

MÔ HÌNH HÓA XU HƯỚNG PHÁT TRIỂN LỚP PHỦ/SỬ DỤNG ĐẤT KHU VỰC ĐÔ THỊ SỬ DỤNG DỮ LIỆU VIỄN THÁM VÀ TRÍ TUỆ NHÂN TẠO

Phạm Văn Tùng¹, Trịnh Lê Hùng^{2,*}, Nguyễn Văn Hưởng³

¹Trường Đại học Tài nguyên và Môi trường thành phố Hồ Chí Minh

²Viện Kỹ thuật công trình đặc biệt, Học viện Kỹ thuật quân sự

³Trường Đại học Tài nguyên và Môi trường Hà Nội

Tóm tắt

Nghiên cứu biến động lớp phủ/sử dụng đất là một trong những nội dung quan trọng phục vụ giám sát, quản lý và quy hoạch sử dụng tài nguyên, đặc biệt đối với khu vực đô thị. Biến động lớp phủ/sử dụng đất là kết quả của mối quan hệ tổng hợp giữa các yếu tố tự nhiên - xã hội. Trong nghiên cứu này, dữ liệu ảnh vệ tinh Landsat năm 2010, 2015 và các lớp thông tin bổ sung được sử dụng để dự báo biến động lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh cho năm 2020, sau đó so sánh với kết quả phân loại lớp phủ/sử dụng đất năm 2020 để đánh giá độ chính xác và hiệu chỉnh mô hình. Các kỹ thuật học máy bao gồm Random Forest (RF), Support Vector Machine (SVM), Classification And Regression Tree (CART) và phương pháp phân loại xác suất cực đại được thử nghiệm để phân loại lớp phủ/sử dụng đất từ ảnh vệ tinh Landsat, sau đó lựa chọn phương pháp có độ chính xác cao nhất. Để dự báo xu hướng phát triển lớp phủ/sử dụng đất trong tương lai, trong nghiên cứu kết hợp sử dụng mô hình toán học Cellular Automata (CA) và các mô hình trí tuệ nhân tạo, bao gồm: Hồi quy logistic (LR) và mạng neural nhân tạo (ANN), sau đó lựa chọn phương pháp có độ chính xác cao hơn để tiến hành dự báo xu hướng phát triển lớp phủ/sử dụng đất năm 2025 và 2030.

Từ khóa: Lớp phủ/sử dụng đất; Viễn thám; GIS; Mô hình hóa; Trí tuệ nhân tạo.

Abstract

Modeling the development trends of Land use/Land cover in urban areas using remote sensing data and artificial intelligence algorithms

Research on Land use/Land cover (LULC) change is one of the important contents serving the monitoring, management, and planning of natural resources, especially for urban areas. LULC change is the result of the synthetic relationship between natural and social factors. In this study, Landsat satellite image data in 2010, 2015 and additional information layers are used to predict LULC change in the Ho Chi Minh City area for 2020, then compared with the LULC classification results in 2020 to evaluate the accuracy and calibrate the model. Machine learning techniques including Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Tree (CART), and maximum likelihood classification method are tested to classify LULC from Landsat satellite images, then select the method with the highest accuracy. To predict the future development trend of LULC, the study combines the use of the Cellular Automata (CA) mathematical model and artificial intelligence

models, including logistic regression (LR) and artificial neural network (ANN), then selects the method with higher accuracy to predict the development trend of LULC in 2025 and 2030.

Keywords: Land use/Land cover; Remote sensing; GIS; Modeling; Artificial intelligence.

Nhận bài:19/9/2024; Phản biện xong: 04/10/2024; Chấp nhận đăng: 20/12/2024

***Tác giả liên hệ, Email:** trinhlehung@lqdtu.edu.vn

DOI: <https://doi.org/10.63064/khtnmt.2024.620>

1. Mở đầu

Biến động lớp phủ/sử dụng đất là một thông tin quan trọng phục vụ công tác quản lý, quy hoạch và phát triển khu vực đô thị (Liu et al., 2020; Xu et al., 2022). Quá trình đô thị hóa nhanh chóng trong thời gian gần đây dẫn đến những thay đổi sâu sắc trong lớp phủ/sử dụng đất khu vực đô thị, trong đó rõ nhất là sự gia tăng các bề mặt không thấm và sự suy giảm diện tích thực vật, mặt nước ở các thành phố.

Bề mặt không thấm là các bề mặt do con người tạo ra, bao gồm các loại bề mặt ngăn chặn quá trình nước không thể xâm nhập vào đất, chẳng hạn như đường giao thông, vỉa hè, bãi đậu xe, mái nhà,... (Qiao et al., 2018). Sự gia tăng bề mặt không thấm sẽ dẫn đến sự gia tăng về quy mô, thời gian và cường độ của dòng chảy trong đô thị, gây ô nhiễm nguồn nước. Sự xuất hiện với mức độ dày đặc các bề mặt không thấm cũng ảnh hưởng đáng kể đến khí hậu đô thị bằng cách thay đổi luồng nhiệt hợp lý và tiềm ẩn nguy cơ gây gia tăng nhiệt độ, dẫn đến hiện tượng đảo nhiệt đô thị.

