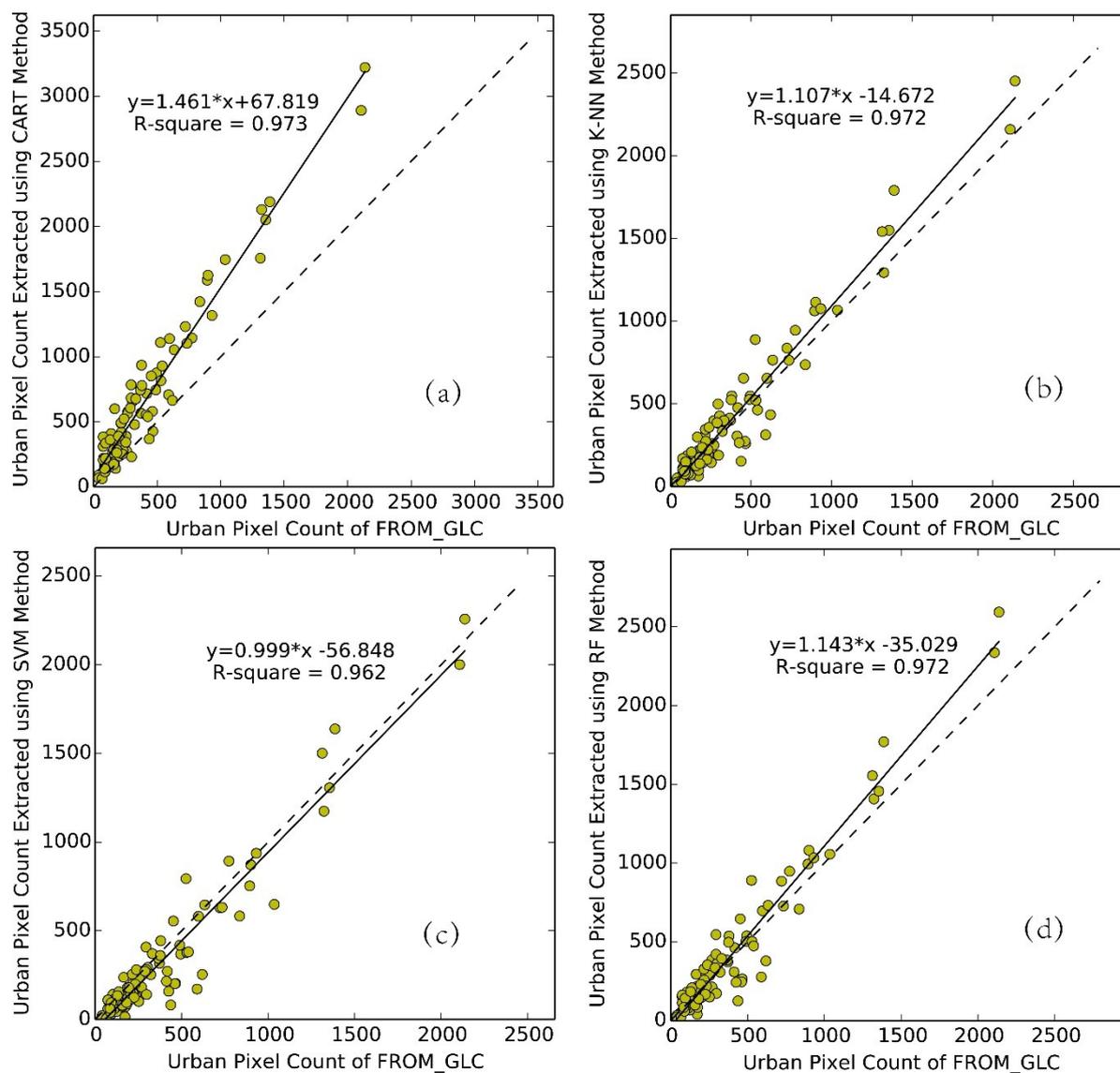


Figure 4. Scatter plot of urban pixel count of each city between predicted results and FROM-GLC for four classifiers: (a) CART, (b) k-NN, (c) SVM, and (d) RF.

Figure 5 shows the spatial distribution of the Kappa coefficient for each city in the study area; distribution of high-agreement and low-agreement regions can be seen clearly. High agreement can be found in most of the cities in the northern part of the study area, including Beijing, Tianjin, Hebei Province, Liaoning Province, and Shandong Province. In Jiangsu Province and Guangdong Province, RF shows higher agreements than the other three; in Zhejiang Province and Fujian Province, RF and k-NN have similar distributions and both are better than SVM and CART. Overall, the high accuracy regions of extracted urban areas show a tendency distributed along the coastline. We postulated a possible explanation that the classification accuracy and NTL DN value are positively correlated, that is to say, the brighter the urban cores, the more likely to be identified as urban pixels. Figure 6 displays the DMSP-OLS NTL image and the mean NTL DN of 99 cities, a spatial distribution type similar to that of Kappa coefficient can be seen from a simple visual comparison.

