



Ứng dụng dữ liệu viễn thám đa cảm biến để thành lập bản đồ lớp phủ đất tại hệ sinh thái rừng ngập mặn Cần Giờ, TP. Hồ Chí Minh

Phạm Thị Mai Thy^{1*}, Trương Nhật Kiều Thi¹, Đặng Phạm Bảo Nghi¹,

Lâm Đạo Nguyên¹, Emiliana Valentini²

¹Trung tâm Ứng dụng công nghệ Vũ trụ TP. Hồ Chí Minh, Trung Tâm Vũ Trụ Việt Nam, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, 18 Hoàng Quốc Việt, Hà Nội, Việt Nam

²Hội đồng Nghiên cứu Quốc gia Ý (CNR), Viện Khoa học Cự, Rome, Ý

Email tác giả liên hệ: ptmthy@vnsc.org.vn

<https://doi.org/10.5281/zenodo.xxxxxxxx>

Tóm tắt:

Nghiên cứu này trình bày cách tiếp cận tích hợp dữ liệu đa phổ Sentinel-2, siêu phổ PRISMA và radar COSMO-SkyMed (CSK) nhằm xây dựng bản đồ sử dụng đất/lớp phủ đất (LULC) cho hệ sinh thái rừng ngập mặn Cần Giờ. Phân tích được bổ sung bằng các biến sinh học (LAI, FAPAR) và thư viện phổ thực địa để tăng cường khả năng tách biệt thảm thực vật và các bề mặt có tính chất phổ tương tự. Mô hình pha trộn phổ tuyến tính (Linear Spectral Mixture Model – LSMM) được áp dụng cho dữ liệu PRISMA nhằm giải quyết hiện tượng trộn phổ phổ biến trong môi trường ven biển. Dữ liệu CSK được dùng để nhận dạng khu vực đô thị và ao nuôi thủy sản. Kết quả phân loại cho độ chính xác tổng thể 88% ($Kappa = 0,83$), cho thấy hiệu quả của cách tiếp cận đa cảm biến trong mô tả cảnh quan dị thể. Phương pháp này có tiềm năng áp dụng cho các nhiệm vụ siêu phổ thế hệ mới.

Từ khóa: Viễn thám đa cảm biến, mô hình pha trộn phổ tuyến tính, sử dụng đất/lớp phủ đất, siêu phổ.

Ngày nhận bài: 22/02/2026 Ngày sửa lại: 25/03/2026 Ngày chấp nhận đăng: 26/03/2026 Ngày xuất bản: 30/04/2026

Multi-sensor remote sensing for land cover mapping in the Can Gio mangrove ecosystem, Ho Chi Minh City

Phạm Thị Mai Thy^{1*}, Trương Nhật Kiều Thi¹, Đặng Phạm Bảo Nghi¹,

Lâm Đạo Nguyên¹, Emiliana Valentini²

¹Ho Chi Minh City Space Technology Application Center (STAC), Vietnam National Space Center (VNSC), Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Ha Noi, Viet Nam

²National Research Council of Italy (CNR-ISP), Rome

Corresponding Author Email: ptmthy@vnsc.org.vn

Abstract:

This study presents an integrated approach utilizing multispectral Sentinel-2, hyperspectral PRISMA, and Synthetic Aperture Radar (SAR) COSMO-SkyMed (CSK) data to map land use/land cover (LULC) patterns in the Can Gio mangrove ecosystem. The analysis is supplemented with biophysical variables (LAI, FAPAR) and an in-situ spectral library to enhance the spectral discriminability of vegetation and spectrally similar surfaces. A Linear Spectral Mixture Model (LSMM) is applied to the PRISMA data to address the severe spectral mixing commonly encountered in coastal environments. Furthermore, CSK data are exploited to delineate urban settlements and aquaculture ponds. The classification results yielded an overall accuracy of 88% ($Kappa = 0.83$), demonstrating the efficacy of the multisensor approach in characterizing heterogeneous landscapes. This methodology exhibits significant potential for application in next-generation hyperspectral missions.

Keywords: *Multi-sensor remote sensing, linear spectral mixture model, land use/land cover, hyperspectral.*

Submission received: 22/02/2026

Revised: 25/03/2026

Accepted: 26/03/2026

Published: 30/04/2026

1. Giới thiệu

Giám sát lớp phủ và sử dụng đất đóng vai trò quan trọng trong quản lý hệ sinh thái ven biển và thích ứng biến đổi khí hậu. Trong những năm gần đây, sự xuất hiện của các vệ tinh siêu phổ như PRISMA mở ra cơ hội mô tả chi tiết đặc tính quang học của các đối tượng bề mặt, vượt trội so với ảnh đa phổ truyền thống [1]. Tuy nhiên, dữ liệu siêu phổ thường đối mặt với thách thức về dung lượng lớn và tính trộn phổ cao tại các khu vực dị thể như rừng ngập mặn.

Cần Giờ là khu dự trữ sinh quyển đầu tiên của Việt Nam, giàu giá trị sinh thái nhưng chịu áp lực mạnh từ đô thị hóa, nuôi trồng thủy sản và du lịch. Sự biến động cấu trúc này đòi hỏi các chiến lược giám sát liên tục và cập nhật nhằm hỗ trợ các cơ quan ra quyết định trong công tác quản lý tài nguyên và bảo tồn đa dạng sinh học [2], [3].