Nhiều nghiên cứu trên thế giới đã minh chứng hiệu quả ứng dụng dữ liệu viễn thám trong trích xuất thông tin về đặc điểm, sự phân bố và sự thay đổi của lớp phủ/sử dụng đất khu vực đô thị. Dữ liệu viễn thám hỗ trợ xác định sự thay đổi

lớp phủ/sử dụng đất trong những khoảng thời gian từ trong quá khứ cho đến hiện tại một cách hệ thống và có tính đồng nhất cao (Sati and Mohan, 2018; Gong et al., 2020; Liu et al., 2020; Yin et al., 2021). Bên cạnh đó, công nghệ GIS với khả năng phân tích dữ liệu không gian mạnh cho phép định lượng, phân tích, mô hình hóa sự biến đổi lớp phủ/sử dụng đất để xác định xu thế phát triển trong tương lai. Trong nghiên cứu tại California (Mỹ), Washburn cho rằng việc phát triển khu vực thương mại dành cho bán lẻ sẽ dẫn đến 86 % lớp phủ đất đô thị trở thành bề mặt không thấm (Washburn et al., 2010). Một số nghiên cứu khác đã sử dụng các mô hình toán học như Markov, Cellular Automata,... để dự báo biến động lớp phủ/sử dụng đất từ dữ liệu viễn thám và GIS (Khawaldah, 2016; Misagh et al., 2018; Chaula, 2019; Asori and Adu, 2023). Cách tiếp cận này cung cấp một phương pháp đơn giản trong dự báo, tuy nhiên các mô hình này cũng có nhược điểm cơ bản khi kết quả dự báo chỉ hiệu quả và đạt độ chính xác cao khi các kịch bản sử dụng đất được thực hiện đầy đủ. Một số nghiên cứu gần đây đã phát triển các mô hình hồi quy giữa tỉ lệ phần trăm bề mặt không thấm và mật độ dân cư để dự báo xu thế phát triển bề mặt không thấm ở các đô thị. Các biến như tăng trưởng dân số, thương mại,... cũng được đưa vào một số mô hình

Nghiên cứu

nhằm nâng cao độ chính xác kết quả dự báo (Azimand et al., 2020; Li et al., 2021; Ramezani et al., 2021).

Thời gian gần đây, các mô hình trí tuệ nhân tạo dựa trên các kỹ thuật học máy đã được áp dụng nhằm nâng cao độ chính xác dự báo xu thế biến động lớp phủ/sử dụng đất từ dữ liệu viễn thám và GIS. Các kỹ thuật học máy như Random Forest (RF), Support Vector Machine (SVM), Classification And Regression Tree (CART),... được sử dụng để nâng cao độ chính xác kết quả phân loại lớp phủ/sử dụng đất, trong khi mạng neural nhân tạo (ANN) và các mô hình hồi quy được áp dụng kết hợp các mô hình toán học để dự báo phân bố lớp phủ/sử dụng đất trong tương lai (Mahyoub et al., 2022; Gharaibeh et al., 2020; Saputra and Li, 2019; Bugday, 2019).

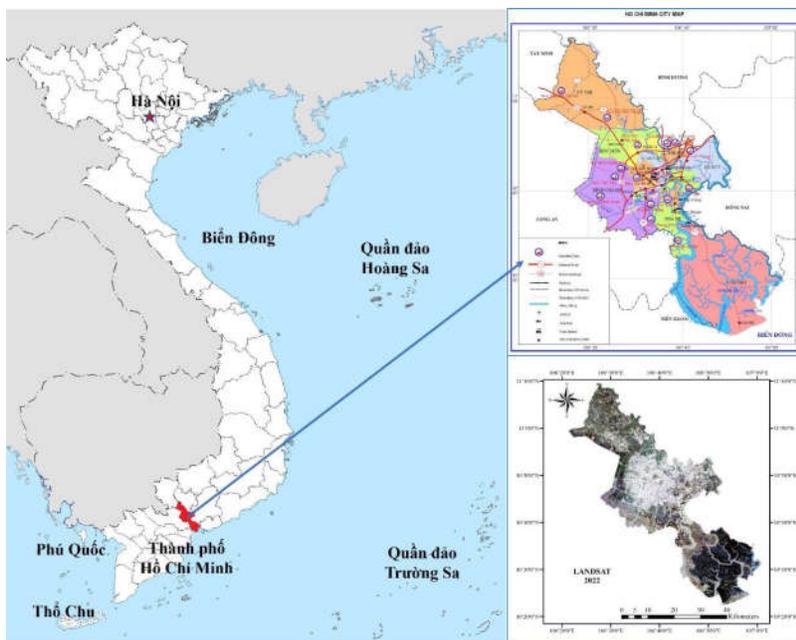
Nghiên cứu này trình bày kết quả mô hình hóa xu hướng biến động lớp phủ/sử

dụng đất khu vực thành phố Hồ Chí Minh từ dữ liệu viễn thám và GIS. Ảnh vệ tinh Landsat các năm 2010, 2015 và các lớp thông tin hỗ trợ được sử dụng để dự báo biến động lớp phủ/sử dụng đất cho năm 2020, sau đó so sánh với kết quả phân loại năm 2020 để đánh giá độ chính xác và hiệu chỉnh mô hình. Các kỹ thuật học máy bao gồm Random Forest (RF), Support Vector Machine (SVM), Classification And Regression Tree (CART) và phương pháp phân loại truyền thống xác suất cực đại được thử nghiệm trong phân loại lớp phủ/sử dụng đất, từ đó lựa chọn phương pháp có độ chính xác cao nhất. Để dự báo biến động lớp phủ/sử dụng đất, trong nghiên cứu thử nghiệm với 02 phương pháp: Hồi quy logistic (LR) và mạng neural nhân tạo (ANN), sau đó lựa chọn phương pháp có độ chính xác cao hơn để tiến hành dự báo cho năm 2025 và 2030.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu và dữ liệu sử dụng

a) Khu vực nghiên cứu



Hình 1: Vị trí địa lý thành phố Hồ Chí Minh

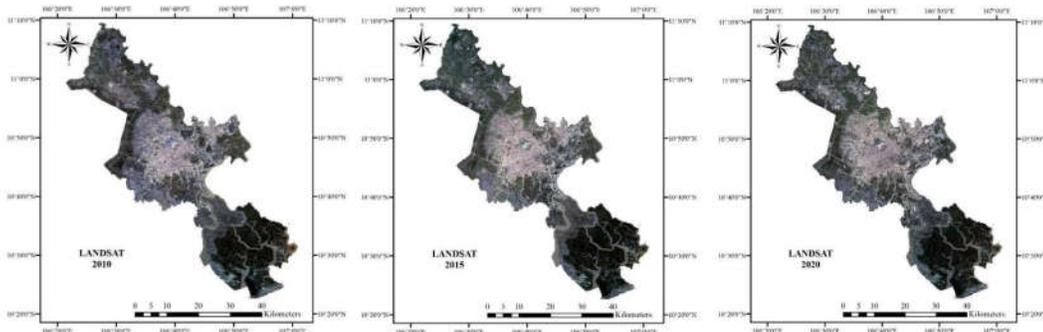
Khu vực nghiên cứu được lựa chọn là thành phố Hồ Chí Minh, một trong hai thành phố lớn nhất của Việt Nam (Hình 1). Thành phố Hồ Chí Minh có diện tích tự nhiên 2.095 km², dân số đạt gần 8,9 triệu người (tính đến tháng 6/2023). Mức độ gia tăng dân số ở thành phố khoảng 1 triệu người sau mỗi 5 năm (bao gồm cả người nhập cư), kéo theo tốc độ đô thị hóa cao.

