Geocarto International

ISSN: 1010-6049 (Print) 1752-0762 (Online) Journal homepage: https://www.tandfonline.com/loi/tgei20

Extraction of urban built-up areas from nighttime lights using artificial neural network

Tingting Xu, Giovanni Coco & Jay Gao

To cite this article: Tingting Xu, Giovanni Coco & Jay Gao (2019): Extraction of urban built-up areas from nighttime lights using artificial neural network, Geocarto International, DOI: 10.1080/10106049.2018.1559887

To link to this article: https://doi.org/10.1080/10106049.2018.1559887

Accepted author version posted online: 26 Dec 2018.
Published online: 21 Mar 2019.

GEOCARTO INTERNATIONAL
https://doi.org/10.1080/10106049.2018.1559887

Taylor & Francis
Taylor & Francis Group

Extraction of urban built-up areas from nighttime lights using artificial neural network

Tingting Xu, Giovanni Coco and Jay Gao

School of Environment, The University of Auckland, Auckland, New Zealand

ABSTRACT
The spatial distribution of urban areas at the national and regional scales is critical for urban planners and governments to design sustainable and environment-friendly future development plans. The nighttime lights (NTL) data provide an effective way to monitor the urban at different scales however is usually achieved by using empirical threshold-based algorithms. This study proposed a novel Artificial Neural Network (ANN) approach, using moderate resolution imageries as NTL, MODIS NDVI and land surface temperature data, to map urban areas. Both random and maximum dissimilarity distance algorithm sampling methods were considered and compared. The validation of the urban areas extracted from MDA-based ANN against the 2011 US national land cover data showed a reasonable quality (overall accuracy = 97.84; Kappa = 0.74) and achieved more accurate result than the threshold method. This study demonstrates that ANN can provide an effective, rapid, and accurate alternative in extracting urban built-up areas from NTL.

ARTICLE HISTORY
Received 4 September 2018
Accepted 29 November 2018

KEYWORDS
Urban area; nighttime light; artificial neural network; maximum dissimilarity distance algorithm

1. Introduction

As a result of economic development and population growth, urbanization has become an unstoppable phenomenon happening on every corner of the Earth (Angel et al. 2011). However, the rapid growth of urban areas has led to a variety of complex problems, such as traffic congestion, air pollution, deforestation, farmland decrease, and massive and chaotic urban settlements with poor infrastructure (Tole 2008; Park et al. 2011; Lagarias 2012). Extracting urban information from various types of data is considered to be one of the essential tasks in managing urbanization. Urban information also reveals the mechanisms and dynamics of irreversible development (Lagarias 2012; Zhang et al. 2015; Aburas et al. 2016). Therefore, explicitly understanding the spatial distribution of urban patterns is critical for urban planners and governments to design a sustainable future. Within this context, monitoring and detection of urban areas are a prerequisite to develop urban plans (Zhao and Murayama 2011; Dahiya 2016).

Remote sensing data at medium to high spatial resolution (Landsat, SPOT, AVHRR, Quick Bird, IKNOS, WorldView…) are commonly used for urban monitoring and detection (Ridd 1995; Herold et al. 2003; Wang et al. 2006; Ma and Xu 2010; Alsharif et al.

CONTACT Tingting Xu txu648@aucklanduni.ac.nz

2015). Extraction of urban built-up areas from such satellite images is essentially a classification process. However, this type of data usually covers only small areas and is ill-suited for large-scale (national and regional scale) land cover mapping. The requirement of a large number of personnel interpreting the data makes the analysis time-consuming and the results subjective. The interpretation task is especially burdensome for a large urban area analyzed over multiple years. To overcome these limitations, the Defence Meteorological Satellite Program's Operational Line-scan System (DMSP – OLS) stable nighttime light data (NTL), a new free data source from a light-detecting sensor, has been developed. The data have become a cost- and time-efficient resource to monitor urban areas and obtain urban information at large spatial and temporal scales (Elvidge et al. 1997, 2012; Henderson et al. 2003; Milesi et al. 2003; Sutton 2003; Gallo et al. 2004; Lu et al. 2008; Small and Elvidge 2011; Zhang and Seto 2011). Imhoff et al. (1997) first introduced the usage of composite NTL data to map urban areas for the conterminous USA. In the same year, Elvidge et al. (1997) mapped human settlements using the same data. Henderson et al. (2003) validated the urban boundaries derived from NTL data at the global scale. Studies using the NTL data bloomed after 2010. For example, Liu et al. (2012), Li et al. (2016), and Xie and Weng (2017) all used enhanced NTL data to extract the dynamics of urban expansion at the national scale.

Nevertheless, the use of NTL data to map urban areas faces technical problems. For example, light saturation and blooming cause overestimation of urban areas while the weak reflectance of nighttime light in small towns causes underestimation (Elvidge et al. 1997; Henderson et al. 2003; Small et al. 2005). To solve these issues and because of its simplicity, the threshold method has been commonly used to delineate urban areas from NTL data (Imhoff et al. 1997; Lawrence et al. 2002; Henderson et al. 2003; Zhang et al. 2013; Ma et al. 2015; Xie and Weng 2016; Zou et al. 2017). With user-produced optimal threshold values, the NTL data can be reclassified as urban and non-urban areas. This method is easy to understand and can be quickly applied. However, its obvious drawback is always the empirically and subjectively determined threshold. Moreover, only one global threshold used for the entire image would decrease the mapping accuracy due to the unbalanced economic and social developments of different areas. This problem can be minimized using the local optimal threshold approach which requires the whole study area to be split into smaller sub-areas (adding another subjective step). Also with this approach, there are the problems of misclassified data and ambiguous connection areas (Cao et al. 2009; Zhou et al. 2014; Xie and Weng 2016). To overcome these limits, machine learning (ML) has been applied to NTL data to extract urban information. Cao et al. (2009) introduced a support vector machine (SVM)-based semi-automatic method to extract urban areas from DMSP-OLS data, and the results showed that the SVM-based algorithm could not only achieve comparable results to the local-optimized threshold method, but also avoid the tedious trial-and-error procedure in setting the threshold. After comparing four ML methods (CART, k-NN, SVM, and random forest (RF)) in mapping urban areas through the integration of NTL and MODIS-NDVI, Jing et al. (2015) found that all of the four provided robust results. Dou et al. (2017) evaluated the performance of the SVM algorithm and local optimal threshold method in extracting urban land using NTL and normalized difference vegetation index (NDVI) data and revealed that SVM was more efficient and accurate. All these studies demonstrated the suitability and potential of ML in extracting urban information from NTL data. However, ANN, a powerful and well-known ML method that has been widely used to map land use, classify land covers, map soil from daytime satellite images (Wang 1994; Zhu 2000; Qi and Zhu 2003), has not been used to map urban areas from NTL data. Moreover, in

most ANN studies, randomness is the most frequent approach to select training samples, the first and probably most important step in developing an ANN. Being a data-driven technique, the classification results of the ANN are heavily determined by the quality of training samples. However, while the "random" method of sample selection inherently avoids the influence of spatial autocorrelation, it still has the possibility to result in imbalanced sampling (Sukhatme and Avadhani 1965; Tong and Liu 2005). To improve the process, the maximum dissimilarity distance cluster sampling algorithm (MDA) is used to select representative training sample sets to train the ANN (Camus et al. 2011). This method follows the Tobler law of geography (Tobler 1970) that features with similar attributes should be closer to each other, and training samples are selected based on different attribute values. The MDA algorithm has been used, for example, to select a subset of 500 deep-water states to provide samples for a dredging strategy based an energetic approach (Reyes-Merlo et al. 2017), and to determine the sample points to estimate the longitudinal dispersion coefficient in straight natural rivers with a small dataset (Wang and Huai 2016). These and other studies (e.g., Rueda et al. 2016; Passarella et al. 2018) demonstrate the utility of MDA as an advanced algorithm for sampling. However, its usage combined with ANN for a geographical study such as urban growth detection has never been tested.

