

SỬ DỤNG GOOGLE EARTH ENGINE ĐỂ PHÂN LOẠI TỰ ĐỘNG RỪNG NGẬP MẶN TỪ ẢNH VỆ TINH SENTINEL-1 VÀ SENTINEL-2 CHO KHU VỰC QUẢNG YÊN, TỈNH QUẢNG NINH

Nguyễn Trọng Cường¹, Trần Quang Bảo², Nguyễn Hải Hoà¹

¹Trường Đại học Lâm nghiệp

²Tổng cục Lâm nghiệp

<https://doi.org/10.55250/jo.vnuf.2022.1.053-065>

TÓM TẮT

Nghiên cứu này sử dụng nền tảng Google Earth Engine (GEE) để thử nghiệm và phân loại tự động rừng ngập mặn cho khu vực Quảng Yên, tỉnh Quảng Ninh. Toàn bộ tư liệu ảnh vệ tinh Radar Sentinel-1 và quang học Sentinel-2 có tỉ lệ mây nhỏ hơn 20% trong năm 2020 của khu vực nghiên cứu đã được sử dụng. Áp dụng hai phương pháp phân loại rừng ngập mặn gồm: (1) sử dụng ngưỡng phân loại rừng ngập mặn từ giá trị tán xạ ngược VH trên ảnh Radar Sentinel-1 và ngưỡng chỉ số rừng ngập mặn CMRI (Combine Mangrove Recognition Index) trên ảnh quang học Sentinel-2; (2) phương pháp rừng ngẫu nhiên (Random Forest - RF) trên cả hai loại tư liệu. Kết quả phân loại rừng ngập mặn của cả hai phương pháp đều có độ chính xác cao. Dữ liệu quang học Sentinel-2 có sự ổn định hơn, khi kết quả đánh giá độ chính xác tổng thể của hai phương pháp sử dụng ngưỡng chỉ số CMRI và phương pháp rừng ngẫu nhiên (Random Forest) lần lượt là 95,9%, (Kappa = 0,95) và 98% (Kappa = 0,97). Dữ liệu SAR Sentinel-1 có độ ổn định thấp hơn với độ chính xác tổng thể 96,2% (Kappa = 0,95) bằng phương pháp ngưỡng giá trị tán xạ ngược VH và 91,4% (Kappa = 0,85) bằng phương pháp rừng ngẫu nhiên. Các kết quả tính diện tích rừng ngập mặn khu vực nghiên cứu cho thấy, dữ liệu Sentinel-2 có sự chênh lệch về diện tích giữa hai phương pháp ít hơn (khoảng 143 ha), trong khi dữ liệu SAR Sentinel-1 có sự chênh lệch lớn hơn (khoảng 287 ha).

Từ khóa: Google Earth Engine, phân loại rừng ngập mặn, Random Forest, rừng ngập mặn, Sentinel-1, Sentinel-2.

1. ĐẶT VẤN ĐỀ

Rừng ngập mặn là vùng đất ngập nước thủy triều nằm ở các vùng nhiệt đới và cận nhiệt đới giữa vĩ độ khoảng 30 độ Bắc và 30 độ Nam (Giri và cộng sự, 2011; Lee và Yeh, 2009), chúng cung cấp một loạt các dịch vụ hệ sinh thái cũng như môi trường sống cho nhiều loài thủy sản biển, lọc nước, ổn định bờ biển, đa dạng sinh học (Rahman và cộng sự, 2013; Abdul Aziz và cộng sự, 2015; Giri và cộng sự, 2015) và cung cấp dịch vụ hệ sinh thái và bảo vệ khu vực ven biển cho các đường bờ biển nhiệt đới và cận nhiệt đới trên thế giới (Valiela và cộng sự, 2001; Veettil và cộng sự, 2018). Trong những năm gần đây, với sự tiến bộ của khoa học công nghệ, đặc biệt là công nghệ viễn thám đã hỗ trợ rất nhiều trong công tác điều tra, giám sát tài nguyên rừng nói chung và rừng ngập mặn nói riêng. (Cường và cộng sự, 2021).