Nằm trong vùng kinh tế trọng điểm phía Nam, thành phố Hồ Chí Minh là trung tâm kinh tế lớn nhất của cả nước, có tốc độ tăng trưởng kinh tế cao và đóng góp lớn vào GDP cả nước. Thành phố có cơ sở hạ tầng đa dạng và hiện đại, là đầu mối giao thông quan trọng của đất nước, bao gồm hệ thống đường cao tốc, cảng biển, sân bay, đường sắt,... Có thể nói, thành

phố có vai trò hết sức quan trọng, là trung tâm kinh tế, tài chính, thương mại và dịch vụ, là vùng động lực cho công cuộc phát triển kinh tế - xã hội của cả nước.

b) Dữ liệu sử dụng

Dữ liệu viễn thám sử dụng trong nghiên cứu là ảnh vệ tinh Landsat chụp khu vực thành phố Hồ Chí Minh, bao gồm 03 cảnh ảnh ngày 11/02/2010 (Landsat 5 TM), 09/02/2015 và 23/02/2020 (Landsat 8 OLI_TIRS) - Hình 2. Các ảnh đều được thu thập vào mùa khô, có chất lượng tốt và không bị ảnh hưởng bởi điều kiện thời tiết. Ảnh được thu thập ở mức độ xử lý L2A, do vậy trong nghiên cứu chỉ tiến hành hiệu chỉnh hình học và cắt theo ranh giới khu vực thực nghiệm.



Hình 2: Dữ liệu ảnh Landsat khu vực nghiên cứu

Để xây dựng bộ dữ liệu bổ sung cho mô hình dự báo phân bố lớp phủ/sử dụng đất, trong nghiên cứu đã phân tích đặc điểm tự nhiên, xã hội và thực trạng dữ liệu khu vực thực nghiệm, từ đó thu thập và xây dựng bộ dữ liệu gồm 09 lớp dữ liệu sau:

- + Mật độ che phủ của thực vật;
- + Nhiệt độ bề mặt;
- + Độ cao;
- + Độ dốc (Slope);
- + Hướng dốc (Aspect);

- + Khoảng cách tới công trình đặc biệt (UBND, trường học, bệnh viện,...);
- + Mật độ dân cư;
- + Khoảng cách tới đường giao thông;
- + Khoảng cách tới thủy hệ.

Các lớp dữ liệu đầu vào này đại diện cho các yếu tố có ảnh hưởng đến sự biến động lớp phủ/sử dụng đất, bao gồm yếu tố tự nhiên như địa hình, lớp phủ, khí hậu và yếu tố xã hội như mật độ dân cư, khoảng cách tới đường giao thông, khoảng cách tới thủy hệ và khoảng cách tới các công trình đặc biệt.

Nghiên cứu

Trong các lớp thông tin này, mật độ che phủ của thực vật và nhiệt độ bề mặt được chiết xuất từ dữ liệu ảnh vệ tinh Landsat 8. Độ cao, độ dốc, hướng dốc được chiết xuất thông qua dữ liệu GIS trên cơ sở sử dụng mô hình số độ cao DEM. Các lớp thông tin về khoảng cách tới đường giao thông, khoảng cách tới thủy hệ, khoảng cách tới công trình đặc biệt được xây dựng dựa trên module Euclidean Distance tích hợp trong phần mềm ArcGIS trên cơ sở các lớp dữ liệu vector về giao thông, thủy hệ, các công trình đặc biệt như trụ sở hành chính, trường học, bệnh viện, di tích lịch sử,... khu vực nghiên cứu.

Dữ liệu giao thông và thủy hệ khu vực nghiên cứu được chiết xuất từ cơ sở dữ liệu (CSDL) nền địa lý quốc gia tỉ lệ 1:50.000. Dữ liệu mô hình số độ cao (DEM) cũng được nội suy từ lớp dữ liệu địa hình của CSDL nền địa lý quốc gia tỉ lệ 1:50.000 khu vực thành phố Hồ Chí Minh tỉ lệ 1:50.000.

Dữ liệu về mật độ dân cư được xây dựng trên nền tảng GIS trên cơ sở số liệu thống kê về dân số, lao động thành phố Hồ Chí Minh năm 2021 của Cục Thống kê thành phố Hồ Chí Minh.

