Ecological Questions 35(2024)3 http://dx.doi.org/10.12775/EQ.2024.030

Mapping impervious surface change from remote sensing and GIS data: A case study in Hochiminh city, Vietnam

Le Hung Trinh[1,*], Van Tung Pham[2], Xuan Bien Tran[3], Van Trung Nguyen[4], Xuan Cuong Vu[2], Thi Hanh Tong[1], Van Phu Le[1]

[1]Le Quy Don Technical University, Hanoi, Vietnam
[2]Ho Chi Minh University of Natural Resources and Environment, Hochiminh city, Vietnam
[3]Hanoi University of Natural Resources and Environmental, Hanoi, Vietnam
[4]Hanoi University of Mining and Geology, Hanoi, Vietnam

*Corresponding author e-mail address: trinhlehung@lqdtu.edu.vn

Received: 25 June 2023 / Accepted: 22 December 2023

Abstract. Impervious surface is artificial surfaces that prevent water from entering the soil. The increase in impervious surface area has led to negative impacts on the urban environment, including an increase in the risk of flooding, a decrease in vegetation cover, and the formation of urban heat islands. This paper presents the results of building a predictive model of impervious surfaces in Hochiminh city from remote sensing and GIS data. Landsat and Sentinel 2 satellite images for the period 2002–2022 are used to classify impervious surfaces and extract input layers about vegetation cover, land surface temperature, combined with GIS data (elevation, slope, aspect, distance to road, distance to hydrology, population density) for modeling and predicting impervious surface changes in future. 03 machine learning algorithms, including Support Vector Machine (SVM), Random Forest (RF), Classification and Regression Trees (CART) and maximum likelihood method are used to classify impervious surfaces from Landsat satellite images, then select the method with the highest accuracy. To predict the future distribution of impervious surface, this study uses Cellular Automata (CA) model and 02 artificial intelligence algorithms (Artificial Neural Network - ANN, Logistic Regression - LR). The results obtained in the study can be effectively used for urban planning, minimizing the impact of the process of increasing the impervious surface on the urban environment.

Keywords: impervious surface, remote sensing, GIS, machine learning, Cellular Automata, Hochiminh city.

1. Introduction

Impervious surfaces are man-made surfaces, including surfaces that prevent water from entering the soil, such as roads, sidewalks, parking lots, roofs and so on (Qiao et al., 2018). In recent years, impervious surfaces have emerged as not only an indicator of the degree of urbanization but also a key indicator of urban environmental quality. An increase in the impervious surface will lead to an increase in the size, duration and intensity of urban runoff. Increasing the impervious surface area will impact and pollute water sources, including pathogens, and toxic substances that contaminate surface water and groundwater. In addition, this increase is also the cause of a decrease in the area of vegetation in urban areas. The appearance of a high spatial density of impervious surfaces can significantly affect urban climate by altering the suitability of heat flows and the potential risk of increasing urban temperature leading to a heat island phenomenon in urban areas. Therefore, information on spatial distribution maps of impervious surface areas is really necessary for designing, planning, managing and protecting urban environmental resources (Liu et al., 2020; Xu et al., 2022).

The satellite remote sensing technique has become a superior method in monitoring and mapping the distribution and estimation of impervious surface areas due to its multi-temporal, multi-spectral, multi-data source and large study area. Researchers around the world and the domestic have proven the success of using remote sensing as an effective tool for extracting information about the characteristics, distribution and changes of impervious surfaces in urban areas. In particular, satellite image data can allow to determine the changes of impervious surfaces in periods from the past to the present in a systematic and highly uniform manner (Sati & Mohan, 2018; Gong et al., 2020; Liu et al., 2020; Yin et al., 2021). In addition, GIS technology with strong spatial data analysis capabilities allows us to quantify, analyze, and model the change of impervious surfaces to give a trend and can determine the rate of change of these factors and make a predict about the development picture of large urban areas.

Based on the results of the impervious surface analysis, many studies around the world have conducted modeling and predicting the future distribution of impervious surfaces. Initial studies often estimate future impervious surface distributions based on the prediction models of land use changes, in which the most common method is to evaluate the weighted conversion to the impervious surface of land use types (Brabec et al., 2002). In a study in California (USA), Washburn reckons that developing a commercial area for retail would lead to 86% of the land cover becoming an impervious surface after construction is complete (Washburn et al., 2010). Some other studies also used the results of predicting land use changes using Markov and CA-Markov models to evaluate the distribution of impervious surfaces (Khawaldah, 2016; Trinh et al., 2017; Misagh et al., 2018; Chaula, 2019; Asori & Adu, 2023). This approach provides a simple method for predicting the future impervious surface distribution. However, these models also have a fundamental disadvantage as the predict results are only effective and highly accurate when the land use scenarios are fully implemented.

Several recent studies have developed regression models between the percentage of impervious surfaces and the population density. Variables such as population growth and trade are also included in some predicting models to improve the accuracy of predictive results of impervious surface distribution (Azimand et al., 2020; Li et al., 2021; Ramezani et al., 2021). Machine learning techniques such as SVM, RF and so on are often used when extracting impervious surfaces from remote sensing data, while artificial neural networks (ANN) and regression techniques are often applied to predictive models of future impervious surface distributions (Mahyoub et al., 2022).

