

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT

PHẠM VĂN TÙNG

**NGHIÊN CỨU CÔNG NGHỆ VIỄN THÁM VÀ GIS TRONG DỰ
BÁO SỰ BIẾN ĐỘNG BỀ MẶT KHÔNG THẨM KHU VỰC
THÀNH PHỐ HỒ CHÍ MINH**

Ngành: Kỹ thuật Trắc địa- bản đồ
Mã số: 9520503

TÓM TẮT LUẬN ÁN TIẾN SĨ KỸ THUẬT

Hà Nội - 2023

Công trình được hoàn thành tại: **Bộ môn Đo ảnh và Viễn thám,
Khoa Trắc địa – Bản đồ và Quản lý đất đai, Trường Đại học Mỏ - Địa chất**

Người hướng dẫn khoa học:

1. PGS.TS Nguyễn Văn Trung

2. PGS.TS Vũ Xuân Cường

Phản biện 1: **PGS.TS Trần Văn Anh**

Trường Đại học Mỏ - Địa chất

Phản biện 2: **PGS.TS Bùi Ngọc Quý**

Đại học Quốc gia Hà Nội

Phản biện 3: **TS Nguyễn Dư Khang**

Cục Viễn thám Quốc gia

Luận án sẽ được bảo vệ trước Hội đồng đánh giá luận án cấp Trường họp
tại Trường

Đại học Mỏ - Địa chất vào hồi giờ ngày tháng năm

Có thể tìm hiểu luận án tại:

- Thư viện Quốc Gia, Hà Nội

- Thư viện Trường Đại học Mỏ - Địa chất

MỞ ĐẦU

1. Tính cấp thiết của đề tài

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, v.v... Trong những năm gần đây, bề mặt không thấm đã nổi lên không chỉ là một chỉ số về mức độ đô thị hóa, mà còn là một chỉ số chính về chất lượng môi trường đô thị.

Cho đến nay, chưa có nghiên cứu nào ở Việt Nam sử dụng công nghệ viễn thám và GIS nhằm mô hình hóa sự biến động bề mặt không thấm và dự báo sự gia tăng bề mặt không thấm ở các đô thị về cả diện tích và không gian. Với thực trạng trên, nghiên cứu sinh đã đề xuất đề tài luận án ***“Nghiên cứu công nghệ viễn thám và GIS trong dự báo sự biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh”***. Kết quả nhận được trong luận án nhằm áp dụng thử nghiệm tại khu vực Thành phố Hồ Chí Minh, ngoài ra mô hình dự báo cũng có thể áp dụng cho các khu vực đô thị khác trong cả nước. Có thể nhận định, đề tài luận án được đề xuất có ý nghĩa khoa học, ý nghĩa thực tiễn và thể hiện sự cần thiết phải nghiên cứu.

2. Mục tiêu nghiên cứu

Mục tiêu nghiên cứu của luận án nhằm lựa chọn được mô hình dự báo biến động bề mặt không thấm trong tương lai tại khu vực Thành phố Hồ Chí Minh sử dụng công nghệ viễn thám và Hệ thống tin địa lý kết hợp các kỹ thuật trí tuệ nhân tạo.

3. Nội dung nghiên cứu

- Tổng quan về vấn đề nghiên cứu;
- Cơ sở khoa học xây dựng mô hình dự báo biến động bề mặt không thấm từ dữ liệu viễn thám và hệ thống tin địa lý;
- Phân tích đặc điểm tự nhiên, kinh tế - xã hội khu vực nghiên cứu;
- Xây dựng các lớp dữ liệu đầu vào cho mô hình dự báo biến động bề mặt không thấm;
- Ứng dụng trí tuệ nhân tạo trong phân loại ảnh viễn thám; Thử nghiệm phân loại bề mặt không thấm từ dữ liệu ảnh vệ tinh Landsat bằng các phương pháp khác nhau: phân loại truyền thống, phân loại dựa trên các kỹ thuật học máy. Lựa chọn phương pháp phân loại có độ chính xác cao nhất áp dụng cho khu vực nghiên cứu;
- Thử nghiệm và lựa chọn mô hình, xây dựng bản đồ dự báo biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh;
- Phân tích, đánh giá kết quả.

4. Đối tượng và phạm vi nghiên cứu

- Đối tượng nghiên cứu là bề mặt không thấm và các mô hình dự báo biến động bề mặt không thấm trên cơ sở kết hợp dữ liệu viễn thám, hệ thống tin địa lý và các kỹ thuật học máy.
- Phạm vi khoa học của đề tài tập trung vào phân tích, đánh giá, thử nghiệm nhằm đề xuất được mô hình dự báo biến động bề mặt không thấm phù hợp với điều kiện cụ thể khu vực nghiên cứu.
- Phạm vi không gian của đề tài là khu vực Thành phố Hồ Chí Minh.
- Phạm vi thời gian: trong luận án sử dụng dữ liệu ảnh vệ tinh Landsat giai đoạn 2010 - 2020, bao gồm 03 cảnh ảnh năm 2010, 2015 và 2020 để mô hình hóa và dự báo biến động bề mặt không thấm các năm 2025, 2030.

5. Phương pháp nghiên cứu

- Phương pháp phân tích, tổng hợp;
- Phương pháp viễn thám;
- Phương pháp GIS;
- Phương pháp mô hình hóa;
- Phương pháp phân tích thống kê;
- Kỹ thuật lập trình.

6. Những điểm mới của luận án

- Đề xuất và xây dựng được bộ dữ liệu đầu vào từ dữ liệu viễn thám và GIS cùng các số liệu thống kê phục vụ xây dựng mô hình dự báo biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh.
- Lựa chọn được mô hình dự báo biến động bề mặt không thấm từ dữ liệu viễn thám và GIS trên cơ sở kết hợp phương pháp mô hình toán học Cellular Automata (CA) và mạng neural nhân tạo (ANN) ứng dụng cho khu vực Thành phố Hồ Chí Minh.

7. Luận điểm bảo vệ

Luận điểm 1: Dữ liệu viễn thám đa thời gian và các lớp thông tin bổ sung (mật độ lớp phủ thực vật, nhiệt độ bề mặt, độ cao, khoảng cách tới đường giao thông, khoảng cách tới thủy hệ, khoảng cách tới công trình đặc biệt, mật độ dân cư) cho phép xây dựng mô hình dự báo biến động bề mặt không thấm phù hợp với điều kiện khu vực Thành phố Hồ Chí Minh.

Luận điểm 2: Mô hình toán học Cellular Automata (CA) kết hợp mạng nơ ron nhân tạo (ANN) giúp dự báo xu thế biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh với độ chính xác cao, trong đó hệ số tương quan so với kết quả phân loại đạt 0,715.

8. Ý nghĩa khoa học và thực tiễn của luận án

Ý nghĩa khoa học: góp phần hoàn thiện cơ sở khoa học và minh chứng tính hiệu quả của phương pháp ứng dụng dữ liệu viễn thám, GIS và phương pháp mô hình hóa trong dự báo biến động bề mặt không thấm khu vực đô thị.

Ý nghĩa thực tiễn: Cung cấp thông tin kịp thời và có độ chính xác cao về xu thế biến động bề mặt không thấm, cả về diện tích và không gian, giúp các nhà quản lý, quy hoạch đô thị, những nhà hoạch định chiến lược phát triển kinh tế xã hội ở phạm vi khu vực đô thị ở nước ta đưa ra những quyết sách hợp lý nhằm làm giảm tối đa các ảnh hưởng tiêu cực tới các hoạt động của con người và môi trường sống trong quá trình phát triển bền vững trong tương lai. Bên cạnh đó, kết quả nhận được trong luận án cũng có thể sử dụng, tham khảo trong công tác nghiên cứu khoa học, giảng dạy ở các trường đại học và viện nghiên cứu.

9. Cơ sở tài liệu thực hiện luận án

- Dữ liệu ảnh viễn thám (ảnh Landsat các năm 2010, 2015 và 2020);
- CSDL nền địa lý quốc gia tỉ lệ 1:50 000, dữ liệu thống kê năm 2021 về mật độ dân số khu vực nghiên cứu;
- Các CSDL mở về vị trí các công trình đặc biệt (<https://opendata.hcmgis.vn>);
- Bản đồ quy hoạch sử dụng đất thành phố Hồ Chí Minh đến năm 2025;
- Các tài liệu, số liệu khác.

10. Cấu trúc luận án

Luận án gồm 03 chương chính, phần mở đầu, kết luận - kiến nghị và tài liệu tham khảo.

CHƯƠNG 1. TỔNG QUAN VỀ VẤN ĐỀ NGHIÊN CỨU

1.1. Tổng quan về bề mặt không thấm

1.1.1. Khái niệm

Bề mặt không thấm (impervious surface) là các bề mặt do con người tạo ra, có đặc điểm ngăn không cho nước mưa ngấm xuống đất một cách tự nhiên, thay vào đó lượng nước mưa được tích tụ và chảy vào các hệ thống thoát nước cũng như sông, suối [94]. Các dạng bề mặt không thấm quen thuộc bao gồm đường giao thông, vỉa hè, bãi đậu xe, mái nhà,....

1.1.2. Ảnh hưởng của bề mặt không thấm đến môi trường

1.1.2.1. Gia tăng nguy cơ ngập lụt

1.1.2.2. Ô nhiễm nước

1.1.2.3. Gia tăng nhiệt độ bề mặt

1.2. Đô thị hóa và sự thay đổi diện tích bề mặt không thấm

1.2.1. Hiện trạng đô thị hóa trên thế giới và ở Việt Nam

Theo số liệu của Ngân hàng thế giới, tính đến năm 2021, tỉ lệ dân cư đô thị của châu Âu đạt 75%, trong đó một số nước có tỉ lệ dân cư đô thị rất cao như Bỉ (98%), Hà Lan (93%), Đan Mạch (88%), Pháp (81%)... [148]. Tỉ lệ dân cư đô thị ở châu Âu liên tục tăng, tuy nhiên tốc độ bắt đầu chậm lại, chẳng hạn trong giai đoạn 2018 – 2021, tốc độ gia tăng dân số đô thị ở châu Âu đạt tương ứng là 0.46%, 0.38%, 0.37% và 0.21%.

Tại châu Á, tốc độ gia tăng dân số đô thị cũng diễn ra mạnh mẽ, có xu hướng nhanh hơn so với toàn thế giới, đặc biệt là ở Trung Quốc và các nước ASEAN [60]. Theo số liệu của Ngân hàng thế giới, tỉ lệ dân cư đô thị tại một số quốc gia châu Á đạt cao như Nhật Bản (91.9%), Hàn Quốc (81.4%), Malaysia (78%), Trung Quốc (62.5%), thậm chí tại Hồng Kông (Trung Quốc), Singapore, tỉ lệ dân cư đô thị đạt 100% [148].

Tại Việt Nam, theo thống kê của Cục Phát triển đô thị, tính đến tháng 12 năm 2022, tổng số đô thị cả nước là 888 đô thị, bao gồm 2 đô thị loại đặc biệt là Hà Nội và Thành phố Hồ Chí Minh, 22 đô thị loại I, 33 đô thị loại II, 47 đô thị loại III, 94 đô thị loại IV, 690 đô thị loại V. Tỷ lệ đô thị hóa cả nước tính theo khu vực nội thành/nội thị là 41.7%; tỷ lệ đô thị hóa cả nước tính theo khu vực toàn đô thị là 53.7% [147]...

1.2.2. Biến động diện tích bề mặt không thấm

Tại Việt Nam, diện tích bề mặt không thấm ở các đô thị lớn như Hà Nội, Thành phố Hồ Chí Minh...có sự gia tăng mạnh mẽ cùng với quá trình đô thị hóa thời gian qua. Sử dụng ảnh vệ tinh quang học Landsat, Hung Q. Ha và Weng đã xác định được diện tích bề mặt không thấm ở Hà Nội tăng 305% từ năm 1988 đến năm 2015 [56].