Hầu hết các nghiên cứu phân loại sử dụng đất/lớp phủ (LULC) tại Việt Nam và các quốc gia nhiệt đới thường phụ thuộc vào phương pháp phân loại dựa trên điểm ảnh (pixel-based) như Random Forest hoặc SVM trên chuỗi dữ liệu Landsat, Sentinel-2 [4], [5], [6]. Các phương pháp này đạt độ chính xác cao đối với việc nhận diện sự hiện diện của rừng ngập mặn nói chung. Tuy nhiên, phương pháp phân loại điểm ảnh thường gặp khó khăn trong việc mô tả sự đa dạng của các kiểu thảm thực vật hoặc hỗn hợp phổ trong cùng một điểm ảnh tại các khu vực ven biển có tính dị thể cao, dẫn đến hạn chế khi phân loại chi tiết các loài ngập mặn hay thảm thực vật đan xen [7]. Sự ra đời của dữ liệu siêu phổ PRISMA cung cấp dải phổ liên tục, mở ra cơ hội phân tách chi tiết bề mặt. Nhằm khắc phục thách thức về sự dư thừa thông tin và tính đa chiều của dữ liệu siêu phổ, nghiên cứu này ứng dụng Mô hình pha trộn phổ tuyến tính (LSMM) kết hợp với thư viện phổ thực địa. Cách tiếp cận này giúp chuyển đổi LSMM từ hướng tiếp cận thống kê đơn thuần sang mô hình học có giám sát [8], rất phù hợp để đánh giá cấu trúc phân mảnh phức tạp tại Cần Giờ, hỗ trợ bóc tách các mảng rừng ngập mặn đan xen với khu vực nuôi trồng thủy sản và hạ tầng đô thị.

Trong nghiên cứu này, dữ liệu PRISMA, Sentinel-2 và CSK được tích hợp theo quy trình đa tầng, kết hợp mô hình giải trộn phổ, chỉ số sinh học và phân tích tán xạ radar. Phương pháp được đánh giá bằng dữ liệu thực địa thu thập trong hai chiến dịch năm 2022 và 2024.



Hình 1. Khu vực nghiên cứu với phạm vi chụp của PRISMA

Khu dự trữ sinh quyển Cần Giờ rộng khoảng 75.000 ha, gồm vùng lõi, vùng đệm và vùng chuyển tiếp. Khu vực có mạng lưới sông ngòi phức tạp, độ đục cao và cấu trúc cảnh quan biến đổi mạnh từ rừng ngập mặn nguyên sinh đến các vùng nông nghiệp, ao nuôi và khu dân cư (Hình 1). Trong giai đoạn 2000–2020, diện tích rừng ngập mặn có xu hướng bị thu hẹp bởi mở rộng sản xuất và hạ tầng [3].

2. Dữ liệu và phương pháp

2.1. Dữ liệu

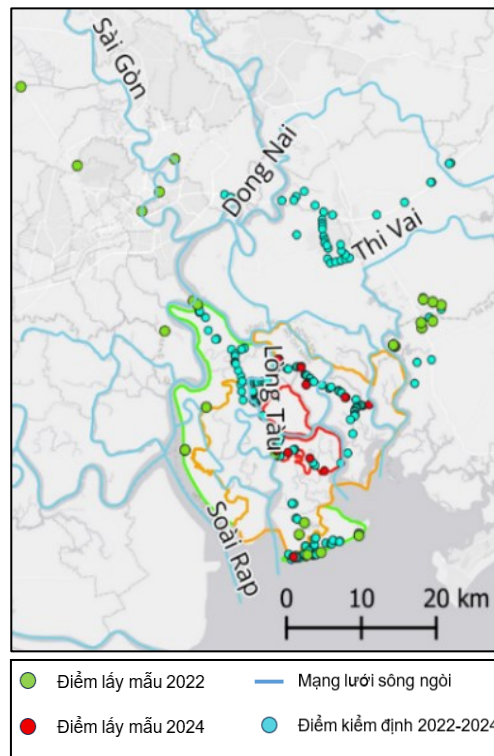
2.1.1. Dữ liệu ảnh vệ tinh

Hệ thống dữ liệu viễn thám đa nền tảng được tích hợp từ ba nguồn cảm biến chính. Nguồn cung cấp dữ liệu chủ đạo là hệ thống ảnh siêu phổ PRISMA, bao gồm 239 kênh phổ liên tục (VNIR-SWIR) với băng thông khoảng 12 nm và độ phân giải không gian 30 m. Các cảnh ảnh PRISMA được tuyển chọn nghiêm ngặt vào cao điểm mùa khô (ngày 14/01/2021, 27/01/2024 và 01/02/2024) nhằm giảm thiểu nhiễu động do mây che phủ. Dữ liệu thực địa giai đoạn 2022–2024 được sử dụng làm thư viện phổ chuẩn (endmembers) để phân tích ảnh năm 2024. Phương pháp này phù hợp do đặc trưng phản xạ phổ của các đối tượng tự nhiên (như cấu trúc tán lá, tính chất đất bùn) tương đối ổn định trong khoảng thời gian ngắn. Dữ liệu siêu phổ này được đồng bộ hóa với ảnh quang học đa phổ Sentinel-2 MSI (Level 2A, độ phân giải 10–20 m) thu nhận cùng thời điểm với PRISMA (ngày 27/01/2024), phục vụ mục đích chiết xuất các thông số sinh lý học thực vật (LAI, FAPAR) và hỗ trợ công tác trực chỉnh hình học. Bên cạnh đó, nhằm khắc phục giới hạn của cảm biến

quang học tại các khu vực phức tạp, nghiên cứu tích hợp thêm dữ liệu radar dải X phân cực đơn (VV) từ vệ tinh COSMO-SkyMed (CSK) với độ phân giải 3 m (ngày 01/03/2023 và 20/07/2023) để nhận dạng cấu trúc đô thị và phân định mạng lưới ao nuôi thủy sản.