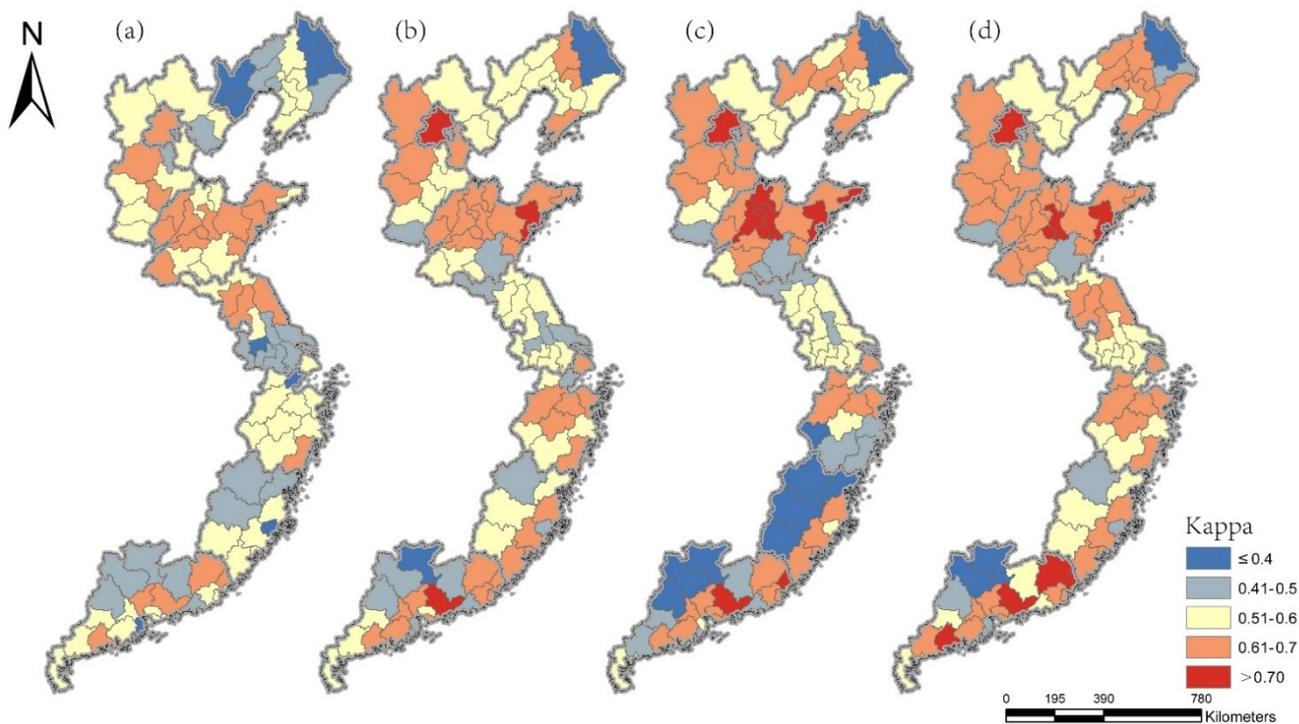


Figure 5. Spatial distribution of Kappa at the city level: (a) CART, (b) k-NN, (c) SVM, and (d) RF.