In this study, we try to address the following objectives: 1) map urban areas of USA applying ANN to NTL data, 2) explore the impacts of two sample selection methods on ANN-based urban area mapping; and 3) evaluate the accuracy of urban mapping with ANN and other methods and demonstrate its utility for urban mapping.

2. Materials and methods

2.1. Study area and data

The contiguous United States has been chosen as the study area because of its vast territory and its tremendous urban expansion over the last two decades, especially along the east and west coasts. The urban area is nearly 250,478 km^2 which is less than 3% of the total land area. Despite being a relatively small percentage, the morphology of urban areas is quite diverse. The geological character of the States is also diverse since most land-covers are mountains, deserts, vegetation, and the Great Plains.

DMSP-OLS Stable Nighttime Lights Annual Time Series (version 4) were sourced from the Earth Observation Group (EOG), NOAA (https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html). The files are cloud-free composites created from all the available archived DMSP-OLS smooth resolution data for the calendar years of 1992 to 2013. The Digital Number (DN) of every pixel for the annual composite data was the average of visible-band DN value of light sources such as cities, towns and other lighting sites, and has been stretched to the unique range 1–63 (0 in the data is the background noise). While the time-series of annual cloud-free composites were produced using the same algorithms and a stringent data selection criterion, the DN values are not strictly comparable from year to year, which prompts the necessity of intercalibration prior to direct comparison of the DN values across the whole time series (Elvidge et al. 1997).

Three MODIS products, land water mask (MOD44w), monthly NDVI (MOD13A3), and eight-day land surface temperature (LST) (MOD11A2), were obtained from the NASA Land Process Distributed Active Archive Center (LP DAAC). All data were resampled at the 1-km spatial resolution and re-projected to the Albers Conical Equal Area projection to be consistent with the NTL and land cover data. Both annual NDVI

Table 1. Coefficients of the second-order regression models for NSL data (2001).

Coefficient	*a*	*b*	*c*	R^2
F12-1999	0	1	0	1
F14-2001	0.0038	0.6334	0.204	0.9500
F15-2001	0.0025	0.7127	0.2717	0.9547
F18-2011	0.002	0.6638	0.0108	0.9092

and LST of 2001 and 2011 were reproduced using the maximum value composite method (MCV) to reduce cloud-contaminated pixels while capturing the changes in spatial scale (Lu et al. 2008; Mildrexler et al., 2009).

The National Land Cover Database (NLCD) has been created by the Multi-Resolution Land Characteristics Consortium across 48 states plus Washington D.C. from 2001 to 2011 (Homer et al. 2007, 2015). NLCD 2001 is based primarily on a decision-tree classification of circa 2001 Landsat satellite data and in this study, it has been aggregated from 30 m to 1 km to provide training samples for the ANN (with impervious area >20% considered as metropolitan urban areas). NLCD 2011 is used as the reference data to validate the ANN after resampling to 1 km spatial resolution.

2.2. Methods

2.2.1. Inter-annual calibration of NTL data

To reduce variation and differences among sensors and improve the comparability, V4 DMSP-OLS NTL data need to be inter-calibrated. The inter-calibration method follows Elvidge et al. (2009) who compared different NTL with reference data to calculate any existing relation. Swain County in North Carolina was selected as the benchmark due to its relatively stable urban expansion process (Li et al. 2016). The NTL data from satellite F2 in 1999 were chosen as the reference dataset because of the highest DN value accumulation. A second-order regression model has applied to inter-calibrate satellite NTL datasets of 2001 and 2011:

$$DN_{calibration} = a \bullet DN^2 + b \bullet DN + c \tag{1}$$

where $DN_{calibration}$ is the inter-calibrated DN value, DN is the original DN value, and a, b, and c are calibration coefficients. The results of inter-calibration are shown in Table 1. Of the two 2001 datasets, F15-2001 was used because of its higher R^2.

2.2.2 Artificial neural network

ANN is a widely used ML model with self-adapting, self-organizing, and self-learning attributes. It enjoys the strengths of independence of particular functions and makes no assumptions regarding the data distribution (Hagan 1996).

We use a feed-forward network (Multi-Layer Perceptron, MLP) configuration which is the most frequently used ANN architecture in previous geo-related papers (Wang 1994; Li and Yeh, 2002; Pijanowski et al. 2002, 2010) and which comprises one input layer, one hidden layer, and one output layer. The structure of the feed-forward MLP involves 3 input nodes, 3 hidden nodes, and 2 output nodes (Figure 1). The three variables in the input layer are reserved for NTL, NDVI, and LST, respectively; the two nodes in the output layer are reserved for the two output classes of urban and non-urban areas. The expected output values have two possibilities of (1, 0) and (0, 0). The former indicates that the cell meets the expectations of urban built-up areas, and the latter signifies non-urban cells. The choice of the number of hidden layers is based on a sensitivity analysis.

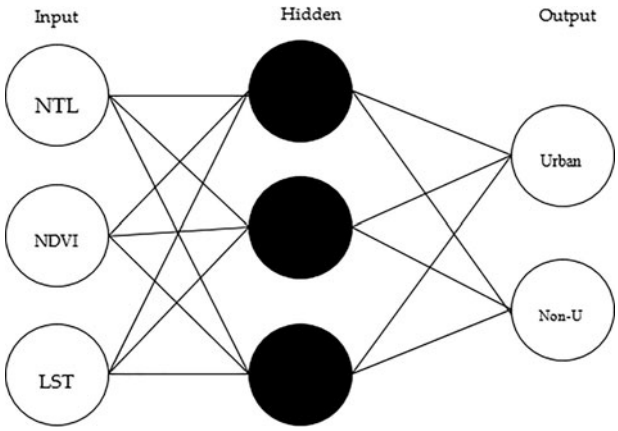

Figure 1. Architecture of the ANN adopted in this study: three nodes are present in both the input and the hidden layers, and two nodes form the output layer.

In general, too few hidden layers cannot construct a robust function that reflects the complex mapping relationships thus causing a more substantial prediction error. Oppositely, too many hidden layers increase the training time and are likely to lead to overfitting problem. After performing several tests and considering both the performance and simplicity of ANN, the number of neurons in the hidden layer was set to 3.

The ANN model must be trained and tested with samples from the study area by applying a supervised learning procedure and a back-propagation (BP) algorithm to control the error-correction process. Every training is implemented iteratively with a target mean square error (MSE). After an initial weight is assigned to each input variable, the ANN starts to "learn" by adjusting weights continuously between neurons in response to the MSE between the modeled and the observed values. The training stops when the MSE threshold is reached and becomes stable. At that moment, the optimized weight for every variable impacting on the decision of urban or non-urban is finalized and the ANN is ready to predict based on the input variables. The output of the ANN, the urban versus nonurban probability, was evaluated for each cell. If the probability value is higher than 0.5 then the cell is considered to be an urban area.

2.2.3 Sample selection

In total, 165,537 sample points representing urban (50,476) and non-urban (115,061) over the contiguous U.S. were selected for ANN modeling, 60% of which were used for training the ANN, 20% for validation and 20% for testing. The training samples were selected using either the random or the MDA algorithm, separately. To compare the impact of different sampling methods, both random and MDA were considered using the same network configuration (3-3-2). The most significant difference between the two methods is that the former uses a stochastic process to select the training database, hence avoiding the influence of spatial autocorrelation. The MDA obeys the Tobler's first law of geography and attempts to generate a more balanced and distinctive training dataset. This is achieved by clustering a representative training subset of size M from a sample database of size N (Camus et al. 2011). To initialize the MDA routine, a single sample is selected as the first centroid and additional points are selected iteratively. The single data point with the maximum distance between itself and the nearest centroid is selected as the next centroid. The MDA routine continues until the user-defined number of centroids is reached.

Table 2. Confusion matrix of Cohen's kappa.

