



Assessment on controlling factors of urbanization possibility in a newly developing city of the Vietnamese Mekong delta using logistic regression analysis

Nguyen Thi Hong Diep^a, Can Trong Nguyen^{a,b,*}, Phan Kieu Diem^a, Nguyen Xuan Hoang^a, Abdulla - Al Kafy^{c,d}

^a College of Environment and Natural Resources, Can Tho University, Viet Nam

^b Joint Graduate School of Energy and Environment (JGSEE), King Mongkut's University of Technology Thonburi (KMUTT). Center of Excellence on Energy Technology and Environment (CEE), PERDO, Ministry of Higher Education, Science, Research and Innovation, Thailand

^c Department of Urban & Regional Planning, Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh

^d ICLEI South Asia, Rajshahi City Corporation, Rajshahi-6203, Bangladesh

ARTICLE INFO

Keywords:

Multiple logistic regression
Mekong delta
Urban expansion
Urbanized driven factors
Urbanization possibility

ABSTRACT

Urban development is dominated by various factors ranging from natural and social factors to accessibility to urban infrastructures. Urbanization in an unfavorable location in terms of the above factors can create difficulties with connecting to the city center, adjacent urban areas, and wastage of land resources. In the context of Can Tho city, a newly developing city, its urban expansion process and factors affecting urbanized possibility were explored by applying multiple logistic regression (MLR) on Landsat imagery and accessible geospatial data sources. The analyses confirmed a significant urban expansion in the entire city between 2003 and 2017, mostly in central districts and along the Hau river. The primary dynamics of urban expansion were explained by an efficient MLR (Area Under Receiver Operating Characteristic – AUROC = 0.803), based on six factors related to accessibility to transportation, developed urban areas, industrial zone, elevation, soil type, and population. A simulation of urbanization probability revealed that most remote areas with low accessibility to urban infrastructures are difficult to urbanize with a probability of less than 40%. In contrast, the high potentially urbanized regions expanded the already built-up areas in riverside districts. Our findings facilitate the understanding of urbanized driven factors in the newly developing delta cities for long-term planning when urbanization remains under control.

1. Introduction

Vietnam has been developing at a high urbanization rate, especially in the last four decades (Chen et al., 2014; Fan et al., 2019). Urbanization facilitates reviving the sluggish economy after political instability, reducing poverty, and increasing household welfare (Glewwe, 2004; Kafy et al., 2021b). It intensively propels the national development and economic growth through converging capital, reinforcing trade between urban and rural areas, creating jobs, and providing multiple services (Sheng, 2017; Sheng and Mohit, 2001; Turok, 2016). Even though fast urbanization provides numerous development opportunities, excessively rapid urbanization and the lack of a master plan can induce several burdens that threaten a city's sustainable development related to

social aspects, overload of urban infrastructures, and environmental degradation (Peng et al., 2010; Sintusingha, 2011). The transformation from agricultural land to built-up areas leads to economic structure changes, unemployment, job transition, urban poverty, and social division, especially to illiterate farmers, older people, and women in rural areas (Heurlin, 2019; Shouhai, 2015). The urban development in Vietnam faces an overload of infrastructures caused by the synergistic impact of intense urbanization, immigration, non-synchronous infrastructure, and inconsequential urban planning (Chu and Nguyen, 2017; Chu and Thi, 2017). For instance, traffic congestion occurs after working hours when a large number of vehicles are present on narrow roads designed for the current population. A shortage of clean water and electricity is another problem, along with industrial development and

* Corresponding author. College of Environment and Natural Resources, Can Tho University, Viet Nam.

E-mail addresses: nthdiep@ctu.edu.vn (N.T. Hong Diep), can.62300800201@mail.kmutt.ac.th, ntcan93@gmail.com (C.T. Nguyen).

<https://doi.org/10.1016/j.pce.2021.103065>

Received 17 October 2020; Received in revised form 29 March 2021; Accepted 9 August 2021

Available online 11 August 2021

1474-7065/© 2021 Elsevier Ltd. All rights reserved.

urban expansion (Nguyen et al., 2021b).

Moreover, the urban environment is vulnerable under dense populations, the bustle of diverse economic sectors, land cover disturbance, and even environmental awareness of urban residents. Most cities in developing countries are facing a weakness of sanitation systems that is likely to affect soil and water environments (McFarlane, 2019). The urban air environment is polluted by transportation and anthropogenic activities (Hien et al., 2020; Pham et al., 2019; Phan et al., 2020). Meanwhile, the urban environment shifts itself through landscape changes. Temperature increase is known as the urban heat island (UHI) effect (Kafy et al., 2021a; US EPA, 2008), and has been broadly investigated in association with urban expansion in the last few years (Can et al., 2019; Estoque et al., 2017; Estoque and Murayama, 2017; Mohan and Kandy, 2015; Son et al., 2017; Tran et al., 2017; Xinmin et al., 2017).

The problems mentioned above originate from incorrect planning and a lack of insight and rationality among the related elements. For example, planning a new urban area may benefit urban spatial planning because it does not depend on the previous construction. Nevertheless, this decision is costly because of land clearance, environmental assessment, and completely new infrastructures; otherwise, the project will be ineffective and asynchronous with other urban areas (Nguyen et al., 2021c). The favorable position for urbanization, therefore, needs to be identified for sustainable urban development.

Can Tho city was established as one of the five national municipalities in 2004. After that, the city has rapidly developed, and the population increased from 611 thousand (2005) to 1,175 thousand (2015). The city's urban resident percentage climbed by 1.1% from 2005 and reached 3.7% in 2015 (United Nations, 2015). In terms of spatial urbanization, a case study in three center districts (Ninh Kieu, Cai Rang, and Binh Thuy) carried out by remote sensing indicated that the built-up areas extended approximately 3,500 ha over 20 years (Son and Thanh, 2018). The annual growth rate of urban areas revealed a fast urban expansion process and this rate soared from 2005 onward (Pham et al., 2010; Trung and Vu, 2018). The city is in the distribution region of acid sulfate soil of the Vietnamese Mekong Delta (VMD), yet alluvial soil that favors agricultural cultivation, a crucial strength of the city, is limited in the area (IRMC, 2003).

Furthermore, cities in the lower Mekong Delta, especially in Can Tho city, are at risk of urban flooding given by the combined impact of low altitude, land subsidence, and sea-level rise (Apel et al., 2016; Chinh et al., 2016; Huong and Pathirana, 2013; Long et al., 2020). Although Can Tho is a newly developing city with a small urban scale compared to Ho Chi Minh and Ha Noi city, it has shown essential signs of urban environmental changes owing to unfavorable planning, such as green space reduction and temperature increase (Diep et al., 2018; Son and Thanh, 2018). For all the reasons mentioned above, conditions in the region for urban building need to be explored.

The logistic regression is utilized to estimate the probability of an event with a categorical dependent variable. More explicitly, the target variable for the logistic regression has a dual aspect, e.g., failure/success, win/loss, and delicious/tasteless. Practically, logistic regression has been successfully applied in studying asthma morbidity caused by urbanization (Ponte et al., 2018), perception of air pollution (Leung et al., 2019), housing abandonment (Gao et al., 2017), and urban growth dynamics with a contribution from Cellular Automata Markov (Siddiqui et al., 2018; Wu et al., 2019). The utilization of logistic regression is for its potential in detecting impact factors of urbanization and non-urbanization regions.

This study aims to contribute a set of substantial factors controlling the urbanization process and detecting potential locations for urban expansion in a developing city in the Vietnamese Mekong Delta. Specifically, this paper attempts to (1) detect the urbanization areas in Can Tho city during 2003–2017 by applying remote sensing-based analysis, (2) assess the impact of considered factors on urbanization ability, and (3) estimate urbanization probability using multiple logistic regression.

In order to present these objectives, this paper is organized as follows: Section 2 below describes the study area, the dataset used for this research, and the methods for, in turn, achieving the stated objectives. The results and discussions are presented in Section 3, and a conclusion ends the paper in section 4.

2. Materials and methods

2.1. Study area description

Can Tho city is located at the center of the VMD on the west side of the Hau river, i.e., a distributary of the Mekong river in Vietnam. The total area is about 1,401 km² stretching from 105°13'38" – 105°50'35" East and 9°55'08" – 10°19'38" North and divided into nine administrative units, consisting of 5 urban districts and 4 rural counties (Fig. 1). Can Tho belongs to a semi-open flooded area and includes three types of geomorphology: natural dike of Hau river, semi-open flood plain of Long Xuyen quadrangle, and flat delta (CTP, 2017; Long et al., 2020; Van Long and Cheng, 2018). The typical elevation is 2–5 m and gradually decreases from Northeast to Southwest. Two main soil groups were found regarding primary soil types: alluvial soil (84%) and acid sulfate soil (16%) (CTP, 2017). The Hau river and dense canal system supply approximately 35 million m³/year, jointly facilitated agricultural cultivation (e.g., 115,000 ha of rice fields and fruit orchards).

Besides this agricultural strength, the city also focuses on industrial production, especially for aquatic processing, rice milling, and agricultural supplies. Presently, there are two industrial areas: Tra Noc industrial area (300 ha) and Hung Phu industrial area (471 ha). The government plans three other industrial parks with a total area of up to 1,600 ha in Thot Not district, O Mon and north of O Mon district. These are all promising developments to boost overall growth and urbanization (CTP, 2017).

2.2. Data

2.2.1. Landsat satellite images

Landsat satellite images were the primary data source for obtaining and assessing urban areas. The research collected two free cloud Landsat scenes of Collection 1 (level 1) at path/row of 125/53. The Landsat 7-ETM and Landsat 8 images were acquired on 27 August 2003 and 29 January 2017, respectively. Although these two scenes were acquired in two different seasons, this seasonal dissimilarity is expected only to influence the identification of paddy fields and similar land cover types. The urban identification is unaffected due to its independence of season. One Landsat 5-TM on 23 November 2003 was also collected to fix the gap error due to trouble on the Scan Line Corrector (SLC) of Landsat 7-ETM (Andreou et al., 2003). The remotely sensed data have a medium resolution (30 m) and cloudlessness in the study area. It was therefore possible to process and extract urban areas for further analyses.