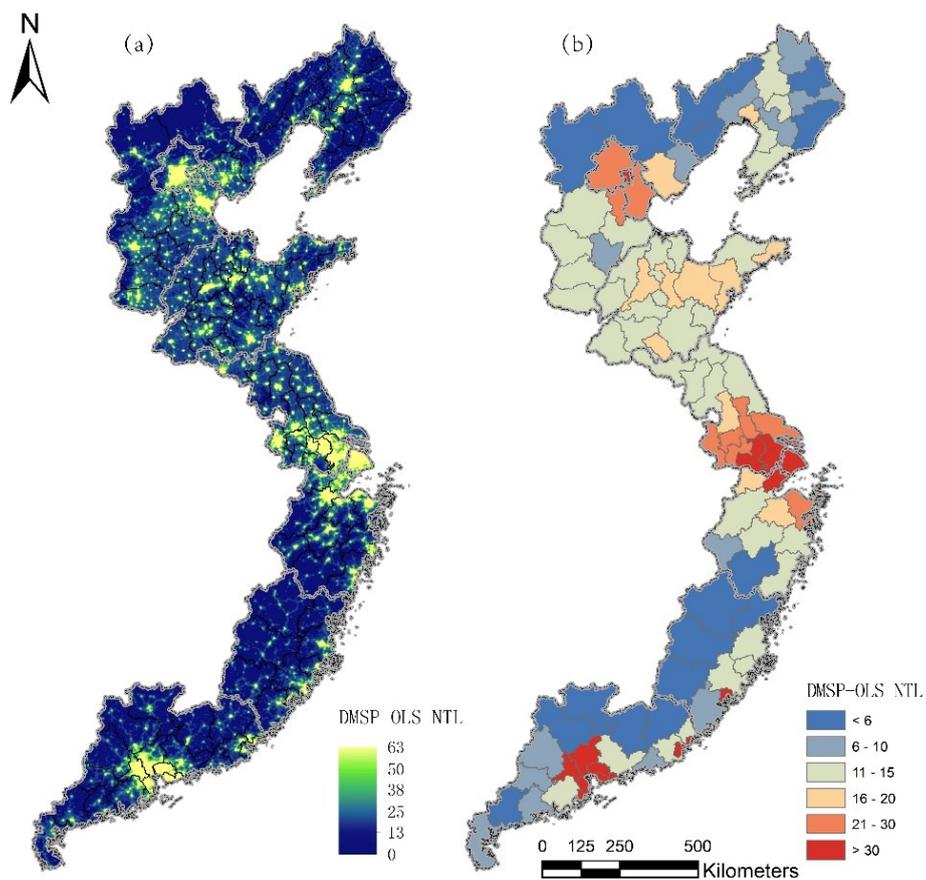


Figure 6. (a) DMSP-OLS NTL image of study area. (b) Mean NTL DN of 99 cities.

We further calculated the standard deviation (STD) of the Kappa coefficient of k-NN, SVM, and RF for each city as

$$STD = \sqrt{\frac{1}{n} \sum_{i=1}^n (k_i - K)^2} \quad (6)$$

where STD is the standard deviation value for each city; $n = 3$ is the number of classifiers; k_i are the Kappa coefficients for k-NN, SVM, and RF; and K is the average Kappa value of the three classifiers.

Figure 7 shows the relationship between Kappa standard deviation and the mean value of pixels with NTL DN greater than 30, for 99 cities. We calculated the mean value of pixels with NTL DN greater than 30 instead of calculating the mean value for all pixels because the pixels having low brightness constitute a major portion of large cities and reduce mean values considerably. We can see that the STD tends to be higher in cities with lower NTL DN values, and brighter cities tend to achieve similar accuracy by using the three classifiers, indicating that the variance of accuracy is positively associated with the NTL DN values. The assumption is valid according to the results of the analyses. The brighter the urban cores, the more likely they are to be identified as urban pixels; urban pixels with lower mean NTL DN values are more likely to be neglected.

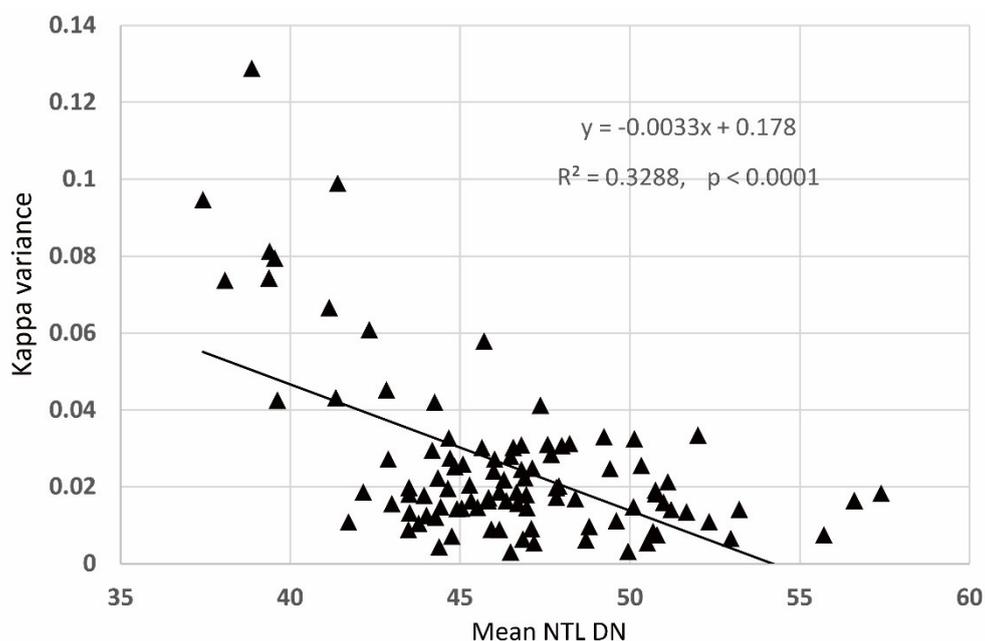


Figure 7. Relationship between the Kappa variance and mean NTL DN of pixels with values greater than 30, for 99 cities.