	ANN modelled result		
Ground truth	Urban	Non-urban	Omission error
Urban	a	b	b / ($a+b$)
Non-urban	c	d	c / ($c+d$)
Commission error	c / ($a+c$)	b / ($b+d$)	
Overall accuracy = ($a+d$) / ($a+b+c+d$)			
$P_e = [(a+b)\times(a+c)]/[(a+b+c+d)^2] + [(c+d)\times(b+d)]/[(a+b+c+d)^2]$			
Kappa = (Overall accuracy − P_e)/(1 − P_e)			

a: number of urban cells in both the modelled results and the ground truth;
b: number of non-urban cells in the modelled results that are urban cells in the ground truth;
c: number of urban cells in the modelled results that are non-urban cells in the ground truth;
d: number of non-urban cells in both modelled results and the ground truth.
P_e: The probability of random agreement.

For example, if the subset is formed by R ($R \leq M$) vectors, the dissimilarity between the vector i of the data sample N–R and the j vectors belonging to the R subset is calculated as:

$$d_{ij} = \ ||x_i - v_j||; \ \ i = 1, \ldots, N-R; \ \ j = 1, \ldots, R \tag{2}$$

where x_i is an n-dimensional vector within the sample database and v_j is a vector of the subset. Subsequently, the dissimilarity $d_{i,\ subset}$ between the vector i and the subset R, is calculated as:

$$d_{i,subset} = \min\{||x_i - v_j||\}; \ \ i = 1, \ldots, N-R; \ \ j = 1, \ldots, R \tag{3}$$

Once the N–R dissimilarities are calculated, the next selected data is the one with the largest value of $d_{i,subset}$.

In this study, 99,322 centroids were selected as the training data. The remaining 66,215 data points were split randomly to be used for validation (33,108) and testing (33,107). In addition, to test the spatial autocorrelation of samples from both methods, Moran's I was calculated. This index calculates the number of samples within different counties and their neighborhoods to describe the spatial distribution trend of sample points (Goodchild 1986). The higher the Moran's I, the more clustered the samples are.

2.2.4. Validation

Model prediction accuracy was assessed calculating the training, validation, and testing accuracy during the ANN training process, respectively. Moreover, we also focus on the size of the urban class so that the Cohen's kappa is used to validate the model results. The validation accuracy is an indicator to assess model quality, while the testing accuracy assesses the robustness and generality of the model. The Cohen's kappa is calculated from a confusion matrix between the modeled result and the reference data. It also provides the commission and omission errors (overestimation and underestimation), the overall accuracy, and the consistency kappa index. Table 2 illustrates how the Cohen's kappa assesses results.

2.2.5. Workflow

As mentioned before, the NTL 2001 data was first inter-calibrated using the 1999 NTL data and integrated with the NDVI and LST data to develop the three input variables for both the random-ANN and the MDA-ANN. Four networks (two for random and two for MDA) with the best validation and testing values were recorded. The training samples of

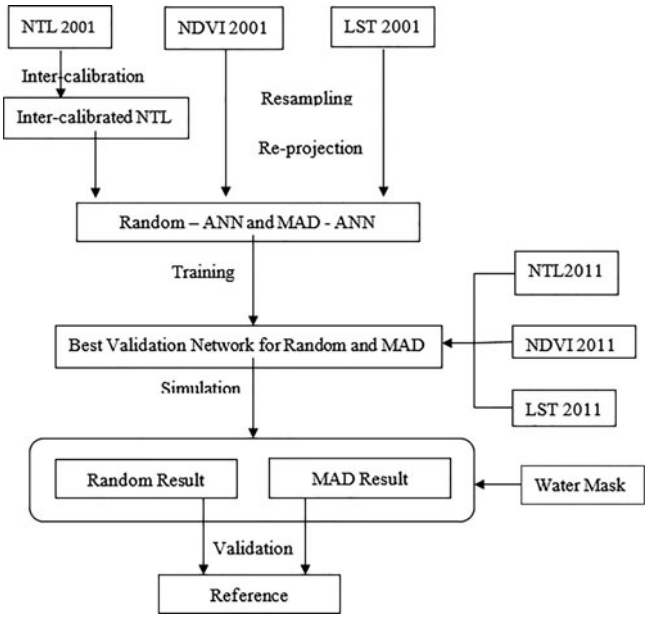

Figure 2. Flowchart of steps necessary to extract urban information from NTL using ANN.

the MDA is always the same, however, the initial weights, validation data, and testing data are different for every MDA-ANN. In order to minimize the bias, sufficient runs for both random-ANN and MDA-ANN are essential. The inter-calibrated NTL 2011, NDVI 2011, and LST 2011 were then applied to the four ANNs built with different training sample selection methods to map the urban areas of the USA in 2011. In addition, the two more outputs were calculated by averaging all the 1000 runs' output values comprised between the mean and plus/minus three standard deviation. Then the water mask was excluded from these six results, and they were compared with NLCD data for validation. The model with a higher Kappa was then used as a potential tool to map urban areas from NTL data at a large scale. Numerical analysis and ANN modeling of this study were undertaken in Matlab@ (Figure 2).

3. Results

3.1. Model accuracy

After performing 1000 runs with both the random-ANN and the MDA-ANN, the accuracy of each approach was evaluated (Table 3) by assessing the performance on the test data as well as the average values of accuracy indicators.

Correlation coefficient (R) and slope value (M) describe the linear regression between the output of ANN and target results. Correlation coefficient and slope values close to 1 indicate a strong similarity between model predictions and ground truth. Table 3 shows that the ANN built using a random sample selection has only a higher training accuracy than the ANN built using MDA, while both validation and testing accuracy is higher when using the MDA. This indicates that even if the training process of MDA is less accurate than in the random case, the predictor is more accurate and general. In addition, the correlation coefficient and slope of regression between simulated and expected result

Table 3. Model accuracy indicators for random-ANN and MDA-ANN (unit: %).

Sampling method	Network	Training	Validation	Testing	R	M
MDA	Best validation	0.911	0.994	0.992	0.974	0.954
	Best testing	0.913	0.993	0.993	0.977	0.932
	Average	0.921	0.992	0.992	0.973	0.953
Random	Best validation	0.950	0.954	0.950	0.909	0.824
	Best testing	0.951	0.951	0.954	0.915	0.833
	Average	0.950	0.950	0.950	0.909	0.827

of MDA show an approximately 7% and 12% increase over the random case at average level. This also confirms that the result based on the MDA is closer to the reference data than the random approach.

3.2. Urban maps from ANN and NTL

NTL, NDVI, and LST of 2011 were applied to the ANN to map urban areas at that time and Figure 3 shows the urban area maps from reference NLCD (red), best validated random-ANN (green), and best validated MDA-ANN (blue).

Figure 3(a) shows the ANN classified urban areas of contiguous U.S. from input samples selected using different methods. The simulated results of six counties in the northwest of North Dakota (Burke, Divide, Dunn, McKenzie, Mountrial, and Williams) were excluded because the Bakken oil field around 2010 affected the urban nighttime lights and mislead the urban mapping procedure. Figure 3(b–m) shows the NTL and the overlay of NLCD, Random-ANN and MDA-ANN results, from huge metropolitan areas such as New York-Philadelphia and Los Angeles on the east and west coasts, to middle-size cities such as Seattle and Phoenix in the north and south, and then to small cities such as Louisville and Omaha in the Midwest. As we can see from these figures, the random selection of training samples leads to over-representation of urban areas compared to the reference data, revealing a strong overestimation (Green). The MDA-ANN (blue) also overestimates urban areas, but to a much less degree than random-ANN. Using these results, we created two Cohan's kappa matrices (Table 4) to validate the model results from different ANNs with the 2011 NLCD data at the national scale.

According to Table 4, the overall accuracy increased by 1.84% from 96.00 to 97.84% when the MDA sample selection method was applied. Since a large amount of non-urban areas influences the OA, Cohan's kappa was calculated and it is more significantly improved from MDA-ANN (0.74) than from random-ANN (0.64). The higher kappa value than 0.7 of MDA-ANN indicates a high spatial consistent accuracy and a good validation result between the predicted result and the reference data. However, we can still see the overestimation (commission) and underestimation (omission) with both MDA-ANN and random-ANN (Figure 3). This is mainly caused by the light blooming, wildfires, oil filed, and the under-representation of small or isolated urban areas when using the ANN.