2.2.2. Spatial data on potential factors

The study gathered available data on the factors expected to influence urbanization. Data formats and sources are shown in Table 1 below. Population density is in the raster format, illustrating the number of people living in an area unit, 1,000 m of pixel size. Altitude data is ALOS PALSAR DEM (Advanced Land Observing Satellite – Phased Array type L-band Synthetic Aperture Radar – Digital elevation model), a secondary elevation data generated by radar satellite image L-band. Transportation network, hydrological system, and industrial areas are vector data presenting the spatial distribution of major roads, rivers and canals, and industrial area locations.

Furthermore, we used the soil paper-map from the Integrated Resources Mapping Center (IRMC) to supply soil type information throughout the study area. The map was generated based on information in 2003 for Can Tho province (i.e., a former province that included Can Tho city and Hau Giang province). Firstly, the image was georeferenced

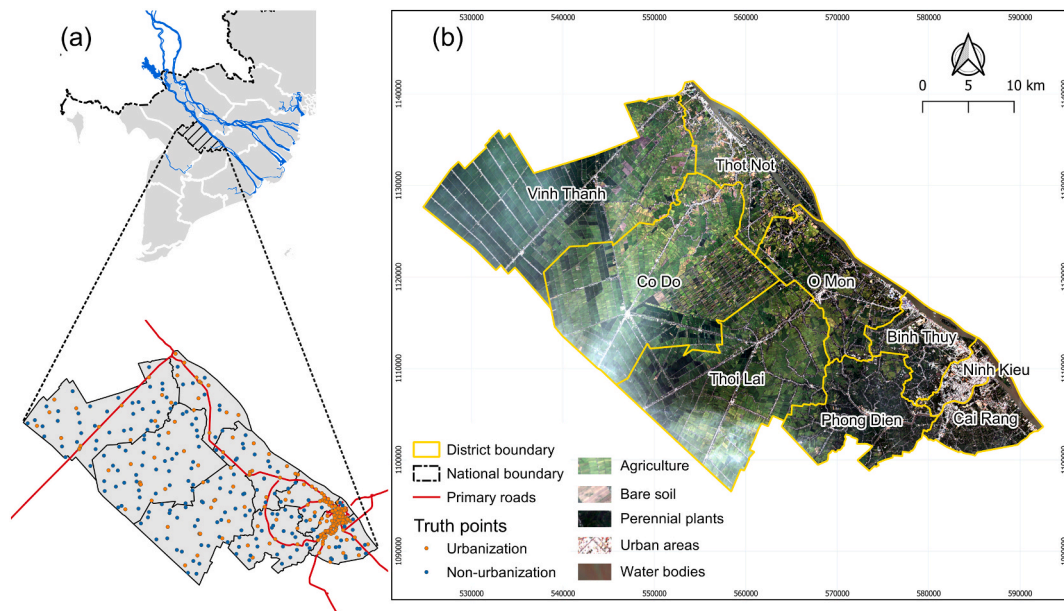


Fig. 1. (a) Can Tho city location in Mekong Delta and truth points for assessing classification accuracy, and (b) Landsat 8-OLI natural color composite (RGB = bands 4, 3, 2) with district boundary. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Spatial data representing impact factors on urbanization.

Data	Factor	Scope	Scale	Format	Source
Administrative map in 2012	Administrative boundary, center location	Province	1:50,000	Image	Environment and Natural Resources Department in Can Tho
ALOS PALSAR Digital elevation model (DEM)	Elevation	Global	12.5 m	Raster	https://search.asf.alaska.edu
Industrial area	Accessibility	Country		Vector	https://vietnam.opendevelopmentmekong.net
Hydrological system	Accessibility	Country		Vector	https://download.geofabrik.de/asia/vietnam.html
Population data in 2003	Population density	Global	1 km	Raster	https://landscan.ornl.gov
Soil classification (paper map)	Soil type	District	1:50,000	Image	Integrated Resources Mapping Center
Transportation network	Accessibility	Country		Vector	https://vietnam.opendevelopmentmekong.net

and digitized on a GIS (geographical information system) platform. There are seven soil types following the Vietnam soil classified system (VSCS), namely Gley-alluvial soil (50.79%), alluvial soil (17.35%), disturbed soil (15.71%), active acid sulfate soil at shallow layers (12.12%), active acid sulfate soil at deep layers (2.39%), alluvial soil (1.39%), and potential acid sulfate soil at deep layers (0.27%). The soil types were subsequently reviewed and scored from 7 to 1 to prioritize urban development when they are not suitable for agricultural cultivation and each soil type's present status. For example, the gley-alluvial soil was given 1 point because it is distributed in the fields with a long tradition of double-triple rice crops, without alum, so it is the most suitable soil for paddy fields. Conversely, the disturbed soil is currently covered by perennial plants, special-use land (e.g., construction, transportation, and residential), which is given a 7 because it seems unfit for agricultural purposes (IRMC, 2003). Likewise, the grades of acid sulfate soil at shallow layers, acid sulfate soil at deep layers, potential acid sulfate soil at deep layers, alluvial soil with various layers, and alluvial soil were classified from 6 to 2, respectively.

2.3. Methodology

2.3.1. Urban classification

Landsat images were first processed through preprocessing steps, including atmospheric correction and reflectance conversion. Digital numbers (DN) are converted to reflectance values, the ratio between reflected radiation from the Earth surface and emitted energy, and approximated using rescaling coefficients as Equation (1). Meanwhile,

the atmospheric correction removes the adverse effects of the atmosphere on satellite images. Landsat 7-ETM+ images contain line-missing data due to Scan Line Corrector (SLC) failure since 2003. In this study, Landsat 7 gaps were filled by another clear-sky scene of Landsat 5 captured on 23 November 2003 without any gap-missing errors to avoid heterogeneity owing to land cover changes.

$$\rho'_\lambda = M_\rho Q_{cal} + A_\rho \quad (1)$$

where ρ'_λ is top of atmosphere reflectance, Q_{cal} is quantized and calibrated pixel value (DN), M_ρ and A_ρ are band-specific multiplicative and additive rescaling factors, respectively.

Next, the multi-band images were analyzed using the principal component analysis (PCA) to reduce data dimensions but enhance the spectral signal. Consequently, it preserves critical information to improve classified accuracy and successfully extract land cover features (Balázs et al., 2018; Jiaju, 1988; Koutsias et al., 2009; Loughlin, 1991; Yang and Du, 2017). The input for PCA is the multiple band image, including visible bands, near-infrared (NIR), and short-wave infrared (SWIR) bands, and the output is three information-rich bands.

Though supervised algorithms (e.g., maximum likelihood and random forest) show various benefits in isolating classes for landcover mapping, classifying all types is not the core goal considered within this research. We mainly concentrated on the robust extraction of urban areas, a central object in the study. The unsupervised classification was successfully utilized in wetland detection, mangrove mapping, vegetation type classification, and crop type mapping (Cissell and Steinberg,

2019; Li et al., 2019; Mandinyenya et al., 2020; Wang et al., 2019). The unsupervised algorithm automatically groups pixels purely based on spectral information. It therefore limits subjectivity due to manual visual interpretation. Furthermore, the accuracy of unsupervised classifiers changes relatively little compared to supervised classifiers (e.g., it does not exceed 10%) (Hazir and Muda, 2020).

Meanwhile, the unsupervised classification requires no field training data (Wang et al., 2019). Therefore, an integration of PCA and unsupervised classification is an efficient approach for urban extraction. We applied an unsupervised classifier of the K-Means algorithm, which is applicable for a target classification, e.g., forest (Oduro Appiah et al., 2020; Tsai et al., 2019) and flood inundation (Borah et al., 2018), with 15 clusters preprocessed for this study. Subsequently, the classes were combined to generate a land cover map based on general land cover types, e.g., agriculture, bare soil, perennial plants, urban areas, and water bodies. This study's primary land cover type is an urban area, which was automatically classified into several classes due to imagery characteristics. For instance, the urban area includes high buildings, dense residential areas, and sparse houses that are visually dissimilar in satellite imagery, color, and texture.

Finally, the classification was evaluated using renowned coefficients of overall accuracy and kappa coefficient (Congalton and Green, 2009; FAO, 2016). In order to focus on the study's main objects and economize field survey expenses, we evaluated the classified maps on two classes of urban areas and non-urban areas (i.e., this class comprises the remaining four land cover types) (Murray et al., 2012; Nguyen et al., 2021a; Son and Thanh, 2018). The evaluating metrics were estimated by the confusion matrix, which was constructed by 200 truth points (i.e., 100 points in the urban area and 100 points in the non-urban area). The truth points were randomly collected from field surveys and very high resolution (VHR) images on Google Earth (Kaliraj et al., 2017; Nguyen et al., 2021a; Rimal et al., 2020; Zhou et al., 2016).

2.3.2. Urban expansion analysis

Due to the diversity of administrative areas between districts and urban density (UD), a normalized ratio of the urban and total area compares the urban concentration among districts (Boori et al., 2015; Makboul et al., 2015). Besides, the annual growth rate (AGR) is an indicator for evaluating urbanized speed, which measures the yearly urban gain ratio and is calculated by equation (2) (Boori et al., 2015; Xiao et al., 2006).

$$AGR = \frac{UA_{n+i} - UA_i}{nTA_{n+i}} = \frac{UD_{n+i} - UD_i}{n} \quad (2)$$

where AGR is the annual growth rate; UA_{n+i} is the urban area at the concerned time; UA_i is the urban area at the initial time; n is time interval; TA_{n+i} is the total area; UD_{n+i} and UD_i are the urban density at the concerned time and initial time, respectively.

2.3.3. Development of logistic model to estimate urbanization probability

2.3.3.1. Data preparation. Random sampling: The administrative area is about 1,400 km², which is approximately 1.56 million pixels on Landsat imagery (30-m resolution). It is considered a limited population. A theoretical sample size for this population with a margin of error ($e = 5\%$) is approximately 400 points, which is relatively less than the total area of urbanization and non-urbanization (Yamane, 1967). Extracting all considered pixels on the entire image is prone to failure because it often exceeds the hardware capacity. Therefore, this study randomly selected 2,000 points for urbanization and non-urbanization and divided them evenly by 1,000 points for each type. This number is higher than the suggested size from the Yamane equation, but it is not oversized for handling. The scattered points were numbered into two code-groups, in which 1 (one) and 0 (zero) represent urbanized and non-urbanized points, respectively.