1.3. Khả năng ứng dụng viễn thám và GIS trong giám sát biến động bề mặt không thấm

Thông tin về các bề mặt không thấm, đặc biệt là mật độ, vị trí, hình dạng, phân bố không gian có ý nghĩa hết sức quan trọng đối với công tác quản lý đô thị cũng như quản lý tài nguyên - môi trường [120]. Các bề mặt không thấm có sự thay đổi thường xuyên, liên tục cùng với quá trình đô thị hóa và phát triển kinh tế - xã hội. Do vậy, phương pháp viễn thám và GIS được xem là cách tiếp cận phù hợp, khả thi và hiệu quả trong nghiên

cứu chiết xuất các bề mặt không thấm, phục vụ công tác giám sát biến động và dự báo sự thay đổi bề mặt không thấm trong tương lai.

1.4. Các phương pháp chiết xuất thông tin bề mặt không thấm từ dữ liệu viễn thám

1.4.1. Phương pháp phân loại truyền thống

1.4.2. Phương pháp sử dụng các chỉ số phổ

1.4.3. Phương pháp phân loại sử dụng trí tuệ nhân tạo

1.5. Tổng quan tình hình nghiên cứu trong và ngoài nước

1.5.1. Trên thế giới

1.5.1.1. Các nghiên cứu về phương pháp chiết xuất bề mặt không thấm

Colstoun và cộng sự đã xây dựng bộ dữ liệu về bề mặt không thấm toàn cầu (GMIS - Global Man-made Impervious Surface) năm 2010 từ dữ liệu ảnh vệ tinh Landsat, trong đó đã xây dựng được bộ bản đồ tỉ lệ phần trăm bề mặt không thấm cho hầu hết các khu vực trên thế giới, ngoại trừ Nam Cực và một số đảo nhỏ bằng phương pháp hồi quy dạng cây Cubist. Đây được xem là bộ dữ liệu đầu tiên ở quy mô toàn cầu về bề mặt không thấm và là bộ dữ liệu hỗ trợ cho bộ dữ liệu về phạm vi định cư và xây dựng của con người (HBASE - Global Human Built-up And Settlement Extent) [44].

Các kỹ thuật học sâu (deep learning) dựa trên mạng neural nhân tạo (Convolution Neural Network, Cellular Neural Networks) cũng được áp dụng trong phân loại bề mặt không thấm từ dữ liệu viễn thám quang học [50], [89], [114]. Mạng neural 3D tích chập (3D CNN) cũng được sử dụng trong nghiên cứu [113] nhằm chiết xuất bề mặt không thấm từ dữ liệu viễn thám độ phân giải cao WorldView và dữ liệu LiDAR hàng không. Các mô hình học sâu như U-Net_SGD_Bands, U-Net_Adam_Bands, U-Net_Adam_Bands+SI và VGG-19_Adam_Bands+SI kết hợp các chỉ số phổ cùng thuật toán SVM cũng được sử dụng trong nghiên cứu của Parekh và cộng sự (2021) nhằm tự động chiết xuất thông tin bề mặt không thấm từ ảnh vệ tinh Sentinel 2 [92].

1.5.1.2. Các nghiên cứu dự báo biến động bề mặt không thấm từ dữ liệu viễn thám và GIS

Jennings và cộng sự (2004) sử dụng dữ liệu lớp phủ theo quốc gia năm 1992 (NLCD 92 – National Land Cover 92), dữ liệu bề mặt không thấm và công nghệ GIS để xác định hệ số chuyển đổi thành bề mặt không thấm đối với từng đối tượng lớp phủ khu vực Trung - Đại Tây Dương. Trên cơ sở đó, các tác giả đã phát triển một mô hình hồi quy tuyến tính giữa dữ liệu NDLC 92 và tỉ lệ phần trăm bề mặt không thấm để dự báo phân bố bề mặt không thấm trong tương lai. Độ chính xác tổng thể của kết quả dự báo được xác định đạt trên 70%, trong đó kết quả dự báo cho các khu vực đô thị có độ chính xác cao hơn so với khu vực nông thôn [62]. Dữ liệu lớp phủ toàn cầu năm 2006 xác định từ ảnh vệ tinh Landsat (NLCD 2006) cũng được sử dụng trong nghiên cứu [96] để dự báo phân bố bề mặt không thấm ở khu vực California (Mỹ). Các tác giả cũng đã xây dựng được công cụ trên nền tảng Excel 2007 cho phép người sử dụng có thể nhập các kịch bản khác nhau để sự báo tốc độ tăng trưởng bề mặt không thấm trong tương lai.

Reilly và cộng sự (2004) đã phát triển một mô hình dự báo phân bố bề mặt không thấm trên cơ sở mối quan hệ giữa mật độ dân số và tỉ lệ bề mặt không thấm xác định từ ảnh vệ tinh. Các dạng hồi quy khác nhau được thử nghiệm để đánh giá mối quan hệ giữa mật độ dân số và tỉ lệ bề mặt không thấm, từ đó tác giả lựa chọn hàm đa thức bậc 2 nhằm dự báo phân bố bề mặt không thấm trong tương lai [99]. Mô hình này rất đơn giản, có thể ước tính nhanh bề mặt không thấm phục vụ công tác quy hoạch, quản lý đô thị, tuy nhiên độ chính xác không cao do mới sử dụng 01 biến (mật độ dân số) để xây dựng mô hình dự báo.

1.5.2. Trong nước

Các nghiên cứu ở Việt Nam cho đến nay chủ yếu tập trung vào chiết xuất bề mặt không thấm từ dữ liệu viễn thám và đánh giá biến động diện tích bề mặt không thấm ở các giai đoạn khác nhau.

Thời gian gần đây, một số nghiên cứu đã sử dụng các chỉ số đất đô thị để nâng cao độ chính xác trong chiết tách thông tin sử dụng đất đô thị như đất xây dựng, đất trống... Nguyễn Hoàng Khánh Linh (2011) sử dụng chỉ số xây dựng (IBI – Index based built-up Index) xác định từ ảnh Landsat TM nhằm phân loại và theo dõi biến động đất xây dựng khu vực thành phố Huế. Độ chính xác kết quả phân loại được đánh giá trên cơ sở 480 mẫu ngẫu nhiên, qua đó kết quả phân loại có độ chính xác tổng thể 91.04%, giá trị chỉ số Kappa đạt 0.904 [16]. Nguyễn Thị Thúy Hạnh đã nghiên cứu đánh giá, so sánh các chỉ số đất đô thị, bao gồm NDBI (Normalized Difference Built-up Index), EBBI (Enhanced Built-up and Bareness Index), UI (Urban Index) và IBI trong chiết tách đất trống và đất xây dựng khu vực Hà Nội từ ảnh vệ tinh Landsat 8. Nghiên cứu của Nguyễn Thị Thúy Hạnh cho thấy, việc sử dụng chỉ số NDBI cho phép phân loại đất xây dựng trên ảnh Landsat 8 với độ chính xác cao hơn so với một số chỉ số đô thị khác như EBBI, IBI, UI [8]. Để nâng cao độ phân giải không gian của các chỉ số đất đô thị, Trịnh Lê Hùng và cộng sự đã tích hợp sử dụng dữ liệu viễn thám đa độ phân giải Landsat 8 và Sentinel 2 MSI trong xác định các chỉ số như NDBaI, EBBI, từ đó nâng cao độ chính xác phân loại đất xây dựng, đất trống đô thị [13], [14].

1.6. Luận giải những vấn đề cần nghiên cứu

Từ phân tích các kết quả đạt được cũng như những hạn chế trong các nghiên cứu trước đó, đặc biệt là các nghiên cứu ở Việt Nam, để khắc phục các hạn chế trên, trong luận án này, nghiên cứu sinh tập trung nghiên cứu một số vấn đề sau:

- Nghiên cứu lựa chọn phương pháp chiết tách thông tin bề mặt không thấm phù hợp từ dữ liệu viễn thám độ phân giải trung bình Landsat trên cơ sở áp dụng các kỹ thuật học máy;
- Nghiên cứu lựa chọn và xây dựng bộ dữ liệu đầu vào cho mô hình dự báo phân bố bề mặt không thấm trong tương lai, bao gồm các lớp dữ liệu về đặc điểm địa hình, lớp phủ, các yếu tố kinh tế - xã hội (mật độ dân cư, khoảng cách đến đường giao thông, khoảng cách đến thủy hệ, khoảng cách đến công trình đặc biệt);
- Nghiên cứu thử nghiệm và lựa chọn mô hình dự báo phân bố bề mặt không thấm phù hợp dựa trên các kỹ thuật hiện đại như hồi quy logistics (LR), mạng neural nhân tạo (ANN), áp dụng với điều kiện cụ thể tại khu vực nghiên cứu ở Thành phố Hồ Chí Minh.

1.7. Tiểu kết chương 1

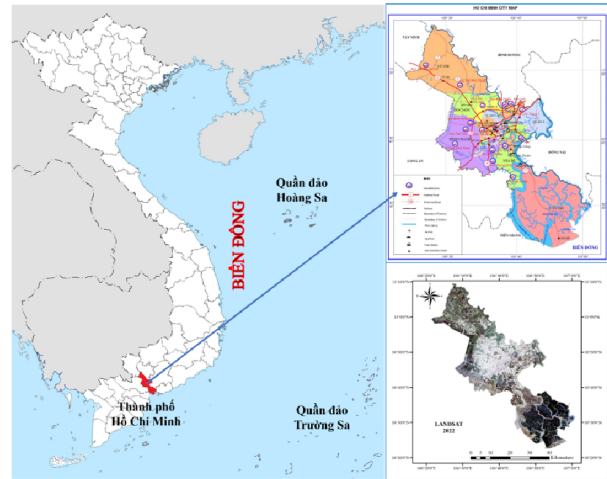
Trong chương 1, nghiên cứu sinh đã nghiên cứu, phân tích các khái niệm cơ bản về đô thị hóa, sự gia tăng bề mặt không thấm trong quá trình đô thị hóa, các phương pháp chiết xuất thông tin bề mặt không thấm từ dữ liệu viễn thám. Từ phân tích tổng quan tình hình nghiên cứu trong và ngoài nước cho thấy, phần lớn các nghiên cứu mới tập trung vào phát triển các phương pháp, thuật toán chiết xuất bề mặt không thấm từ dữ liệu viễn thám; các nghiên cứu dự báo biến động bề mặt không thấm thường sử dụng kết quả dự báo biến động lớp phủ/sử dụng đất từ ảnh viễn thám đa thời gian. Một số nghiên cứu đã xây dựng các mô hình hồi quy trên cơ sở mối quan hệ giữa bề mặt không thấm và phản xạ phổ, dữ liệu ánh sáng, tăng trưởng thương mại, mật độ dân số. Mặc dù vậy, số lượng biến giải thích trong các mô hình này là rất hạn chế, trong đó nhiều mô hình chỉ sử dụng duy nhất một biến. Trong khi đó, chưa có nghiên cứu nào ở Việt Nam quan tâm đề xuất, xây dựng mô hình phục vụ dự báo phân bố bề mặt không thấm trong tương lai. Trên cơ sở phân tích các kết quả đạt được,

các hạn chế còn tồn tại, nghiên cứu sinh đã đề xuất các định hướng nghiên cứu trong luận án, làm cơ sở để lựa chọn phương pháp và thực nghiệm dự báo biến động bề mặt không thấm ở khu vực nghiên cứu.