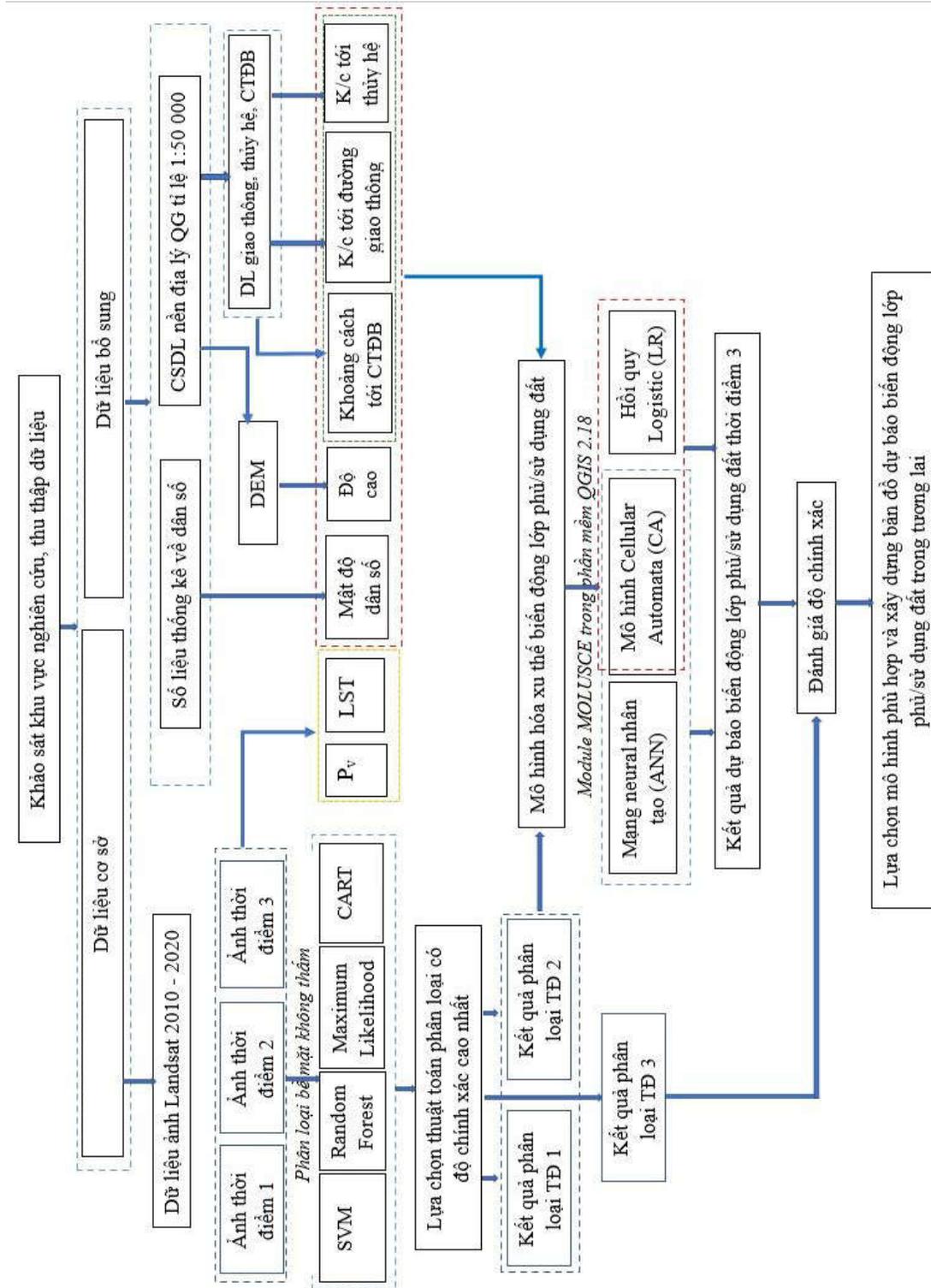
Dữ liệu về các công trình đặc biệt được khai thác từ CSDL, trong đó có vị trí các công trình đặc biệt như trụ sở các cơ quan, di tích lịch sử, trường học,...

2.2. Phương pháp nghiên cứu

Sơ đồ phương pháp nghiên cứu sử dụng trong bài báo được trình bày trên Hình 3. Trong mô hình đề xuất, bộ dữ liệu đầu vào bao gồm dữ liệu cơ sở (ảnh Landsat giai đoạn 2010 - 2020) và bộ dữ liệu bổ sung (gồm các lớp thông tin về

điều kiện tự nhiên, kinh tế - xã hội khu vực nghiên cứu). Để xây dựng bộ dữ liệu bổ sung, 9 lớp thông tin (mật độ che phủ của thực vật, nhiệt độ bề mặt, độ cao, độ dốc, hướng sườn, khoảng cách đến công trình đặc biệt, mật độ dân cư, khoảng cách tới đường giao thông và khoảng cách tới thủy hệ) chiết xuất từ các nguồn dữ liệu khác nhau được đánh giá mức độ tương quan với phân bố lớp phủ/sử dụng đất trên cơ sở xác định hệ số tương quan Pearson giữa các lớp thông tin này và chỉ số đất xây dựng NDBI (Normalized Difference Vegetation Index - Chỉ số thực vật khác biệt chuẩn hóa). Thông qua so sánh giá trị hệ số tương quan, trong nghiên cứu lựa chọn các lớp thông tin bổ sung có vai trò quan trọng đến sự thay đổi lớp phủ/sử dụng đất ở khu vực nghiên cứu.

Ở bước tiếp theo, lớp phủ/sử dụng đất được phân loại từ ảnh Landsat năm 2010 và 2015 bằng các thuật toán khác nhau (RF, SVM, CART, maximum likelihood), sau đó lựa chọn thuật toán phân loại có độ chính xác cao nhất. Mô hình toán học Cellular Automata (CA) được sử dụng để dự báo sự thay đổi về không gian của lớp phủ/sử dụng đất năm 2020, trong khi 02 kỹ thuật học máy (mạng neural nhân tạo - ANN và hồi quy logistic) được sử dụng để xây dựng ma trận khả năng chuyển đổi. Kết quả dự báo và kết quả phân loại lớp phủ/sử dụng đất năm 2020 được so sánh, đánh giá để lựa chọn phương pháp dự báo có độ chính xác cao nhất. Cuối cùng, phương pháp được lựa chọn sẽ được áp dụng để dự báo phân bố lớp phủ/sử dụng đất cho năm 2025 và 2030. Quá trình xử lý dữ liệu được thực hiện trên module MOLUSCE của phần mềm QGIS 2.18.