Besides multi-temporal remote sensing data, additional information layers such as cover, surface temperature, topography, distance to roads, and so forth are also used for modeling urban growth trends. The Cellular Automata (CA) mathematical model has been widely used in these studies based on the combination of machine learning techniques (Gharaibeh et al., 2020; Saputra & Li, 2019; Bugday, 2019).

This paper presents modeling and predicting results of impervious surface changes in Hochiminh City from remote sensing and GIS data. Landsat satellite images for the years 2002 and 2012, combined with additional information layers were used to predict impervious surface fluctuations for 2022, then compare with the impervious surface classification results in 2022 to evaluate the accuracy and calibrate the model. Machine learning techniques including Random Forest, Support Vector Machine (SVM), Classification and Regression Tree (CART) and the traditional maximum probability classifier have been tested in the classification of impervious surfaces, from which the method with the highest accuracy was selected. To predict the impermeability surface variation, this study experimented with 02 methods: logistic regression (LR) and artificial neural network (ANN), then selected the method with higher accuracy to predict the distribution of impervious surfaces for 2032. Data processing was performed on the MOLUSCE module of QGIS 2.18 software.

2. Materials and methods

2.1. Study area description

The selected research area is Hochiminh city, one of the two largest cities in Vietnam. Hochiminh city has a natural area of 2,095 km^2, and a population of 8.96 millions (according to the 2019 census results). As of January 2023, according to World Population Review, the population of Hochiminh City is 9,320,866 peoples. In fact, the population of Hochiminh City reaches nearly 13 million people, which includes about 3 millions immigrants. The population growth rate in the city is about 1 million people every 5 years, leading to a high urbanization rate.

Located in the southern key economic region, Hochiminh city is the largest economic centre of the country, has a high economic growth rate and contributes greatly to the country's GDP (about one-third of the country's total GDP). The city has a diversified and modern infrastructure and is an important transport hub of the country, including a system of highways, seaports, airports, railways, etc. It can be said that the city plays a very important role, as an economic, financial, commercial and service centre, as a driving force for the socio-economic development of the country. The geographical location of the study area is shown in Figure 1.