Các tư liệu viễn thám có độ phân giải trung bình đóng một vai trò quan trọng trong việc lập bản đồ các rừng ngập mặn trên các vùng địa lý rộng lớn. Trong các nghiên cứu về rừng ngập mặn, ảnh viễn thám quang học có độ phân giải

không gian trung bình thường được sử dụng thường xuyên và phổ biến hơn. Các kỹ thuật phân loại tư liệu ảnh có độ phân giải trung bình rất thích hợp cho việc lập bản đồ các hệ sinh thái (trừ cấp độ loài), giám sát các thay đổi quy mô lớn, phân tích các mối quan hệ môi trường khu vực và đánh giá tình trạng của rừng ngập mặn (sức khỏe, mật độ...) Biến động rừng ngập mặn trên toàn cầu dễ nhận thấy từ việc phân tích dữ liệu có độ phân giải trung bình (Cường và cộng sự, 2021).

Hai phát triển gần đây trong lĩnh vực quan sát trái đất có khả năng cải thiện đáng kể hiệu quả của việc giám sát rừng ngập mặn trên toàn cầu. Đầu tiên, các vệ tinh miễn phí Sentinel-1 và Sentinel-2 của châu Âu. Hai loại tư liệu này với sự vượt trội về độ phân giải không gian (10 m) cũng như độ phân giải thời gian cao đang góp phần hữu hiệu vào việc phân biệt rừng ngập mặn theo nhiều cấp độ (Manna và Raychaudhuri, 2018). Trong đó, ảnh Sentinel-2 có thể lập bản đồ chính xác về phạm vi rừng ngập mặn và phân biệt cơ bản các quần xã loài

ngập mặn (Wang và cộng sự, 2018). Thứ hai, nền tảng đám mây của GEE cho phép khả năng tính toán hiệu suất cao để xử lý dữ liệu không gian địa lý (Amani và cộng sự, 2020; Gorelick và cộng sự, 2017). GEE chứa một kho lưu trữ dữ liệu Sentinel-1 và Sentinel-2 đã được xử lý trước và cập nhật liên tục đã cho phép phát triển công cụ xử lý hiệu quả dữ liệu trên quy mô toàn cầu (Chen và cộng sự, 2017; Gorelick và cộng sự, 2017).

Về phương diện phân loại rừng ngập mặn, một số phương pháp thường sử dụng để thành lập rừng ngập mặn như giải đoán bằng mắt, sử dụng các chỉ số thực vật, phân loại dựa trên các pixel ảnh gồm phân loại có kiểm định và không có kiểm định, ngoài ra còn một số phương pháp khác như phân loại cây quyết định, phương pháp dựa trên đối tượng, hoặc sử dụng các dữ liệu đầu vào để phân loại ngẫu nhiên... hoặc có thể sử dụng một phương pháp hoặc kết hợp các phương pháp với nhau (Cương và cộng sự, 2021). Trong các phương pháp phân loại có giám sát, các nghiên cứu thường sử dụng các chỉ số thực vật. Các chỉ số thực vật khác biệt chuẩn hóa (NDVI), chỉ số thực vật điều chỉnh của đất (SAVI) và chỉ số diện tích lá (LAI) được sử dụng rộng rãi nhất. Các chỉ số này có thể làm nổi bật các đặc tính nội tại của thực vật có liên quan đến lá, độ xanh tươi và sức sống của đối tượng. Mỗi chỉ số có một biểu thức cụ thể, có thể thể hiện các đặc tính của thảm thực vật tốt hơn so với việc sử dụng các kênh ảnh riêng lẻ. Tuy vậy, các chỉ số thực vật nói chung không dành riêng cho rừng ngập mặn và khó có thể phân biệt rừng ngập mặn với thảm thực vật trên cạn (Gathot Winarso, 2014). Rừng ngập mặn và rừng ngập mặn trên cạn dày đặc có thể tạo ra các giá trị NDVI giống nhau, do đó việc tách chúng khỏi tư liệu viễn thám rất khó khăn trừ khi có thêm các dữ liệu đầu vào khác (Alvin B và cộng sự, 2020).