4.3. Comparison with the Contextual Classification

Table 5 shows the classification accuracies using the contextual classification method. The best performer is k-NN, followed by SVM. Comparing Tables 4 and 5, we can see that all results are improved by the new method proposed in this study. The best performer is k-NN, with OA improvement of 0.016 and Kappa coefficient improvement of 0.084 over the best contextual classification results.

Table 5. Accuracy assessments of the four algorithms for extracting urban areas by province/municipality.

City\Province	FROM-GLC Urban Pixels	CART			k-NN			SVM			RF		
		OA	Kappa	Urban Pixels									
Beijing	2107	0.859	0.450	2686	0.901	0.551	2029	0.863	0.463	2858	0.865	0.463	2636
Tianjin	1323	0.643	0.203	2234	0.779	0.339	2562	0.703	0.253	2493	0.699	0.248	2175
Shanghai	2137	0.859	0.641	1787	0.863	0.652	1794	0.866	0.670	2008	0.861	0.644	1784
Hebei	296.4	0.865	0.266	772.1	0.938	0.453	336.8	0.899	0.337	395.8	0.870	0.270	370.0
Shandong	244.2	0.948	0.539	339.2	0.962	0.612	323.5	0.953	0.568	330.6	0.948	0.539	339.0
Liaoning	249.0	0.905	0.286	435.3	0.958	0.459	286.7	0.941	0.382	349.6	0.907	0.289	319.3
Jiangsu	302.1	0.944	0.520	522.8	0.944	0.522	337.4	0.941	0.518	370.9	0.948	0.541	396.7
Zhejiang	399.3	0.960	0.515	247.1	0.961	0.528	254.1	0.962	0.545	274.1	0.961	0.519	244.9
Fujian	199.8	0.981	0.550	188.0	0.981	0.547	190.9	0.980	0.565	230.8	0.981	0.550	188.1
Guangdong	408.2	0.929	0.481	625.0	0.930	0.482	425.8	0.934	0.516	458.0	0.926	0.481	425.0
Average	374.1	0.928	0.451	509.4	0.948	0.514	428.6	0.939	0.491	505.1	0.929	0.456	449.6

4.4. Sensitivity Analysis

An important step for accurate estimation of urban areas by using machine learning methods is the selection of training samples. A simple threshold strategy was used to select the potential urban and non-urban pixels in this study. We used sensitivity analysis to evaluate the influence of the initial thresholds upon the final classification results by changing each threshold at a time while maintaining the initial value of other thresholds. We assigned a step value of ± 1 and a range of 30 to 55 for the NTL DN value, and a step value of ± 0.01 and a range of 0.3 to 0.5 for the NDVI value. The best performers for each method were used respectively (RF for the per-pixel classification and k-NN for the contextual classification).

Figures 8 and 9 show the Kappa coefficient of each changed value of the initial thresholds for urban area extraction results using the two methods. From Figure 8, we can see that the Kappa coefficient increases with an increase in the NTL DN value from 30 to 40, it decreases when the NTL DN value is greater than 50, but it does not vary when the NTL DN values change from 40 to 50, indicating that classification outputs are not sensitive to initial threshold changes of the NTL DN value in the range of 40 to 50. From Figure 9, we can see that accuracies show more significant fluctuations when the initial NDVI threshold changes, indicating that the classification results are more sensitive to initial NDVI thresholds for both methods. However, when NDVI is in the range of 0.4 to 0.45, a consistent Kappa coefficient can be achieved by using per-pixel classification; meanwhile, accuracies vary when using contextual classification.

The initial thresholds of NTL and NDVI give only the basic conditions to define the potential training samples with high certainties, and the above sensitivity analysis indicates that it is safe to choose the initial thresholds of the NTL DN value in the range 40 to 50, and the NDVI value in the range 0.4 to 0.45 using the new method.