We selected 12 cities from these sub-areas in Figure 3 and provided detailed information on overestimation, underestimation, overall accuracy, and Kappa (Table 5).

Irrespective of the city size, the four indexes fluctuated widely, which would indicate no straight relationship between the city size and the model accuracy. The overall accuracy experienced a slight increase trend from nearly 87%, to 96% and the Kappa kept consistent around 0.80 for all except OKL close to 0.7. The lowest commission error was for Seattle at 8.36%, which also has the highest omission error (19.95%) among the 12 cities. Phoenix was the least underestimated (1.94%) but also was significantly overestimated

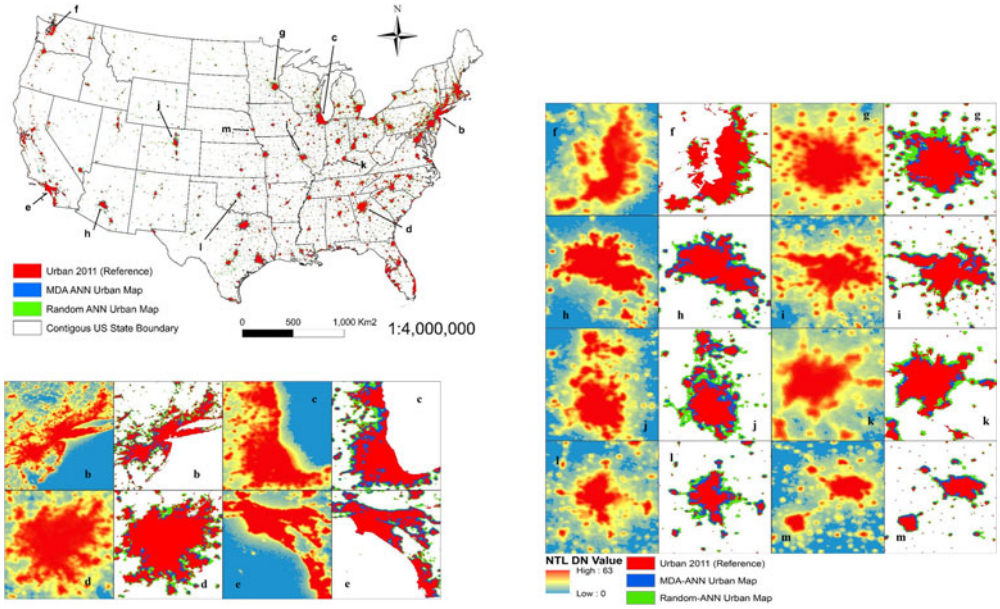

Figure 3. Urban maps from ANN and NTL. (a) Contiguous U.S.A. (b-h) different city-metropolitan areas. First and third columns are the NTL data, second and fourth columns are the overlay comparison among urban area (reference) from NLCD (red), the urban area map extracted from random-ANN (green), and the urban area map extracted from MDA-ANN (blue). The results are ordered by area size from big to small: (b). New York – Philadelphia metropolitan area, (c). Chicago – Milwaukee metropolitan area, (d) Atlanta metropolitan area, (e). Los Angles – San Diego metropolitan area, (f). Seattle, (g). Minneapolis, (h). Phoenix, (i). St. Louis, (j). Denver, (k). Louisville, (l). Oklahoma City, (m). Omaha-Lincoln.

Table 4. Validation results of random and MDA ANN for contiguous U.S.

ANN type	Overall accuracy (%)	Kappa
Random ANN	96.00	0.64
MDA ANN	97.84	0.74

Table 5. Accuracy indexes for 12 cities.

	Size (km^2)	Pop.	Commission (%)	Omission (%)	OA (%)	Kappa
NY	8683	17,800,000	14.00	12.46	86.95	0.74
CHI	5498	8,308,000	17.53	5.10	89.44	0.79
ATL	5083	3,500,000	12.10	13.25	91.52	0.81
LA	4320	11,789,000	13.18	5.06	93.03	0.86
Sea	2470	2,712,000	8.36	19.95	95.01	0.83
Min	2316	2,389,000	26.55	12.98	92.12	0.75
STL	2147	2,078,000	23.41	12.83	93.74	0.78
Phoenix	2069	2,907,000	32.33	1.94	95.23	0.77
Denver	1292	1,985,000	25.85	3.78	96.58	0.82
Louisville	1013	864,000	16.41	17.95	94.82	0.80
OKL	835	747,000	37.84	6.70	93.49	0.71
Omaha	586	627,000	31.25	11.58	95.85	0.75

(32.33%). Generally speaking, the commission error increased with the city size among the 12 cities, from around 10% to more than 30%. Omission error showed no clear trend with city size.

According to Figure 4, we compared the MDA ANN urban mapping results of Seattle and Phoenix to explore the underestimation issue. Most underestimated cells from the figure were small clusters or single cells not far away from either the metropolitan area or

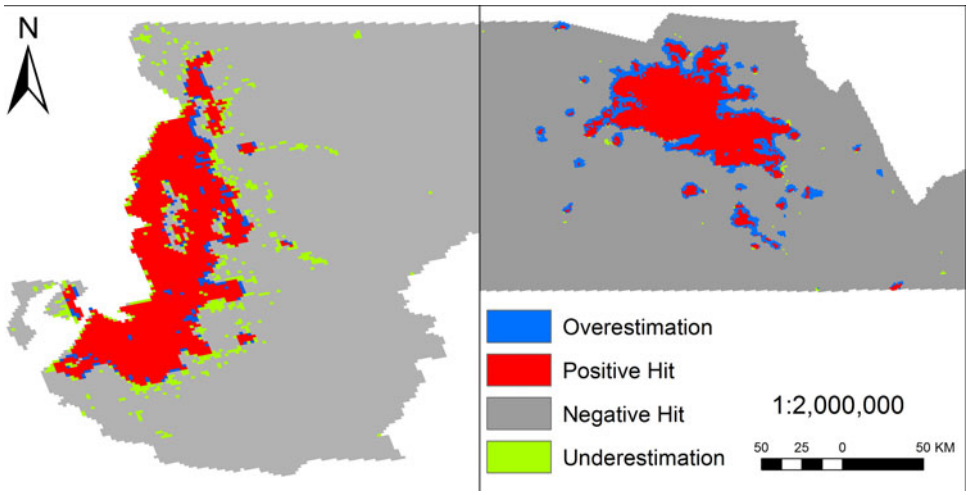

Figure 4. Underestimation of Seattle and Phoenix.

Table 6. Accuracy indexes for state case.

State	City patch number	Average urban area	Commission (%)	Omission (%)	Kappa	Total edge	Edge density
SD	28	2557.882	48.91	26.57	0.6	1,260,070	17.5936
NB	44	3623.271	39.83	25.70	0.66	1,843,580	11.5640
CO	64	6276.976	37.02	14.31	0.72	4,599,570	11.4495
TX	306	7539.973	34.99	13.20	0.73	23,163,700	10.0396

the satellite cities around it, which could be treated as urban-rural or country living areas. Since Seattle has more of these small clusters, it has a higher omission error than Phoenix.

Regarding the size issue, the results were further evaluated at the state scale (Table 6). The characters of variables might be too specific at the city scale but they could be more uniform at the state scale. We extracted the urban areas for four states: Texas, Colorado, Nebraska, and South Dakota. The major city size, average city size, and number of cities all ranged from large to small.

A higher edge density combined with a smaller total edge indicates relatively smaller urban patches in a state, which will raise its underestimation. As the number of cities and the average urban size decrease, this state will have both a higher commission and omission error and lower kappa. This would be very obvious with some less populated states such as Nebraska and South Dakota because the overestimation will still exist with less positive hit urban cells and much more small clusters with more scattered spatial distributions in these states.