Euclidean distance calculation and rasterization: To analyze factors affecting urbanization, raster data can be directly used for data extraction to serve further analysis. The vector data, however, needs to be transformed into secondary data. We used Euclidean distance (ED) for the effects of transportation networks, natural hydrological systems, industrial zones, town centers (district People's Committee) and already developed urban areas. More explicitly, the distance to already developed urban areas is defined as the shortest distance from a considered point to the nearest urban areas at a base time. The ED was broadly applied in GIS-related research, such as sustainability analysis and moveable time optimization (Alharbi, 2015; Alzamili et al., 2015; Rimal et al., 2019; Villacreses et al., 2017; Zhou et al., 2019). The Euclidean distance was calculated by the vector layers and converted to raster form with a pixel size of 30-m.

2.3.3.2. Data extraction. The concerned raster layers of eight considered factors (Fig. A2) were then extracted as values at sampling points using the Point Sampling Tool. The tool sequentially retrieves DN values of raster layers at the point location and adds values into the attribute table. The attribute table was then exported to the data table (*.csv) for logistic analysis.

2.3.3.3. Logistic analysis and validation. Logistic regression was utilized to explore contributing factors to urbanization in a Mekong Delta city. Foremost, the simple logistic regression on each factor was tested to detect the influence tendency. These simple logistic regressions were evaluated by probability and AIC (Akaike information criterion), which estimates prediction error and provides a standard for model selection. The low AIC value represents a better model with less prediction error. A multiple logistic regression (Equation (3)) was then applied to 70% of the dataset, while 30% of the remaining data was supplied for model validation.

$$P = \frac{e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}}{1 + e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}} \quad (3)$$

where P is the probability; e is the natural logarithm (~ 2.718); α is the interception; β is the weight of the development factor; and x is the variables in the multiple logistic regression.

The logistic model was practically validated by collating prediction values from the proposed logistic regression and testing dataset. Area Under Receiver Operating Characteristic (AUROC) is a criterion for model evaluation. The ROC curve is generated by the true positive rate (TPR) against the false positive rate (FPR). A best-fit model has AUROC approaching 1.0. The model's performance might vary between different training and test datasets. The model validation was repeated 1,000 times on different training and testing datasets to limit this bias and increase subjectivity. Mean values of AUROC were then utilized for model assessment.

3. Results and discussions

3.1. Accuracy of urban classification

The land cover maps and urban areas were obtained through an integrated PCA approach and unsupervised classification (Fig. A1), which proved the capability of this combined method in urban feature extraction to produce a highly accurate performance. All accuracy standards were higher than 90% for both years, in which the highest percentage reached 98.9% (Table 2). The Kappa coefficients of urban class in 2003 and 2017 achieved 0.89 and 0.94, respectively. The result in 2003 had lower reliability than the result in 2017 due to data missing from SLC-failure rather than seasonal differences between two Landsat images. Even though the image quality was much improved by the gap-filling task, the data quality and the existence of bare soil jointly reduces classification efficiency. The misclassification was attributed to the

Table 2
Mapping accuracy assessment of urban classification in 2003 and 2017.

Year	Classification	Reference		\sum User	User's accuracy (%)
		Urban area	Non-urban area		
2003	Urban area	99	10	109	90.83
	Non-urban area	1	90	91	98.90
	\sum Producer	100	100	200	
	Producer's accuracy (%)	99.0	90.0		
	Overall accuracy (%)				94.50
	Kappa coefficient				0.89
2017	Urban area	96	2	98	97.96
	Non-urban area	4	98	102	96.08
	\sum Producer	100	100	200	
	Producer's accuracy (%)	96.0	98.0		
	Overall accuracy (%)				97.00
	Kappa coefficient				0.94

spectral similarity between bare-soil (non-urban areas) and urban areas, especially in the peri-urban areas (Nguyen et al., 2021a). These land covers reflect more radiation and even have a high temperature in the satellite imagery. Yet the classifications generally attained a strong agreement level, with over 90% of data being reliable (McHugh, 2012).

3.2. Urbanization trends in the 2003–2017 period

The classification indicated that urban areas covered 3,639 ha (2003) and 13,044 ha (2017) corresponding to 2.5% and 9.1% of the provincial area, respectively. The increase was equivalent to AGR = 18.5% per year in the entire city. Generally, the urban area is spatially distributed in two primary hubs along the Hau river (a tributary of the Mekong river) and follows major roads (Fig. 2). The first enormous center was in the district cluster of Ninh Kieu – Binh Thuy – Cai Rang,

located in the southeastern city with a dense concentration of downtown, educational institutions, and administrative agencies. The aggregated urban area in this region accounted for 44.7% of the total urban area, but this proportion fell to 33.6% in 2017. The remaining center was mainly in the northern Thot Not district, where the Vam Cong bridge project went through. This project was aimed to link developing centers in the MKD, support the national highway project (NHP), and break the traffic monopoly of Can Tho city in connecting to western provinces. Since its master planning in 2010, the project has led to the dynamics of land transformation and the real estate market in Thot Not.

Table 3 presents parameters for conveniently evaluating urban concentration. The urban area indeed expanded in all districts, in which the most increased districts were Thot Not (2,025 ha), Co Do (1,331 ha), and Vinh Thanh (1,229 ha). These places are mostly peri-urban areas,

Table 3
Urbanization parameters of built-up area, urban density in 2003 and 2017, and annual growth rate (percent per year) over the observed period.

District	Area (Hectares)		Density (Percent)		Increase		AGR (% per year)
	2003	2017	2003	2017	Area	Density	
Thoi Lai	216	1,045	0.8	3.9	828	3.1	0.22
Phong Dien	75	579	0.6	4.6	504	4.0	0.29
Vinh Thanh	545	1,775	1.8	5.8	1,229	4.0	0.29
Co Do	329	1,660	1.0	5.2	1,331	4.2	0.30
O Mon	335	1,069	2.6	8.2	734	5.7	0.40
Cai Rang	282	1,145	4.2	16.9	864	12.8	0.91
Binh Thuy	538	1,531	7.7	21.9	993	14.2	1.01
Thot Not	513	2,537	4.3	21.1	2,025	16.8	1.20
Ninh Kieu	806	1,702	27.4	57.9	897	30.5	2.18
Entire city	3,639	13,044	2.5	9.0	9,405	6.5	0.46

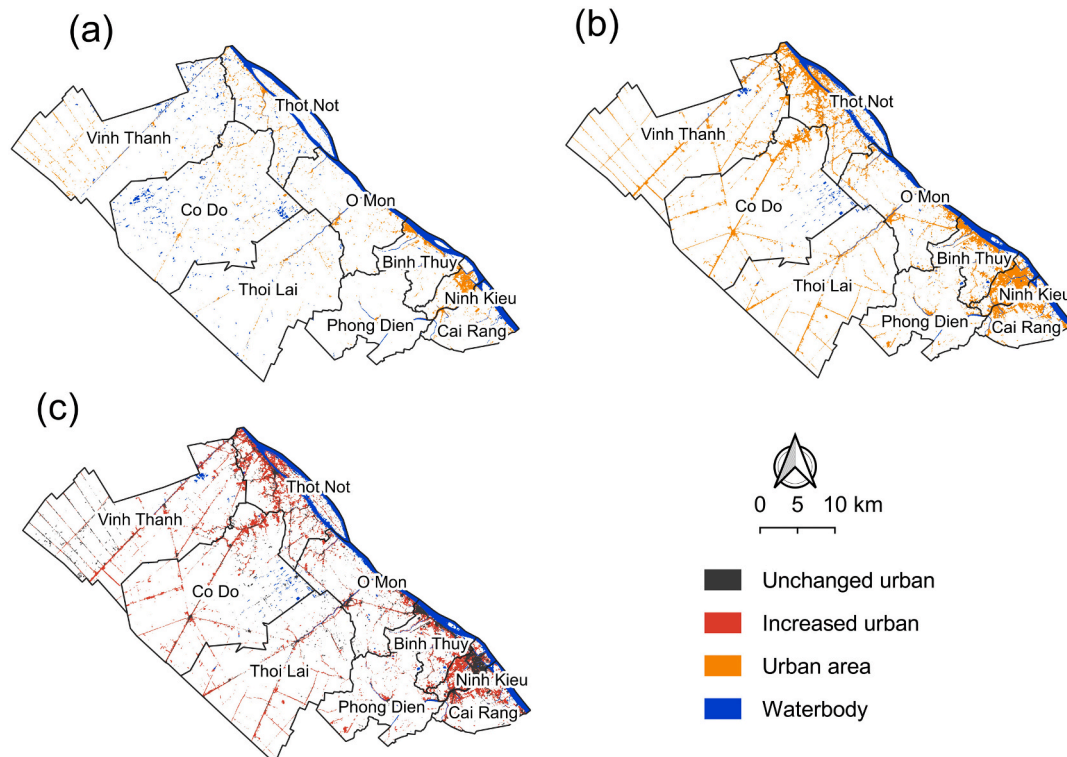


Fig. 2. Spatial distribution of urban area in Can Tho in (a) 2003, (b) 2017, and (c) urban changes during 2003–2017.

with favorable conditions for urban sprawl. Contrarily, dense urban areas in both time points were found in modern towns such as Ninh Kieu, Binh Thuy, Cai Rang, Thot Not, O Mon (only in 2003), in which Ninh Kieu was always the densest urban district with 27.4% and 57.9% in turn. The values were compared to the entire city's overall value to evaluate whether a district's indicator is low or high. Likewise, urbanized speed was assessed by using the annual rate. Ninh Kieu was still the fastest urbanization with 2.18% per year. Thot Not achieved 1.2% per year. Binh Thuy and Cai Rang yearly grew approximately 1.01% and 0.91%. The riverside counties developed faster since they are located on the main transport route connecting the southern provinces of the Hau river with the ability to trade goods easily. Conversely, the western district complex was slowly urbanized. It was less than 0.3% per year, with its immense paddy fields.

3.3. Logistic probability model

3.3.1. Descriptive statistics

Fig. 3 shows the principal differences between urbanization and non-urbanization points on eight significant elements expected to influence urbanization. The distance-related elements of urbanization were smaller than the non-urbanized values. On the contrary, the factors of elevation, population, and soil score had higher values for urbanization compared to non-urbanization. More clearly, the statistical parameters from the descriptive statistics, including mean, standard deviation (SD), and maximum (Max) values, significantly differed from the rest of the group. The minimum (Min) values of the two groups were not substantially altered for a few factors (e.g., distance to transportation, river, developed urban areas, population, and soil score). However, the distinction came from a frequency distribution, in which urbanization concentrated tightly toward the minimum value for distance factors and vice versa for non-urbanization.