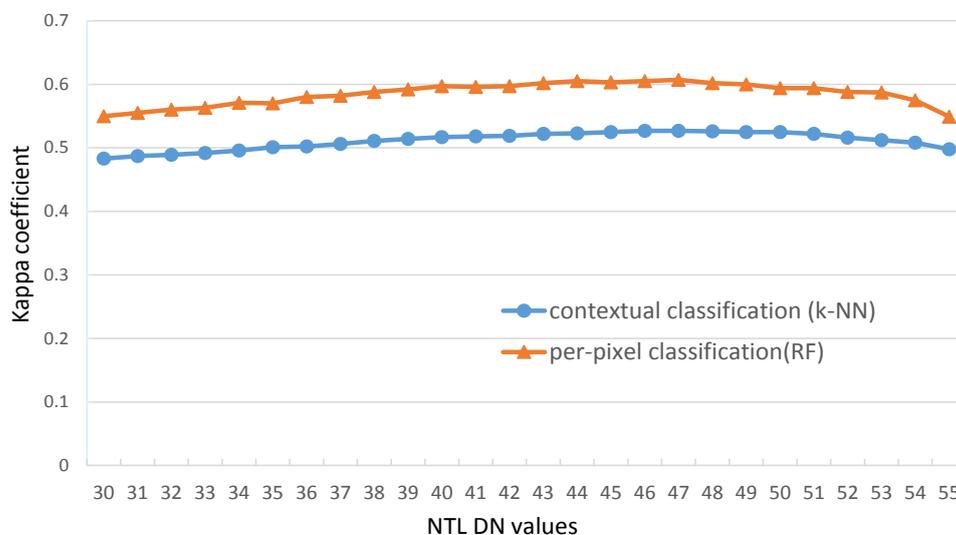


Figure 8. Kappa coefficient for the changed initial threshold of the NTL DN value for the urban extraction results.

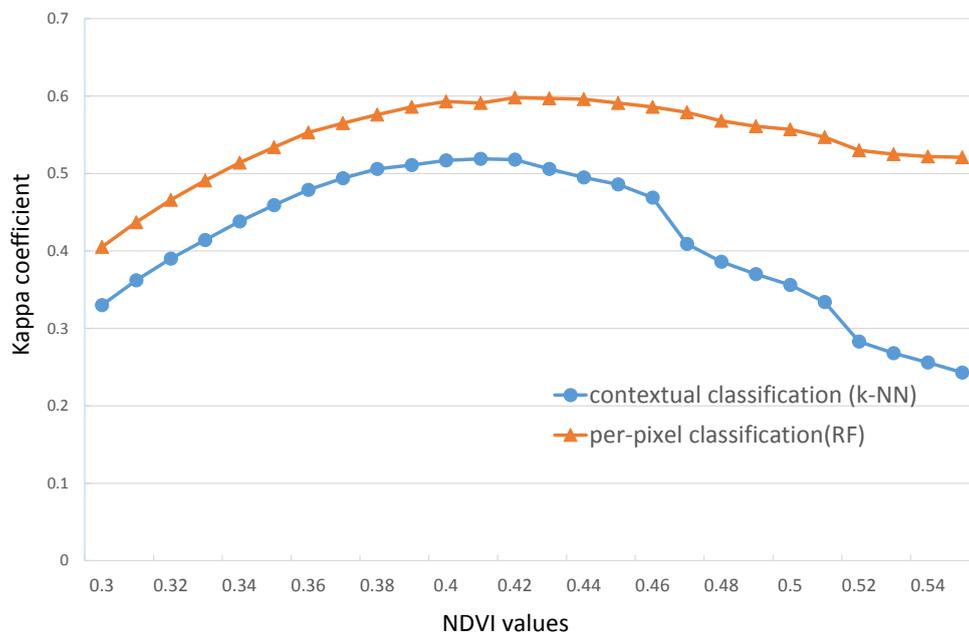


Figure 9. Kappa coefficient for the changed initial threshold of the NDVI value for the urban extraction results.

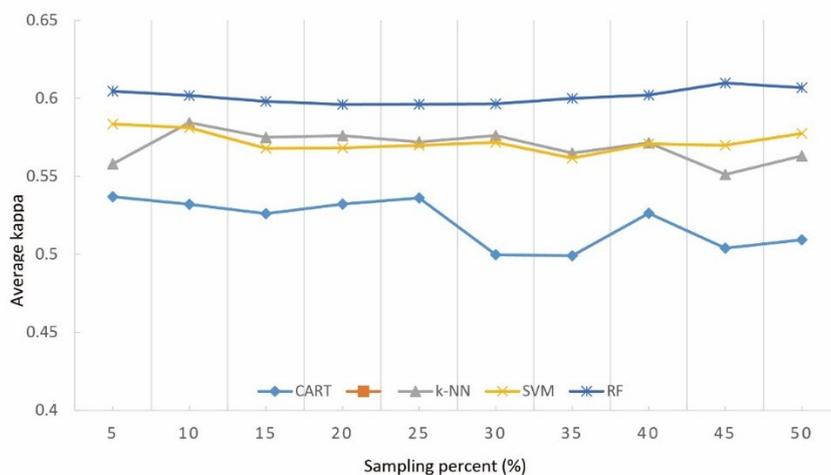


Figure 10. Average kappa coefficient with different sample percentages.