3.3. Comparison with the thresholding method

If implemented with NTL data and other environment variables, as a data-driven technology ANN will automatically classify urban and non-urban only on the basis of the training samples, thus it is more objective than the threshold method. Using overall accuracy and Kappa, the MDA-ANN is compared with the widely used threshold method for extracting urban areas at the city, regional, and national levels. At the city scale, the MDA-based ANN simulation results of Atlanta and Phoenix metropolitan areas were

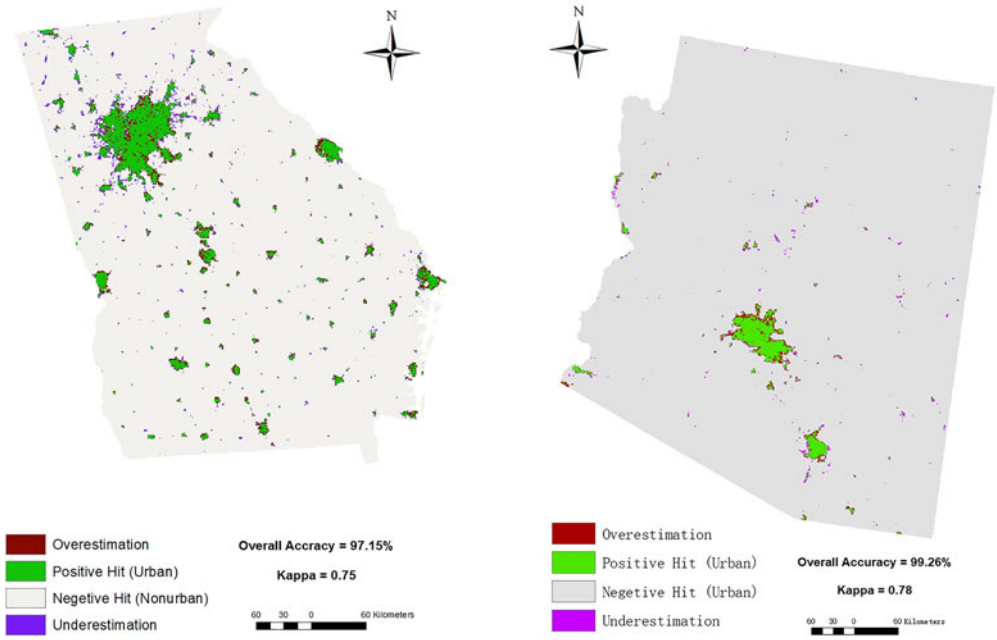

Figure 5. MDA ANN 2011 Urban Map of Georgia (a) and Arizona (b) (city-state level).

compared with those obtained by Xie and Weng (2017). They mapped these urban areas from spatiotemporally enhanced NTL data using the local threshold method that achieved the best result (93.2% accurate, Kappa = 0.65), less accurate than our results at the city scale (95.2/0.77). We also tested these areas with the same data that two specific MDA-ANNs at city-state scale were built and run with Atlanta - Georgia (Figure 5a) and Phoenix - Arizona (Figure 5b). The results have outperformed the threshold method with the significate higher kappa (0.75 for Georgia/0.78 for Arizona) demonstrating that MDA-ANN has the capability to map urban areas also at the state scale.

At the regional level MDA-based ANN also performed well. Taking the Southeast U.S.A. as an example, Li et al. (2016) combined NTL with NDVI to generate the Vegetation Adjusted NTL Urban Index (VANUI) and used it as a classification index to map urban expansion. They achieved an overall accuracy of 84.98% and a kappa of 0.58, with a commission error larger than 65%. After applying the MDA-based ANN to the same Southeast U.S.A using the same input variables (Figure 6a), the overall accuracy increased by more than 10% to 95.29% and Kappa jumped to 0.66 while the commission error dropped to less than 45%. This is nearly 20% lower than the threshold-based result. The same approach was also applied to the Midwest region (Figure 6b) that the Kappa increased to 0.73 with an overall accuracy at 97.71% and the commission error reduced to 30.23%. Judged against all accuracy indexes, the urban areas extracted using the proposed machine-learning approach have improved accuracy.

The above outperformed findings of ANN approaches are also confirmed at the national level. For instance, the cluster-based method proposed by Zhou et al. (2014) achieved 91% of overall accuracy and 0.69 for kappa in mapping the urban extent of U.S.A from NTL data. Using the same study area and the same input data, the Kappa result of MDA-based ANN is raised to 0.74 and the overall accuracy is 7% higher (97.84%). This result was also comparable to that of Xie and Weng et al. (2017) with the spatiotemporally enhanced DMSP/OLS NTL data with the contiguous U.S.A (99.2%/0.71).

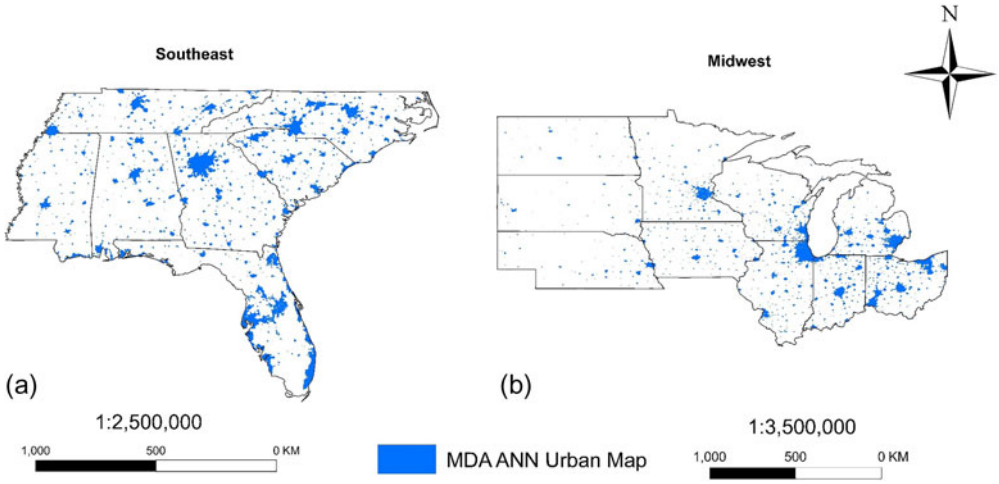

Figure 6. MDA ANN Urban Map of Southeast and Midwest U.S.A in 2011.

4. Discussion

4.1. Random vs. MDA

MDA is a clustering (attribute not spatially) method, especially useful for sampling two classes such as urban and nonurban. It has already been shown to be an effective sample selection method in other geophysical studies. This study also confirms its potential for urban mapping from NTL data. Based on 1000 ANNs runs for both random and MDA methods, the average validation and testing accuracy of the random sample selection ANN was 0.9504 and 0.9503, respectively, while for MDA, the accuracy rose to 0.9918. The higher regression values also showed that the target result from MDA-ANN is closer to the real data than for the random-ANN.

For this study, urban areas usually have a higher NTL value and hotter LST, but lower NDVI values. In contrast, non-urban areas are characterized by opposite values. Taking the state of Colorado (CO) as an example (because of its rectangle shaped boundary) and using the same number of training sample points (500), the differences between the two sampling methods can be highlighted (Figure 7). In the MDA sample selection method, 47.6% (238 of 500) selected samples were distributed inside existing urban areas whereas the remaining points were either very close to urban areas or located far away in other land cover types. The distribution of the selected samples ensures that both urban and non-urban locations are equally represented in the training dataset which in turn reduces overestimation in the classification process. However, in the random sample selection method, only 18.6% (93 of 500) selected points fell inside existing urban areas. This underrepresentation of non-urban areas in the training samples is likely to cause overfitting over non-urban samples, leading to overestimation and underestimation. In addition, to explore more explicitly the spatial distribution trend of these two sample datasets, the Moran's I and the average NTL, NDVI, LST of selected points inside each county of CO were also calculated. Results indicate a compact clustering spatial distribution of MDA (Morans' $I = 0.34$, $z = 4.83$, and $p = 0$), compared to the results of random selection (Morans' $I = 0.14$, $z = 2.01$, and $p = 0.045$). The results of Moran's I confirm that the spatial autocorrelation of different input variables is implicitly accounted for in the MDA algorithm, ultimately providing a more accurate ANN prediction than the random sample selection method.