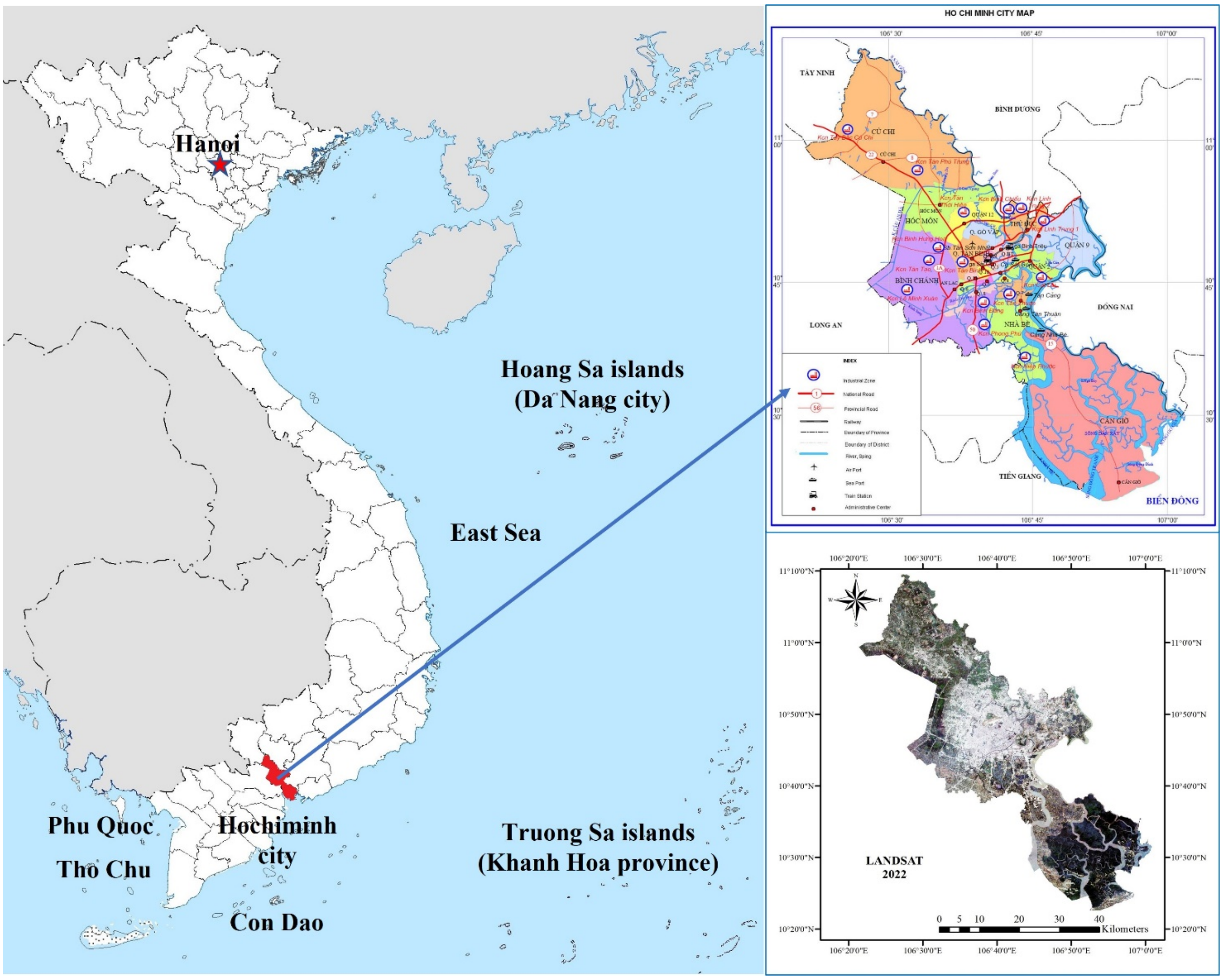


Figure 1. Geographical location of Hochiminh city

2.2. Materials

Remote sensing data used in the study are Landsat satellite images taken in Hochiminh city on March 10, 2002; March 5, 2012; and February 28, 2022. The photos are all collected in the dry season, have good quality and are not affected by weather conditions. Images were collected at the L2A processing level, so in this study, only geometric correction and clipping were carried out according to the boundary of the experimental area.

Based on the analysis of natural and social characteristics and data status in the Hochiminh city area, as well as on the results obtained in previous studies, in this article, we collect and build an additional input data set of the impervious surface volatility prediction model with 08 data layers, including (1) Vegetative cover density; (2) land surface temperature; (3) Elevation; (4) Slope; (5) Aspect; (6) Population density; (7) Distance to the road; (8) Distance to hydrology.

In these information layers, vegetation cover density and surface temperature were extracted from remote sensing data (Sentinel 2 MSI and Landsat 8/9 optical satellite images). Slope, aspect, and elevation are extracted through GIS data based on using the digital elevation model DEM. The information layers about the distance to the road and the distance to the river are built based on the Euclidean Distance module integrated into ArcGIS software based on vector data layers on traffic and hydrology in the study area.

Traffic and hydrological data in the study area are mined from the Open Development Mekong Database (https://data.opendevelopmentmekong.net). Map data to extract hydrological and traffic information was obtained from the website International Steering Committee for Global Mapping version 2 (V.2) published by the Department of Survey, Mapping and GeoInformation of Viet Nam - Ministry of Natural Resources and Environment of Vietnam. The data has some additional attributes and is assigned the Vietnamese national standard object code according to the national basic geographic list standard promulgated by the Ministry of Natural Resources and Environment in 2012.

2.3. Methodology

Based on studying the characteristics of the experimental area, experiences from studies in the world and Vietnam, as well as the current status of data sources, the article proposes the selection of a model to predict the fluctuations of the impervious surface in the Hochiminh city from remote sensing and GIS data based on the Cellular Automata model combined with machine learning techniques as shown in Figure 2.

Step 1: Survey the study area, collect data

In this step, after researching, surveying and analyzing the characteristics of the test area as well as the current state of the data source, the input data includes satellite images (Landsat images, Sentinel 2 MSI), digital elevation model (DEM) and other data layers are collected to build input information layers for the impervious surface variation prediction model.

Step 2: Build additional data layers

With 8 additional data layers of the impervious surface variation prediction model from remote sensing and GIS data are extracted and built, specifically as follows :

+ The information layer about vegetation cover density is extracted from optical satellite images Sentinel 2 MSI through NDVI vegetation index.

+ The surface temperature information layer is extracted from Landsat 8 satellite image.

+ Terrain information layers, including elevation, slope, and aspect are extracted from SRTM digital elevation model (DEM).

+ The population density information layer is built on the WorldPop database.

All these input layers are built in a raster data structure, then interpolated to the same resolution (30m pixel size) to synchronize the input data layers of the impervious surface variation prediction model as well as to fit the base dataset (Landsat image has a spatial resolution of 30 m).