4.5. Impact of Different Training Set Percent

To obtain a comprehensive view of the impact of different training sets upon the classification outputs by using the new method proposed in this paper, we repeated the four categories of classification algorithms with a series of sampling percentages (*i.e.*, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, and 50%), carrying out an accuracy assessment for each sample percentage (Figure 10). The overall decreasing trend can be detected from Figure 10 with the increase in sample percentage, and a very low accuracy can be reached at 45% for CART and k-NN. SVM has been used in some studies (*e.g.*, [7,23]) to extract urban areas from DMSP-OLS NTL images and achieved good results in this study, when the sampling percentage ranges from 10% to 40%. SVM and k-NN produced similar accuracy. For RF, when

sampling percentage increases, the kappa line charts display a more stable trend and better accuracy than the other three methods, indicating a stable and better performance with RF. These analyses indicate that RF, as an ensemble learning algorithm, is stable and exhibits good performance in extracting urban information from coarse resolution data and at a large scale.

5. Discussion

We proposed a per-pixel classification method to extract urban areas from DMSP-OLS NTL and MODIS NDVI data; a simple threshold strategy and probability of selection were used to select the training sets, and the training sets were input into machine learning classifiers to classify the unknown pixels. Comparable accuracy was obtained compared to the existing contextual classification method. Because the contextual classification is a region-growing procedure, the initial urban seeds are defined as the pixels with maximum NTL DN values in each patch having NTL DN greater than 40. NTL saturation may cause bias in the classification, and the urban pixels are growing from the initial urban seeds, we must ensure that initial urban seeds include all the potential urban patches by inspecting the NTL values of many urban patches [7]. In contrast, the method proposed in this study collects random samples on the basis of the probability of selection, and all the pixels are classified at one time; compared to the existing method, it is easier to implement. Further, sensitivity analysis was also addressed in our study: for the two categories of methods, classification results were more sensitive to changes in the initial NDVI value than to those of NTL DN value; with the same conditions, the new method produced results with more stable accuracy and were less sensitive to the changes of initial thresholds. In general, this new approach demonstrates the following two advantages: (1) it automatically obtains a training set with a simple threshold strategy without any other reference data and with less human participation or interaction; (2) it uses a probability of selection to select the training set through a random procedure; final classification results are not sensitive to the initial thresholds within a safe range. Even so, uncertainties of initial thresholds may arise when applying this method to other regions or other years, and sensitivity analysis or selection of optimal thresholds is needed when employing the approach.

Among the four classifiers that were implemented in this study, RF produced classification results with the highest average Kappa, and when the sample set changed, it produced results with stable accuracies. However, choosing the proper algorithm is also significant when using the method, because although RF obtained the best accuracy on average, we can see from Figure 4, that in some regions (e.g., Beijing, Tianjin and Shanghai) SVM performed better. Additionally, in this study we adopted the algorithm parameters on the basis of recommendations by developers or empirical values; this would be another issue that affects the final results. Selecting good parameters while the application conditions vary from one environment to another and one data type to another is still a challenging problem; additional research in this area is required.

Moreover, because the method did not perform well in some cities, and the distribution of the cities with higher accuracies showed a similar tendency as the cities with higher NTL DN values, we postulated an assumption that the classification accuracy and NTL DN values are positively correlated. Preliminary analysis was carried out, and results indicated that the assumption was valid, indicating that uncertainties may be introduced in cities with low NTL illumination. Whereas, one may note that it is just an assumption based on the results in this study, and the analysis is also preliminary; details of the

relationships between DMSP-OLS NTL data and urban morphology and urban extraction accuracy are complex that we still cannot draw conclusions, and it calls for our further inquiry and research.

6. Conclusions

Mapping urban areas or human settlements at regional or global scales is often based on the DMSP-OLS NTL data [23,34–36], and threshold-based algorithms are widely used for extracting urban areas or human settlements [7,13,37]. However, biased estimated (overestimated or underestimated) problems can be a limiting factor when compared across cities with different levels of development. Although local threshold algorithms can be used to meet inter-region variances, the determination of suitable threshold values is empirical and difficult [14,15].

In this paper, we presented a machine-learning-based approach to derive urban areas from DMSP-OLS NTL and MODIS NDVI data; four classification algorithms were employed for comparison, and a region-by-region strategy was utilized. VANUI, which is an urban index to reduce the effects of NTL saturation and increase variation of the NTL signal, was implemented as an independent variable. Extracted urban areas were validated against the FROM-GLC image. Results showed that on average, RF achieved the best extraction results among the four classifiers. Meanwhile, CART produced highly overestimated results, compared to the three other classifiers. Although k-NN and SVM tended to produce similar accuracy, less-bright areas around the urban cores seemed to be ignored when using SVM, which resulted in the underestimation of urban areas. However, quantity assessment results showed that the results produced by SVM exhibited better agreement (Kappa coefficient) in large cities such as Beijing, Tianjin, and Shanghai.