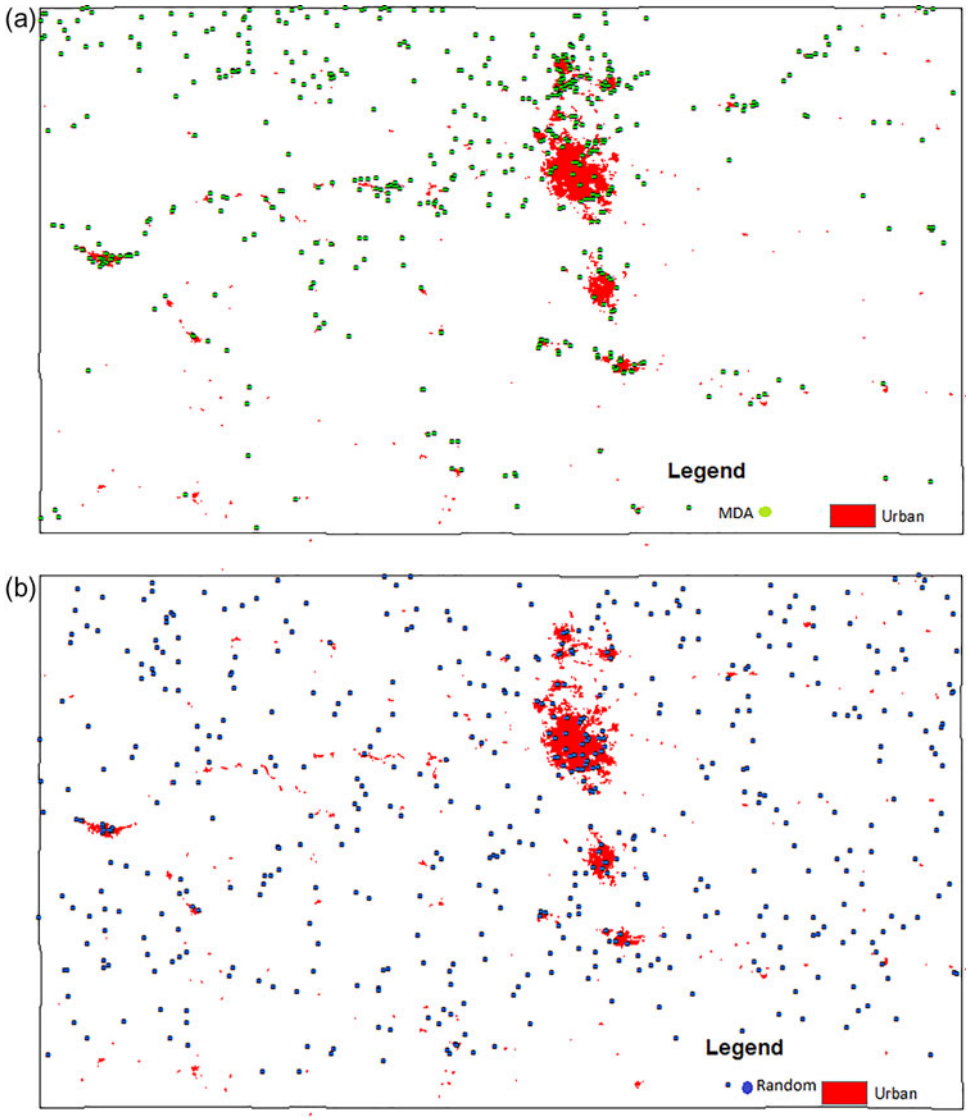

Figure 7. Samples from (a) MDA and (b) Random sampling for the state of Colorado. Green dots are training samples using the MDA method and blue dots are randomly selected training samples.

4.2. Uncertainty

At the national, regional, state and city scales, the MDA-based ANN all outperforms or is as good as the threshold method, but is a more rapid and objective procedure (Jing et al. 2015; Dou et al. 2017). Also, suitable threshold values are usually difficult to determine and certainly not universal (Milesi et al. 2003; Liu et al. 2017). Thus, the more objective MDA-based ANN has the potential to become a novel tool for extracting and mapping urban information from conventional NTL data. Although the MDA-based ANN has been shown to perform well in extracting urban information on a diversity scales, there are some sources of uncertainty that need to be discussed.

We assumed that our algorithm would overestimate larger cities but underestimate small cities. However, according to the analysis of results of 12 cities, we found that there

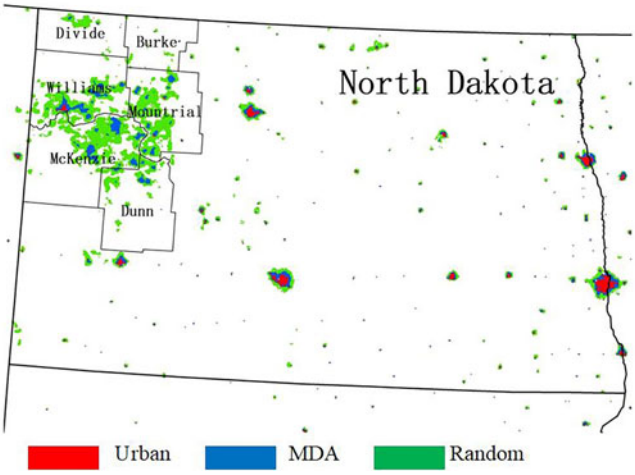

Figure 8. Map results including the oil field in North Dakota, 2011.

was no obvious relation between the city size and the model result. Smaller cities also had a relative large commission error and small omission error, despite its size. This is due to the city diversity such as the city shape, road network, and population density, which would all affect the light blooming and saturation to the overestimation (Townsend and Bruce 2010; Bagan and Yamagata 2015; Shi et al. 2015). However, at the state level, the result accuracy decreased as the number of cities and average urban size decreased, yielding more uniform trainings. To further address this issue, more predictive variables could be added to improve the training process. For example, DEM, road density, population density, and distance to existing urban areas are all reasonable candidates/variables to improve the predictors. In addition, ANN could be scaled down to the city-state level to make the method more specific to a target area.

Another issue is the underestimation of small clusters or single urban pixels which is typical of the MDA-ANN. It is more obvious within some less populated states as they have a large number of small clusters or single cells of residential areas separately distributed over an extensive area. This issue is likely to be mainly caused by the mismatching of the spatial resolution between NTL and NLCD, with limited spatial (1 km) and radiometric (8 bits) resolutions of DMSP - OLS NTL data. Most of the single or low density urban pixels are tiny residential areas which may not have digital values as strong as those of major urban areas and were lost during the training process. MDA uses every attribute value to cluster samples into different categories in feature space. It does not control the spatial domain but will affect the spatial distribution of the sample because of the Tobler's Law. It aims to provide the most differentiate samples for training and has good performance defining the boundaries of the data space. However, if the data values between these unclear urban cells and nonurban samples are not too distinctive to be distinguished through NTL, NDVI, and LST, they will be missed. For example, the rural residential areas and the isolated urban cell may have very close temperature and vegetation distribution as adjacent non-urban cells with a relatively dark tone so that the attribute values of them could be quite close. This will confuse the ANN during the training process. This issue could be resolved by the set up two training processes. One is for mapping metropolitan areas and the other one is specifically for relatively small urban areas and isolated cells.