CHƯƠNG 2. CƠ SỞ KHOA HỌC PHƯƠNG PHÁP DỰ BÁO BIẾN ĐỘNG BỀ MẶT KHÔNG THẤM KHU VỰC THÀNH PHỐ HỒ CHÍ MINH SỬ DỤNG CÔNG NGHỆ VIỄN THÁM VÀ GIS

2.1. Đặc điểm khu vực nghiên cứu

Thành phố Hồ Chí Minh là một trong hai thành phố lớn nhất của nước ta, là trung tâm kinh tế, văn hoá - du lịch, giáo dục - khoa học kỹ thuật - y tế lớn của cả nước. Thành phố nằm trong toạ độ địa lý khoảng $10^{\circ}10' - 10^{\circ}38'$ vĩ độ bắc và $106^{\circ}22' - 106^{\circ}54'$ kinh độ đông. Phía Bắc giáp tỉnh Bình Dương, Tây Bắc giáp tỉnh Tây Ninh, Đông và Đông Bắc giáp tỉnh Đồng Nai, Đông Nam giáp tỉnh Bà Rịa - Vũng Tàu, Tây và Tây Nam giáp tỉnh Long An và Tiền Giang (Hình 2.1).



Hình 2.1. Vị trí địa lý Thành phố Hồ Chí Minh

2.2. Nghiên cứu lựa chọn các nhân tố ảnh hưởng đến sự biến động bề mặt không thấm

2.2.1. Lựa chọn bộ dữ liệu bổ sung

Trên cơ sở phân tích đặc điểm tự nhiên, xã hội và thực trạng dữ liệu ở khu vực Thành phố Hồ Chí Minh cùng kết quả đạt được trong các nghiên cứu trước đó, trong luận án này, nghiên cứu sinh tiến hành thu thập và xây dựng bộ dữ liệu đầu bổ sung của mô hình dự báo biến động bề mặt không thấm với 09 lớp dữ liệu, bao gồm: 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.

2.2.2. Đánh giá tương quan giữa bộ dữ liệu bổ sung và dữ liệu cơ sở

Trong luận án, hệ số tương quan Pearson (hệ số r) được sử dụng để lượng hóa mức độ chặt chẽ của mối quan hệ tuyến tính giữa dữ liệu về bề mặt không thấm (thông qua chỉ số đất xây dựng NDBI) và bộ dữ liệu bổ sung (09 lớp dữ liệu đầu vào). Kết quả xác định hệ số tương quan của các lớp dữ liệu bổ sung và chỉ số NDBI được thể hiện trong bảng 2.1.

Bảng 2.1. 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

STT	Lớp dữ liệu bổ sung	Hệ số tương quan	
		Chỉ số NDBI 2010	Chỉ số NDBI 2015
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

Phân tích kết quả đạt được cho thấy, các yếu tố địa hình như độ dốc và hướng dốc có mối tương quan yếu với bề mặt không thấm. Do vậy trong luận án lựa chọn bộ dữ liệu bổ sung cuối cùng cho mô hình dự báo biến động bề mặt không thấm khu vực thành phố Hồ Chí Minh bao gồm 07 lớp dữ liệu sau: (1) Mật độ che phủ của thực vật; (2) Nhiệt độ bề mặt; (3) Độ cao; (4) Khoảng cách tới đường giao thông; (5) Khoảng cách tới thủy hệ; (6) Mật độ dân cư; (7) Khoảng cách tới công trình đặc biệt.

2.3. Cơ sở khoa học phương pháp chiết xuất các lớp dữ liệu bổ sung cho mô hình dự báo biến động bề mặt không thấm

2.3.1. Mật độ che phủ thực vật

Trong luận án, nghiên cứu sinh lựa chọn sử dụng chỉ số thực vật NDVI để xác định mật độ che phủ thực vật (P_v) thông qua công thức sau [85], [133]:

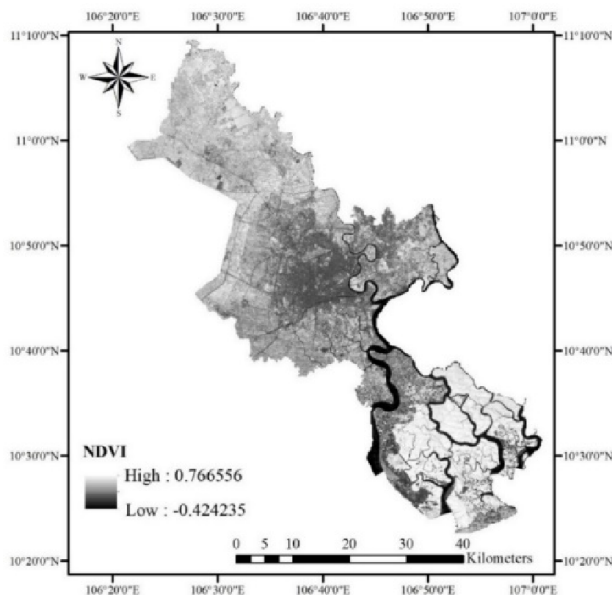
$$P_v = \frac{NDVI - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}} \quad (2.3)$$

Trong đó:

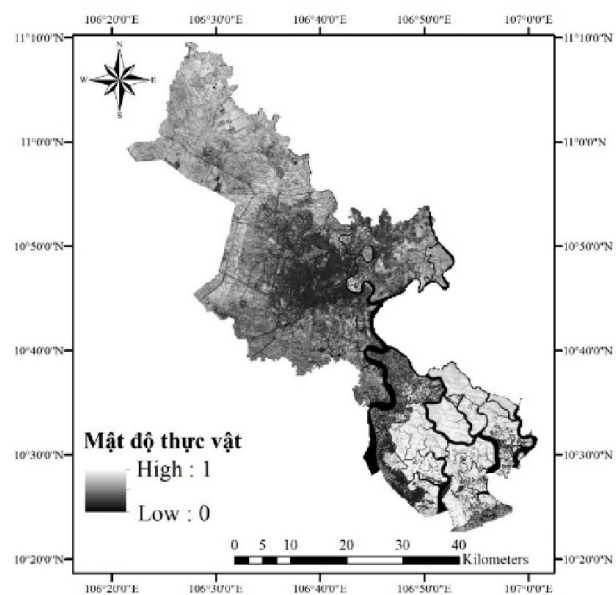
$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (2.4)$$

RED và NIR là giá trị phản xạ phổ tại các kênh đỏ và cận hồng ngoại [102].

Kết quả xác định chỉ số thực vật NDVI và mật độ che phủ thực vật khu vực Thành phố Hồ Chí Minh từ ảnh Landsat 8 ngày 23/2/2020 được thể hiện trên các hình 2.4, 2.5.



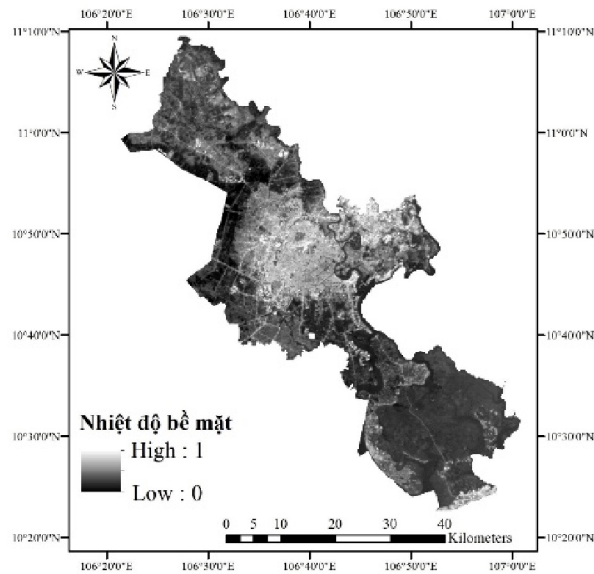
Hình 2.4. Chỉ số thực vật NDVI khu vực Thành phố Hồ Chí Minh, chiết xuất từ dữ liệu ảnh Landsat 8 ngày 23/2/2020



Hình 2.5. Kết quả xác định lớp dữ liệu mật độ che phủ thực vật khu vực Thành phố Hồ Chí Minh từ ảnh Landsat 8 ngày 23/2/2020

2.3.2. Nhiệt độ bề mặt

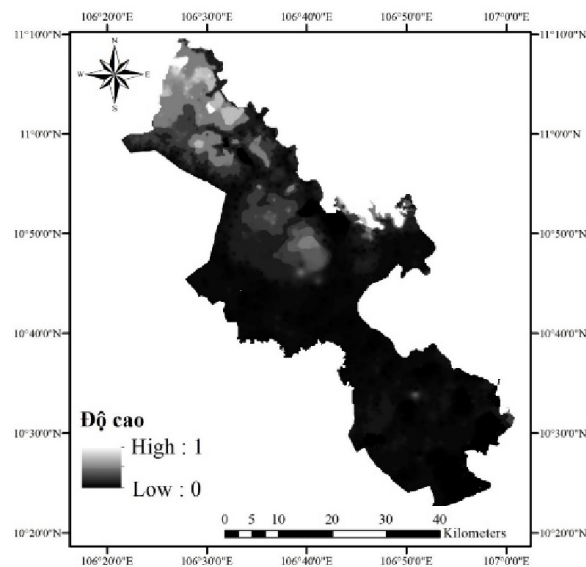
Kết quả xác định lớp thông tin nhiệt độ bề mặt khu vực Thành phố Hồ Chí Minh từ ảnh vệ tinh Landsat 8 ngày 23/2/2020 được thể hiện trên hình 2.6.



Hình 2.6. Kết quả xác định lớp dữ liệu nhiệt độ bề mặt khu vực TP. Hồ Chí Minh từ ảnh vệ tinh Landsat 8 ngày 23/2/2020 sau khi chuẩn hóa về thang giá trị 0-1

2.3.3. Độ cao

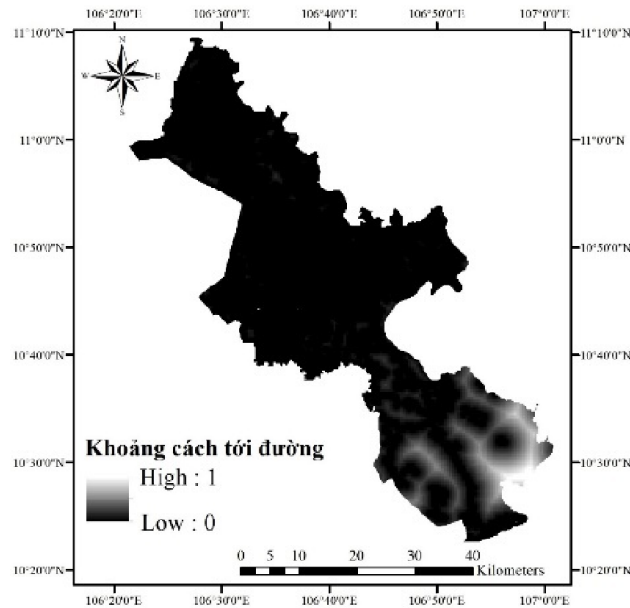
Kết quả xây dựng lớp dữ liệu về độ cao khu vực Thành phố Hồ Chí Minh được trình bày trên hình 2.7.



Hình 2.7. Lớp dữ liệu độ cao xây dựng từ CSDL nền địa lý quốc gia tỉ lệ 1:50 000 khu vực TP. Hồ Chí Minh sau khi chuẩn hóa về thang giá trị 0-1

2.3.4. Khoảng cách tới đường giao thông

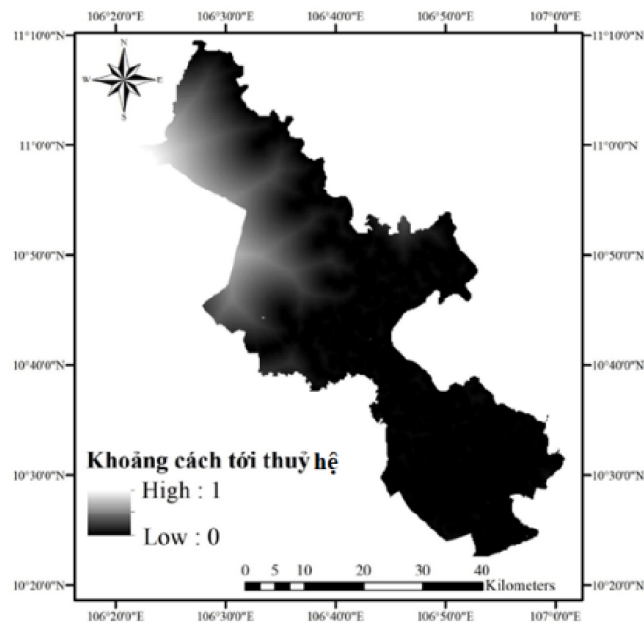
Sử dụng module Euclidean Distance tích hợp trong phần mềm ArcGIS, nghiên cứu sinh đã xây dựng lớp dữ liệu về khoảng cách tới đường giao thông khu vực Thành phố Hồ Chí Minh như hình 2.8.