3.3.2. Development of urbanization probability by multiple logistic regression

The independent effects on the urbanization of variables and their tendency (i.e., increasing or decreasing) were initially tested using

simple logistic regression and evaluated based on probability and AIC (Akaike's Information Criteria). The smaller the AIC value, the more significant the factor is. Fig. 4 shows that the most influential models were isolated contribution from distance to developed urban areas (AIC = 2,417.6), elevation (AIC = 2,477.5), distance to industrial areas (AIC = 2,536.3), and distance to transportation (AIC = 2,578.9). Conversely, the probability model of soil score presented the least impact on urbanization. Regarding the impact tendency, the distances were inversely proportional to the urbanization probability, about 70% except for distance to river, where the highest probability was approximately 60%. In other words, the highly urbanized regions are inclined to be adjacent to the old built-up area, transportation, and near the industrial areas. Elevation, population, and soil positively contributed to the probability, in which the highest urbanization probability of elevation and population was about 80%. The difference between urbanized and non-urbanized possibilities built by soil score was relatively small about 35% (non-urbanization) and 60% (urbanization). That means urbanization favorably developed in a place with high elevation, dense population, and reasonable soil type.

The isolated impacts are integrated into a multiple logistic regression for more accurately explaining the urbanized possibility. Table 4 presents the result of regression analysis, in which the best fit model has a contribution of 6 out of 8 variables. The contributors include distance to transportation, distance to developed urban areas, distance to industrial areas, elevation, population, and soil score. Distance to the river was dismissed in the target model because it has multicollinearity with road systems. Owing to an intertwined canal system, waterway has been an essential means of transportation. The roads then have been built to connect centers and remote regions, mainly following the former roads and pathways along the rivers, therefore roads and canals are often in parallel. Similarly, the urbanization ability was unexplainable by distance to the center. The significant urbanization form throughout the city was linear spreading along the streets rather than urban sprawl from a specific center, which went on mostly in the great city center at Ninh Kieu district.

With this integration, the multiple regression shows that the 6-variable model is the best fit model with $AIC = 1,398 \pm 21.58$. This value

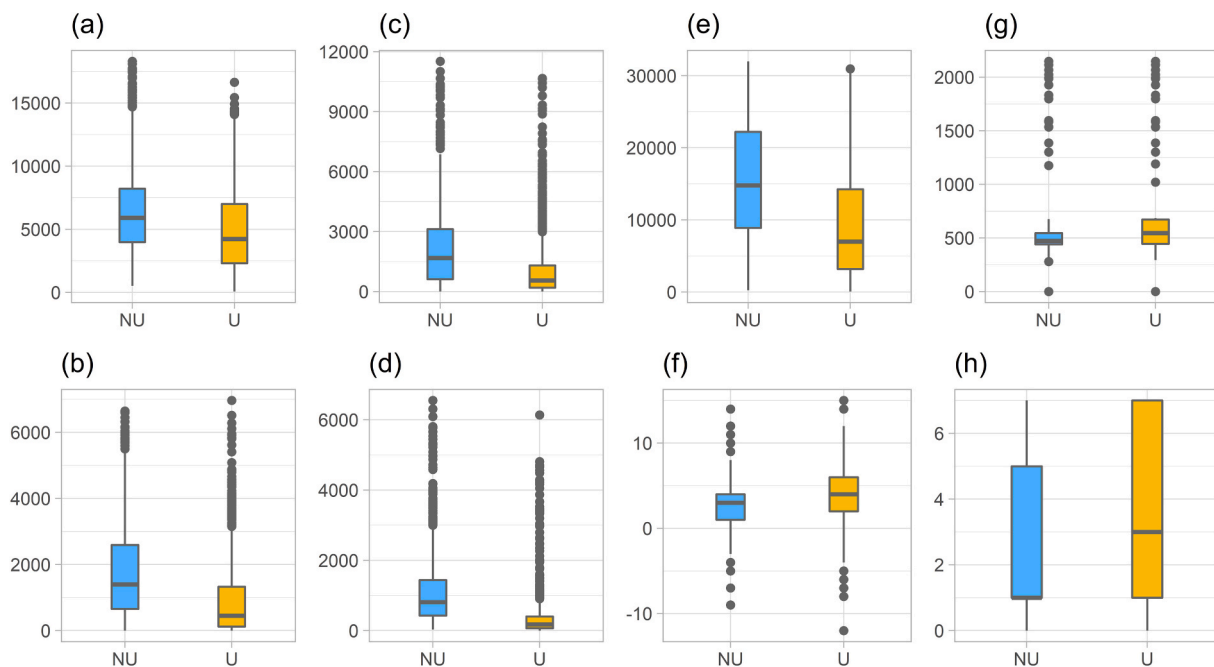


Fig. 3. Boxplot illustrates the overall differences between urbanization and non-urbanization data points in investigated factors (a) distance to center, (b) distance to transportation, (c) distance to river, (d) distance to developed urban areas, (e) distance to industrial areas, (f) elevation, (g) population, and (h) soil score. Note: NU = Non-urbanization, U = Urbanization.

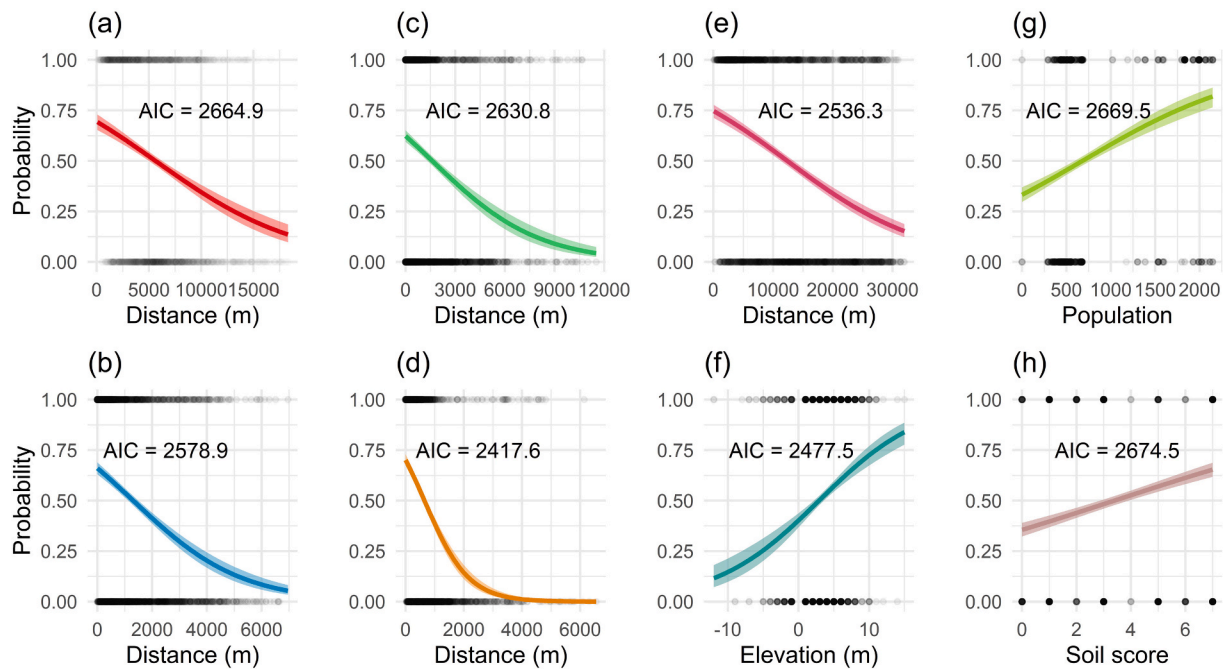


Fig. 4. Simple logistic regression between urbanization probability and considered factor, (a) distance to center, (b) distance to transportation, (c) distance to river, (d) distance to developed urban areas, (e) distance to industrial areas, (f) elevation, (g) population, and (h) soil score. Note: 0 represents non urbanization probability and 1 represents urbanization probability.

Table 4

Estimated coefficients and significance level of examined variables in the multiple logistic regression.

Considered factors	Estimated coefficient	Statistical significance
Intercept	$3,794 \times 10^{-4}$	***
Distance to town center	-0.1487×10^{-4}	ns
Distance to transportation	-2.904×10^{-4}	***
Distance to river	-80.49×10^{-4}	ns
Distance to developed urban areas	-5.346×10^{-4}	***
Distance to industrial areas	-48.03×10^{-4}	***
Elevation	890.9×10^{-4}	***
Population	3.991×10^{-4}	*
Soil score	$1,974 \times 10^{-4}$	***

Note: *** and * symbols present the statistical significance level at 0.001 and 0.01; ns is not statistically significant and eliminated from the final multiple logistic regression.

was relatively smaller against the values of separated variables. Similar to the simple logistic regression, the three distance-related variables adversely affected the urbanization expressed through negative coefficients. This regression can be interpreted as an area that will become an urban area when it has dense population density, high altitude, near an industrial area, road, previous urban area, and favorable soil type. Fig. 5 depicts AUROC lines generated from training, testing data, and all considered data points, which achieved 0.823, 0.803, and 0.818, respectively. Generally, the receiver operating characteristic (ROC) curves are smooth and AUROC > 0.8. It implies the model's efficiency at obtaining relatively high accuracy. Therefore, the obtained multiple logistic regression can be applied to estimate urbanization probability.