Để giải quyết thách thức trong việc lập bản đồ rừng ngập mặn với các chỉ số thực vật, một số nhà nghiên cứu đã đề xuất các chỉ số thực vật đặc trưng cho rừng ngập mặn bằng cách sử dụng các băng tần đầu vào khác nhau và dữ liệu vệ tinh, ví dụ các chỉ số CMRI, MI, MFI, MVI.... Mặc dù vậy, vẫn còn nhiều phương pháp khác được phát triển, điển hình là hệ thống phân loại

rừng ngẫu nhiên Random Forest được cung cấp trên nền tảng GEE đã tạo rất nhiều thuận lợi cho các hoạt động nghiên cứu về rừng ngập mặn hiện nay (Ghorbanian và cộng sự, 2021; Ghorbanian và cộng sự, 2020). RF là một bộ phân loại không tham số, sử dụng kỹ thuật tổng hợp để kết hợp các kết quả phân loại của các cây quyết định ngẫu nhiên độc lập khác nhau và để dự đoán (Breiman, 2001). Mỗi cây quyết định ngẫu nhiên này được huấn luyện bởi một tập hợp con các mẫu huấn luyện, được gọi là mẫu trong túi và sử dụng phần còn lại, được gọi là mẫu ngoài túi, để xác nhận chéo nội bộ. Sau đó, kết quả của chúng được tích hợp để đưa ra kết quả phân loại cuối cùng. Hơn nữa, RF đã chứng minh khả năng xử lý dữ liệu chiều cao bằng cách tạo ra các bản đồ đầy tiềm năng (Ghorbanian và cộng sự, 2021).

Mục tiêu của nghiên cứu này là đánh giá khả năng sử dụng hai tư liệu viễn thám (ảnh Radar Sentinel-1 và ảnh quang học Sentinel-2) để phân loại rừng ngập mặn khu vực Quảng Yên, tỉnh Quảng Ninh. Bằng cách sử dụng phương pháp ngưỡng phân loại từ chỉ số CMRI (với ảnh Sentinel-2), ngưỡng phân loại từ giá trị VH (với ảnh Sentinel-1) và phương pháp Random Forest vào cả hai tư liệu để tính toán tự động trên nền tảng GEE. Kết quả của nghiên cứu này sẽ là cơ sở khoa học quan trọng để lựa chọn tư liệu và phương pháp phù hợp phục vụ giám sát rừng ngập mặn cho khu vực nghiên cứu nói riêng, tỉnh Quảng Ninh và các tỉnh ven biển phía Bắc nói chung.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Khu vực nghiên cứu

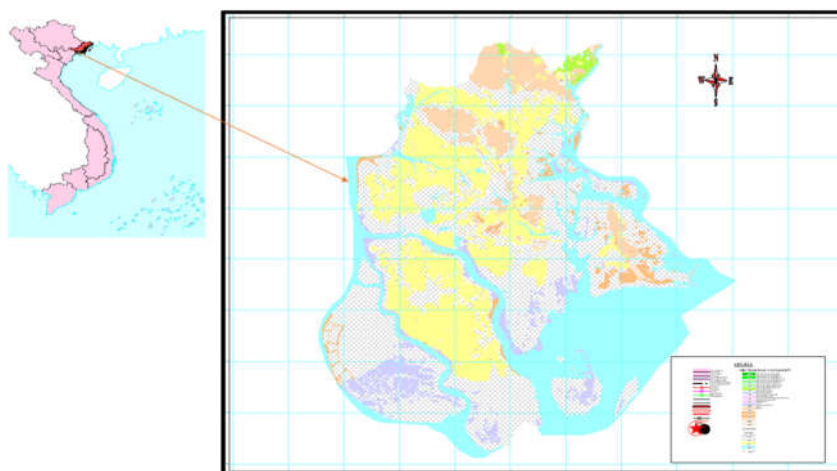
Khu vực nghiên cứu bao gồm toàn bộ rừng ngập mặn thuộc thị xã Quảng Yên, tỉnh Quảng Ninh (hình 01). Đặc điểm chính của ba trạng thái như sau (1) Rừng ngập mặn: Gồm rừng ngập mặn khép kín xen kẽ và rừng ngập mặn tán thưa, rừng ngập mặn tự nhiên và rừng trồng; 2) vùng nước gồm khu vực sông, cửa sông, hồ, kênh, rạch, các vùng nước nhỏ, và ở cửa biển, các trại nuôi tôm; (3) đối tượng khác bao gồm thực vật không phải rừng ngập mặn, ao nuôi trồng thủy sản bị bỏ hoang, ao muối bị bỏ hoang, đất trống và ẩm ướt, bãi bồi, đất khô và các khu dân cư

2.2. Vật liệu và phương pháp nghiên cứu

ngày 01/01/2020 đến 31/12/2020) của tư liệu Sentinel-1 và Sentinel-2.