2.1.2. Dữ liệu khảo sát thực địa

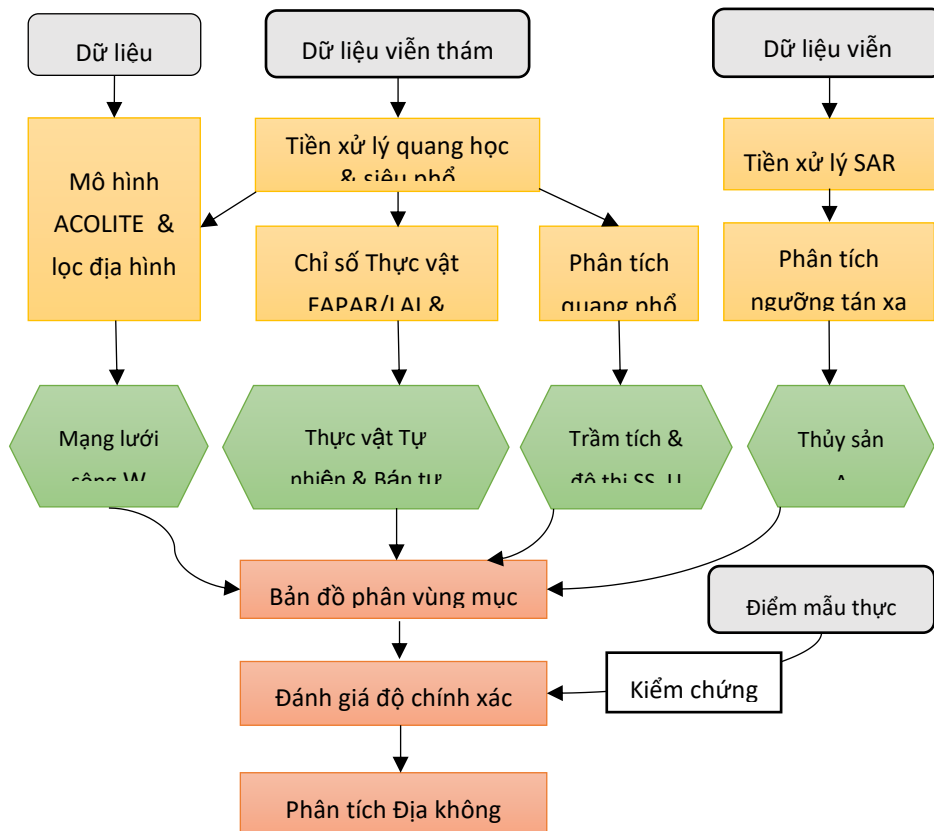
Công tác thu thập dữ liệu hiện trường được triển khai qua hai đợt khảo sát vào tháng 10/2022 và tháng 04/2024 (Hình 2). Hệ thống thiết bị chuyên dụng được sử dụng bao gồm máy quang phổ SR3500+ để đo phản xạ bề mặt, thiết bị WISP-3 đo phổ nước, cùng các thiết bị đo đặc chỉ số sinh học (ống kính Fish-eye đo chỉ số diện tích lá (LAI), cảm biến LI-180 đo bức xạ quang hợp hữu hiệu (PAR)) và đầu dò Trime-Pico64 xác định độ ẩm đất.



Hình 2. Sơ đồ vị trí của các dữ liệu mẫu tại thực địa

Dữ liệu đo phổ tại thực địa (tháng 10/2022) được sử dụng làm các mẫu chuẩn (endmembers) để huấn luyện mô hình LSM. Đồng thời, thông tin thu thập từ hai đợt khảo sát (tổng cộng 343 điểm) được dùng làm tập dữ liệu kiểm định và xây dựng ma trận nhầm lẫn. Các điểm khảo sát được xác định trên thiết bị di động, đảm bảo định vị chính xác qua hệ thống GPS. Tại mỗi điểm, nhóm ghi nhận các thông tin gồm: tọa độ địa lý, điều kiện thời tiết tại thời điểm khảo sát, loại lớp phủ, ngày khảo sát, tình hình biến động gần đây. Dữ liệu thu được đóng vai trò quan trọng trong việc hiệu chỉnh và đánh giá độ chính xác của kết quả phân loại ảnh viễn thám.

2.2. Phương pháp thực hiện



Hình 3. Sơ đồ tổng thể quy trình nghiên cứu

2.2.1. Tiền xử lý và phân loại

Chuỗi dữ liệu thô trải qua quy trình tiền xử lý: ảnh PRISMA được hiệu chỉnh hình học theo hệ quy chiếu của Sentinel-2, tiến hành ghép khảm và áp dụng mặt nạ lọc mây. Tương tự, dữ liệu radar CSK được hiệu chỉnh bức xạ, khử nhiễu đóm và trực chỉnh địa hình bằng phương pháp Range-Doppler Terrain Correction. Ảnh Sentinel-2 được khai thác bằng công cụ SNAP S2 Toolbox (ESA) nhằm tính toán hai chỉ số LAI và FAPAR, thiết lập các ngưỡng không gian quan trọng.