Fig. 6 depicts the urbanized possibility throughout Can Tho city except for the already built-up area. The most urbanized area has a probability of about 96.6% adjacent to the developed urban area. The potentially urbanized regions (higher than 80%) account for 5,994 ha

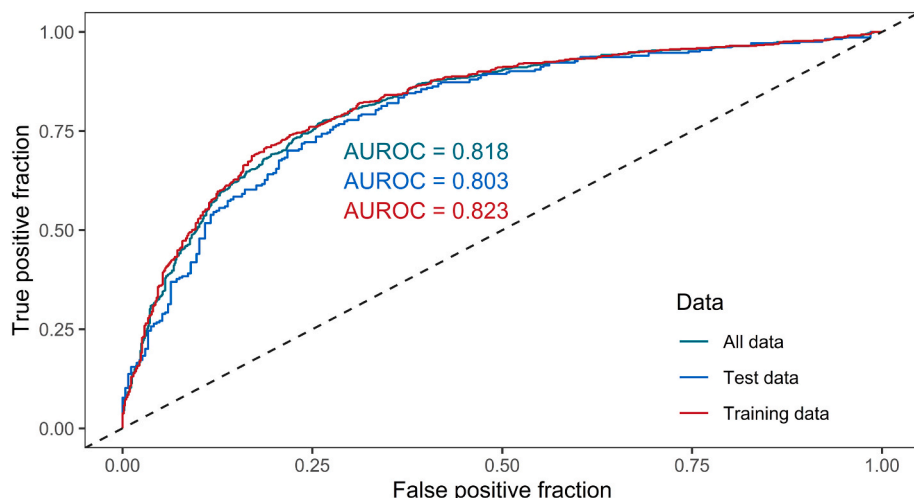


Fig. 5. A plot presenting AUROC lines from a typical training-validation data and all considered data points.

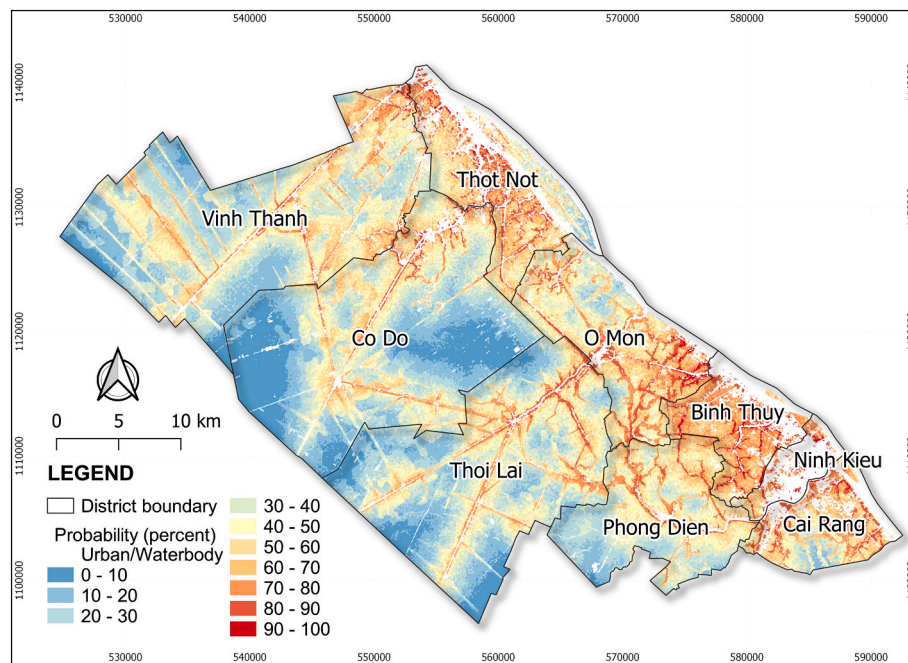


Fig. 6. Spatial distribution of urbanization probability aggregated by the multiple logistic regression and affecting factors. Urban areas and waterbodies were excluded from this map.

divided into two spatial patterns: expansion along the riverbank and inland linear spreading. Contrarily, the distant regions in Co Do and Thoi Lai districts are restricted in urbanization, at only 0.2%. The ratio between low urbanized areas is much greater than the high urbanization possibility. For instance, areas with a 20% probability are seven times higher than the potential urbanization area mentioned above.

3.4. Discussions

3.4.1. Urban expansion dynamics

The factors affecting urban expansion were investigated and divided into three main categories. These are natural conditions (elevation, soil type), social condition (population density), and infrastructures (approaching roads, industrial factories, and developed urban areas). The topographic aspects such as elevation, slope, and aspect strongly influence urban sprawl in high altitudes, plateaus, and mountainous areas (Li et al., 2018; Pijanowsld et al., 2009). The terrain of MKD is low, flat, and relatively homogeneous. It was therefore simplified by one factor called elevation. It did not much significantly affect urban distribution compared to other elements. The elevation differences, although not much significant, has a particular effect on residential area distribution. Urban development in high terrain is indeed vital since the MKD urban areas have been experiencing urban flooding. The situation will be exacerbated in the rainy season by the integrated contribution from the complete dike system upstream of the Vietnam Mekong river (An Giang province), land subsidence, sea-level rise, and heavy rain (Huong and Pathirana, 2013; Tri et al., 2013; Takagi et al., 2015, 2016). Besides, the urban areas of Can Tho city have developed under the balancing and priority of qualified soil for agricultural strength such as fruit specialties in Phong Dien district and rice in Co Do and Vinh Thanh district linking to the Long Xuyen Quadrangle (the enormous Vietnamese rice bowl). The preferred soil types for agriculture are fertile alluvial or acid sulfate soil in the deep layers. Meanwhile, disturbance soil and acid sulfate soil in shallow layers are unprofitably used for agricultural cultivation, a potential land fund for residential areas.

Population density and proximity to major roads have been broadly studied in urbanization research (Deng et al., 2008; Longyu et al., 2009) relating to urbanization on demographic aspects such as population

growth and accessibility to the essential urban infrastructures. The infrastructures, including electricity networks, water supply systems, and connecting roads, are performed at the first stage of a new residential project. Therefore, a favorable region for urban expansion has easy accessibility to these infrastructures. Besides, we additionally considered proximity to the already built-up area and industrial areas, and the results revealed a positive association. The urban area developed based on inheritance and gradually expanded outside. The neighborhood of industrial parks was examined for whether it would be more urbanized compared to further areas. The formation of industrial areas encourages general economic development and local transformation through land cover change to serve the industrial infrastructures, factories, warehouses, and internal roads. The concentrated industry project attracts numerous migrant workers. Hence, the surrounding areas rapidly develop to meet workers' needs, such as rental rooms, markets, food restaurants, and entertainment services. The economic factors (e.g., Gross Domestic Product - GDP) were not considered in this study due to data limitation, yet the proximity to industrial parks partly represents the economic growth.

3.4.2. Implication for urban planning, land assessment, and sustainable cultivation

By adopting the logistic urbanization model with six proposed variables, we will be able to predict the distribution of potential urbanization areas. Despite the low quality of variables caused by data availability, the model effectiveness was high and was expected to be more accurate with qualified data like soil maps within this study. An example of urbanization in 2025 and 2030 was achieved by simple prediction using cumulative area estimation (Fig. 7). More explicitly, the urban area will be expanded to a region holding 73.53% (18,204 ha) in 2025 and 70.67% (21,534 ha) in 2030.

Implication for urbanization prediction: the research confirmed the logistic model's capability to assess dual-result events. This will be a foundation for urbanized prediction to serve master planning and urban planning. High potential urbanization areas will be determined to develop a built-up area and low potential areas will be prioritized for agriculture after considering all potential barriers. The logistic model can also explore sustainable regions for plants (e.g., rice, rubber, coffee,

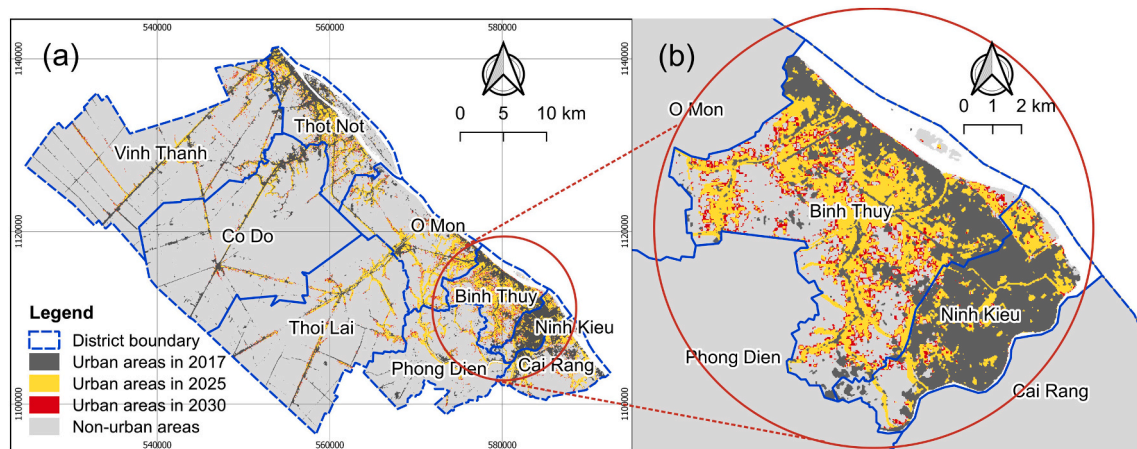


Fig. 7. Simulation of urbanization area in 2025 and 2030 adopted by the logistic probability model in (a) entire Can Tho city, (b) enlarged map of Ninh Kieu-Binh Thuy districts.

and sugar cane). Moreover, the traditional land evaluation methods identify the land elements (e.g., flooding depth, soil type, and cultivation farming) and overlap to find the land unit and suitable regions. This research framework is expected to go beyond these applications to adjust the land evaluation approach and for faster and more efficient decision making.

5. Conclusions

Our study investigated the urban expansion in 2003 and 2017 and built a probability model to simulate urbanized potential based on multiple logistic regression. The primary input data for regression is the urban area obtained by image processing, which achieved relatively high accuracies. After 14 years, the urban area increased by 2.5 times in 2003, which was about 3,639 ha. The chain of riverside counties is densely concentrated, and it developed faster than those inland districts. The yearly urbanized speed of Ninh Kieu during this period was 2.18% per year, and the urban area exceeded half of the total area at the end of the period.

Among the eight examined variables, the proximity to the center and river were not meaningful in the multiple logistic regression for explaining urbanized dynamics in the entire Can Tho city. The distance-related factors inversely contributed to the probability model and vice versa for the remaining factors. This study additionally added the distance to the previous urban and industrial areas for consideration. The concentrated industry park encouraged urbanization through a socio-economic synthesis process, including worker migration. The relatively high AUROC standard ($AUROC = 0.803$) revealed the efficiency of a six-variable probability model. The application of this model shows that the majority of the city area is medium-low urbanized. There is about 44,253 ha of low urbanization probability (lower than 30%) against 5,994 ha of high potential (greater than 80%). The research suggests a research framework for potentially adjusting traditional land evaluation habits and predicting urbanization to serve sustainable urban planning.