Last but not least, the illumination light not from urban areas must be treated carefully because it is a significant source of overestimation. Taking this study for example, a megacity approximately with the size of Chicago is detected in the North West of the North Dakota State from 2011 urban map, which obviously is not a city but a Bakken oil field (US Energy Information Administration, https://www.eia.gov/todayinenergy/detail.php?id=3750). The training process of ANN was not affected by the Bakken oil field because it had not been discovered at the training stage of year 2001. However, for the simulation and validation processes, this error must be taken into account when using the 2011 NTL data. Six counties of North Dakota (Burke, Divide, Dunn, McKenzie, Mountrial, and Williams) were significantly affected by the Bakken oil fields so should be eliminated from the simulation process. Their influence on random ANN is higher than on MDA ANN in the original simulation results (Figure 8), which may indicate that MDA ANN would be less affected by error (noise) data than random ANN during the training process.

5. Conclusions

In this paper, a machine learning approach was used to derive urban areas from DMSP - OLS NTL data in conjunction with MODIS NDVI, and LST data. We used artificial neural networks and applied two sample selection methods (random and MDA) to assess the impact of training samples on the "learning" process. The extracted urban areas were validated against the NLCD data and compared with the threshold-based method. The validation and comparison results indicate that ANN can be used to extract urban areas through NTL data. This approach led to reasonable results for the urban simulation of the contagious US at city-state, regional, and national scales. Therefore, ANN can be considered as an effective and successful approach for rapid and accurate extraction of urban land areas especially if combined with NDVI and LST data.

In addition, the impact of training samples was also tested by applying a random approach and a clustering (MDA) algorithm. Overall, MDA-ANN provides a better result with greater overall accuracy and Cohan's Kappa than random. Random training does not account for effects related to spatial autocorrelation but the imbalanced training samples will lead to inefficient training hence that increases the overestimation problem. On the other hand, MDA can distinguish the distinct urban area from close non-urban area to reduce the overestimation. However, the algorithm underperforms when trying to identify small towns as well as isolated small urban cells overall amplifying underestimation. This result shows that the ANN results are indeed affected by the distribution of the training samples so that sampling methods should be carefully considered before embarking on a mapping exercise.

Although the ANN could provide timely and reasonable results, problems of over/underestimation still persist and the per-pixel-based process requires further refinement. Apart from environmental variables, economic factors such as population density, road network density, and GDP could also be considered as independent input variables in the ANN. Such additional variables can enrich the unique characters of training samples and hence improve the predictive skills of ANN (Ghosh et al. 2010, 2013). In addition, the new NTL products from NOAA – VIIRS, with a finer spatial resolution and more quantization levels, may provide a more detailed data source to improve the accuracy of urban area mapping with ANN (Schueler et al. 2013; Shi et al. 2014).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Tingting Xu http://orcid.org/0000-0002-4468-1413

References

Aburas MM, Ho YM, Ramli MF, Ash'aari ZH. 2016. The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: a review. Int J Appl Earth Obs Geoinf. 52: 380–389. DOI: http://dx.doi.org/10.1016/j.jag.2016.07.007

Alsharif AAA, Pradhan B, Mansor S, Shafri HZM. 2015. Urban expansion assessment by using remotely sensed data and the relative Shannon Entropy Model in GIS: a case study of Tripoli, Libya. Theor Emp Res Urban Manag. 10(1):55–71.

Angel S, Parent J, Civco DL, Blei A, Potere D. 2011. The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. Progr Plan. 75(2):53–108. DOI: http://dx.doi.org/10.1016/j.progress.2011.04.001

Bagan H, Yamagata Y. 2015. Analysis of urban growth and estimating population density using satellite images of nighttime lights and land-use and population data. GISci Remote Sens. 52(6):765–780.

Camus P, Mendez FJ, Medina R, Cofiño AS. 2011. Analysis of clustering and selection algorithms for the study of multivariate wave climate. Coast Eng. 58(6):453–462. DOI: doi.org/10.1016/j.coastaleng.2011.02.003

Cao X, Chen J, Imura H, Higashi O. 2009. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. Remote Sens Environ. 113(10):2205–2209. DOI: http://dx.doi.org/10.1016/j.rse.2009.06.001

Dahiya BS. 2016. Impact of economic development on regional structure of urban systems in India spatial diversity and dynamics in resources and urban development. Dordrecht, Netherland: Springer; ISBN: 978-94-017-9785-6

Dou Y, Liu Z, He C, Yue H. 2017. Urban land extraction using VIIRS nighttime light data: an evaluation of three popular methods. Remote Sens. 9(2):175. DOI: https://doi.org/10.3390/rs9020175

Elvidge C, Baugh K, Hobson V, Kihn E, Kroehl H, Davis E, Cocero D. 1997. Satellite inventory of human settlements using nocturnal radiation emissions: A contribution for the global toolchest. Global Change Biol. 3(5):387–395. −DOI: https://doi.org/10.1046/j.1365-2486.1997.00115.x

Elvidge CD, Ziskin D, Baugh KE, Tuttle BT, Ghosh T, Pack DW, Erwin EH, Zhizhin M. 2009. A fifteen year record of global natural gas flaring derived from satellite data. Energies. 2(3):595–622.

Elvidge CD, Baugh KE, Anderson SJ, Sutton PC, Ghosh T. 2012. The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite. Soc Geogr. 7(1):23–35.

Gallo KP, Elvidge CD, Yang L, Reed BC. 2004. Trends in night-time city lights and vegetation indices associated with urbanization within the conterminous USA. Int J Remote Sens. 20:2003–2007. DOI: https://doi.org/10.1080/01431160310001640964

Ghosh T, Anderson SJ, Elvidge CD, Sutton PC. 2013. Using nighttime satellite imagery as a proxy measure of human well-being. Sustainability. 5(12):4988–5019. DOI: https://doi.org/10.3390/su5124988

Ghosh T, Powell RL, Elvidge CD, Baugh KE, Sutton PC, Anderson A. 2010. Shedding light on the global distribution of economic activity. Open Geogr. 3:147–160. DOI:

Goodchild MF. 1986. Spatial autocorrelation (Volume 47 of CATMOG Series, Issue 47 of Concepts and techniques in modern geography). Geo Books: The University of California

Hagan MT, Demuth HB, Beale MH. 1996. Neural Network Design, Boston, MA: PWS Publishing.

Henderson M, Yeh ET, Gong P, Elvidge C, Baugh K. 2003. Validation of urban boundaries derived from global night-time satellite imagery. Int J Remote Sens. 24(3):595–609.

Herold M, Goldstein NC, Clark KC. 2003. The spatiotemporal form of urban growth: measurement, analysis and modeling. Remote Sens Environ. 86(3):286–302.

Homer C, Dewitz J, Fry J, Coan M, Hossain N, Larson C, Herold N, McKerrow A, VanDriel JN, Wickham J. 2007. Completion of the 2001 National Land Cover Database for the Conterminous United States. Photogramm Eng Remote Sensing. 73(4):337–341.

Homer CG, Dewitz JA, Yang L, Jin S, Danielson P, Xian G, Coulston J, Herold ND, Wickham JD, Megown K. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. Photogramm Eng Remote Sensing. 81(5):345–354.

Imhoff ML, Lawrence WT, Stutzer DC, Elvidge DC. 1997. A technique for using composite DMSP/OLS "City Lights" satellite data to map urban area. Remote Sens Environ. 61(3):361–370.

Jing W, Yang Y, Yue X, Zhao X. 2015. Mapping Urban Areas with Integration of DMSP/OLS Nighttime Light and MODIS Data Using Machine Learning Techniques. Remote Sens. 7(9):12419–12439.

Lagarias A. 2012. Urban sprawl simulation linking macro-scale processes to micro-dynamics through cellular automata, an application in Thessaloniki, Greece. Appl Geogr. 34:146–160. DOI: http://dx.doi.org/10.1016/j.apgeog.2011.10.018

Lawrence WT, Imhoff ML, Kerle N, Stutzer D. 2002. Quantifying urban land use and impact on soils in Egypt using diurnal satellite imagery of the Earth surface. Int J Remote Sens. 23(19):3921–3937.