2.2.2. Mô hình LSMM

Để mô tả sự đan xen phức tạp về mặt không gian của cảnh quan thông qua dữ liệu siêu phổ PRISMA, nghiên cứu ứng dụng mô hình hỗn hợp phổ tuyến tính (Linear Spectral Mixture Model - LSMM). Phương pháp này tích hợp phân tích thống kê không gian phổ trộn với thư viện phổ thực địa được thu thập trong giai đoạn 2022–2024. Cụ thể, tập hợp các tín hiệu gốc (endmember - EM) được trích xuất thông qua phép phân tích thành phần chính (PCA) nhằm đảm bảo sự hiện diện của các đặc trưng phổ tối kết hợp cùng phổ đo đạc tại hiện trường. Các EM này được phân làm ba nhóm bề mặt đại diện: (1) thảm thực vật tự nhiên (đặc trưng bởi các loài *Avicennia alba*, *Rhizophora apiculata*, *Nypa fruticans*); (2) thảm thực vật bán tự nhiên (bao gồm đồng cỏ và cây trồng *Manihot esculenta*); và (3) nhóm nền đất trống (gồm bãi bùn, bãi cát và đất nội đồng).

Quá trình giải trộn tạo ra các bản đồ độ phủ phân số (Fractional Abundance Maps – FAMs), sau đó được chuẩn hóa và phân loại rời rạc hóa dựa trên các tổ hợp phổ đạt tần suất xuất hiện cao hơn 10%. Việc ứng dụng trực tiếp thư viện phổ thực địa đối với dữ liệu siêu phổ không chỉ khắc phục được hạn chế về sự dư thừa thông tin và tính đa chiều của phổ, mà còn giảm thiểu biến thiên phổ của các EM. Điều này góp phần gia tăng tính khách quan trong việc thiết lập ngưỡng phân loại cho các đối tượng sử dụng đất – lớp phủ (LULC).

2.2.3. Phân loại LULC

Nghiên cứu sử dụng thuật toán phân loại dựa trên LSMM cho dữ liệu quang học và phân ngưỡng tương quan pha cho dữ liệu radar. Tập mẫu huấn luyện được trích xuất trực tiếp từ thư viện phổ bao gồm hàng trăm phổ đo đạc tại thực địa trong giai đoạn 2022-2024. Kết quả giải trộn phổ được kết hợp với các công cụ phân tích không gian (như tạo vùng đệm, lọc nhiễu) để chiết xuất và tổng hợp thành 5 lớp phủ sử dụng đất chính ngoài ra bổ sung thêm 3 mức độ đục của nước. Mặc dù dữ liệu siêu phổ có khả năng bóc tách chi tiết hơn (lên tới 16 đặc trưng bề mặt), các đối tượng đã được gộp lại thành 5 nhóm sinh thái cơ bản để phù hợp với mục tiêu thành lập bản đồ hiện trạng cấp vùng và giảm thiểu sai số phân loại. Nhằm tinh chỉnh và chi tiết hóa ranh giới không gian của các lớp phủ, hệ thống tiêu chí không gian đặc thù đã được tích hợp. Cụ thể, ngưỡng độ lệch chuẩn của chỉ số diện tích lá (LAI) ở mức dưới 0,2 (tính toán trên cửa sổ trượt 3×3 điểm ảnh) được thiết lập để nhận diện các diện tích cây công nghiệp và ruộng lúa. Trong đó, lúa nước được phân định rạch ròi bằng cách giới hạn phạm vi phân bố trong bán kính 6 km tính từ mạng lưới sông ngòi. Đồng thời, vùng đệm (buffer) 100 m dọc theo bờ sông và bờ biển được ứng dụng làm tiêu chí phân tách giữa đất trồng nội đồng và bãi bồi ven bờ.

Đối với khu vực đô thị, chỉ số tương quan pha (coherence) được trích xuất từ hai cảnh ảnh radar SAR CSK (thu nhận ngày 01/03 và 20/07/2023) với ngưỡng phân định là 0,5, kết hợp cùng bộ lọc trượt (moving filter) kích thước 3×3 nhằm triệt tiêu nhiễu đốm. Phương pháp phân ngưỡng coherence này đặc biệt hữu hiệu trong việc phân tách các bề mặt nhân tạo có tính ổn định cao ra khỏi nền cảnh quan tự nhiên thường xuyên biến động [9]. Riêng ranh giới của lớp nuôi trồng thủy sản được kế thừa trực tiếp từ các kết quả nghiên cứu tiền đề.

2.2.4. Kiểm định độ chính xác

Để việc đánh giá độ chính xác được khách quan và đầy đủ căn cứ khoa học, các mẫu thực địa được thu thập đảm bảo tính đại diện cho đối tượng phân loại. Các mẫu cần được thu thập ngẫu nhiên và phân bố bao phủ khu vực nghiên cứu.

Công tác đánh giá độ chính xác của mô hình được thực hiện thông qua ma trận nhầm lẫn, tập trung vào sáu nhóm vùng mục tiêu chính (loại trừ các lớp có tỷ lệ diện tích phân bố dưới 2%). Dữ liệu tham chiếu để kiểm định bao gồm 343 điểm thu thập độc lập từ thực địa cùng 5 điểm tham chiếu bổ sung. Cuối cùng, để khẳng định độ tin cậy của phương pháp, kết quả phân loại LULC được đối chiếu song song với bản đồ sử dụng đất/lớp phủ Việt Nam năm 2020 (độ phân giải 10 m) do JAXA phát hành, qua đó xác định độ chính xác toàn cục (OA) cũng như độ chính xác của nhà sản xuất (PA) cho từng lớp phủ tương ứng.

Nghiên cứu dùng ma trận sai số phân loại và hệ số Kappa nhằm thống kê kiểm tra và đánh giá sự phù hợp giữa kết quả phân loại và kết quả khảo sát thực tế kết hợp với thông tin từ ảnh từ Google Earth Pro. Ma trận sai số là một ma trận vuông trong đó số hàng và số cột tương ứng với số loại lớp phủ đã phân loại trước đó. Từ bảng ma trận sai số cung cấp cho ta các chỉ số đo lường của việc tính toán chính xác như độ chính xác tổng thể, phần trăm của sai số bỏ sót và sai số nhầm lẫn, sai số người sản xuất và người sử dụng.