Hình 2.8. Lớp dữ liệu khoảng cách tới đường giao thông khu vực TP. Hồ Chí Minh sau khi chuẩn hóa về thang giá trị 0-1

2.3.5. Khoảng cách tới thủy hệ

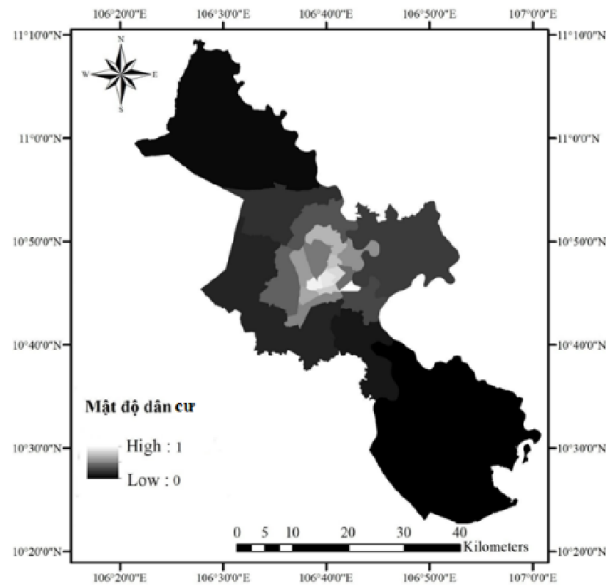
Kết quả xây dựng lớp dữ liệu khoảng cách tới thủy hệ khu vực Thành phố Hồ Chí Minh được trình bày trên hình 2.10.



Hình 2.10. Lớp dữ liệu khoảng cách tới thủy hệ khu vực TP. Hồ Chí Minh sau khi chuẩn hóa về thang giá trị 0-1

2.3.6. Dữ liệu mật độ dân cư

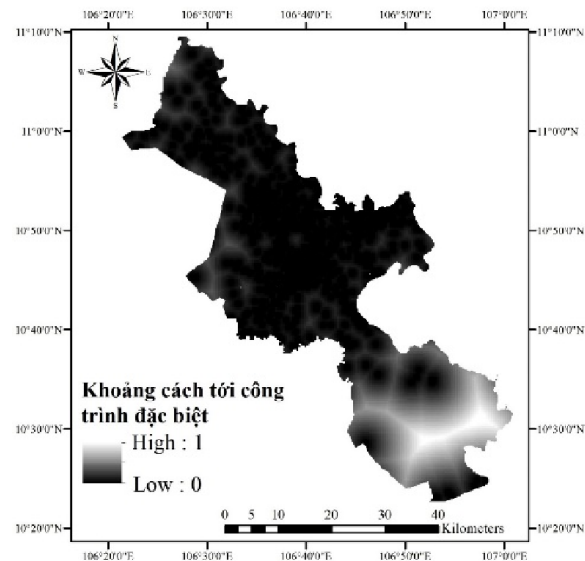
Kết quả xây dựng lớp dữ liệu về mật độ dân cư khu vực Thành phố Hồ Chí Minh được trình bày trên hình 2.12.



Hình 2.12. Lớp dữ liệu mật độ dân cư khu vực TP. Hồ Chí Minh sau khi chuẩn hóa về thang giá trị 0-1

2.3.7. Khoảng cách tới công trình đặc biệt

Khoảng cách tới công trình đặc biệt cũng được xây dựng bằng công cụ Euclidean Distance trong phần mềm ArcGIS 10 (hình 2.13).



Hình 2.13. Lớp dữ liệu khoảng cách tới công trình đặc biệt khu vực Thành phố Hồ Chí Minh sau khi chuẩn hóa về thang giá trị 0-1

2.4. Cơ sở khoa học phương pháp phân loại bề mặt không thấm từ dữ liệu viễn thám

2.4.1. Thuật toán xác suất cực đại

2.4.2. Thuật toán Random Forest

2.4.3. Thuật toán Support Vector Machine

2.4.4. Thuật toán Classification and Regression Tree

2.5. Mô hình Cellular Automata và các kỹ thuật học máy mô phỏng sự thay đổi bề mặt không thấm

2.5.1. Mô hình Cellular Automata

Mô hình CA được phát triển đầu tiên bởi Stanislaw Ulam vào những năm 1940 khi nghiên cứu các tinh thể và John Von Neumann ở lĩnh vực các hệ thống tự sao chép [118], [137]. Đến những năm 1950, mô hình CA bắt đầu được sử dụng trong việc mô phỏng hệ sinh thái. Về bản chất, mô hình CA đưa ra các dự báo về không gian với đầu vào là các xác suất chuyển đổi lớp phủ/sử dụng đất và trạng thái hiện tại của chúng. Một hệ CA bao gồm các ô (cell) không gian riêng biệt, trong đó mỗi cell đều có các trạng thái đặc trưng. Nhìn chung, một mô hình CA bao gồm 04 thành phần cơ bản sau [83]:

- (1): không gian lưới L mà mô hình hoạt động; (2): trạng thái cell Q trong không gian lưới.
- (3): quy tắc chuyển đổi f nhằm xác định quá trình chuyển đổi không gian của các đối tượng;
- (4): trạng thái của vùng lân cận Δ ảnh hưởng đến ô trung tâm.

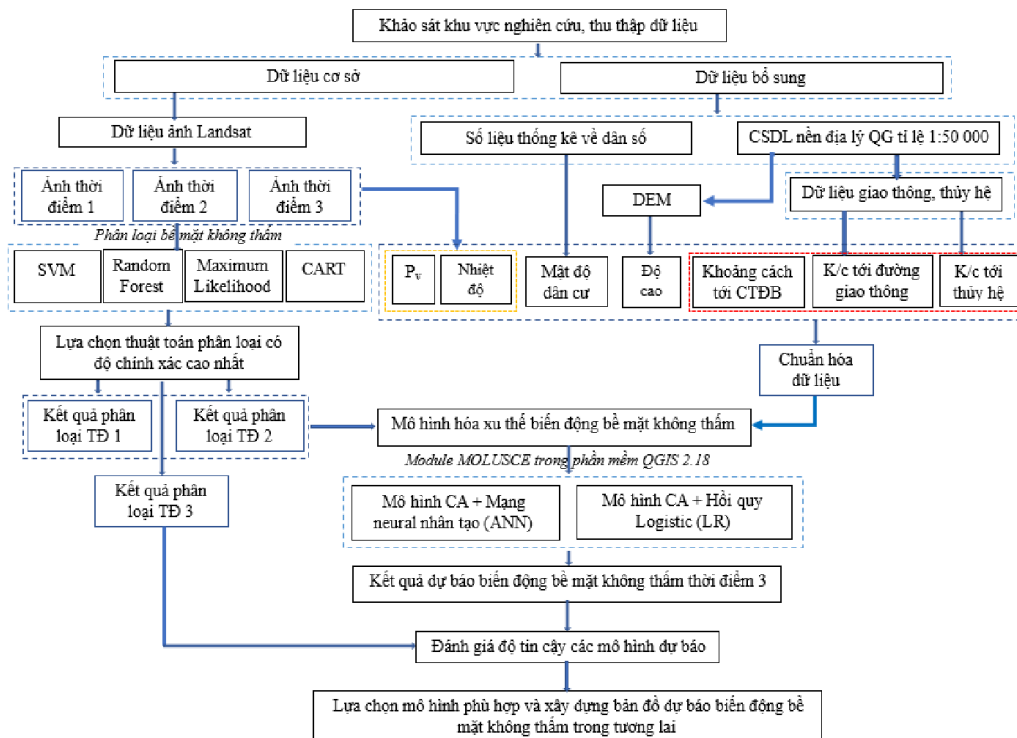
2.5.2. Mạng nơ ron nhân tạo (ANN)

Mạng nơ ron nhân tạo (Artificial Neural Network) là một mô hình tính toán được truyền cảm hứng từ bộ não con người [141]. Trong thời gian gần đây, ANN đã được các nhà nghiên cứu sử dụng trong các lĩnh vực trí tuệ nhân tạo như nhận dạng giọng nói, nhận dạng hình ảnh. ANN đã chứng minh tính hiệu quả trong dự báo các kết quả đầu ra khi so sánh với các mô hình thống kê trong các bài toán dựa trên hồi quy [38], [144], dự báo chuỗi thời gian [142].

2.6.3. Hồi quy Logistic (LR)

Hồi quy Logistic được sử dụng trong nhiều lĩnh vực khác nhau bao gồm học máy, các lĩnh vực y tế và khoa học xã hội. Mặc dù hồi quy Logistic được sử dụng chủ yếu với các biến phụ thuộc, nhưng kỹ thuật này thường được mở rộng cho các tình huống liên quan đến các biến mục tiêu có 3 kết quả trở lên [121].

2.6. Sơ đồ mô hình dự báo sự biến động bề mặt không thấm



Hình 2.232. Sơ đồ quy trình công nghệ lựa chọn mô hình dự báo sự biến động bề mặt không thấm từ dữ liệu viễn thám và GIS

2.7. Tiểu kết chương 2

Trong chương 2, nghiên cứu sinh đã phân tích, đánh giá đặc điểm khu vực thực nghiệm (Thành phố Hồ Chí Minh) và lựa chọn bộ dữ liệu đầu vào cho mô hình dự báo biến động bề mặt không thấm thông qua so sánh hệ số tương quan giữa các lớp dữ liệu bổ sung và chỉ số đất xây dựng NDBI xác định từ ảnh vệ tinh Landsat năm 2010 và 2015. Trên cơ sở đó, luận án đã lựa chọn 03 cảnh ảnh Landsat ở 3 thời điểm làm bộ dữ liệu cơ sở và 07 lớp dữ liệu bổ sung để mô hình hóa quá trình phát triển bề mặt không thấm, bao gồm: mật độ dân số, mật độ che phủ thực vật, nhiệt độ bề mặt, độ cao, khoảng cách tới công trình đặc biệt, khoảng cách tới đường giao thông và khoảng cách tới thủy hệ.

Tất cả các lớp dữ liệu bổ sung được biên tập và nội suy về độ phân giải không gian 30m để thống nhất với bộ dữ liệu cơ sở (ảnh Landsat có độ phân giải không gian 30m), sau đó chuẩn hóa về giá trị 0 – 1 để đưa vào mô hình dự báo xu hướng phát triển bề mặt không thấm.

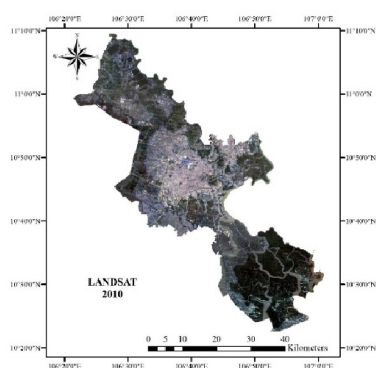
Để dự báo phân bố bề mặt không thấm trong tương lai, trong luận án đã lựa chọn và đề xuất kết hợp mô hình toán học Cellular Automata và 02 kỹ thuật học máy (ANN, hồi quy Logistic). Ảnh vệ tinh thời điểm 1 và 2 cùng bộ dữ liệu bổ sung được sử dụng để dự báo phân bố bề mặt không thấm cho thời điểm 3, sau đó so sánh với kết quả phân loại từ ảnh Landsat thời điểm 3 để lựa chọn mô hình phù hợp và có độ chính xác cao nhất. Mô hình sau khi hoàn thiện được sử dụng để dự báo biến động bề mặt không thấm trong tương lai.