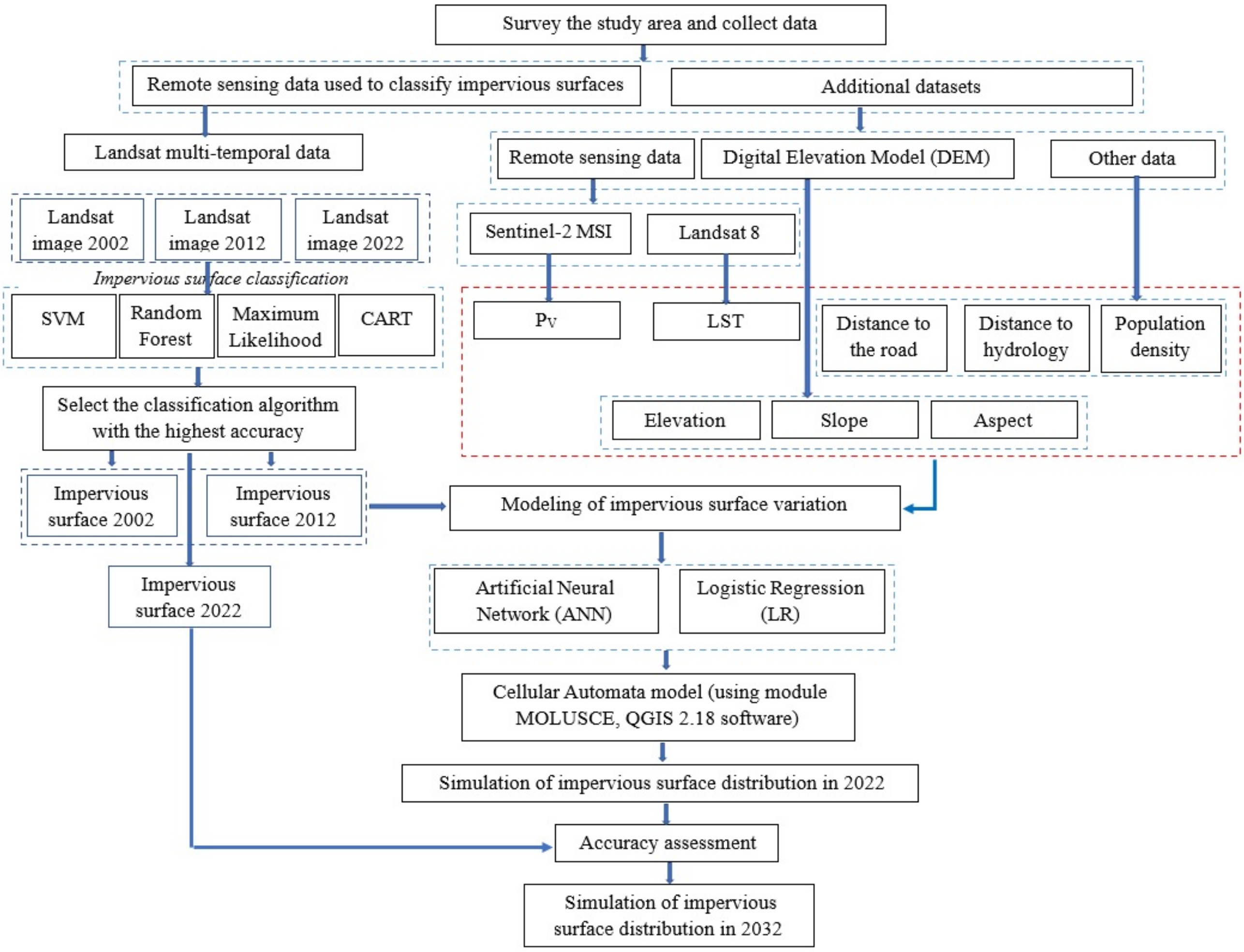


Figure 2. The process of predicting the variation of impervious surface in Hochiminh City

Step 3: Classify impervious surfaces from multi-temporal Landsat satellite images

This study conducted experiments with the maximum probability algorithm - This algorithm is considered as the standard classification method when classifying based on pixels and 03 popular machine learning algorithms (RF, SVM, CART) to classify the impervious surface of Hochiminh City area from multi-temporal Landsat satellite image data, thereby choosing a method suitable for experimental area conditions.

Step 4: Prediction of impervious surface variation using the CA model and machine learning techniques

The results of the classification of impervious surfaces at time 1 (2002) and time 2 (2012) and additional data sets are used to model and predict impervious surface variation based on the CA model incorporating machine learning techniques. Experimental research with 02 machine learning techniques that are commonly used in the predicting models of land cover/land use trends (ANN, Logistic regression) to predict the distribution of impervious surfaces in future (time 3, 2022).

Step 5: Evaluate the accuracy and choose the appropriate model

The prediction results of the impervious surface distribution are compared with the results of the impervious surface classification at time 3 (in 2022) to evaluate the accuracy and select the appropriate model.

Step 6: Building a predict map of impervious surface changes

After selecting the most appropriate and accurate model, the study proceeds to predict the future distribution of impervious surfaces. The interval between the selected impervious surface volatility predicting periods is 10 years. From the results of the impervious surface classification in 2002 and 2012, combined with additional data sets, predict for 2022 and compare with the classification results in 2022 to complete the model.

3. Results and discussion

Landsat satellite images of Hochiminh city area in 2002, 2012 and 2022 after preprocessing and cutting along the boundary of the study area are shown in Figure 3 (natural colour combinations).

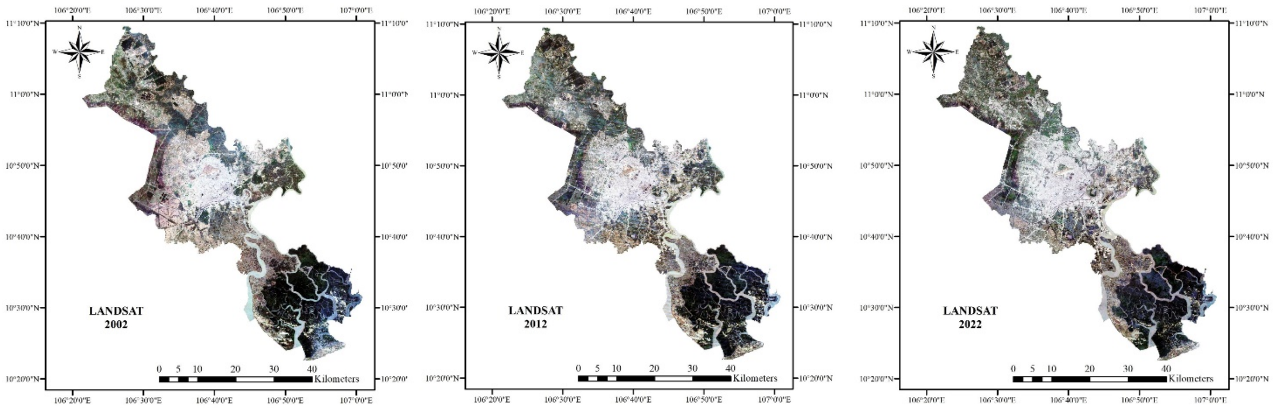


Figure 3. Landsat image data of the study area

The results of building additional information layers (08 data layers) in the study area are presented in Figure 4. These information layers are also edited according to the boundary of Hochiminh city area and interpolated to a spatial resolution of 30m to be consistent with the spatial resolution of Landsat satellite images.