Chỉ số Kappa (K) nhằm thống kê kiểm tra và đánh giá sự phù hợp giữa những nguồn dữ liệu khác nhau hoặc khi áp dụng các thuật toán khác nhau. Cách xác định chỉ số Kappa được thể hiện ở công thức như sau:

Hệ số Kappa sẽ có giá trị 0 - 1; giá trị càng cao thì kết quả giải đoán ảnh càng tốt. Đối với dữ liệu ảnh viễn thám có độ phân giải thấp và trung bình, nếu hệ số Kappa đạt trên 0,7 thì kết quả giải đoán ảnh có thể chấp nhận và sử dụng được.

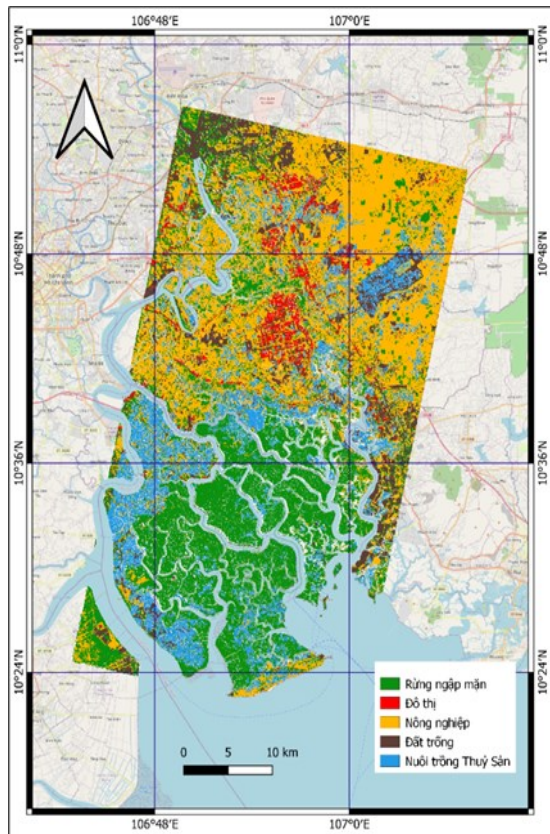
3. Kết quả nghiên cứu và thảo luận

3.1. Kết quả nghiên cứu

3.1.1. Cấu trúc lớp phủ

Dựa trên kết quả phân loại (Hình 4), cấu trúc lớp phủ của Cần Giờ năm 2024 phản ánh rõ rệt đặc trưng của một hệ sinh thái ven biển chuyển tiếp.

Cụ thể, diện tích mặt nước (W) chiếm tỷ lệ lớn nhất khi bao phủ 46% toàn bối cảnh (tương đương khoảng 122.000 ha). Tỷ lệ này không chỉ minh chứng cho mật độ mạng lưới sông ngòi dày đặc mà còn khẳng định vai trò điều tiết thủy văn quan trọng của hệ thống cửa sông Sài Gòn – Đồng Nai. Bên cạnh đó, việc phân định các mức độ đục nước từ ảnh PRISMA cho thấy sự khác biệt đáng kể giữa các nhánh sông, qua đó hỗ trợ nhận diện quá trình vận chuyển vật chất lơ lửng và động lực dòng chảy triều.



Hình 4. Kết quả phân loại

Đối với cấu trúc bề mặt thực vật, mảng thực vật bán tự nhiên (SN) ghi nhận mức độ bao phủ đáng kể với 25% (tương đương 67.000 ha), cao gấp đôi so với diện tích thảm thực vật tự nhiên (N) ở mức 14% (khoảng 36.000 ha). Sự phân bố này thể hiện đặc tính pha trộn phổ điển hình trong khu vực; trong khi các dải rừng ngập mặn trưởng thành hiện diện rõ rệt tại vùng lõi, thì các thảm cỏ và cây

trồng nông nghiệp lại phân bố đan xen ở vùng chuyển tiếp. Thực tế này khẳng định năng lực của mô hình LSMM trong việc mô tả chính xác dải gradient chuyển tiếp thông qua cơ chế bóc tách tỷ lệ phổ của từng tín hiệu gốc (endmember) bên trong một điểm ảnh (pixel). Cuối cùng, các không gian chịu tác động sinh thái sâu sắc từ hoạt động nhân sinh, bao gồm nhóm đất tràm tích – khu vực đô thị (SS+U) và vùng chuyên canh nuôi trồng thủy sản (A), chiếm tỷ trọng lần lượt là 8% (21.000 ha) và 7% (20.000 ha).

3.1.2. Phân bố độ đục nước

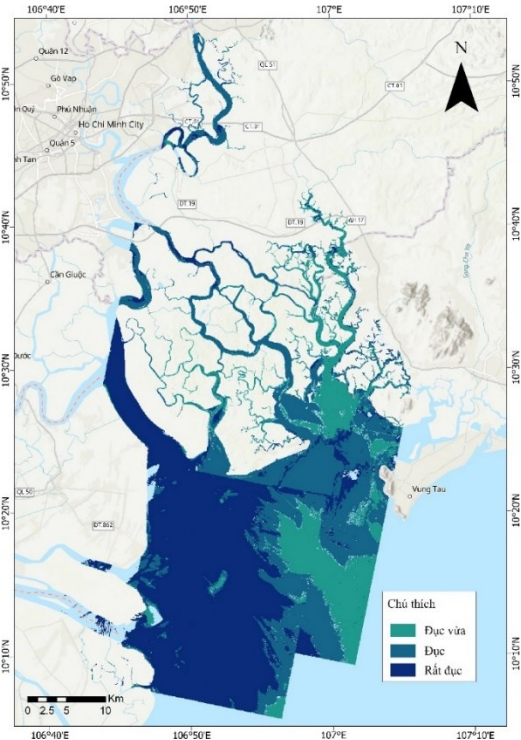
Việc khai thác năng lực phân giải quang phổ cao từ cảm biến PRISMA cho phép định lượng và phân tách bề mặt thủy vực thành ba cấp độ đục rõ rệt: đục vừa ($SPM \approx 30 \text{ g/m}^3$), đục ($\approx 100 \text{ g/m}^3$) và rất đục ($>100 \text{ g/m}^3$).