Funding

This project was funded by the Technological Cooperation Project “Building capacity for Can Tho University to be an excellent institution of education, scientific research and technology transfer” of Japan

Appendix

International Cooperation Agency (JICA) and the Can Tho University Improvement Project VN14-P6 funded by the Japanese ODA (No.: ODA-E8).

CRediT authors contribution statement

Nguyen Thi Hong Diep: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition. **Can Trong Nguyen:** Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Phan Kieu Diem:** Writing – original draft and Writing – review & editing. **Nguyen Xuan Hoang:** Project administration, Writing – original draft and Writing – review & editing. **Abdulla - Al Kafy:** Writing – review & editing. All authors have read and agreed to the published version of the manuscript

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors thank the Department of Environment and Natural Resources in Can Tho city for providing valuable materials and supports during the survey. We highly appreciate being funded in part by the Technological Cooperation Project “Building capacity for Can Tho University to be an excellent institution of education, scientific research and technology transfer” of JICA. We are also grateful to the Can Tho University Improvement Project VN14-P6, which was funded by the Japanese ODA loan, and to all those who provided us the possibility to complete this research through the ODA-E8 project. Our gratitude extends to Assoc. Prof. Dr. Amnat Chidthaisong and Ho Nguyen for their critical comments throughout the research completion. We owe a high debt to the project management board and colleagues at the College of Environment and Natural Resources for supporting us during the project accomplishment. Lastly, we also thank the U.S Geological Survey and all data providers for providing freely accessible satellite images and data used in this study.

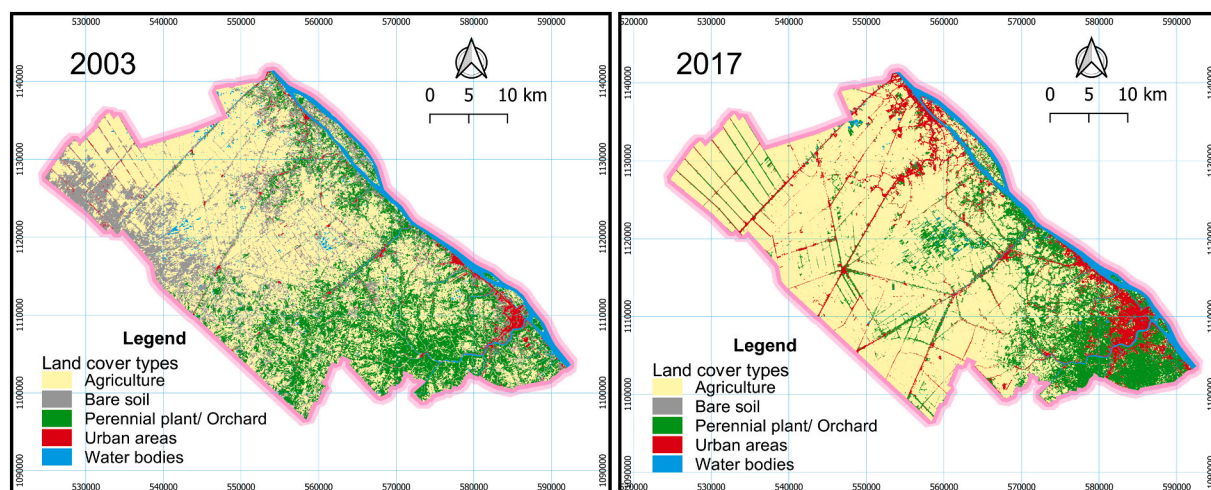


Fig. A1. Can Tho land cover maps in 2003 and 2017 obtained by Landsat imagery classification.

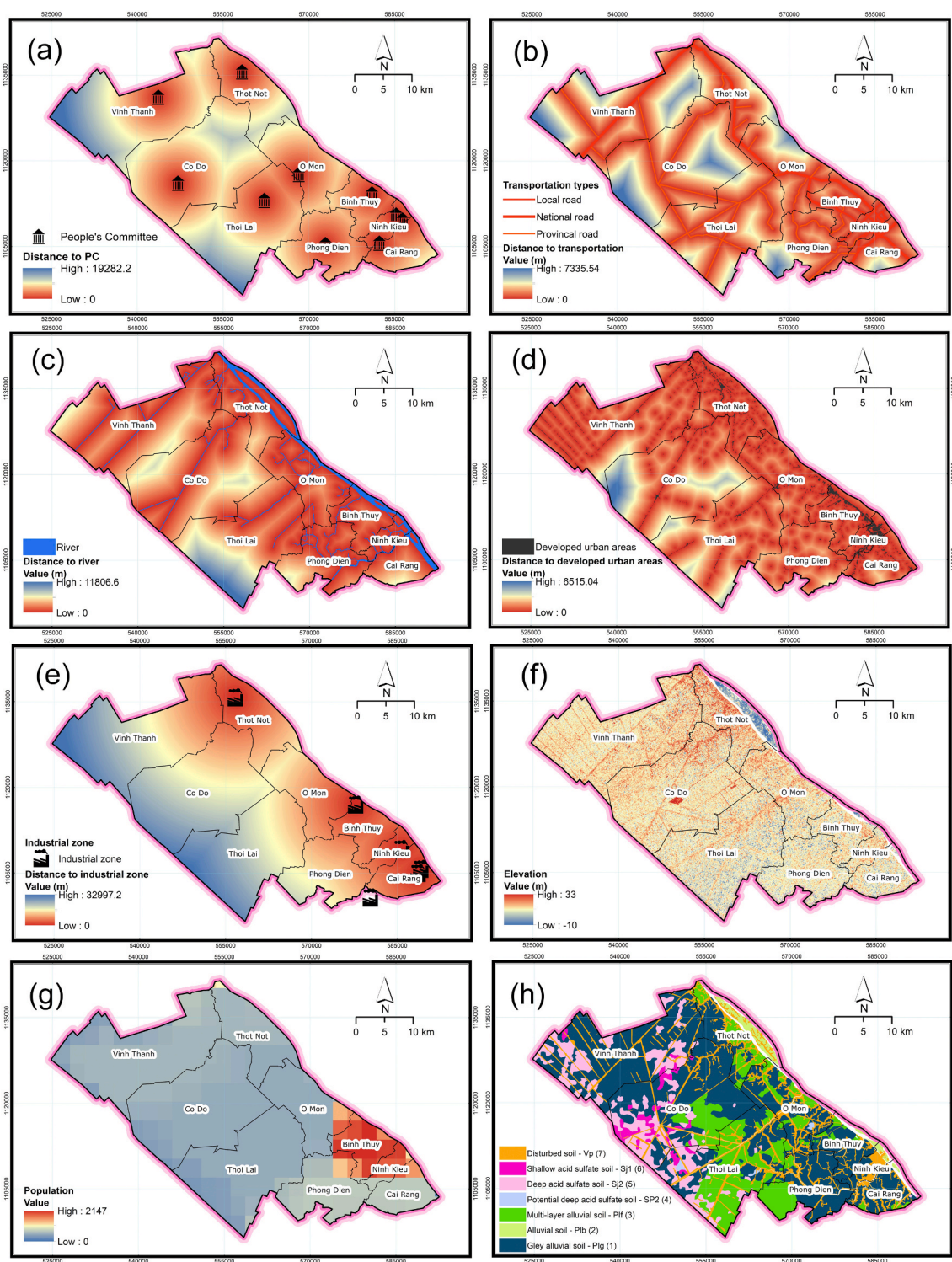


Fig. A2. Distribution maps of (a) distance to ward/district People's Committee, (b) distance to transportation, (c) distance to river, (d) distance to developed urban areas, (e) distance to industrial zones, (f) elevation, (g) population, (h) soil type and score.

References

- Alharbi, R., 2015. A GIS-based decision support system for reducing air ambulance response times: a case study on public schools in jeddah city. *J. Geogr. Inf. Syst.* 7, 384–391. <https://doi.org/10.4236/jgis.2015.74030>.
- Alzamili, H.H., El-mewafi, M., Beshr, A.M., Awad, A., 2015. GIS-Based-Multi-Criteria-Dicesion-Analysis-for-Industrial-Site-Selection-in-Al-Nasiriyah-City-in-Iraq.doc. *Int. J. Sci. Eng. Res.* 6, 1330–1337. <https://doi.org/10.13140/RG.2.1.4883.2480>.
- Andreoufouet, S., Bindschadler, R., Brown De Colstoun, E.C., Choate, M., Chomentowski, W., Christopherson, J., Doorn, B., Hall, D.K., Holifield, C., Howard, S., Kranenburg, C., Lee, S., Masek, J.B., Moran, M., Mueller-Karger, F., Ohlen, D., Palandro, D., Price, J., Qi, J., Reed, B., Samek, J., Scaramuzza, P., Skole, D., Schott, J., Storey, J., Thome, K., Torres-Pulliza, D., Vogelmann, J., Williams, D.L., Woodcock, C., Wylie, B., 2003. Preliminary assessment of the value of landsat 7 ETM+ data following scan line corrector malfunction.
- Apel, H., Martínez Trepát, O., Nghia Hung, N., Thi Chinh, D., Merz, B., Viet Dung, N., 2016. Combined fluvial and pluvial urban flood hazard analysis: concept