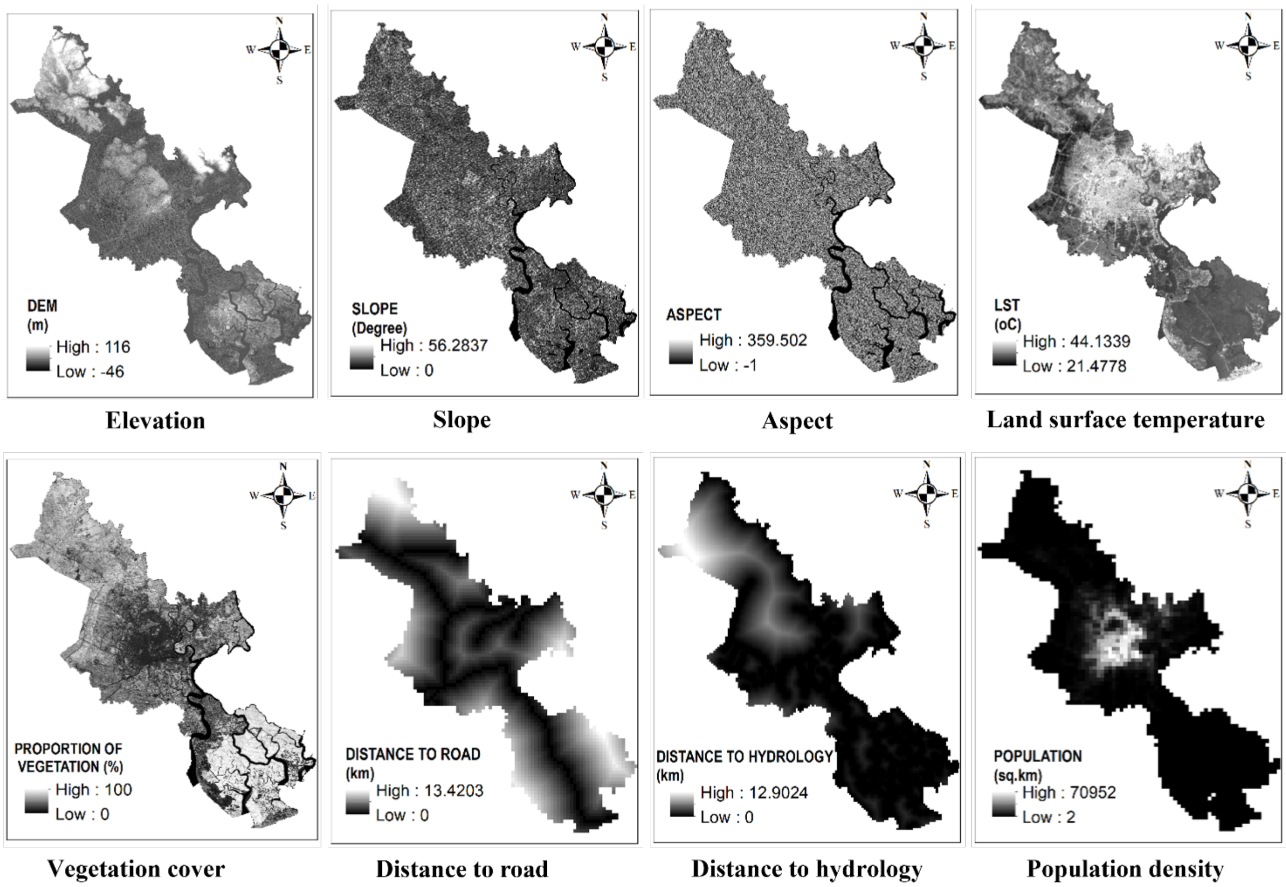


Figure 4. Supplementary dataset (8 layers) of impervious surface variability prediction model

The surface coating classification results, including the impervious surface in Hochiminh city in 2002, 2012 and 2022 are shown in Figure 5, in which the surface coating is classified into 6 classes: Water body, Impervious surface, Shrubland, Agricultural land, Bare land and Forest.