CHƯƠNG 3. THỰC NGHIỆM DỰ BÁO BIẾN ĐỘNG BỀ MẶT KHÔNG THẤM KHU VỰC THÀNH PHỐ HỒ CHÍ MINH

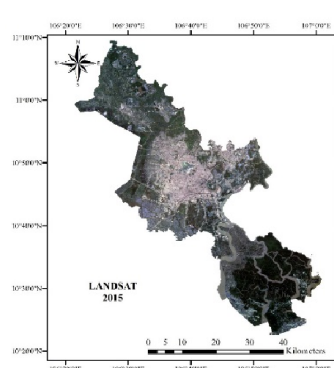
3.1. Đặc điểm dữ liệu sử dụng

3.1.1. Dữ liệu viễn thám

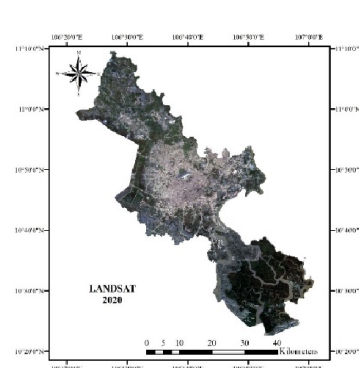
Dữ liệu viễn thám sử dụng trong nghiên cứu là ảnh vệ tinh Landsat chụp khu vực Thành phố Hồ Chí Minh, bao gồm 03 cảnh ảnh ngày 11/02/2010 (Landsat 5 TM), 9/02/2015 và 23/02/2020 (Landsat 8 OLI_TIRS).



Hình 3.1. Ảnh vệ tinh Landsat TM ngày 11/2/2010

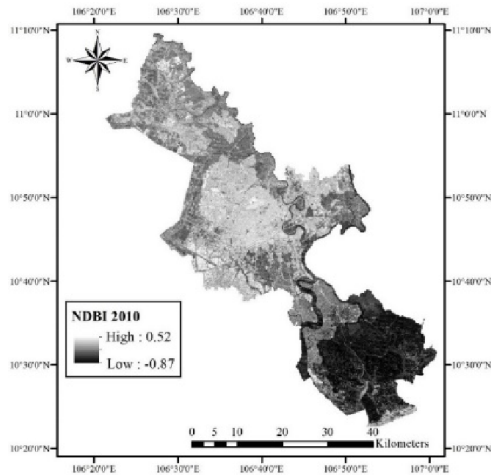


Hình 3.2. Ảnh vệ tinh Landsat 8 OLI_TIRS ngày 9/02/2015

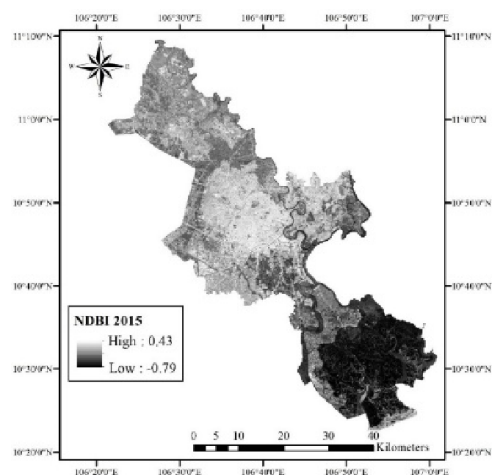


Hình 3.3. Ảnh vệ tinh Landsat 8 OLI_TIRS ngày 23/2/2020

Để hỗ trợ lựa chọn các lớp dữ liệu đầu vào của mô hình dự báo biến động bề mặt không thấm, ảnh Landsat năm 2010 và 2015 được sử dụng để tính chỉ số đất xây dựng NDBI, sau đó xác định hệ số tương quan với bộ dữ liệu bổ sung. Kết quả xác định chỉ số NDBI từ ảnh Landsat năm 2010 và 2015 được trình bày trên các hình 3.4 và 3.5.



Hình 3.4. Chỉ số NDBI xác định từ ảnh Landsat năm 2010



Hình 3.5. Chỉ số NDBI xác định từ ảnh Landsat năm 2015

3.1.2. Dữ liệu khác

Ngoài dữ liệu viễn thám, trong luận án sử dụng CSDL nền địa lý quốc gia tỉ lệ 1:50 000 khu vực Thành phố Hồ Chí Minh để xây dựng mô hình số độ cao (DEM) và chiết xuất dữ liệu về hệ thống đường giao thông, thủy hệ.

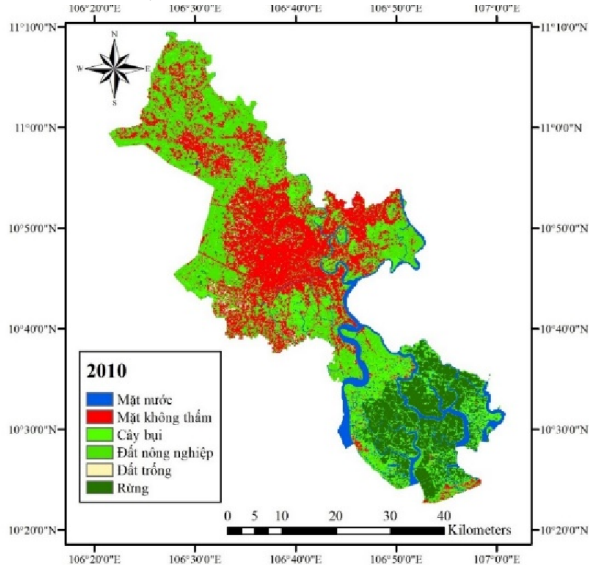
Dữ liệu mật độ dân số được thu thập từ kết quả điều tra dân số, lao động năm 2021 của Cục Thống kê thành phố Hồ Chí Minh [146], sau đó xây dựng lớp dữ liệu dạng raster để đưa vào mô hình.

Dữ liệu về vị trí các công trình đặc biệt được thu thập tại CSDL <https://opendata.hcmgis.vn/> [150].

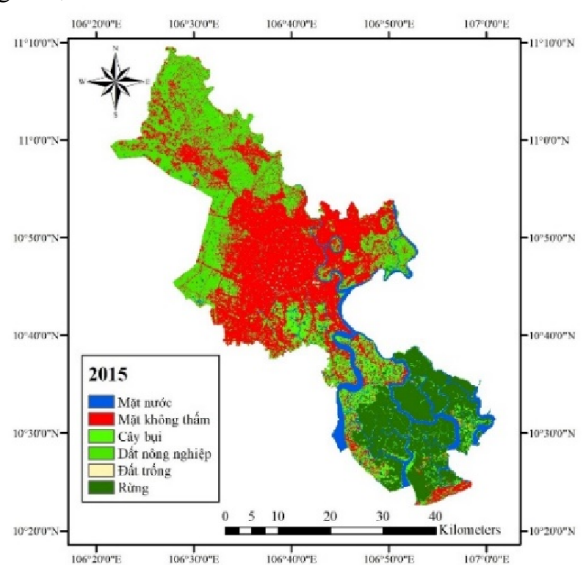
3.2. Kết quả phân loại bề mặt không thám từ ảnh vệ tinh Landsat đa thời gian

3.2.1. Kết quả phân loại sử dụng thuật toán RF

Trên hình 3.6 và hình 3.7 trình bày kết quả phân loại lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh từ ảnh vệ tinh Landsat năm 2010 và năm 2015 bằng thuật toán RF



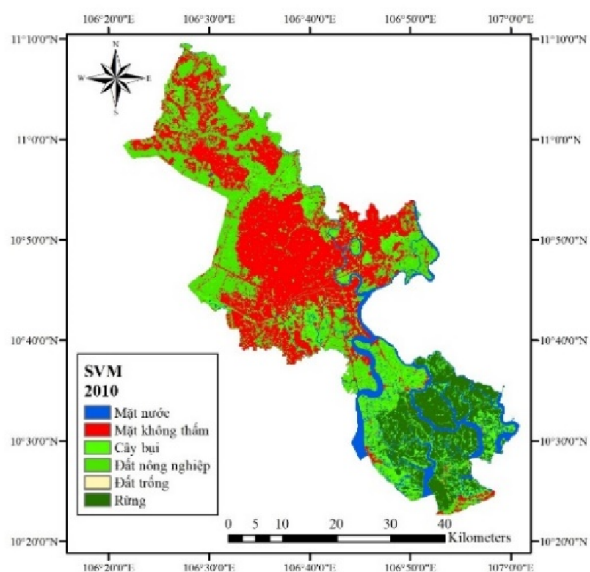
Hình 3.6. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2010 bằng thuật toán RF



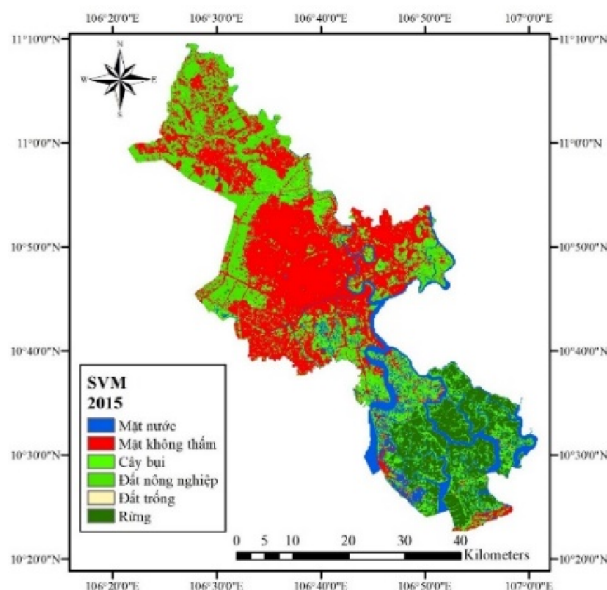
Hình 3.7. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2015 bằng thuật toán RF

3.2.2. Kết quả phân loại sử dụng thuật toán SVM

Kết quả phân loại lớp phủ/sử dụng đất khu vực thành phố Hồ Chí Minh từ ảnh vệ tinh Landsat năm 2010 và năm 2015 bằng thuật toán SVM được trình bày trên hình 3.8 và hình 3.9



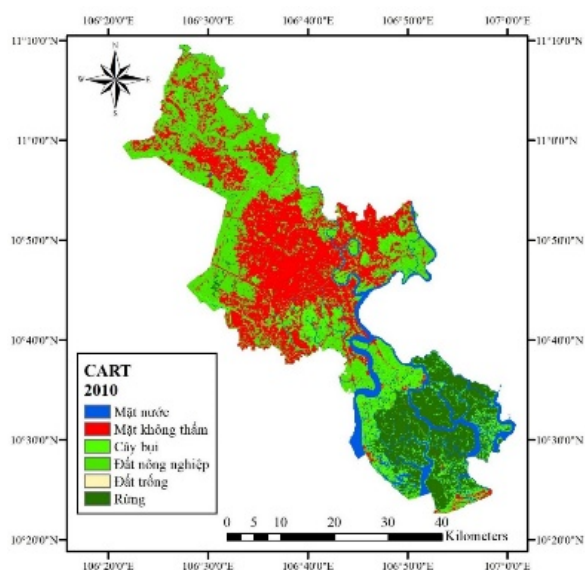
Hình 3.8. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2010 bằng thuật toán SVM



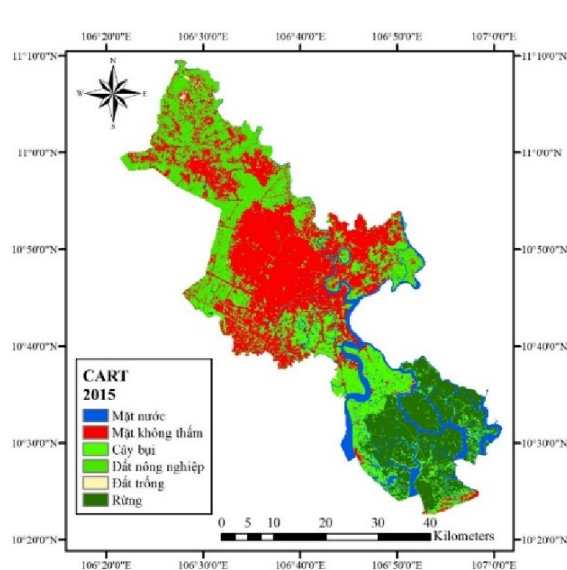
Hình 3.9. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2015 bằng thuật toán SVM

3.2.3. Kết quả phân loại sử dụng thuật toán CART

Bên cạnh 02 kỹ thuật học máy quen thuộc (RF và SVM), trong luận án cũng thử nghiệm phân loại lớp phủ/sử dụng đất bằng thuật toán Cây phân loại và Hồi quy (Classification and Regression Tree - CART). Kết quả phân loại lớp phủ/sử dụng đất khu vực thực nghiệm từ ảnh vệ tinh Landsat năm 2010, năm 2015 được trình bày trên hình 3.10 và hình 3.11.