Sự phân hóa không gian trên bản đồ (Hình 5) cho thấy sự chênh lệch lớn về nồng độ chất rắn lơ lửng (SPM) giữa các hệ thống sông chính. Nổi bật nhất là nhánh sông Soài Rạp ở phía Tây, nơi duy trì trạng thái "rất đục" ở mức cực đại kéo dài từ khu vực nội đồng ra đến tận cửa biển. Ngược lại, khu vực trung tâm (sông Lòng Tàu) và phía Đông (sông Thị Vải) ghi nhận nồng độ bùn cát thấp hơn, chủ yếu phân bố ở ngưỡng "đục" và "đục vừa". Sự đối lập không gian này phức tạp hóa chân thực động lực thủy văn của khu vực; trong đó Soài Rạp đóng vai trò là hành lang vận chuyển trầm tích chủ đạo. Cường độ lan truyền bùn cát mạnh mẽ này là hệ quả trực tiếp của động lực dòng chảy xiết kết hợp với dao động thủy triều biên độ lớn đặc trưng trong giai đoạn mùa khô, có đóng góp quan trọng vào quá trình bồi tụ và duy trì sự cân bằng địa mạo cho toàn bộ hệ sinh thái ven bờ.

3.1.3. Hiệu quả giải trộn phổ

Mô hình LSMM đã chứng minh ưu thế vượt trội trong việc bóc tách tỷ lệ pha trộn quang học phức tạp giữa các mảng rừng, dải đất trồng và hệ sinh thái ven bờ. Các bản đồ độ phủ phân số được kiến tạo từ mô hình cho thấy sự tương thích không gian mang tính nhất quán cao so với các hệ quy chiếu ghi nhận trực tiếp tại thực địa.

Kết quả nghiên cứu đã khẳng định ưu thế vượt trội của mô hình LSMM trong việc xử lý bài toán điểm ảnh hỗn hợp, một thách thức đặc thù tại các hệ sinh thái ven biển có tính không đồng nhất không gian cao như Cần Giờ. Bằng cách bóc tách chi tiết tỷ lệ pha trộn quang học phức tạp bên trong từng điểm ảnh, mô hình đã phân



Hình 4. Phân loại độ đục quang học: đục vừa ($SPM \approx 30 \text{ g/m}^3$), đục ($\approx 100 \text{ g/m}^3$) và rất đục ($>100 \text{ g/m}^3$)

định rõ ràng các ranh giới chuyển tiếp vốn rất mờ nhạt giữa dải rừng ngập mặn rậm rạp, các bãi đất trống và hệ sinh thái thủy vực ven bờ. Đặc biệt, các FAM được kiến tạo từ LSM không chỉ thể hiện sự phân bố không gian hợp lý về mặt sinh thái, mà còn đạt độ tương thích và nhất quán rất cao khi đối soát chéo với các hệ quy chiếu thu thập trực tiếp tại thực địa. Điều này chứng minh phương pháp tiếp cận dưới điểm ảnh là công cụ đắc lực để khắc phục những hạn chế của phương pháp phân loại truyền thống khi lập bản đồ tại các vùng cảnh quan đan xen phức tạp.

3.1.4. Độ chính xác phân loại

Kiểm định định lượng thông qua ma trận nhầm lẫn cho thấy mô hình đạt độ chính xác toàn cục (OA) ở mức 88% cùng hệ số Kappa là 0,83. Đối với một khu vực cảnh quan có mức độ dị thể mạnh và chịu tác động đan xen liên tục giữa tự nhiên – nhân sinh như Cần Giờ, đây là một ngưỡng độ chính xác khả quan và đáng tin cậy.

Bảng 1. Bảng ma trận nhầm lẫn

| Thực tế \ Dự đoán | Rừng ngập mặn/thực vật | Đô thị | Nông nghiệp | Đất trống | Nuôi trồng TS | Tổng hàng |
|------------------------|------------------------|--------|-------------|-----------|---------------|-----------|
| Rừng ngập mặn/thực vật | 164 | 0 | 9 | 2 | 2 | 177 |
| Đô thị | 1 | 20 | 0 | 0 | 0 | 21 |
| Nông nghiệp | 3 | 6 | 76 | 2 | 2 | 89 |
| Đất trống | 1 | 4 | 0 | 15 | 0 | 20 |
| Nuôi trồng TS | 4 | 1 | 0 | 2 | 29 | 36 |
| Tổng cột | 173 | 31 | 85 | 21 | 33 | |

Độ chính xác toàn cục (OA):

$$\text{Overall Accuracy} = (164+20+76+15+29) / 343 \approx 0,886$$

Độ chính xác kỳ vọng (EA):

$$\text{Expected Accuracy} = \sum_i (\text{Tổng hàng}_i \times \text{Tổng cột}_i) / 3432 = 0,344$$

Hệ số Kappa:

$$K = (OA - EA) / (1 - EA) = (0,886 - 0,344) / (1 - 0,344) \approx 0,827$$

Kết quả độ chính xác toàn cục và hệ số Kappa đều đạt chuẩn với độ chính xác > 85% và Kappa > 0,8. Với điều kiện nguồn ảnh PRISMA có độ phân giải 30m thì kết quả trên đã đạt yêu cầu cho phép.