Li Q, Lu L, Weng Q, Xie Y, Guo H. 2016. Monitoring Urban Dynamics in the Southeast U.S.A. Using Time-Series DMSP/OLS Nightlight Imagery. Remote Sens. 8(7):578.

Li X, Yeh AG-O. 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. Int J Geogr Inf Sci. 16(4):323–343.

Liu Z, He C, Zhang Q, Huang Q, Yang Y. 2012. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. Landsc Urban Plan. 106(1):62–72. DOI: 10.1016/j.landurbplan.2012.02.013

Liu Y, Yang Y, Jing W, Yao L, Yue X, Zhao X. 2017. A New Urban Index for Expressing Inner-City Patterns Based on MODIS LST and EVI Regulated DMSP/OLS NTL. Remote Sens. 9(8):777.

Lu D, Tian H, Zhou G, Ge H. 2008. Regional mapping of human settlements in southeastern China with multisensor remotely sensed data. Remote Sens Environ. 112(9):3668–3679. DOI: 10.1016/j.rse.2008.05.009

Ma T, Zhou Y, Zhou C, Haynie S, Pei T, Xu T. 2015. Night-time light derived estimation of spatio-temporal characteristics of urbanization dynamics using DMSP/OLS satellite data. Remote Sens Environ. 158:453–464. DOI: 10.1016/j.rse.2014.11.022

Ma Y, Xu R. 2010. Remote sensing monitoring and driving force analysis of urban expansion in Guangzhou City, China. Habitat Int. 34(2):228–235. DOI: 10.1016/j.habitatint.2009.09.007

Mildrexler DJ, Zhao M, Running SW. 2009. Testing a MODIS global disturbance index across North America. Remote Sens Environ. 113(10):2103–2117.

Milesi C, Elvidge CD, Nemani RR, Running SW. 2003. Assessing the impacts of urban land development on net primary productivity in the southeastern United States. Remote Sens Environ. 86(3):401–410. –

Park S, Jeon S, Kim S, Choi C. 2011. Prediction and comparison of urban growth by land suitability index mapping using GIS and RS in South Korea. Landsc Urban Plan. 99(2):104–114. DOI: 10.1016/j.landurbplan.2010.09.001

Passarella M, Goldstein EB, De Muro S, Coco G. 2018. The use of genetic programming to develop a predictor of swash excursion on sandy beaches. Nat Hazards Earth Syst Sci. 18(2):599–611.

Pijanowski BC, Brown DG, Shellito BA, Manik GA. 2002. Using neural networks and GIS to forecast land use changes: a Land Transformation Model. Comput Environ Urban Syst. 26:552–575.

Pijanowski BC, Tayyebi A, Delavar MR, Yazdanpanah MJ. 2010. Urban expansion simulation using geospatial information system and artificial neural networks. Int J Environ Res. 3(4):493–502.

Qi F, Zhu A. 2003. Knowledge discovery from soil maps using inductive learning. Int J Geogr Inf Sci. 17(8):771–795.

Reyes-Merlo MÁ, Ortega-Sánchez M, Díez-Minguito M, Losada MA. 2017. Efficient dredging strategy in a tidal inlet based on an energetic approach. Ocean Coast Manage. 146(Supplement C):157–169. DOI: 10.1016/j.ocecoaman.2017.07.002

Ridd MK. 1995. Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. Int J Remote Sens. 16(12):2165–2185.

Rueda A, Gouldby B, Méndez FJ, Tomás A, Losada IJ, Lara JL, Díaz-Simal P. 2016. The use of wave propagation and reduced complexity inundation models and metamodels for coastal flood risk assessment. J Flood Risk Manage. 9(4):390–401.

Small C, Pozzi F, Elvidge CD. 2005. Spatial analysis of global urban extent from DMSP-OLS night lights. Remote Sens Environ. 96(3-4):277–291.

Small C, Elvidge CD. 2011. Mapping decadal change in anthropogenic night light. Procedia Environ Sci. 7:353–358.

Schueler CF, Lee TF, Miller SD. 2013. VIIRS constant spatial-resolution advantages. Int J Remote Sens. 34(16):5761–5777.

Shi K, Huang C, Yu B, Yin B, Huang Y, Wu J. 2014. Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas. Remote Sens Lett. 5(4):358–366.

Shi K, Yu B, Hu Y, Huang C, Chen Y, Huang Y, Chen Z, Wu J. 2015. Modeling and mapping total freight traffic in China using NPP-VIIRS nighttime light composite data. GIScience Remote Sens. 52(3):274–289.

Sutton P. 2003. A scale-adjusted measure of "urban sprawl" using nighttime satellite imagery. Remote Sens Environ. 86(3):353–369. –

Sukhatme BV, Avadhani MS. 1965. Controlled selection a technique in random sampling. Ann Inst Stat Math. 17(1):15–28.

Tobler W. 1970. A computer movie simulating urban growth in the Detroit region. Econ Geogr. 46(Supplement):234–240.

Tong F, Liu X. 2005. Samples selection for artificial neural network training in preliminary structural design. Tinshhua Sci Technol. 10(2):233–239.

Townsend AC, Bruce DA. 2010. The use of night-time lights satellite imagery as a measure of Australia's regional electricity consumption and population distribution. Int J Remote Sens. 31(16):4459–4480.

Tole L. 2008. Changes in the built vs. non-built environment in a rapidly urbanizing region: a case study of the Greater Toronto Area. Comput Environ Urban Syst. 32(5):355–364.

Wang F. 1994. The use of artificial neural networks in a geographical information system for agricultural land-suitability assessment. Environ Plan A. 26(2):265–284.

Wang G, Chen J, Li Q, Ding H. 2006. Quantitative assessment of land degradation factors based on remotely-sensed data and cellular automata: a case study of Beijing and its neighboring areas. Environ Sci. 3(4):239–253.

Wang Y, Huai W. 2016. Estimating the longitudinal dispersion coefficient in straight natural rivers. J Hydraul Eng. 142(11):04016048.

Xie Y, Weng Q. 2016. Updating urban extents with nighttime light imagery by using an object-based thresholding method. Remote Sens Environ. 187:1–13. DOI: 10.1016/j.rse.2016.10.002

Xie Y, Weng Q. 2017. Spatiotemporally enhancing time-series DMSP/OLS nighttime light imagery for assessing large-scale urban dynamics. ISPRS J Photogramm Remote Sens. 128:1–15. DOI: 10.1016/j.isprsjprs.2017.03.003

Zhang Q, Seto KC. 2011. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. Remote Sens Environ. 115(9):2320–2329.

Zhang Q, Schaaf C, Seto KC. 2013. The vegetation adjusted NTL urban index: A new approach to reduce saturation and increase variation in nighttime luminosity. Remote Sens Environ. 129:32–41.

Zhang H, Jin X, Wang L, Zhou Y, Shu B. 2015. Multi-agent based modeling of spatiotemporal dynamical urban growth in developing countries: simulating future scenarios of Lianyungang city, China. Stoch Environ Res Risk Assess. 29(1):63–78.

Zhao Y, Murayama Y. 2011. Urban Dynamics Analysis Using Spatial Metrics Geosimulation. In Spatial Analysis and Modelling in Geographical Transformation Process, Murayama, Yuji, Thapa, Rajesh Bahadur (Eds.) Springer, ISBN: 978-94-007-0671-2

Zhou Y, Smith SJ, Elvidge CD, Zhao K, Thomson A, Imhoff M. 2014. A cluster-based method to map urban area from DMSP/OLS nightlights. Remote Sens Environ. 147:173–185. DOI: 10.1016/j.rse.2014.03.004

Zhu AX. 2000. Mapping soil landscape as spatial continua: The Neural Network Approach. Water Resour Res. 36(3):663–677.

Zou Y, Peng H, Liu G, Yang K, Xie Y, Weng Q. 2017. Monitoring urban clusters expansion in the middle reaches of the Yangtze River, China, using time-series nighttime light images. Remote Sens. 9(10):1007.