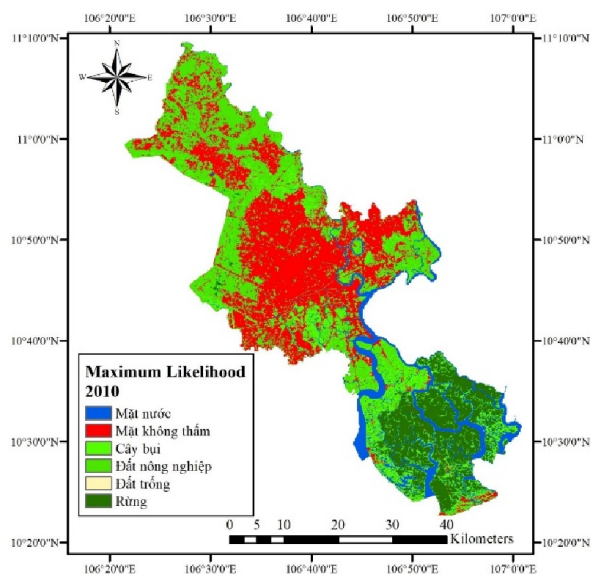
Hình 3.10. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2010 bằng thuật toán CART



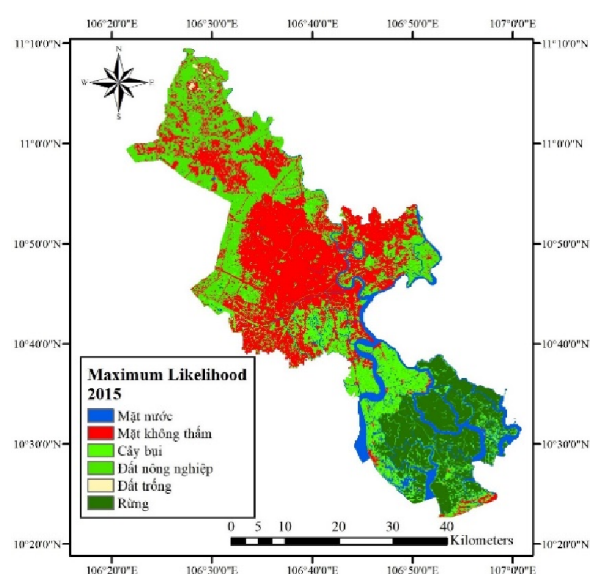
Hình 3.11. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2015 bằng thuật toán CART

3.2.4. Kết quả phân loại sử dụng thuật toán Maximum Likelihood

Bên cạnh 03 thuật toán học máy, trong luận án cũng thử nghiệm phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh bằng thuật toán phân loại có giám sát quen thuộc – xác suất cực đại. Đây được xem là thuật toán phân loại chuẩn trong phân loại dựa trên điểm ảnh, có độ chính xác cao khi so sánh với các thuật toán khác như khoảng cách ngắn nhất, hình hộp. Kết quả phân loại lớp phủ/sử dụng đất khu vực nghiên cứu từ ảnh Landsat năm 2010 và 2015 bằng thuật toán xác suất cực đại được thể hiện trên các hình 3.12 và 3.13.



Hình 3.12. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2010 bằng thuật toán Maximum Likelihood



Hình 3.13. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2015 bằng thuật toán Maximum Likelihood

3.2.5. Đánh giá và lựa chọn phương pháp phân loại phù hợp

Để đánh giá và lựa chọn phương pháp phân loại lớp phủ/sử dụng đất phù hợp với đặc điểm khu vực nghiên cứu, ngoài phân tích, so sánh độ chính xác phân loại từng đối tượng lớp phủ riêng biệt, trong luận án cũng sử dụng các thông số khác bao gồm độ chính xác tổng thể và chỉ số Kappa. Kết quả so sánh, đánh giá độ chính xác phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh từ ảnh Landsat năm 2010 và 2015 được trình bày trong các bảng 3.17 và 3.18.

Bảng 3.17. Độ chính xác tổng thể và chỉ số Kappa của các thuật toán phân loại đối với ảnh Landsat khu vực nghiên cứu năm 2010

Thuật toán	SVM	RF	CART	ML
Độ chính xác tổng thể	86.90%	93.44%	90.77%	85.98%
Chỉ số Kappa	0.871	0.919	0.887	0.829

Bảng 3.18. Độ chính xác tổng thể và chỉ số Kappa của các thuật toán phân loại đối với ảnh Landsat khu vực nghiên cứu năm 2015

Thuật toán	SVM	RF	CART	ML
Độ chính xác tổng thể	88.72%	93.54%	89.81%	87.29%
Chỉ số Kappa	0.863	0.921	0.876	0.845

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 các kết quả trên, trong luận án tiến hành lựa chọn thuật toán RF để phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh từ ảnh Landsat đa thời gian, làm cơ sở để xây dựng các lớp thông tin đầu vào cho mô hình dự báo phân bố bề mặt không thấm trong tương lai sử dụng mô hình Cellular Automata và các kỹ thuật học máy. Kết quả phân loại các năm 2010 và 2015 được dùng để dự báo phân bố bề mặt không thấm năm 2020, sau đó so sánh với kết quả phân loại lớp phủ/sử dụng đất năm 2020 bằng thuật toán RF để hiệu chỉnh và lựa chọn mô hình phù hợp với khu vực thực nghiệm.

3.3. Kết quả dự báo phân bố bề mặt không thấm khu vực Thành phố Hồ Chí Minh

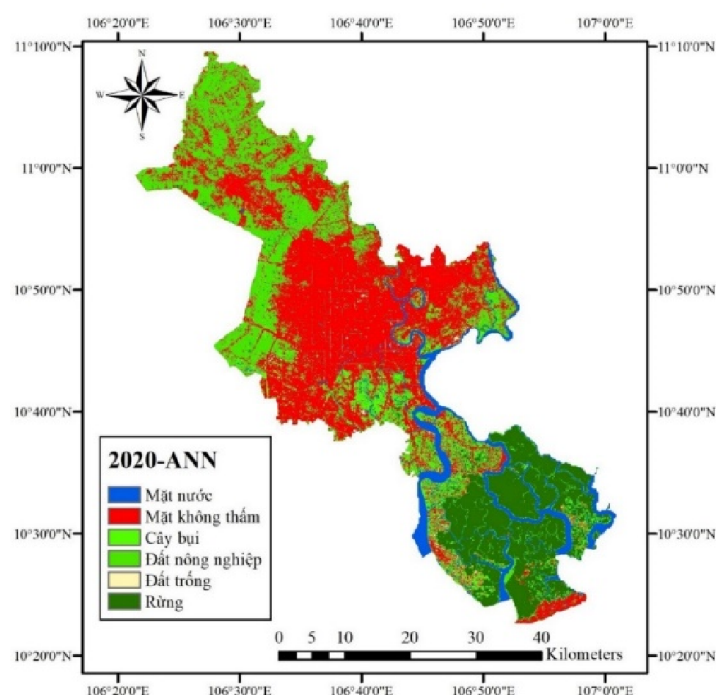
3.3.1. Kết quả dự báo biến động mặt không thấm năm 2020

3.3.1.1. Sử dụng mô hình CA kết hợp với thuật toán ANN

Để dự báo phân bố bề mặt không thấm khu vực Thành phố Hồ Chí Minh năm 2020 từ dữ liệu viễn thám và GIS, luận án đã sử dụng mô hình CA kết hợp thuật toán ANN trong module MOLUSCE của phần mềm QGIS 2.18.

Trong nghiên cứu tiến hành khởi tạo 2000 mẫu ngẫu nhiên và thiết lập bộ tham số bao gồm : vùng lân cận - neighbourhood (1 px), tỉ lệ học - learning rate (0.001), số vòng lặp - iterations (1000), lớp ẩn - hidden layers (12) và quán tính - momentum (0.05). Với bộ tham số thiết lập trong luận án, quá trình học tập dữ liệu đạt giá trị Kappa là 0.754 khi so sánh tương quan với tập dữ liệu huấn luyện.

Kết quả mô hình hoá quá phân bố lớp phủ/sử dụng đất khu vực nghiên cứu năm 2020 bằng mô hình CA kết hợp với thuật toán ANN được trình bày trong hình 3.16



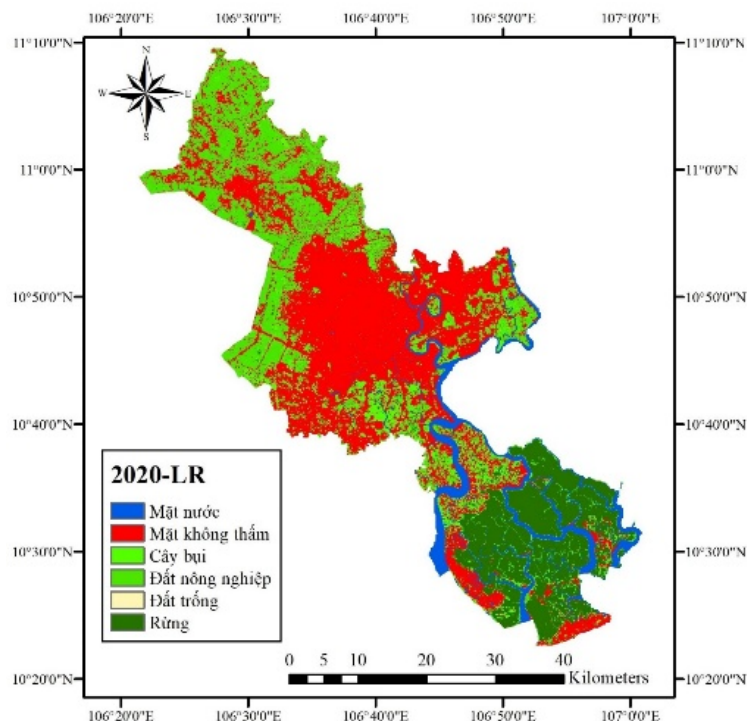
Hình 3.16. Kết quả dự báo phân bố bề mặt không thấm năm 2020 khu vực Thành phố Hồ Chí Minh bằng mô hình CA kết hợp với thuật toán ANN

3.3.1.2. Sử dụng mô hình CA kết hợp với thuật toán hồi quy Logistic (LR)

Với thuật toán hồi quy Logistic, bộ tham số đầu vào được thiết lập trong luận án bao gồm : số lượng mẫu - number of samples (2000), vùng lân cận - neighbourhood (1 px), số vòng lặp tối đa - maximum iterations (100). Bộ tham số này cũng được lựa chọn khi thử nghiệm với nhiều giá trị tham số khác nhau, sau đó lựa chọn bộ tham số có độ chính xác trong học tập dữ liệu (huấn luyện) cao nhất. Với bộ tham số thiết lập

trong luận án, quá trình học tập dữ liệu đạt giá trị Kappa là 0.731 khi so sánh tương quan với tập dữ liệu huấn luyện.