The classification results were also compared with an existing contextual classification method, and sensitivity analysis was carried out on the two methods by changing the initial thresholds. According to the results, the new method achieved higher Kappa coefficients and more robust results, and RF as an ensemble learning algorithm produced a more stable accuracy. As a result, this approach is proved to be successful for mapping urban areas through combined use of MODIS NDVI and DMSP-OLS NTL images. In addition, DMSP-OLS NTL and MODIS NDVI can be freely downloaded and have a global coverage and time series; thus the approach proposed in this paper can be expanded to other regions and other years.

However, per-pixel-based urban area mapping is not sufficient, and areas and spatial information can be lost because of mixed pixels [7]. More research is needed for sub-pixel-based urban or human settlement mapping at coarse spatial resolution at regional and global scales. In addition, new nighttime light image instruments, such as the Visible Infrared Imager Radiometer Suite (VIIRS), with finer spatial resolution and higher quantization levels, may provide a more detailed data source to improve the accuracy of urban extent mapping [7,15,38].

Acknowledgments

This research was supported by the China Knowledge Center for Engineering Sciences and Technology (No. CKCEST-2015-1-4) and the National Special Program on Basic Science and Technology Research of China (No. 2013FY110900). The authors are indebted to the National Aeronautics and Space Administration (<http://modis.gsfc.nasa.gov/>) and the National Centers for

Environmental Information (<http://ngdc.noaa.gov/eog>) for providing the MODIS and DMSP-OLS data that were used in this study, and to the Center for Earth System Science Tsinghua University (<http://www.cess.tsinghua.edu.cn/>) for providing land cover data. We also thank the Data Sharing Infrastructure of Earth System Science for providing the administrative boundaries data (<http://www.geodata.cn/>). In addition, we would like to thank the three anonymous reviewers for their helpful comments and suggestions in enhancing this manuscript.

Author Contributions

Wenlong Jing drafted the manuscript and was responsible for the research design, experiment, and analysis. Yaping Yang reviewed the manuscript and was responsible for the research design and analysis. Xiafang Yue and Xiaodan Zhao supported the data preparation and the interpretation of the results. All of the authors contributed to editing and reviewing the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Seto, K.C.; Fragkias, M.; Guneralp, B.; Reilly, M.K. A meta-analysis of global urban land expansion. *PLoS ONE* **2011**, *6*, e23777.
2. Grimm, N.B.; Faeth, S.H.; Golubiewski, N.E.; Redman, C.L.; Wu, J.; Bai, X.; Briggs, J.M. Global change and the ecology of cities. *Science* **2008**, *319*, 756–760.
3. Foley, J.A.; DeFries, R.; Asner, G.P.; Barford, C.; Bonan, G.; Carpenter, S.R.; Chapin, F.S.; Coe, M.T.; Daily, G.C.; Gibbs, H.K.; *et al.* Global consequences of land use. *Science* **2005**, *309*, 570–574.
4. Hahs, A.K.; McDonnell, M.J.; McCarthy, M.A.; Vesk, P.A.; Corlett, R.T.; Norton, B.A.; Clemants, S.E.; Duncan, R.P.; Thompson, K.; Schwartz, M.W.; Williams, N.S.G. A global synthesis of plant extinction rates in urban areas. *Ecol. Lett.* **2009**, *12*, 1165–1173.
5. Seto, K.C.; Kaufmann, R.K.; Woodcock, C.E. Landsat reveals China's farmland reserves, but they're vanishing fast. *Nature* **2000**, *406*, 121.
6. Rotem-Mindali, O.; Michael, Y.; Helman, D.; Lensky, I.M. The role of local land-use on the urban heat island effect of Tel Aviv as assessed from satellite remote sensing. *Appl. Geogr.* **2015**, *56*, 145–153.
7. Cao, X.; Chen, J.; Imura, H.; Higashi, O. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. *Remote Sens. Environ.* **2009**, *113*, 2205–2209.
8. Li, C.; Li, C.; Wang, J.; Wang, L.; Hu, L.; Gong, P. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. *Remote Sens.* **2014**, *6*, 964–983.
9. Lo, C.P.; Choi, J. A hybrid approach to urban land use/cover mapping using Landsat 7 enhanced thematic mapper plus (ETM+) images. *Int. J. Remote Sens.* **2004**, *25*, 2687–2700.