2.2.1. Tư liệu ảnh vệ tinh

Sử dụng toàn bộ ảnh trong năm 2020 (từ



Hình 1. Bản đồ vị trí khu vực nghiên cứu và bản đồ hiện trạng rừng năm 2015 thị xã Quảng Yên

Sentinel-1 là vệ tinh Radar khẩu độ tổng hợp (SAR) của châu Âu, thu thập dữ liệu băng tần C ở chế độ phân cực kép trong mọi điều kiện thời tiết với độ phân giải theo thời gian 6 ngày. Đã sử dụng hình ảnh phát hiện trên phạm vi mặt đất cấp-1C (GRD) với độ phân giải không gian 10 m ở cả chế độ tăng dần và giảm dần. Có tổng cộng, 31 ảnh Sentinel-1 trong các phân cực VV (truyền và nhận dọc) và VH (truyền dọc và nhận ngang) có được vào năm đã được sử dụng (bảng 1).

Spectral Instrument (MSI). Cảm biến này ghi lại thông tin bề mặt Trái đất trong 13 dải quang phổ từ vùng nhìn thấy được đến vùng hồng ngoại sóng ngắn (SWIR), với các độ phân giải không gian khác nhau trong khoảng từ 10 m đến 60 m. Trong nghiên cứu này, bài báo sử dụng toàn bộ các giải của Sentinel-2 cho phương pháp Random Forest và 3 giải xanh (Green), đỏ (RED) và hồng ngoại gần (NIR) với độ phân giải không gian 10 m để tính cho chỉ số thực vật. Tổng cộng, 12 ảnh Sentinel - 2 được thu thập vào năm 2020 đã được sử dụng để lập bản đồ hệ sinh thái rừng ngập mặn (bảng 1).

Sentinel -2 cũng là tư liệu do Cơ quan Vũ trụ châu Âu cung cấp và mang cảm biến Multi

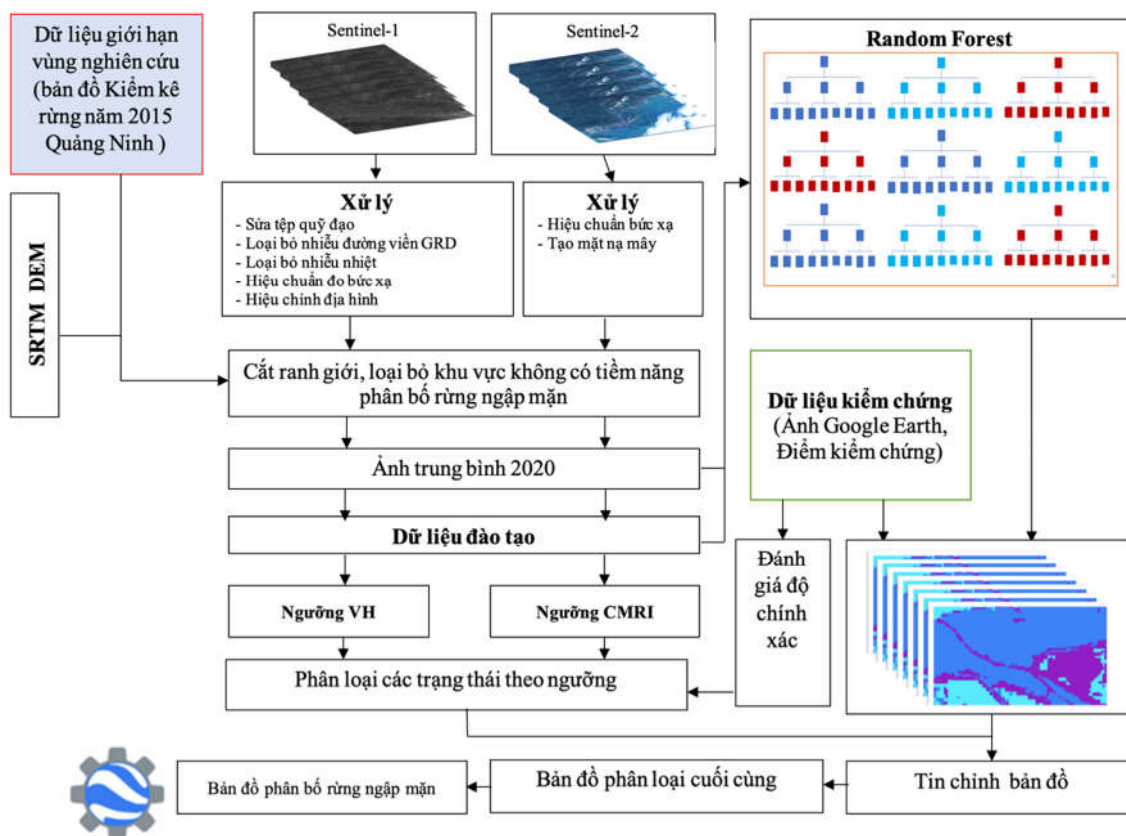
Bảng 1. Thông tin số lượng các tư liệu ảnh sử dụng