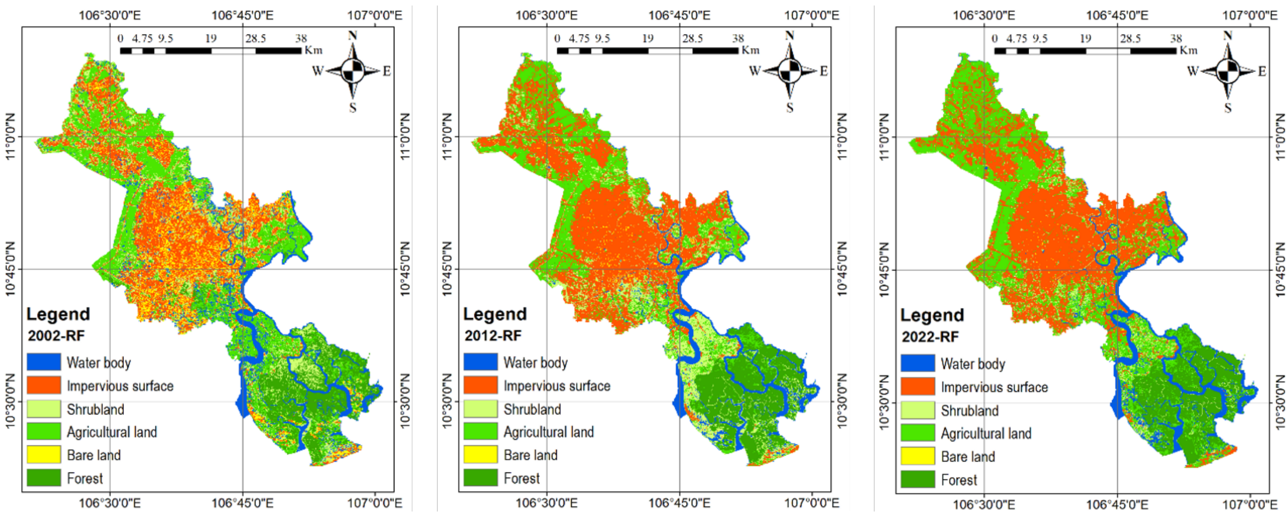


Figure 5. The results of the classification of land cover/land use in Hochiminh city in 2002, 2012 and 2022 by RF algorithm

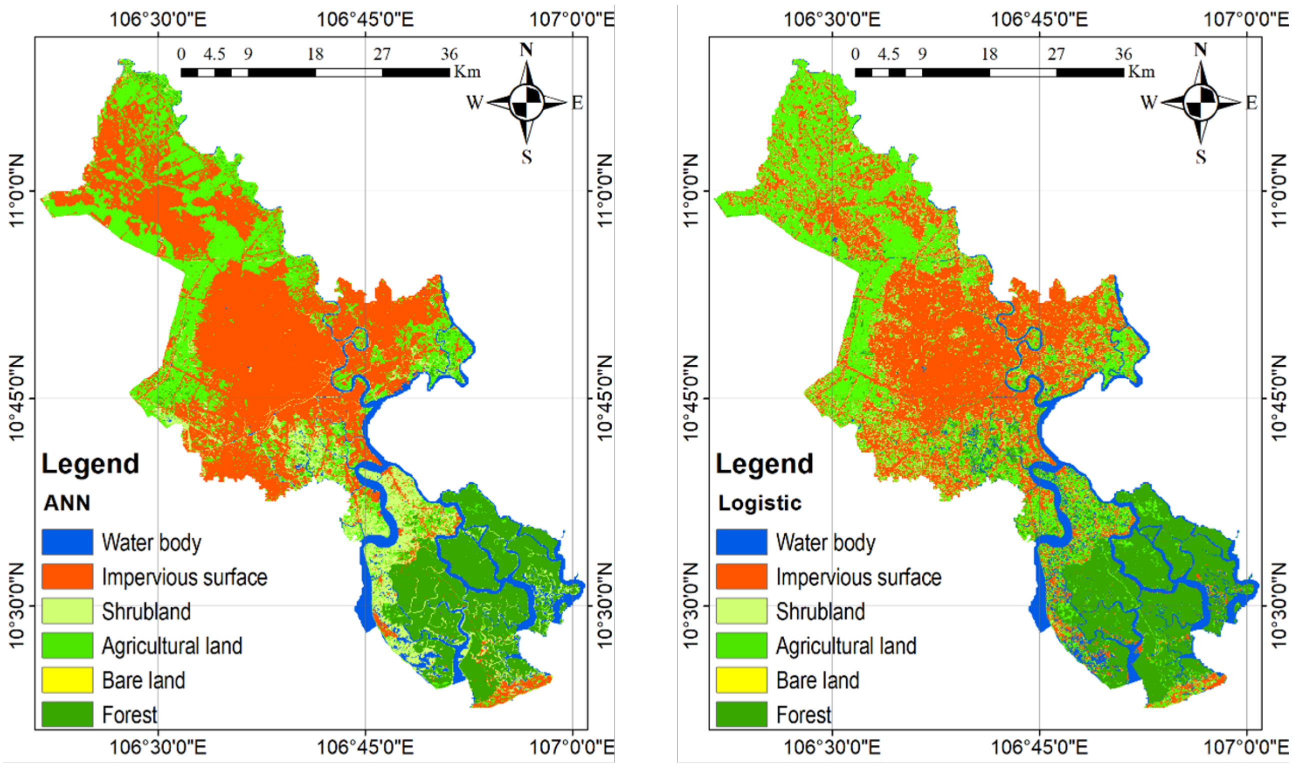
Based on the results of surface coating classification in Hochiminh city area by 04 methods: maximum probability, RF, SVM and CART, in the study, the overall accuracy and the Kappa index were determined, then select the classification method with the highest accuracy. From the results in Table 1, it can be seen that the RF classification method gives the highest accuracy, shown both in the overall accuracy parameter and the Kappa index in the surface cover classification of Hochiminh city for all 3 years: 2002, 2012, and 2022. Therefore, this study chooses to use the RF method to classify surface coatings in Hochiminh city area, as a basis for building a predictive model of impervious surface fluctuations.

Table 1. Overall accuracy and Kappa index of classification algorithms for Landsat images of the study area in 2002.