Nổi bật nhất trong bộ phân loại là nhóm Thảm thực vật tự nhiên (rừng ngập mặn) và Khu vực đô thị, cả hai đều ghi nhận mức độ chính xác của người dùng (UA) đạt đỉnh ở mức trên 90%. Kết quả tích cực này là minh chứng rõ nét cho tính hiệu quả của chiến lược tích hợp đa nguồn dữ liệu: trong khi các biến số sinh học (như LAI) trích xuất từ ảnh quang học Sentinel-2 phát huy tối đa khả năng phân biệt, khoanh vi vùng thảm thực vật có cấu trúc tán lá dày; thì tín hiệu radar từ vệ tinh CSK lại thể hiện độ nhạy cao trong việc nhận dạng hình học, tách biệt các cấu trúc nhân tạo và khu dân cư ra khỏi nền cảnh quan chung.

Ngược lại, nhóm Đất trống bộc lộ mức độ sai số lớn nhất với chỉ số UA chỉ đạt 75%. Hạn chế này xuất phát từ bản chất vật lý không ổn định của bề mặt khu

vực nghiên cứu; đặc trưng quang phổ của các bãi bùn cát ven biển và đất nội đồng thường xuyên bị biến đổi mạnh mẽ dưới tác động của sự thay đổi độ ẩm và chế độ thủy triều. Sự nhầm lẫn phổ phức tạp này đặt ra yêu cầu thiết thực cho các hướng nghiên cứu mở rộng, qua đó cho thấy sự cần thiết phải tích hợp thêm mô hình độ cao số (DEM) hoặc khai thác dữ liệu radar phân cực kép (dual-polarization SAR) nhằm tăng cường độ nhạy trong việc phân tách ranh giới các bãi trầm tích trong tương lai.

3.2. Thảo luận

Việc tích hợp PRISMA – Sentinel-2 – COSMO SKyMed mang lại khả năng phân loại vượt trội trong môi trường dị thể. LSMM phát huy hiệu quả trong nhận dạng các vùng chuyên tiếp – nơi thuật toán pixel-based thường gặp khó khăn [10]. Điểm nổi bật của phương pháp này là khả năng phân biệt chi tiết các lớp cây trồng nông nghiệp. Khác với các sản phẩm LULC truyền thống thường nhầm lẫn cây nông nghiệp với rừng ngập mặn hoặc rừng trồng, mô hình đề xuất đã bóc tách thành công cây công nghiệp (ví dụ: cao su, trà) và các loại cây trồng khác (ví dụ: sắn). Sentinel-2 đóng vai trò quan trọng trong việc ràng buộc sinh học thông qua LAI và FAPAR, giúp phân biệt rõ cây trồng và thảm thực vật tự nhiên. Đồng thời, dữ liệu CSK tăng cường khả năng nhận dạng khu vực đô thị và mạng lưới ao nuôi nhờ tín hiệu tán xạ radar ổn định, giúp khắc phục triệt để những hạn chế của cảm biến quang học tại các hệ sinh thái ven biển thường xuyên bị mây che phủ [11].

Dựa trên sự tương quan giữa cấu trúc lớp phủ và hiệu quả bóc tách không gian, có thể thấy cấu trúc liên tục nguyên thủy của Khu dự trữ sinh quyển Cần Giờ đang trải qua quá trình phân mảnh rõ rệt. Sự chênh lệch tỷ lệ giữa thảm thực vật tự nhiên (chỉ còn 14%) so với thảm thực vật bán tự nhiên (25%), kết hợp với sự hiện diện dày đặc của hạ tầng đô thị và ao nuôi thủy sản (tổng 15%), là minh chứng cho sức ép từ các hoạt động chuyển đổi mục đích sử dụng đất. Quá trình chuyển đổi mục đích sử dụng đất này đã phá vỡ cấu trúc liên mạch của rừng ngập mặn, chia cắt thảm thực vật tự nhiên thành nhiều mảng nhỏ. Xu hướng gia tăng mật độ rìa và phân mảnh cảnh quan này phản ánh rõ nét tác động kéo dài của các hoạt động nhân sinh trong hai thập kỷ qua, hệ quả là các tuyến đường kết nối tự nhiên của sinh vật bị đứt gãy, khiến cho vùng lõi ngày càng bị thu hẹp [12]. Ứng dụng mô hình giải trộn phổ tuyến tính (LSMM) là thực sự cần thiết trong bối cảnh này, bởi nó có khả năng phát hiện các ranh giới chuyên tiếp mờ nhạt mà phương pháp phân loại dựa trên điểm ảnh (pixel-based) thông thường chưa thể nhận diện chính xác.

Một số hạn chế được ghi nhận, gồm: (i) PRISMA chưa có sản phẩm sinh học được hiệu chỉnh tốt; (ii) độ phân giải 30 m vẫn gây khó khăn với các bề mặt cỡ nhỏ; (iii) lớp đất-trầm tích có phổ biến đổi mạnh theo độ ẩm và thành phần trầm tích.

Các hướng phát triển tiếp theo bao gồm khai thác dữ liệu siêu phổ CHIME, EnMAP; sử dụng phân cực kép SAR để mô tả cấu trúc khu vực đô thị và ao nuôi; kết hợp mô hình học sâu cho dữ liệu siêu phổ khi có bộ nhãn đầy đủ [13].