- development and application to Can Tho city, Mekong Delta. Vietnam. Nat. Hazards Earth Syst. Sci. 16, 941–961. <https://doi.org/10.5194/nhess-16-941-2016>.
- Balázs, B., Bíró, T., Dyke, G., Singh, S.K., Szabó, S., 2018. Extracting water-related features using reflectance data and principal component analysis of Landsat images. Hydrol. Sci. J. 63, 269–284. <https://doi.org/10.1080/02626667.2018.1425802>.
- Boori, M.S., Netzbund, M., Choudhary, K., Vozenflek, V., 2015. Monitoring and modeling of urban sprawl through remote sensing and GIS in Kuala Lumpur, Malaysia. Ecol. Processes 4, 1–10. <https://doi.org/10.1186/s13717-015-0040-2>.
- Borah, S.B., Sivasankar, T., Ramya, M.N.S., Raju, P.L.N., 2018. Flood inundation mapping and monitoring in Kaziranga National Park, Assam using Sentinel-1 SAR data. Environ. Monit. Assess. 190, 1–11. <https://doi.org/10.1007/s10661-018-6893-y>.
- Can, N.T., Diep, N.T.H., Iabchoon, S., Varnakovid, P., Minh, V.Q., 2019. Analysis of factors affecting urban heat island phenomenon in bangkok metropolitan area, Thailand. VNU J. Sci. Earth Environ. Sci. 31, 53–62. <https://doi.org/10.25073/2588-1094/vnuees.4355>.
- Chen, M., Zhang, H., Liu, W., Zhang, W., 2014. The global pattern of urbanization and economic growth: evidence from the last three decades. PLoS One 9, e103799. <https://doi.org/10.1371/journal.pone.0103799>.
- Chinh, D.T., Gain, A.K., Dung, N.V., Haase, D., Kreibich, H., 2016. Multi-variate analyses of flood loss in Can Tho city, Mekong delta. Water 8, 1–21. <https://doi.org/10.3390/w8010006>.
- Chu, X.N., Nguyen, V.T., 2017. Vietnamese urbanization: actual situation and solutions for sustainable development. Adv. Nat. Appl. Sci. 11, 41–48.
- Chu, X.N., Thi, H.D., 2017. Actual situation and solutions for reducing the traffic jams and congestion in Vietnam. Adv. Nat. Appl. Sci. 11, 26–33.
- Cissell, J.R., Steinberg, M.K., 2019. Mapping forty years of mangrove cover trends and their implications for flats fisheries in Ciénaga de Zapata, Cuba. Environ. Biol. Fish. 102, 417–427. <https://doi.org/10.1007/s10641-018-0809-0>.
- Congalton, R.G., Green, K., 2009. Assessing the accuracy of remotely sensed data: principles and practices. In: Second ed. (Ed.), The Photogrammetric Record. Taylor & Francis Group. <https://doi.org/10.1111/j.1477-9730.2010.00574.2.x>.
- CTP, 2017. Overview of can Tho city [WWW document]. CanTho portal. <https://www.cantho.gov.vn/wps/portal/> accessed 5.25.20.
- Deng, X., Huang, J., Rozelle, S., Uchida, E., 2008. Growth, population and industrialization, and urban land expansion of China. J. Urban Econ 63, 96–115. <https://doi.org/10.1016/j.jue.2006.12.006>.
- Diep, N.T.H., Biet, N., Van, Can, N.T., 2018. Urban land use mapping and greenhouse gases estimation in Can Tho city. Can Tho Univ. J. Sci 54, 30–39. <https://doi.org/10.22144/ctu.jvn.2018.036>.
- Estoque, R.C., Murayama, Y., 2017. Monitoring surface urban heat island formation in a tropical mountain city using Landsat data (1987–2015). ISPRS J. Photogrammetry Remote Sens. 133, 18–29. <https://doi.org/10.1016/j.isprsjprs.2017.09.008>.
- Estoque, R.C., Murayama, Y., Myint, S.W., 2017. Effects of landscape composition and pattern on land surface temperature: An urban heat island study in the megacities of Southeast Asia. Sci. Total Environ 577, 349–359. <https://doi.org/10.1016/j.scitotenv.2016.10.195>.
- Fan, P., Ouyang, Z., Nguyen, D.D., Nguyen, T.T.H., Park, H., Chen, J., 2019. Urbanization, economic development, environmental and social changes in transitional economies: Vietnam after Doimoi. Landsc. Urban Plan 187, 145–155. <https://doi.org/10.1016/j.landurbplan.2018.10.014>.
- FAO, 2016. Map Accuracy Assessment and Area Estimation: A Practical Guide (No. No.46/E). National forest monitoring assessment working paper, Rome.
- Gao, X., Xu, A., Liu, L., Deng, O., Zeng, M., Ling, J., Wei, Y., 2017. Understanding rural housing abandonment in China's rapid urbanization. Habitat Int. 67, 13–21. <https://doi.org/10.1016/j.habitatint.2017.06.009>.
- Glewwe, P., 2004. An overview of economic growth and household welfare in Vietnam in the 1990s. In: Glewwe, P., Agrawal, N., Dollar, D. (Eds.), Regional and Sectoral Studies Economic Growth, Poverty, and Household Welfare in Vietnam. The World Bank, Washington D.C, pp. 1–26.
- Hazir, M.H.M., Muda, T.M.T., 2020. The viability of remote sensing for extracting rubber smallholding information: a case study in Malaysia. Egypt. J. Remote Sens. Sport Sci. 23, 35–47. <https://doi.org/10.1016/j.ejrs.2018.05.001>.
- Heurlin, C., 2019. Unemployment among land-losing farmers in China: evidence from the 2010 census. J. Contemp. China 28, 434–452. <https://doi.org/10.1080/10670564.2018.1542223>.
- Hien, P.D., Men, N.T., Tan, P.M., Hangartner, M., 2020. Impact of urban expansion on the air pollution landscape: a case study of Hanoi, Vietnam. Sci. Total Environ 702 (134635). <https://doi.org/10.1016/j.scitotenv.2019.134635>.
- Huong, H.T.L., Pathirana, A., 2013. Urbanization and climate change impacts on future urban flooding in Can Tho city. Vietnam. Hydrol. Earth Syst. Sci. 17, 379–394. <https://doi.org/10.5194/hess-17-379-2013>.
- IRMC, 2003. Soil map in can Tho province. Integrated Resources Mapping Center (IRMC).
- Jiaju, L., 1988. Development of principal component analysis applied to multitemporal Landsat TM data. Int. J. Rem. Sens. 9, 1895–1907. <https://doi.org/10.1080/01431168808954988>.
- Kafy, A.A., Faisal, A. Al., Rahman, M.S., Islam, M., Rakib, A. Al., Islam, M.A., Khan, M.H. H., Sikdar, M.S., Sarker, M.H.S., Mawa, J., Sattar, G.S., 2021a. Prediction of seasonal urban thermal field variance index using machine learning algorithms in Cumilla, Bangladesh. Sustain. Cities Soc. 4, 102542 <https://doi.org/10.1016/j.scs.2020.102542>.
- Kafy, A.A., Naim, M.N.H., Subramanyam, G., Faisal, A.A., Ahmed, N.U., Rakib, A. Al., Kona, M.A., Sattar, G.S., 2021b. Cellular Automata approach in dynamic modeling of land cover changes using RapidEye images in Dhaka, Bangladesh. Environ. Challenges 64, 100084. <https://doi.org/10.1016/j.envc.2021.100084>.
- Kaliraj, S., Chandrasekar, N., Ramachandran, K.K., Srinivas, Y., Saravanan, S., 2017. Coastal landuse and land cover change and transformations of Kanyakumari coast, India using remote sensing and GIS. Egypt. J. Remote Sens. Sport Sci. 20, 169–185. <https://doi.org/10.1016/j.ejrs.2017.04.003>.
- Koutsias, N., Mallinis, G., Karteris, M., 2009. A forward/backward principal component analysis of Landsat-7 ETM+ data to enhance the spectral signal of burnt surfaces. ISPRS J. Photogramm. Rem. Sens. 64, 37–46. <https://doi.org/10.1016/j.isprsjprs.2008.06.004>.
- Leung, A., Phuong, T., Le, L., 2019. The perception of air pollution exposure from commuting in Ho Chi Minh City, Vietnam. Transp. Res. Procedia, 0–16.
- Li, G., Sun, S., Fang, C., 2018. The varying driving forces of urban expansion in China: insights from a spatial-temporal analysis. Landsc. Urban Plan 174, 63–77. <https://doi.org/10.1016/j.landurbplan.2018.03.004>.
- Li, N., Li, L., Lu, D., Zhang, Y., Wu, M., 2019. Detection of coastal wetland change in China: a case study in Hangzhou Bay. Wetl. Ecol. Manag. 27, 103–124. <https://doi.org/10.1007/s11273-018-9646-3>.
- Long, N., Van, Cheng, Y., Le, T.D.N., 2020. Flood-resilient urban design based on the indigenous landscape in the city of Can Tho, Vietnam. Urban Ecosyst. <https://doi.org/10.1007/s11252-020-00941-3>.
- Longyu, S., Guofan, S., Shenghui, C., Xuanqi, L., Tao, L., Kai, Y., Jingzhu, Z., 2009. Urban three-dimensional expansion and its driving forces - a case study of Shanghai, China. Chin. Geogr. Sci. 19, 291–298. <https://doi.org/10.1007/s11769-009-0291-x>.
- Loughlin, W.P., 1991. Principal component analysis for alteration mapping. Photogramm. Eng. Rem. Sens. 57, 1163–1169.
- Makboul, Y., Hakdaoui, M., Ghafiri, A., Elmoutaki, S., 2015. Monitoring urban evolution between 1975 and 2015 using GIS and remote sensing technics: case of Laayoune City (Morocco). Int. J. Adv. Res. 3, 331–342.
- Mandinyenya, B., Monks, N., Mundy, P.J., Sebata, A., Chirima, A., 2020. Habitat choices of African buffalo (*Syncerus caffer*) and plains zebra (*Equus quagga*) in a heterogeneous protected area. Wildl. Res. 47, 137–145. <https://doi.org/10.1071/WR18201>.
- McFarlane, C., 2019. The urbanization of the sanitation crisis: Placing Waste in the City. Dev. Change 50, 1239–1262. <https://doi.org/10.1111/dech.12533>.
- McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochem. Méd. 22, 276–282. <https://doi.org/10.11613/bm.2012.031>.
- Mohan, M., Kandya, A., 2015. Impact of urbanization and land-use/land-cover change on diurnal temperature range: a case study of tropical urban airshed of India using remote sensing data. Sci. Total Environ. 506–507, 453–465. <https://doi.org/10.1016/j.scitotenv.2014.11.006>.
- Murray, N.J., Phinn, S.R., Clemens, R.S., Roelfsema, C.M., Fuller, R.A., 2012. Continental scale mapping of tidal flats across east Asia using the landsat archive. Rem. Sens. 4, 3417–3426. <https://doi.org/10.3390/rs4113417>.
- Nguyen, C.T., Chidhaisong, A., Diem, P.K., Huo, L., 2021a. A modified bare soil index to identify bare land features during agricultural fallow-period in southeast asia using landsat 8. Land 10, 1–17. <https://doi.org/10.3390/land10030231>.
- Nguyen, C.T., Diep, N.T.H., Sanwit, L., 2021c. Direction of urban expansion in the Bangkok Metropolitan Area, Thailand under the impacts of a national strategy. Vietnam J. Earth Sci. 43 (3) <https://doi.org/10.15625/2615-9783/16313>.
- Nguyen, C.T., Nguyen, D.T.H., Phan, D.K., 2021b. Factors affecting urban electricity consumption: a case study in the Bangkok Metropolitan Area using an integrated approach of earth observation data and data analysis. Environ. Sci. Pollut. Res. 28, 12056–12066. <https://doi.org/10.1007/s11356-020-09157-6>.
- Oduro Appiah, J., Opio, C., Donnelly, S., 2020. Measuring forest change patterns from oil and gas land use dynamics in northeastern British Columbia, 1975 to 2017. Environ. Monit. Assess. 192 <https://doi.org/10.1007/s10661-019-7958-2>.
- Peng, X., Chen, X., Cheng, Y., 2010. Urbanization and its consequences. In: Zheng, Y. (Ed.), Demography. Encyclopedia of Life Support Systems. Encyclopedia of Life Support Systems (EOLSS), UK, pp. 210–235.
- Pham, T.D., Pham, V.H., Luu, Q.T., Ngo, X.T., Nguyen, T.N.T., Bui, Q.H., 2019. Analyzing the impacts of urban expansion on air pollution in Vietnam using the SEAP platform. IOP Conf. Ser. Earth Environ. Sci. 266 <https://doi.org/10.1088/1755-1315/266/1/012008>.
- Pham, T.M.T., Raghavan, V., Pawar, N.J., 2010. Urban expansion of can Tho city, Vietnam: a study based on multi-temporal satellite images. Geoinformatics 21, 147–160. <https://doi.org/10.6010/geoinformatics.21.147>.
- Phan, C.C., Nguyen, T.Q.H., Nguyen, M.K., Park, K.H., Bae, G.N., Seung-bok, L., Bach, Q. V., 2020. Aerosol mass and major composition characterization of ambient air in Ho Chi Minh City, Vietnam. Int. J. Environ. Sci. Technol. <https://doi.org/10.1007/s13762-020-02640-0>.
- Pijanowsld, B.C., Tayyebi, A., Delavar, M.R., Yazdanpanah, M.J., 2009. Urban expansion simulation using geospatial information system and artificial neural networks. Int. J. Environ. Res. 3, 495–502.
- Ponte, E.V., Cruz, A.A., Athanasio, R., Carvalho-Pinto, R., Fernandes, F.L.A., Barreto, M. L., Stelmach, R., 2018. Urbanization is associated with increased asthma morbidity and mortality in Brazil. Clin. Res. J. 12, 410–417. <https://doi.org/10.1111/crj.12530>.
- Rimal, B., Keshkar, H., Sharma, R., Stork, N., Rijal, S., Kunwar, R., 2019. Simulating urban expansion in a rapidly changing landscape in eastern Tarai. Nepal. Environ. Monit. Assess 191. <https://doi.org/10.1007/s10661-019-7389-0>.
- Rimal, B., Rijal, S., Kunwar, R., 2020. Comparing support vector machines and maximum likelihood classifiers for mapping of urbanization. J. Indian Soc. Remote Sens 48, 71–79. <https://doi.org/10.1007/s12524-019-01056-9>.
- Sheng, Y.K., 2017. Economic development and poverty reduction in southeast asia. In: Development and Poverty: China and the World. Guangzhou.