Year	**Classification accuracy**	**SVM**	**RF**	**CART**	**MD**
2002	Overall accuracy	90.16%	92.56%	89.13%	88.33%
	Kappa index	0.881	0.909	0.867	0.858
2012	Overall accuracy	89.69%	92.61%	88.80%	87.01%
	Kappa index	0.874	0.906	0.863	0.841
2022	Overall accuracy	91.26%	93.02%	87.65%	88.10%
	Kappa index	0.892	0.915	0.858	0.847

To model the development process of urban surfaces, this paper uses the Cellular Automata (CA) mathematical model that combines two machine learning techniques: ANN networks and Logistic regression. First of all, for the ANN algorithm, we set up a set of parameters including neighbourhood (1 px), learning rate (0.001), number of loops (1000), hidden layer (12) and inertia (0.05). After training, the data learning process reaches the Kappa coefficient value of 0.768 with the validation dataset. With the Logistic regression algorithm, the set of input parameters is set including the number of samples (2000), neighbourhood (1 px) and maximum number of loops (100).

The results of modeling urban growth in Hochiminh City in 2022 from Landsat satellite images in 2002 and 2012, combined with additional data sets using ANN algorithm and Logistic regression are presented in Figure 6.



Prediction by ANN algorithm **Prediction by Logistic regression algorithm**

Figure 6. Prediction results of impervious surface distribution in 2022 by ANN and Logistic regression algorithms

In order to evaluate the accuracy of models to predict urban surface changes in Hochiminh city in 2022, the study uses the results of land cover/land use classification from Landsat images in 2022 by RF algorithm for comparison. The obtained results show that the value of the Kappa coefficient reaches 0.7888 for the ANN algorithm and 0.7471 for the Logistic regression algorithm, respectively. Thus, the ANN algorithm has a higher accuracy than the Logistic regression when predicting the variation of the urban cover in Hochiminh city. From this result, this article uses the ANN algorithm to build a predict map of the impervious surface changes in Hochiminh city in 2032 (Fig. 7).

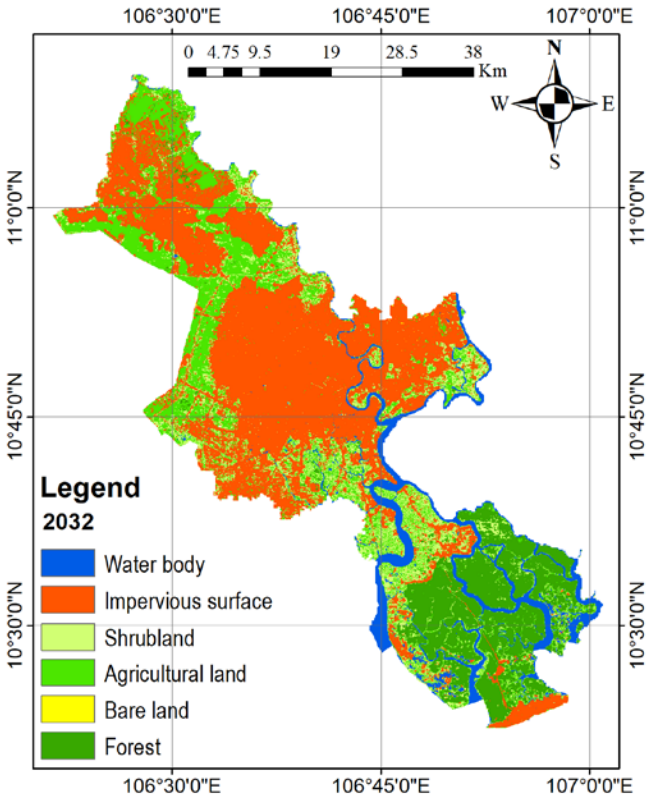


Figure 7. Predict of impervious surface changes in 2032 in Hochiminh city area

Table 2. Area of objects in the period 2002–2032.

Year / Area (km^2)	2002	2012	2022	2032
Water body	179.1558	174.9651	171.1934	169.7266
Impervious surface	565.3362	788.7048	890.7503	938.8395
Shrub	373.608	296.7072	214.2327	206.815
Agricultural land	655.438	509.3893	468.1101	428.7168
Bare land	124.4662	14.24865	0.544801	0.523848
Forest	197.3857	311.375	350.5587	350.7683

The results of determining the area of land cover/land use objects in Hochiminh city in the period 2002–2032 are shown in Table 2. It can be seen that the area of water body, shrubs, agricultural land, and bare land all tend to decrease in the period 2002–2032, in which agricultural land and bare land/shrub are reduced the most. Meanwhile, the water body area fluctuated but not significantly. With forest cover, the forest area in 2032 tends to be stable, not much different from 2022 due to the conservation policies of the Can Gio mangrove area. The impervious surface area continues to increase significantly, it is predicted that by 2032, it will reach over 938 km^2, nearly double that of

2002 and increase by about 5.4% compared to 2022. The impermeability surface has slowed down in the period of 2022–2032 compared to the period of 2002–2022 due to the stability in urban development in Hochiminh city. The impervious surface is still concentrated mainly in the central area of Hochiminh city, and according to the predict results, by 2032 there will be significant development in the area of Can Gio district, the section adjacent to the sea.

4. Conclusion

The study proposed to develope a the impervious surface prediction model in Hochiminh city from remote sensing and GIS data based on the Cellular Automata model and machine learning techniques. Three Landsat image scenes taken in the study area in 2002, 2012, and 2022 are used to classify land cover/land use by 04 different algorithms (maximum likelihood, RF, SVM, CART), from which to choose the algorithm with the highest classification accuracy. To predict the trend of impervious surface changes, this paper also uses an additional dataset including 08 layers extracted from remote sensing data, GIS and socio-economic databases.