10. Lu, D.; Hetrick, S.; Moran, E. Land cover classification in a complex urban-rural landscape with QuickBird imagery. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 1159–1168.
11. Myint, S.W.; Gober, P.; Brazel, A. Per-pixel vs. Object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **2011**, *115*, 1145–1161.
12. Weng, Q. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sens Environ.* **2012**, *117*, 34–49.
13. Zhang, Q.; Seto, K.C. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sens. Environ.* **2011**, *115*, 2320–2329.
14. Zhang, Q.; Schaaf, C.; Seto, K.C. The vegetation adjusted NTL urban index: A new approach to reduce saturation and increase variation in nighttime luminosity. *Remote Sens. Environ.* **2013**, *129*, 32–41.
15. Small, C.; Pozzi, F.; Elvidge, C. Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sens. Environ.* **2005**, *96*, 277–291.
16. Imhoff, M.L.; Lawrence, W.T.; Stutzer, D.C.; Elvidge, C.D. A technique for using composite DMSP/OLS “City Lights” satellite data to accurately map urban areas. *Remote Sens. Environ.* **1997**, *61*, 361–370.
17. Henderson, M.; Yeh, E.T.; Gong, P.; Elvidge, C.; Baugh, K. Validation of urban boundaries derived from global night-time satellite imagery. *Int. J. Remote Sens.* **2003**, *16*, 595–609.
18. Manuel Fernandez-Delgado, E.C.; Barro, S. Do we need hundreds of classifiers to solve real world classification problems. *J. Mach. Learn. Res.* **2014**, *49*, 3133–3181.
19. Bureau, C.S. *China Statistical Year Book 2010*; China Statistics Press: Beijing, China, 2010.
20. Yu, L.; Wang, J.; Li, X.; Li, C.; Zhao, Y.; Gong, P. A multi-resolution global land cover dataset through multisource data aggregation. *Sci. China Earth Sci.* **2014**, *57*, 2317–2329.
21. Gong, P.; Wang, J.; Yu, L.; Zhao, Y.; Zhao, Y.; Liang, L.; Niu, Z.; Huang, X.; Fu, H.; Liu, S; *et al.* Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654.
22. Yu, L.; Wang, J.; Gong, P. Improving 30m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: A segmentation-based approach. *Int. J. Remote Sens.* **2013**, *34*, 5851–5867.
23. Pandey, B.; Joshi, P.K.; Seto, K.C. Monitoring urbanization dynamics in India using DMSP/OLS night time lights and SPOT-VGT data. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 49–61.
24. Shao, Z.; Liu C. The integrated use of DMSP-OLS nighttime light and MODIS data for monitoring large-scale impervious surface dynamics: A case study in the Yangtze River Delta. *Remote Sens.* **2014**, *6*, 9359–9378.
25. Rokach, L.; Maimon, O. *Data Mining with Decision Trees: Theory and Applications*; World Scientific Pub Co Inc: Singapore, 2008.
26. Breiman, L.; Freidman, J.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees (Wadsworth Statistics/Probability)*; Chapman and Hall/CRC: Boca Raton, FL, USA, 1984.
27. Hand, D.; Mannila, H.; Smyth, P. *Principles of Data Mining*; The MIT Press: Cambridge, MA, USA, 2001.
28. Altman, N.S. An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* **1992**, *46*, 175–185.

29. Harrington, P. *Machine Learning in Action*; Manning Publications: Shelter Island, NY, USA, 2012.
30. Brian, S.; Everitt, S.L.; Leese, M.; Stahl, D. *Miscellaneous Clustering Methods, in Cluster Analysis*, 5 ed.; John Wiley & Sons, Ltd: Chichester, UK, 2011.
31. Kotsiantis, S.B. Supervised machine learning—A review of classification techniques. *Informatica* **2007**, *31*, 249–268.
32. Fabian Pedregosa, G.V.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, V.; Blondel, M.; Peter Prettenhofer, R.W.; Dubourg, V.; Vanderplas, J.; Passos, A.; *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
33. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
34. Fan, J.; Ma, T.; Zhou, C.; Zhou, Y.; Xu, T. Comparative estimation of urban development in China's cities using socioeconomic and DMSP/OLS night light data. *Remote Sens.* **2014**, *6*, 7840–7856.
35. Lu, D.; Tian, H.; Zhou, G.; Ge, H. Regional mapping of human settlements in southeastern China with multisensor remotely sensed data. *Remote Sens. Environ.* **2008**, *112*, 3668–3679.
36. Ma, T.; Zhou, C.; Pei, T.; Haynie, S.; Fan, J. Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: A comparative case study from China's cities. *Remote Sens. Environ.* **2012**, *124*, 99–107.
37. Small, C.; Elvidge, C.D. Mapping decadal change in anthropogenic night light. *Procedia Environ. Sci.* **2011**, *7*, 353–358.
38. Li, X.; Xu, H.; Chen, X.; Li, C. Potential of NPP-VIIRS nighttime light imagery for modeling the regional economy of China. *Remote Sens.* **2013**, *5*, 3057–3081.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).