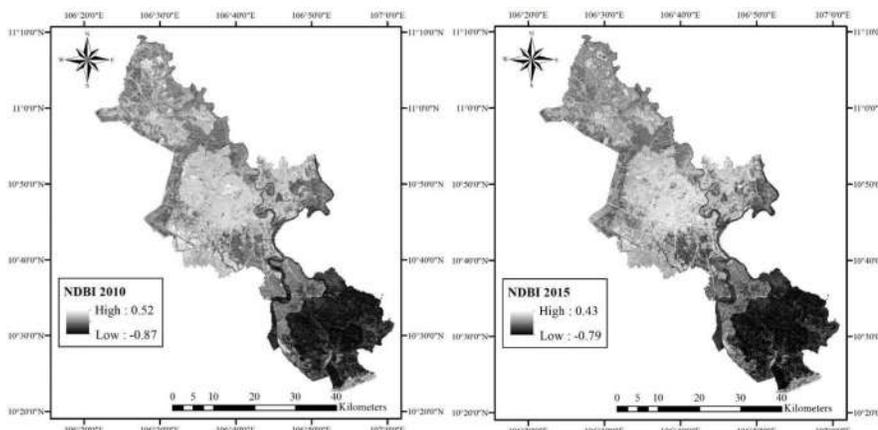


Hình 3: Sơ đồ mô hình dự báo xu hướng phát triển lớp phủ/sử dụng đất

3. Kết quả và thảo luận

Để đánh giá và lựa chọn bộ dữ liệu đầu vào cho mô hình dự báo, nghiên cứu tiến hành xác định hệ số tương quan giữa 09 lớp dữ liệu (mật độ che phủ của thực vật, nhiệt độ bề mặt, độ cao, độ dốc, hướng sườn, khoảng cách đến công trình đặc biệt, mật độ dân cư, khoảng cách tới

đường giao thông và khoảng cách tới thủy hệ) và chỉ số đất xây dựng NDBI (đại diện cho lớp phủ/sử dụng đất khu vực nghiên cứu). Kết quả xác định chỉ số NDBI đối với ảnh Landsat năm 2010 và 2015 khu vực thành phố Hồ Chí Minh được thể hiện trên Hình 4. Kết quả xác định hệ số tương quan của các lớp dữ liệu bổ sung và chỉ số NDBI được thể hiện trong Bảng 1.



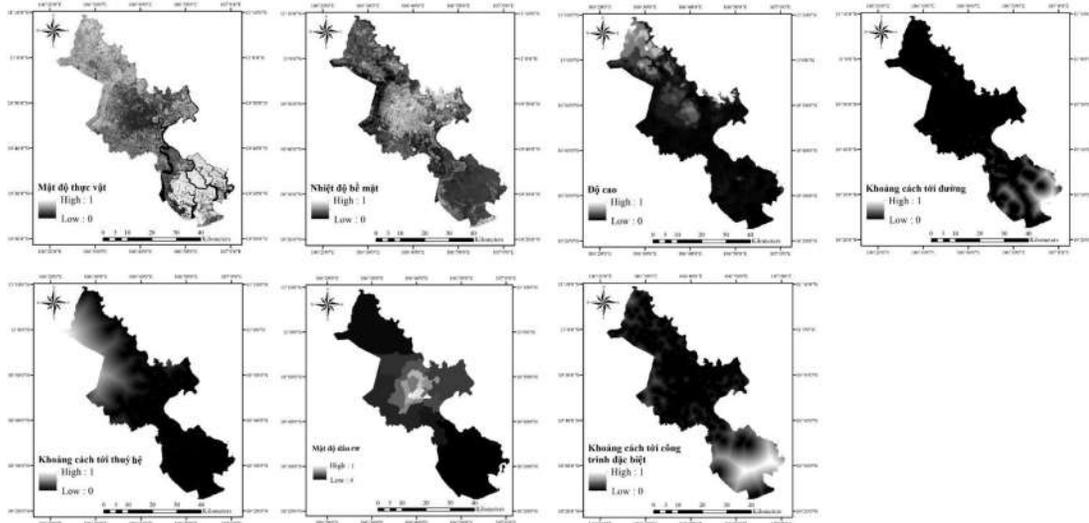
Hình 4: Chỉ số đất xây dựng NDBI xác định từ ảnh Landsat năm 2010 và 2015

Bảng 1. Giá trị tương quan của các lớp dữ liệu bổ sung

STT	Lớp dữ liệu bổ sung	Hệ số tương quan	
		Chỉ số NDBI 2010	Chỉ số NDBI 2015
1	Độ cao	0,321	0,325
2	Độ dốc	0,129	0,129
3	Hướng dốc	-0,112	-0,119
4	Mật độ che phủ thực vật	-0,358	-0,405
5	Nhiệt độ bề mặt	0,550	0,626
6	Mật độ dân số	0,398	0,456
7	Khoảng cách tới công trình đặc biệt	-0,661	-0,666
8	Khoảng cách tới đường giao thông	-0,572	-0,566
9	Khoảng cách tới thủy hệ	0,310	0,287

Từ kết quả đánh giá mối tương quan giữa các lớp thông tin bổ sung và chỉ số NDBI, trong nghiên cứu đã lựa chọn bộ dữ liệu bổ sung cho mô hình dự báo xu thế biến động lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh bao gồm 07 lớp dữ liệu, gồm: Mật độ che phủ của thực vật, nhiệt

độ bề mặt, độ cao, khoảng cách đến công trình đặc biệt, mật độ dân cư, khoảng cách tới đường giao thông và khoảng cách tới thủy hệ - Hình 5. Các lớp dữ liệu bổ sung này được chuẩn hóa về thang đo (0 - 1) và kích thước pixel 30 m để thống nhất với bộ dữ liệu cơ sở (ảnh Landsat).