The results of land cover/land use classification in 2002, 2012 and additional data sets are used to predict the urban surface cover development trend in 2022 by 02 methods (Artificial neural networks - ANN and Logistic regression), and then compare with the results of the land cover/land use classification from the Landsat image in 2022 to assess the accuracy. The obtained results show that the ANN algorithm has higher accuracy in predicting the trend of urban surface development. From this result, in the study, the ANN algorithm was used to predict the distribution of urban surface objects, including the impervious surface of Hochiminh city in 2032. The results obtained in the research provide objective and reliable information, helping managers in urban planning and development.

References

Asori M. & Adu P., 2023, Modeling the impact of the future state of land use land cover change patterns on land surface temperatures beyond the frontiers of greater Kumasi: A coupled cellular automaton (CA) and Markov chains approaches. Remote Sensing Applications: Society and Environment 29, 100908.

Azimand K., Aghighi H. & Matkan A., 2020, Classification and prediction of spatio-temporal Change of impervious urban surfaces and its impacts on urban heat intensity. Journal of Climate Research 11(41): 15–34.

Brabec E., Schulte S. & Richards P., 2002, Impervious surfaces and water quality: A review of current literature and its implications for watershed planning. Journal of Planning Literature 16(4): 499–514.

Bugday E. & Bugday S., 2019, Modeling and simulating land use/cover change using Artificial neural network from remote sensing data. CERNE 25(2), https://doi.org/10.1590/01047760201925022634

Chaula J., 2019, Ca-Markov Model for simulating Land use land cover dynamics in Rufiji delta of Tanzania. American Journal of Scientific Research and Essays 4, 27, p. 1–15.

Gharaibeh A., Shaamala A., Obeidat R. & Kofahi S., 2020, Improving land-use change modeling by integrating ANN with Cellular Automata-Markov Chain model. Heliyon 6(9), e05092. Doi:10.1016/j.heliyon.2020.e05092

Gong P., Li X., Wang J., Bai Y., Chen B., Hu T., Liu X., Xu B., Yang J., Zhang W. & Zhou Y., 2020, Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. Remote Sensing of Environment 236, 111510.

https://data.opendevelopmentmekong.net.

Khawaldah H., 2016, A prediction of future land use/land cover in Amman area using GIS-based Markov model and remote sensing. Journal of Geographic Information System 8(3): 412–427. Doi: 10.4236/jgis.2016.83035

Li F., Li E., Zhang C., Samat A., Liu W., Li C. & Atkinson P., 2021, Estimating artificial impervious surface percentage in Asia by fusing multi-temporal MODIS and VIIRS nighttime light data. Remote Sensing 13, 212. https://doi.org/10.3390/rs13020212

Liu F., Zhao Y., Rizwan M., Liu X. & Chen M., 2020, Impervious surface expansion: a key indicator for environment and urban agglomeration - a case study of Guangdong-Hong Kong-Macao greater bay area by using Landsat data. Journal of Sensor 3, p. 1 –21.

Mahyoub S., Rhinane H., Mansour M., Fadil A. & Okaishi W., 2022, Impervious surface prediction in Marrakech city using Artificial Neural Network. International Journal of Advanced Computer Science and Applications (IJACSA) 13(7): 183–189.

Misagh N., Samani N. & Tomanain A., 2018, Simulation of urban development in Tabriz using CA-Markov model and multi-criteria decision making. Human Geography Research Quarterly 50(1): 217–231.

Qiao K., Zhu W., Hu D., Hao M., Chen S. & Cao S., 2018, Examining the distribution and dynamics of impervious surface in different function zones in Beijing. Journal of Geographical Sciences 28: 669–684.

Ramezani M., Yu B. & Che Y., 2021, Prediction of total imperviousness from population density and land use data for urban areas (case study: South East Queensland, Australia). Applied Sciences 11(21), 10044. https://doi.org/10.3390/app112110044

Saputra M. & Lee H., 2019, Prediction of land use and land cover changes for North Sumatra, Indonesia, using an Artificial Neural-Network based Cellular Automaton. Sustainability 11, 3024, p. 1-16.

Sati A. & Mohan M., 2018, The impact of urbanization during half a century on surface meteorology based on WRF model simulations over National Capital Region, India. Theoretical and Applied Climatology 134: 309–323.

Trinh L.H., Nguyen T.T.N., Vu D.T. & Bui T.P., 2017, Assessement and prediction of urban land use changes of Hanoi city using remote sensing and GIS techniques. Hochiminh city University of Education Journal of Science, Natural Sciences and Technology 14(3): 176 – 187.

Xu T., Li E., Samat A., Li Z., Liu W. & Zhang L., 2022, Estimating large-scale interannual dynamic impervious surface percentages based on regional divisions. Remote Sensing 14, 3786.

Yin Z., Kuang W., Bai Y. Dou Y., Chi W., Ochege F. & Pan T., 2021, Evaluating the dynamic changes of urban land and its fractional covers in Africa from 2000–2020 using time series of remotely sensed images on the big data platform. Remote Sensing 13(21), 4288, https://doi.org/10.3390/rs13214288

Washburn B., Yancey K. & Mendoza J., 2010, User's guide for the California impervious surface coefficients. Office of Environmental Health Hazard Assessment, California Environmental Protection Agency. http://oehha.ca.gov/ ecotox/iscug123110.html