4. Kết luận

Nghiên cứu đã chứng minh hiệu quả vượt trội của việc tích hợp dữ liệu viễn thám đa nền tảng (Sentinel-2, PRISMA, COSMO-SkyMed) cùng mô hình hỗn hợp

phổ tuyến tính (LSMM) trong lập bản đồ LULC tại Khu dự trữ sinh quyển Cần Giờ. Phương pháp này giải quyết thành công bài toán "điểm ảnh hỗn hợp", mang lại độ chính xác toàn cục (OA) đạt 88% ($Kappa = 0,83$). Đặc biệt, sự kết hợp giữa các biến sinh học từ ảnh quang học và tín hiệu tán xạ radar đã giúp nhận diện rừng ngập mặn và khu vực đô thị với độ tin cậy (UA) cao lên đến 93%.

Bên cạnh giá trị phương pháp luận, kết quả bản đồ còn định lượng rõ nét sự phân mảnh cảnh quan và đứt gãy sinh thái, thể hiện qua việc thảm thực vật tự nhiên bị thu hẹp chỉ còn 14% dưới sức ép của đô thị và ao nuôi thủy sản. Cách tiếp cận đề xuất không chỉ cung cấp công cụ đắc lực hỗ trợ các nhà quản lý trong chiến lược bảo tồn đa dạng sinh học ven biển, mà còn thiết lập khung tham chiếu quan trọng cho các sứ mệnh vệ tinh siêu phổ thế hệ mới trong tương lai.

Lời cảm ơn

Nghiên cứu này thực hiện với sự hỗ trợ của đề tài “Tăng cường không gian – mạng lưới khoa học để giám sát nhanh chóng sự dịch chuyển cảnh quan” mã số QTIT01.03/23-24 thuộc Nhiệm vụ Hợp tác Quốc tế của Viện Hàn lâm Khoa học và Công nghệ Việt Nam với Hội đồng Nghiên cứu Quốc gia Italia giai đoạn 2023-2024.

Cam kết của các tác giả

Tất cả các tác giả có tên trong bài báo cam kết sự đồng thuận và không có xung đột lợi ích trong công bố khoa học tại bài báo này.

Tài liệu tham khảo

- [1] R. Loizzo et al., "PRISMA: The Italian hyperspectral mission", *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 175-178, 2018. DOI: 10.1109/IGARSS.2018.8518512
- [2] L. T. H. Pham, T. Q. Vo, T. D. Dang, and U. T. N. Nguyen, "Monitoring mangrove association changes in the Can Gio Biosphere Reserve and implications for management", *Remote Sensing Applications: Society and Environment*, vol. 13, pp. 298-305, 2019. DOI: 10.1016/j.rsase.2018.11.009
- [3] B. K. Veettil et al., "Mangroves of Vietnam: Historical development, current state of research and future threats", *Estuarine, Coastal and Shelf Science*, vol. 218, pp. 212-236, 2019. DOI: 10.1016/j.ecss.2018.12.021
- [4] H. T. Le et al., "Characterizing spatiotemporal patterns of mangrove forests in Can Gio biosphere reserve using Sentinel-2 imagery", *Applied Sciences*, vol. 10, no. 12, pp. 4058, 2020. DOI: 10.3390/app10124058
- [5] M. H. Phan, M. J. F. Stive, "Managing mangroves and coastal land cover in the Mekong Delta", *Ocean & Coastal Management*, vol. 219, pp. 106013, 2022. DOI: 10.1016/j.ocecoaman.2021.106013
- [6] Truong Nhat Kieu Thi et al., "Analysis of changes in the agricultural and forest areas and the impact of socio-economic activities in the Ba River Basin, Vietnam", *IOP Conference Series: Earth and Environmental Science*, vol. 652, 012014, 2021.
- [7] M. Belgiu, L. Drăguț, "Random forest in remote sensing: A review of applications and future directions", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24-31, 2016. DOI: 10.1016/j.isprsjprs.2016.01.011



- [8] C. Small, D. Sousa, "The standardized spectroscopic mixture model", *Remote Sensing*, vol. 16, no. 20, pp. 3768, 2024. DOI: 10.3390/rs16203768
- [9] M. Chini, R. Pelich, L. Pulvirenti, N. Pierdicca, R. Hostache, and P. Matgen, "Sentinel-1 InSAR coherence to detect floodwater in urban areas", *Remote Sensing*, vol. 11, no. 1, pp. 107, 2019. DOI: 10.3390/rs11020107
- [10] B. Somers et al., "Endmember variability in spectral mixture analysis: A review", *Remote Sensing of Environment*, vol. 115, no. 6, pp. 1603-1616, 2011. DOI: 10.1016/j.rse.2011.03.003
- [11] E. Valentini, M. Righini, S. Sapio, S. Liburdi, V. Cima, S. V. Nghiem, and A. Taramelli, "SAR data for monitoring rapidly changing ecosystems: Aquaculture and urbanization in the Càn Giò Mangrove Biosphere Reserve, South Vietnam", *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1424-1427, 2023. DOI: 10.1109/IGARSS52108.2023.10282114
- [12] A. N. T. Do, H. D. Tran, M. Ashley, and A. T. Nguyen, "Monitoring landscape fragmentation and aboveground biomass estimation in Can Gio Mangrove Biosphere Reserve over the past 20 years", *Ecological Informatics*, vol. 70, pp. 101743, 2022. DOI: 10.1016/j.ecoinf.2022.101743
- [13] A. Signoroni et al., "Deep learning meets hyperspectral image analysis: A multidisciplinary review", *Journal of Imaging*, vol. 5, no. 5, pp. 52, 2019. DOI: 10.3390/jimaging5050052