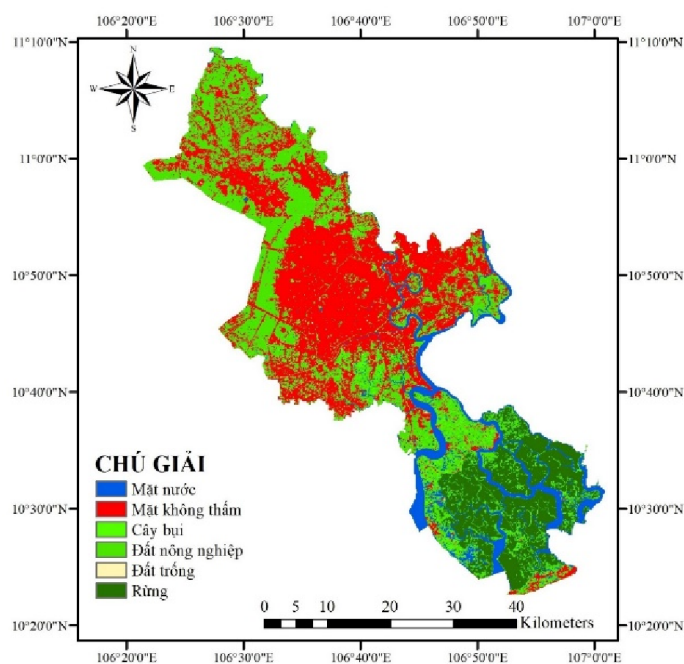
Kết quả mô hình hoá lớp phủ/sử dụng đất đô thị khu vực Thành phố Hồ Chí Minh năm 2020 sau khi sử dụng mô hình CA kết hợp với thuật toán Logistic Regression được trình bày trong hình 3.17.



Hình 3.17. Kết quả dự báo phân bố bề mặt không thấm năm 2020 khu vực TP. Hồ Chí Minh bằng mô hình CA kết hợp với thuật toán hồi quy Logistic

3.3.2. Đánh giá và hoàn thiện mô hình dự báo biến động bề mặt không thấm

Nhằm đánh giá độ chính xác của các mô hình mô phỏng biến động bề mặt đô thị khu vực Thành phố Hồ Chí Minh 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 (hình 3.18) để so sánh mức độ tương quan với kết quả mô hình hóa năm 2020.



Hình 3.18. Kết quả phân loại lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh năm 2020 bằng thuật toán RF

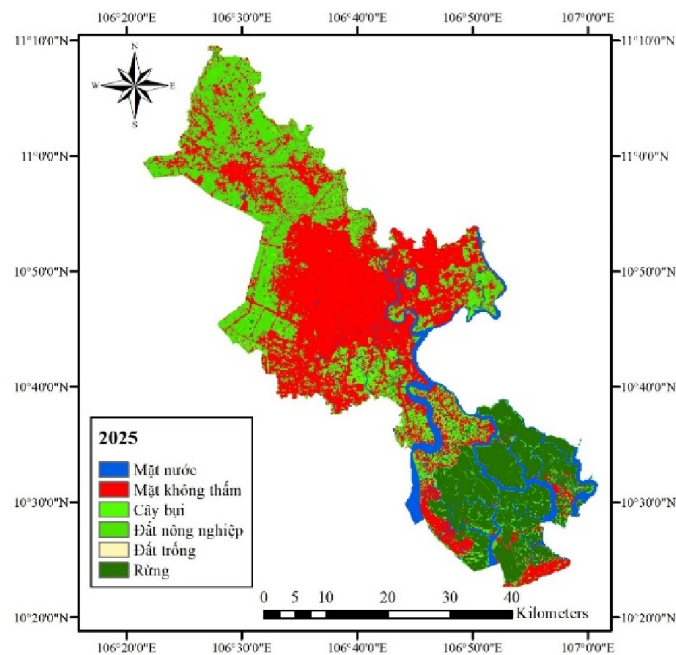
Kết quả nhận được cho thấy, hệ số tương quan giữa lớp phủ/sử dụng đất được mô hình hóa bằng mô hình CA kết hợp với thuật toán ANN và kết quả phân loại năm 2020 đạt 0.715. Trong khi đó, khi sử dụng mô hình CA kết hợp với thuật toán hồi quy Logistic, hệ số tương quan khi so sánh với kết quả phân loại đạt 0.666.

Như vậy, phương pháp kết hợp mô hình CA và thuật toán ANN cho độ chính xác cao hơn trong mô phỏng lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh khi so sánh với thuật toán hồi quy Logistic. Từ kết quả này, trong luận án lựa chọn phương pháp kết hợp mô hình CA và thuật toán ANN để mô phỏng sự thay đổi bề mặt không thấm khu vực Thành phố Hồ Chí Minh trong tương lai.

3.3.3. Kết quả dự báo biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh trong tương lai

3.3.3.1. Dự báo phân bố bề mặt không thấm đến năm 2025

Bản đồ mô phỏng lớp phủ/sử dụng đất, trong đó có bề mặt không thấm khu vực Thành phố Hồ Chí Minh năm 2025 được trình bày trên hình 3.19.

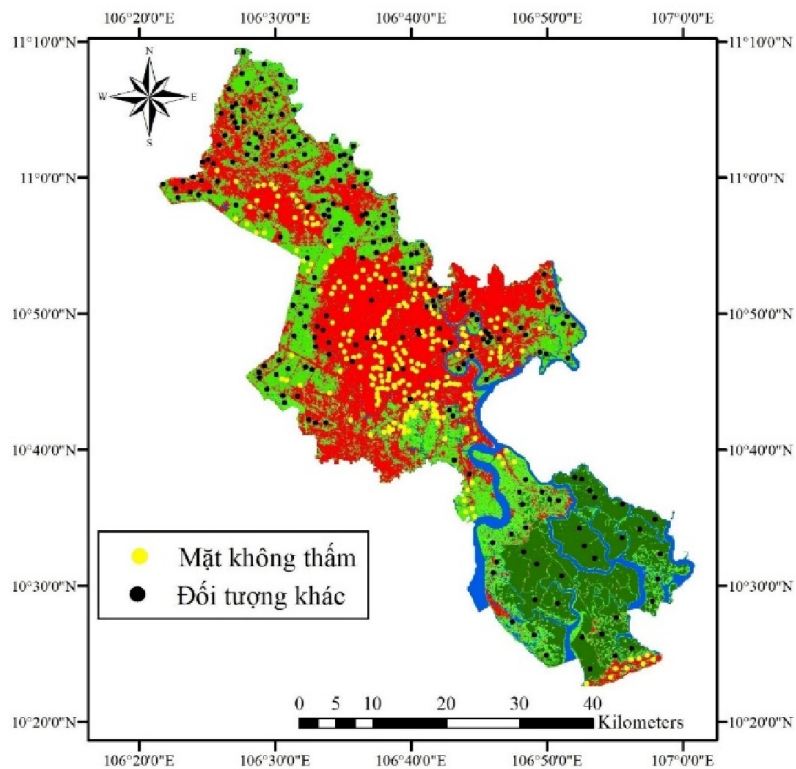


Hình 3.19. Kết quả dự báo phân bố bề mặt không thấm khu vực thành phố Hồ Chí Minh năm 2025 bằng mô hình CA kết hợp với thuật toán ANN

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 3.21.

Bảng 3.21. Diện tích của các đối tượng giai đoạn 2010 - 2025

Năm Diện tích (km²)	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



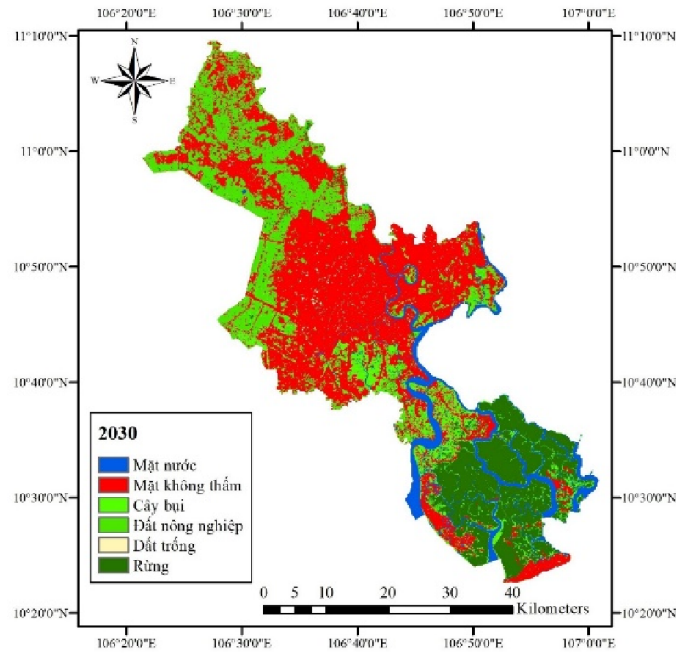
Hình 3.22. Vị trí các điểm lấy mẫu nhằm so sánh kết quả dự báo phân bố bề mặt không thấm năm 2025 và bản đồ quy hoạch năm 2025

Để xác định mức độ tương quan giữa kết quả dự báo và bản đồ quy hoạch, nghiên cứu sinh 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) – Hình 3.22. Kết quả nhận được cho thấy, trong 242 điểm bề mặt không thấm, có 211 điểm được dự báo 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 luận án cho phép dự báo xu hướng phát triển bề mặt không thấm khu vực Thành phố Hồ Chí Minh với độ tin cậy cao.

3.3.3.2. Dự báo phân bố bề mặt không thấm đến năm 2030

Sau khi dự báo phân bố bề mặt không thấm năm 2025, kết quả này tiếp tục được sử dụng để mô hình hóa và dự báo phân bố bề mặt không thấm khu vực Thành phố Hồ Chí Minh đến năm 2030 bằng mô hình CA kết hợp với thuật toán ANN (hình 3.23).

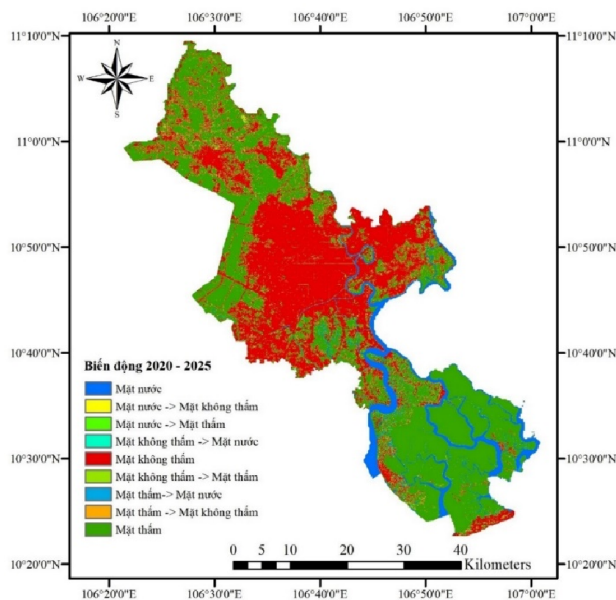
Theo kết quả mô hình hóa đến năm 2030, diện tích bề mặt không thấm khu vực thành phố Hồ Chí Minh đạt 833.241 km², tăng 8.165 km² so với năm 2025 (tương đương 1.00%), tốc độ gia tăng bề mặt không thấm cơ bản giống như với giai đoạn 2020 – 2025 và không nhiều do mật độ xây dựng có sự bão hòa sau năm 2025.