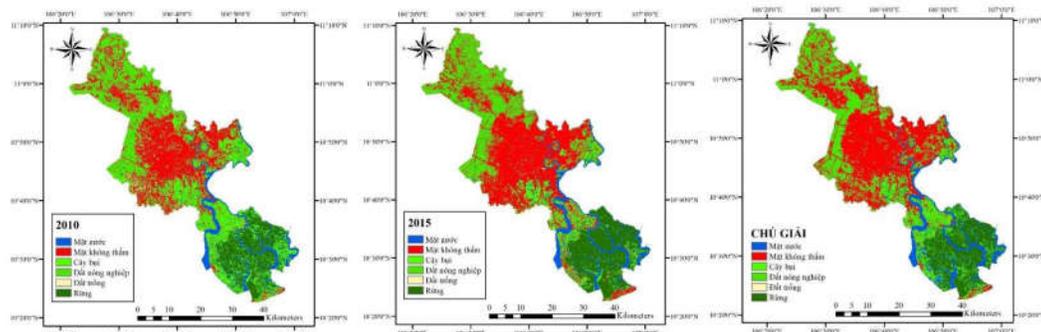
Hình 5: Bộ dữ liệu bổ sung (7 lớp) của mô hình dự báo biến động bề mặt không thám

Để đánh giá và lựa chọn phương pháp phân loại lớp phủ/sử dụng đất phù hợp với đặc điểm khu vực nghiên cứu, trong bài báo thử nghiệm với 04 phương pháp: 03 thuật toán học máy (RF, SVM, CART) và thuật toán phân loại xác suất

cực đại (ML). Kết quả đánh giá độ chính xác tổng thể, chỉ số Kappa khi phân loại lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh từ ảnh Landsat năm 2010, 2015 bằng 04 thuật toán được thể hiện trên Bảng 2.

Bảng 2. Độ chính xác tổng thể và chỉ số Kappa của các thuật toán phân loại lớp phủ/sử dụng đất

Năm	Độ chính xác	SVM	RF	CART	MD
2010	Độ chính xác tổng thể	86,90 %	93,44 %	90,77 %	85,98 %
	Chỉ số Kappa	0,871	0,919	0,887	0,829
2015	Độ chính xác tổng thể	88,72 %	93,54 %	89,81 %	87,29 %
	Chỉ số Kappa	0,863	0,921	0,876	0,845
2020	Độ chính xác tổng thể	91,26 %	93,02 %	87,65 %	88,10 %
	Chỉ số Kappa	0,892	0,915	0,858	0,847



Hình 6: Kết quả phân loại lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh bằng thuật toán RF

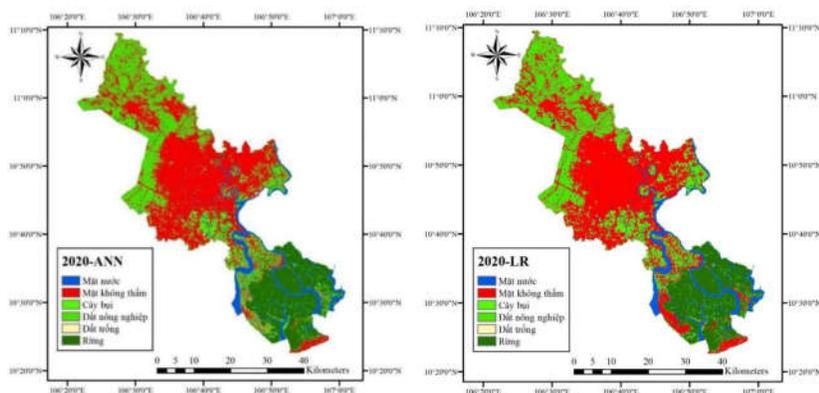
Nghiên cứu

Phân tích kết quả đạt được cho thấy, ở cả 2 thông số sử dụng để đánh giá độ chính xác kết quả phân loại, bao gồm độ chính xác tổng thể và chỉ số Kappa, thuật toán RF đều cho kết quả tốt nhất. Từ kết quả này, trong nghiên cứu lựa chọn thuật toán RF để phân loại lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh bằng ảnh Landsat giai đoạn 2010 - 2020. Kết quả phân loại lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh các năm 2010, 2015 và 2020 bằng thuật toán RF được thể hiện trên Hình 6, trong đó lớp phủ bề mặt được phân loại thành 6 lớp: Mặt nước, bề mặt không thấm, cây bụi, đất nông nghiệp, đất trồng và rừng.

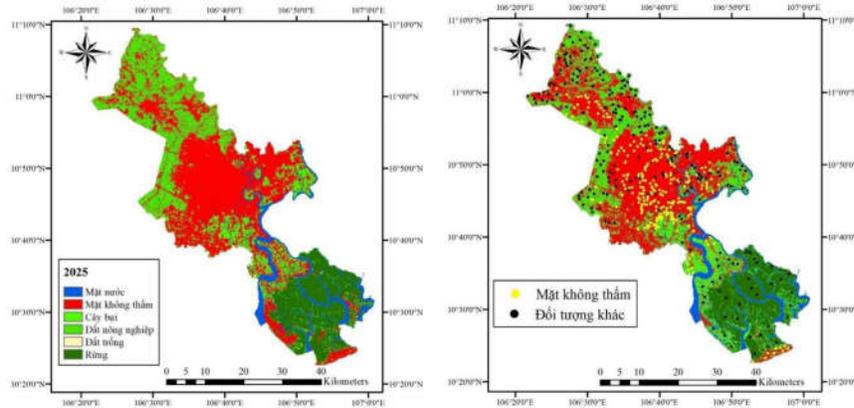
Để mô hình hóa quá trình phát triển lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh, trong bài báo sử dụng mô hình toán học Cellular Automata (CA) kết hợp 02 kỹ thuật học máy: Mạng ANN và hồi quy Logistic. Trước hết, đối với thuật toán ANN, trong nghiên cứu thiết lập bộ tham số bao gồm: Vùng lân cận (1 px), tỉ lệ học (0,001), số vòng lặp (1000), lớp ẩn (12) và quán tính (0,05). Sau khi huấn luyện, quá trình học tập dữ liệu đạt giá trị hệ số Kappa là 0,768 với bộ dữ liệu xác thực. Với thuật toán hồi quy Logistic, bộ tham số đầu vào được thiết lập bao gồm: Số lượng mẫu (2000), vùng lân cận (1 px), số vòng lặp

tối đa (100). Kết quả dự báo xu thế biến động lớp phủ bề mặt khu vực thành phố Hồ Chí Minh năm 2020 từ ảnh vệ tinh Landsat năm 2010, 2015 và bộ dữ liệu bổ sung sử dụng thuật toán ANN và hồi quy Logistic được trình bày trên Hình 7.