- Sheng, Y.K., Mohit, R.S., 2001. Employment and migration in the urban future of southeast asia. In: Virchow, D., Braun, J., von (Eds.), *Villages in the Future: Crops, Jobs and Livelihood*. Springer, Verlag Berlin Heidelberg, pp. 306–377.
- Shouhai, D., 2015. Employment in township urbanization in China. *Soc. Sci. China* 36, 152–167. <https://doi.org/10.1080/02529203.2015.1029675>.
- Siddiqui, Asfa, Siddiqui, Almas, Maithani, S., Jha, A.K., Kumar, P., Srivastav, S.K., 2018. Urban growth dynamics of an Indian metropolitan using CA Markov and logistic regression. *Egypt. J. Remote sens. Sport Sci.* 21, 229–236. <https://doi.org/10.1016/j.ejrs.2017.11.006>.
- Sintusingha, S., 2011. Bangkok's urban evolution: challenges and opportunities for urban sustainability. In: Sorensen, A., Okata, J. (Eds.), *Megacities. Urban Form, Governance, and Sustainability*. Springer, Tokyo, pp. 133–161. https://doi.org/10.1007/978-4-431-99267-7_7.
- Son, N.T., Chen, Chi-farn, Chen, Cheng-ru, Thanh, B., Vuong, T.H., 2017. Assessment of urbanization and urban heat islands in Ho Chi Minh City, Vietnam using Landsat data. *Sustain. Cities Soc.* 30, 150–161. <https://doi.org/10.1016/j.scs.2017.01.009>.
- Son, N.T., Thanh, B.X., 2018. Decadal assessment of urban sprawl and its effects on local temperature using Landsat data in Cantho city. *Vietnam. Sustain. Cities Soc.* 36, 81–91. <https://doi.org/10.1016/j.scs.2017.10.010>.
- Takagi, H., Thao, N.D., Anh, L.T., 2016. Sea-level rise and land subsidence: impacts on flood projections for the Mekong Delta's largest city. *Sustainability* 8, 1–15. <https://doi.org/10.3390/su8090959>.
- Takagi, H., Ty, T.V., Thao, N.D., Esteban, M., 2015. Ocean tides and the influence of sea-level rise on floods in urban areas of the Mekong Delta. *J. Flood Risk Manag* 8, 292–300. <https://doi.org/10.1111/jfr3.12094>.
- Tran, D.X., Pla, F., Latorre-Carmona, P., Myint, S.W., Caetano, M., Kieu, H.V., 2017. Characterizing the relationship between land use land cover change and land surface temperature. *ISPRS J. Photogramm. Remote Sens* 124, 119–132. <https://doi.org/10.1016/j.isprsjprs.2017.01.001>.
- Tri, V.P.D., Trung, N.H., Thanh, V.Q., 2013. Vulnerability to flood in the Vietnamese Mekong delta: mapping and uncertainty assessment. *Former. Part J. Environ. Sci. Eng.* 2, 229–237.
- Trung, L.V., Vu, N.N., 2018. Application of remote sensing and GIS for assessing the urbanization trend in Can Tho city. *Sci. Technol. Dev. J. Sci. Earth Environ* 2, 57–62.
- Tsai, Y.H., Stow, D.A., López-Carr, D., Weeks, J.R., Clarke, K.C., Mensah, F., 2019. Monitoring forest cover change within different reserve types in southern Ghana. *Environ. Monit. Assess.* 191 <https://doi.org/10.1007/s10661-019-7450-z>.
- Turok, I., 2016. Getting urbanization to work in Africa: the role of the urban land-infrastructure-finance nexus. *Area Dev. Policy* 1, 30–47. <https://doi.org/10.1080/23792949.2016.1166444>.
- United Nations, 2015. World Urbanization Prospects, UNDESA. Department of Economic and Social Affairs. Population Department, United Nations, New York. <https://doi.org/10.4054/DemRes.2005.12.9>.
- US EPA, 2008. Urban Heat Island Basics. In: *Urban Heat Islands, Reducing (Ed.), Compendium of Strategies*. U.S. Environmental Protection Agency, Washington, DC.
- Van Long, N., Cheng, Y., 2018. Urban landscape design adaption to flood risk: a case study in can Tho city, Vietnam. *Environ. Urban. Asia* 9, 138–157. <https://doi.org/10.1177/0975425318783587>.
- Villacreses, G., Gaona, G., Martínez-Gómez, J., Jijón, D.J., 2017. Wind farms suitability location using geographical information system (GIS), based on multi-criteria decision making (MCDM) methods: the case of continental Ecuador. *Renew. Energy* 109, 275–286. <https://doi.org/10.1016/j.renene.2017.03.041>.
- Wang, S., Azzari, G., Lobell, D.B., 2019. Crop type mapping without field-level labels: random forest transfer and unsupervised clustering techniques. *Remote Sens. Environ. Times* 222, 303–317. <https://doi.org/10.1016/j.rse.2018.12.026>.
- Wu, W., Zhao, S., Henebry, G.M., 2019. Drivers of urban expansion over the past three decades: a comparative study of Beijing, Tianjin, and Shijiazhuang. *Environ. Monit. Assess.* 191 <https://doi.org/10.1007/s10661-018-7151-z>.
- Xiao, J., Shen, Y., Ge, J., Tateishi, R., Tang, C., Liang, Y., Huang, Z., 2006. Evaluating urban expansion and land use change in Shijiazhuang, China, by using GIS and remote sensing. *Landsc. Urban Plann.* 75, 69–80. <https://doi.org/10.1016/j.landurbplan.2004.12.005>.
- Xinmin, Z., Estoque, R.C., Murayama, Y., 2017. An urban heat island study in Nanchang City, China based on land surface temperature and social-ecological variables. *Sustain. Cities Soc.* 32, 557–568. <https://doi.org/10.1016/j.scs.2017.05.005>.
- Yamane, T., 1967. *Statistics, an Introductory Analysis*, Second. ed. Harper and Row, New York.
- Yang, J., Du, X., 2017. An enhanced water index in extracting water bodies from Landsat TM imagery. *Spatial Sci.* 23, 141–148. <https://doi.org/10.1080/19475683.2017.1340339>.
- Zhou, D., Lin, Z., Lim, S.H., 2019. Spatial characteristics and risk factor identification for land use spatial conflicts in a rapid urbanization region in China. *Environ. Monit. Assess.* 191, 1–22. <https://doi.org/10.1007/s10661-019-7809-1>.
- Zhou, D., Zhao, S., Zhang, L., Liu, S., 2016. Remotely sensed assessment of urbanization effects on vegetation phenology in China's 32 major cities. *Remote Sens. Environ. Times* 176, 272–281. <https://doi.org/10.1016/j.rse.2016.02.010>.