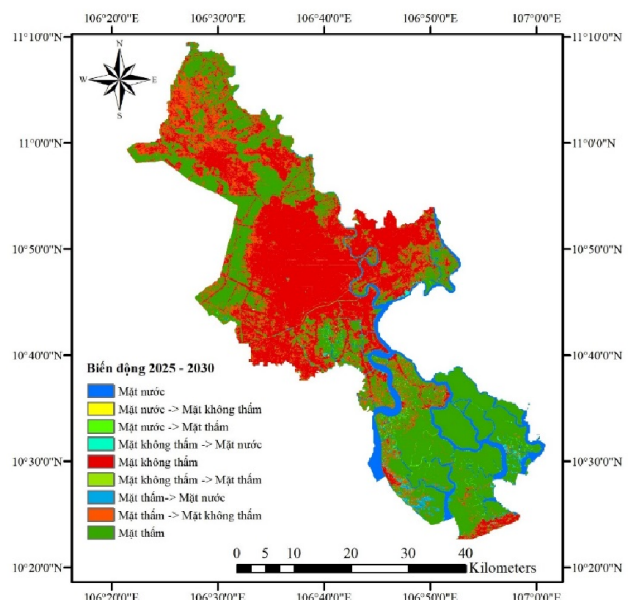
Hình 3.23. Kết quả dự báo phân bố bề mặt không thấm khu vực Thành phố Hồ Chí Minh năm 2030 bằng mô hình CA kết hợp với thuật toán ANN

3.3.3.3. Đánh giá biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh

Trên hình 3.25 và 3.26 trình bày kết quả đánh giá biến động bề mặt không thấm khu vực thành phố Hồ Chí Minh các giai đoạn 2020 – 2025 và 2025 – 2030, trong đó xem xét sự chuyển đổi qua lại giữa 03 đối tượng: bề mặt không thấm, bề mặt thấm và mặt nước. Kết quả phân tích biến động diện tích qua lại giữa 03 đối tượng này được thể hiện trong bảng 3.22.



Hình 3.25. Biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh giai đoạn 2020 – 2025



Hình 3.26. Biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh giai đoạn 2025 – 2030

Bảng 3.22. Diện tích biến động của các đối tượng giai đoạn 2020 - 2025 và 2025 - 2030

Biến động (km²)	2020 - 2025	2025 - 2030
Mặt nước -> Mặt không thấm	1.177	0.381
Mặt nước -> Mặt thấm	90.795	25.789
Mặt không thấm -> Mặt nước	1.384	2.668
Mặt không thấm -> Mặt thấm	18.477	39.492
Mặt thấm -> Mặt nước	13.410	20.736
Mặt thấm -> Mặt không thấm	101.954	179.820

3.4. Tiểu kết chương 3

Trong chương 3, nghiên cứu sinh đã thử nghiệm với các mô hình khác nhau trong mô phỏng biến động lớp phủ/sử dụng đất, trong đó có bề mặt không thấm 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. Kết quả nhận được cho thấy, thuật toán RF có độ chính xác cao nhất khi phân loại lớp phủ/sử dụng đất khu vực nghiên cứu.

Để dự báo xu thế biến động bề mặt không thấm, trong nghiên cứu 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 mô phỏng sự phát triển lớp phủ bề mặt đô thị năm 2020 bằng mô hình CA và 02 kỹ thuật học máy: 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 phương pháp kết hợp mô hình CA với 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 kết hợp mô hình CA 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 dự báo phân bố bề mặt không thấm năm 2025 cũng được so sánh với bản đồ quy hoạch sử dụng đất khu vực Thành phố Hồ Chí Minh đến năm 2025, trong đó với 242 điểm đại diện cho bề mặt không thấm, có 211 điểm được dự báo chính xác (87,19%), độ chính xác tổng thể và chỉ số Kappa đạt tương ứng 85,70% và 0,715. Biến động giữa bề mặt không thấm và các đối tượng khác (bề mặt thấm, mặt nước) cũng được đánh giá thông qua các bản đồ biến động bề mặt không thấm giai đoạn 2020 - 2025 và 2025 - 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.

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết luận

Qua quá trình nghiên cứu, trong luận án đã đề xuất và lựa chọn được mô hình phù hợp nhằm mô phỏng xu hướng phát triển lớp phủ/sử dụng đất, trong đó có bề mặt không thấm khu vực Thành phố Hồ Chí Minh trong tương lai từ dữ liệu viễn thám đa thời gian và GIS cùng các kỹ thuật trí tuệ nhân tạo. Từ những kết quả đạt được trong luận án, nghiên cứu sinh rút ra các kết luận sau:

1. Dữ liệu viễn thám và GIS có thể sử dụng hiệu quả phục vụ xây dựng các lớp thông tin đầu vào cho mô hình dự báo biến đổi bề mặt không thấm. Trong luận án, ngoài dữ liệu viễn thám đa thời gian (03 cảnh ảnh

vệ tinh Landsat các năm 2010, 2015 và 2020), 7 lớp thông tin bổ sung bao 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 được sử dụng để mô hình hóa sự thay đổi trong lớp phủ/sử dụng đất. Các lớp thông tin này được chiết xuất và xây dựng từ dữ liệu viễn thám và GIS cùng các dữ liệu về tự nhiên, kinh tế - xã hội, sau đó được chuẩn hóa trước khi đưa vào mô hình dự báo biến động bề mặt không thấm.

2. Trong luận án đã thử nghiệm với 04 phương pháp phân loại lớp phủ/sử dụng đất, bao gồm phương pháp xác suất cực đại và 03 kỹ thuật học máy (RF, SVM, CART) để phân tích, đánh giá và lựa chọn phương pháp có độ chính xác cao nhất. Kết quả nhận được cho thấy, với khu vực thử nghiệm tại Thành phố Hồ Chí Minh, thuật toán RF cho độ chính xác cao nhất khi phân loại lớp phủ/sử dụng đất, thể hiện qua so sánh độ chính xác tổng thể và chỉ số Kappa (đều đạt trên 0,9). Từ kết quả này, thuật toán RF được lựa chọn để phân loại lớp phủ/sử dụng đất khu vực nghiên cứu bằng ảnh vệ tinh Landsat giai đoạn 2010 - 2020, làm cơ sở để mô hình hóa lớp phủ/sử dụng đất năm 2025 và 2030 khu vực Thành phố Hồ Chí Minh.

3. Phân tích kết quả nhận được cho thấy, phương pháp kết hợp mô hình toán học Cellular Automata (CA) và kỹ thuật học máy (ANN) cho phép mô phỏng lớp phủ/sử dụng đất khu vực Thành phố Hồ Chí Minh với độ chính xác cao, trong đó hệ số tương quan khi so sánh kết quả mô hình hóa và kết quả phân loại lớp phủ/sử dụng đất năm 2020 đạt 0.715. Từ kết quả này, trong luận án đã tiến hành mô phỏng biến động bề mặt không thấm khu vực Thành phố Hồ Chí Minh năm 2025 và 2030. Theo đó, diện tích bề mặt không thấm có sự gia tăng đáng kể, trong đó kết quả mô phỏng đến năm 2025 đạt 825.076 km², gần hơn 1.3 lần so với năm 2010 (616.983 km²). Diện tích bề mặt không thấm tiếp tục gia tăng khoảng 1.00% đến năm 2030 theo kết quả dự báo, đạt 833.241 km². Ngoài khu vực trung tâm Thành phố và huyện Củ Chi, bề mặt không thấm năm 2025 và 2030 có sự phát triển đáng kể ở khu vực huyện Cần Giờ (đoạn giáp biển và các xã phía bắc của huyện như Bình Khánh, An Thới Đông, Tam Thôn Hiệp).

4. Mô hình dự báo biến động lớp phủ/sử dụng đất, trong đó có bề mặt không thấm từ dữ liệu viễn thám và GIS trên cơ sở các kỹ thuật học máy xây dựng trong luận án là một công cụ hiệu quả, cung cấp thông tin khách quan và chính xác giúp các nhà quản lý trong công tác quy hoạch, phát triển bền vững khu vực đô thị ở Thành phố Hồ Chí Minh. Mô hình này có thể áp dụng cho các khu vực khác có điều kiện tự nhiên, kinh tế - xã hội tương tự.

2. Kiến nghị

1. Do điều kiện về mặt dữ liệu, một số lớp dữ liệu đầu vào của mô hình dự báo biến động bề mặt không thấm vẫn còn được khai thác từ các CSDL quốc tế. Ngoài ra, số lượng các lớp dữ liệu đầu vào vẫn còn chưa thực sự đầy đủ các yếu tố ảnh hưởng đến sự thay đổi bề mặt không thấm. Để hoàn thiện hơn mô hình, trong tương lai nghiên cứu sinh sẽ tiếp tục nghiên cứu, thu thập các lớp dữ liệu đầu vào sát với điều kiện thực tiễn khu vực thực nghiệm, từ đó nâng cao độ chính xác của mô hình dự báo.

2. Áp dụng các mô hình dự báo biến động lớp phủ/sử dụng đất, trong đó có bề mặt không thấm cho các khu vực đô thị để hỗ trợ các nhà quản lý trong công tác quy hoạch và phát triển bền vững đô thị.

CÁC CÔNG TRÌNH KHOA HỌC ĐÃ CÔNG BỐ

1. Phạm Văn Tùng, Nguyễn Văn Trung, Nguyễn Hữu Long, Nguyễn Đức Hùng (2018). *Quan trắc sự mở rộng bề mặt không thám sử dụng dữ liệu ảnh vệ tinh SPOT-5 và Sentinel-2 ở khu vực thành phố Hồ Chí Minh*. Tạp chí khoa học kỹ thuật Mỏ - Địa chất, 59, 2, 69-76
2. Phạm Văn Tùng, Nguyễn Văn Trung, Vũ Xuân Cường, Nguyễn Văn Sơn (2020). *Thành lập bản đồ bề mặt không thám sử dụng dữ liệu ảnh vệ tinh Sentinel-2 ở khu vực thành phố Hồ Chí Minh*. Hội nghị Toàn quốc Khoa học Trái đất và Tài nguyên với Phát triển bền vững - ERSD, 86-93,
3. Nguyễn Văn Trung, Phạm Văn Tùng, Nguyễn Thanh Bình, Phạm Ngọc Quân, Phan Văn Khoái, Đỗ Thanh Phong, Nguyễn Thanh Tuấn, Huỳnh Tấn Phước, Nguyễn Thị Thùy Linh (2021). *Mối quan hệ giữa hiện tượng đảo nhiệt đô thị và mật độ dân số các quận và huyện ở thành phố Hồ Chí Minh lấy từ dữ liệu ảnh vệ tinh hồng ngoại nhiệt*. Hội nghị khoa học Quốc gia về Công nghệ Địa không gian trong Khoa học trái đất và Môi trường, 231-240
4. Phạm Văn Tùng (2021). *Thành lập bản đồ bề mặt không thám sử dụng dữ liệu ảnh vệ tinh sentinel-2 ở khu vực Thành phố Hồ Chí Minh*. Tạp chí Khoa học Tài nguyên và Môi trường, Số 38, 30-38.
5. Lê Thị Thu Hà, Nguyễn Văn Trung, Nguyễn Ngọc Khoa, Nguyễn Đăng Phương, Võ Thị Tuyết, Nguyễn Giang Thọ, Phạm Văn Tùng và Lê Văn Định (2021), *Nghiên cứu mối quan hệ giữa đảo nhiệt đô thị và mật độ dân số các quận, huyện ở Thành phố Hồ Chí Minh bằng dữ liệu ảnh Sentinel - 3 SLSTR*. Tạp chí Khoa học kỹ thuật Mỏ - Địa chất, 62, 5, 67-75.
6. Phạm Văn Tùng (2023). *Giám sát không gian phát triển đô thị khu đông Thành phố Hồ Chí Minh bằng dữ liệu ảnh viễn thám*. Tạp chí Giao thông Vận tải, Số tháng 4/2023 (năm thứ 64), 130 - 133.
7. Phạm Văn Tùng, Trịnh Lê Hùng, Nguyễn Văn Trung, Vũ Xuân Cường (2023). *Nghiên cứu dự báo biến động bề mặt không thám khu vực thành phố Hồ Chí Minh từ dữ liệu viễn thám và GIS*. Tạp chí Đo đạc và Bản đồ, Số 56, 25 – 33.