Để đánh giá độ chính xác của các mô hình dự báo biến động lớp phủ/sử dụng đất năm 2020, trong nghiên cứu sử dụng kết quả phân loại lớp phủ/sử dụng đất từ ảnh Landsat năm 2020 bằng thuật toán RF để so sánh. Kết quả nhận được cho thấy, hệ số tương quan giữa lớp phủ/sử dụng đất năm 2020 được mô hình hóa bằng mô hình CA kết hợp với thuật toán ANN và kết quả phân loại đạt 0,715. Trong khi đó, khi sử dụng mô hình CA kết hợp với thuật toán hồi quy Logistic, hệ số tương quan khi so sánh với kết quả phân loại đạt 0,666. Như vậy, thuật toán ANN có độ chính xác cao hơn so với hồi quy Logistic khi dự báo biến động lớp phủ khu vực đô thị thành phố Hồ Chí Minh. Từ kết quả này, trong bài báo sử dụng thuật toán ANN để xây dựng bản đồ dự báo biến động lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh các năm 2025 và 2030 (Hình 8). So sánh diện tích các đối tượng lớp phủ/sử dụng đất giai đoạn 2010 - 2025 được thể hiện trong Bảng 3.



Hình 7: Kết quả dự báo năm 2020 bằng thuật toán ANN và hồi quy Logistic



Hình 8: Kết quả dự báo năm 2025 và bản đồ quy hoạch sử dụng đất năm 2025 khu vực TP. HCM

Bảng 3. Diện tích của các đối tượng giai đoạn 2010 - 2025

Diện tích (km ²) \ Năm	Năm			
	2010	2015	2020	2025
Mặt nước	170.149	167.176	161.879	159.705
Mặt không thấm	616.983	786.107	816.806	825.076
Cây bụi	408.248	268.444	273.423	267.405
Đất nông nghiệp	551.434	458.123	437.473	441.749
Đất trồng	31.479	19.683	13.706	10.834
Rừng	263.942	342.702	338.948	337.470

Kết quả xác định diện tích các đối tượng lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh giai đoạn 2010 - 2025 được thể hiện trên Bảng 3. Có thể nhận thấy, diện tích mặt nước, cây bụi, đất nông nghiệp, đất trồng đều có xu thế giảm dần trong giai đoạn 2010 - 2025, trong đó giảm mạnh nhất là đất nông nghiệp và đất trồng/cây bụi. Trong khi đó, diện tích mặt nước có biến động nhưng không đáng kể. Với lớp phủ rừng, diện tích rừng năm 2025 có xu thế ổn định, không khác nhiều so với năm 2020 do các chính sách bảo tồn khu vực rừng ngập mặn Cần Giờ. Diện tích bề mặt không thấm tiếp tục có sự gia tăng đáng kể, theo dự báo đến năm 2025 đạt trên 825 km². Tốc độ gia tăng bề mặt không thấm có chậm dần trong giai đoạn 2020 - 2025 so với giai đoạn 2010 - 2020 do sự ổn định trong phát triển đô thị

ở thành phố Hồ Chí Minh. Bề mặt không thấm vẫn tập trung chủ yếu ở khu vực trung tâm thành phố Hồ Chí Minh và theo kết quả dự báo đến năm 2025 có sự phát triển đáng kể ở khu vực huyện Cần Giờ đoạn giáp biển.

Nhằm đánh giá mức độ hiệu quả của kết quả dự báo xu hướng phát triển lớp phủ/sử dụng đất bằng mô hình đề xuất, nghiên cứu tiến hành so sánh kết quả dự báo năm 2025 và bản đồ quy hoạch sử dụng đất thành phố Hồ Chí Minh năm 2025. Để xác định mức độ tương quan giữa kết quả dự báo và bản đồ quy hoạch, trong nghiên cứu lấy ngẫu nhiên 242 điểm đại diện cho bề mặt không thấm và 228 điểm cho các đối tượng khác (không phải bề mặt không thấm). Kết quả nhận được cho thấy, trong 242 điểm bề mặt không thấm, có 211 điểm được dự báo

Nghiên cứu

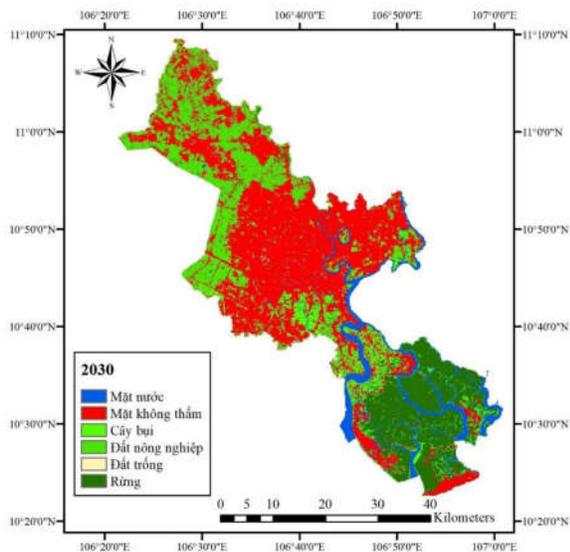
đúng (87,19 %) và 31 điểm không đúng (12,81 %). Với các điểm đại diện cho các đối tượng không phải bề mặt không thấm, có 192/228 điểm được dự báo đúng (84,21 %). Độ chính xác tổng thể và chỉ số Kappa đạt được là 85,70 % và 0,715. Như vậy, có thể nhận định, mô hình đề xuất trong nghiên cứu cho phép dự báo xu hướng phát triển lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh với độ tin cậy cao. Từ kết quả đạt được, trong nghiên cứu cũng xây dựng bản đồ dự báo phân bố lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh năm 2030 (Hình 9).

Phân tích kết quả đạt được cho thấy, theo kết quả mô hình hóa đến năm 2030, diện tích bề mặt không thấm khu vực thành phố Hồ Chí Minh đạt 833.241 km², tăng 8.165 km² so với năm 2025 (tương đương 1,00 %), tốc độ gia tăng bề mặt không thấm cơ bản giống như với giai đoạn 2020 - 2025 và không nhiều do mật độ xây dựng có sự bão hòa sau năm 2025.

Diện tích một số đối tượng khác có sự ổn định so với năm 2025 như diện tích mặt nước (158.457 km² năm 2030 so với 159.705 km² năm 2025), rừng (336.787 km² năm 2030 so với 337.470 km² năm 2025) do sự ổn định trong quy hoạch các đối tượng lớp phủ này.

Diện tích cây bụi, đất trống, đất nông nghiệp tiếp tục có sự suy giảm so với năm 2025, trong đó cây bụi giảm từ 267.405 km² (năm 2025) xuống còn 262.358 km² năm 2030; Đất nông nghiệp còn 440.056 km² năm 2030, giảm nhẹ so với 441.479 km² năm 2025. Diện tích đất trống đến năm 2030 chỉ còn 8.336 km² do phần lớn đất trống đã chuyển sang đất xây dựng do quá trình phát triển đô thị ở thành phố Hồ Chí Minh.

Từ Hình 9 cũng cho thấy, diện tích bề mặt không thấm gia tăng chủ yếu ở khu vực huyện Củ Chi và Cần Giờ, trong khi ở trung tâm thành phố Hồ Chí Minh không có sự thay đổi lớn khi mật độ bề mặt không thấm đã đạt mức bão hòa.



Hình 9: Dự báo bề mặt không thấm tại Thành phố Hồ Chí Minh đến năm 2030

4. Kết luận

Nghiên cứu đã đề xuất lựa chọn mô hình dự báo biến động lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh từ dữ liệu viễn thám và GIS trên cơ sở mô hình toán học CA và các kỹ thuật học máy. 03 cảnh ảnh Landsat chụp khu vực nghiên cứu các năm 2010, 2015, 2020 được sử dụng để phân loại lớp phủ/sử dụng đất bằng 04 thuật toán khác nhau (xác suất cực đại, RF, SVM, CART) và lựa chọn thuật toán có độ chính xác phân loại cao nhất. Để dự báo xu thế biến động lớp phủ/sử dụng đất, trong bài báo cũng sử dụng bộ dữ liệu bổ sung bao gồm 07 lớp được chiết xuất từ dữ liệu viễn thám, GIS và các CSDL về kinh tế, xã hội.

Kết quả phân loại lớp phủ/sử dụng đất các năm 2010, 2015 và bộ dữ liệu bổ

sung được sử dụng để dự báo xu thế phát triển lớp phủ bề mặt đô thị năm 2020 bằng 02 phương pháp mạng ANN và hồi quy Logistic, sau đó so sánh với kết quả phân loại lớp phủ/sử dụng đất từ ảnh Landsat năm 2020 để đánh giá độ chính xác. Kết quả nhận được cho thấy, thuật toán ANN có độ chính xác cao hơn trong dự báo xu thế phát triển bề mặt đô thị. Từ kết quả này, trong nghiên cứu đã sử dụng thuật toán ANN để dự báo phân bố các đối tượng bề mặt đô thị, trong đó có bề mặt không thấm khu vực thành phố Hồ Chí Minh năm 2025 và 2030. Kết quả nhận được trong nghiên cứu cung cấp thông tin khách quan và tin cậy, giúp các nhà quản lý trong công tác quy hoạch và phát triển đô thị.

TÀI LIỆU THAM KHẢO

- [1]. 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 approach*. Remote Sensing Applications: Society and Environment, 29, 100908.
- [2]. 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.
- [3]. 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.
- [4]. Bugday E., Bugday S., (2019). *Modeling and simulating Land use/cover change using Artificial neural network from remotely sensing data*. CERNE 25(2), <https://doi.org/10.1590/01047760201925022634>.
- [5] Chaula J., (2019). *Ca-Markov Model for simulating Land use/Land cover dynamics in the Rufiji delta of Tanzania*. American Journal of Scientific Research and Essays, 4:27, 1 - 15.
- [6]. Gharaibeh A., Shyamala 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.
- [7]. 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.
- [8]. Khawaldeh H., (2016). *A prediction of future land use/land cover in the Amman area using GIS-based Markov model and remote sensing*. Journal of Geographic Information System, Vol.8(3), 412 - 427. Doi: 10.4236/jgis.2016.83035.
- [9]. 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>.
- [10]. Liu F., Zhao Y., Rizwan M., Liu X., Chen M., (2020). *Impervious surface expansion: a key indicator for the environment and urban agglomeration - a case study of Guangdong-Hong Kong-Macao greater bay area by using Landsat data*. Journal of Sensors, 3, 1 - 21.
- [11]. Mahyoub S., Rhinane H., Mansour M., Fadil A., Okashi 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.
- [12]. 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.
- [13]. Qiao K., Zhu W., Hu D., Hao M., Chen S., Cao S., (2018). *Examining the*

Nghiên cứu

distribution and dynamics of impervious surfaces in different function zones in Beijing. Journal of Geographical Sciences, 28, 669 - 684.

[14]. 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>.

[15]. 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, 16 pp.

[16]. Sati A., Mohan M., (2018). *The impact of urbanization during half a century on surface meteorology based on WRF model simulations over the National Capital Region, India.* Theoretical and Applied Climatology, 134, 309 - 323.

[17]. 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.

[18]. 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 a time series of remotely sensed images on the big data platform.* Remote Sensing, 13(21), 4288, <https://doi.org/10.3390/rs13214288>.

[19]. 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>.

[20]. <https://data.opendevopmentmekong.net>.

[21]. <http://www.pso.hochiminhcity.gov.vn/>.

[22]. <https://opendata.hcmgis.vn/>