Dữ liệu	Mùa trong năm 2020				Tổng	Thời gian	Tỉ lệ mây (%)
	Xuân	Hè	Thu	Đông			
Sentinel-1	7	8	7	9	31	01/01/2020 –	
Sentinel-2	2	1	5	4	12	31/12/2020	<20

2.2.2. Phương pháp xử lý số liệu

Quá trình xử lý dữ liệu Sentinel-1 và Sentinel-2 được thực hiện trong GEE. GEE là một nền tảng dựa trên đám mây để phân tích dữ liệu môi trường quy mô toàn cầu. Nó kết hợp một danh mục nhiều Petabyte gồm hình ảnh vệ tinh và bộ dữ liệu không gian địa lý, cơ sở hạ tầng tính toán của Google được tối ưu hóa để xử lý song song dữ liệu không gian địa lý, Giao diện lập trình ứng dụng (API) cho Java Script và Python và môi trường phát triển tích hợp dựa

trên nền web để tạo mẫu nhanh và trực quan hóa các phân tích không gian phức tạp. GEE được thiết kế để người dùng hầu như không phải lo lắng về các phép chiếu bản đồ khi thực hiện tính toán ở khu vực rộng lớn, các tham số chiếu được yêu cầu trong phép chiếu đầu ra. Theo mặc định, GEE thực hiện lấy lại mẫu theo thuật toán láng giềng gần nhất (Nearest Neighbor - NN) trong quá trình tham chiếu. Tóm tắt quá trình xử lý ảnh và phân loại rừng ngập mặn được thể hiện trong hình 2, bao gồm các bước như sau:



Hình 2. Mô tả quá trình nghiên cứu

1) Tạo các mẫu đào tạo và mẫu kiểm chứng

Trước tiên, tạo ra 1326 điểm mẫu ngẫu nhiên trên ảnh độ phân giải cao có sẵn trong Qgis và Google Earth của toàn bộ khu vực nghiên cứu (trong đó rừng ngập mặn 757 điểm, mặt nước 261 điểm và đối tượng khác 308 điểm). Trong tổng số điểm mẫu đã tạo ra, sử dụng ngẫu nhiên 60% trong số đó để tạo dữ liệu tập huấn, 40% số điểm còn lại được dùng để kiểm chứng kết quả sau phân loại. Việc chia tách ngẫu nhiên dẫn đến độ lệch thấp trong việc thực hiện các kết quả phân loại cuối cùng. Toàn bộ các điểm này được đưa vào GEE để lập trình và tính toán tự động cho cả phương pháp tính bằng các chỉ số thực vật và phương pháp rừng ngẫu nhiên.

2) Tạo vùng tiềm năng có phân bố rừng ngập mặn

Để tạo các vùng tiềm năng phân bố rừng ngập mặn, đồng thời loại bỏ các khu vực có đặc điểm giống với rừng ngập mặn nằm sâu trong đất liền (các khu vực ven ao hồ ngập nước quang năm) và các vùng có độ cao, độ dốc lớn là nơi không có khả năng xuất hiện rừng ngập mặn, bài báo sử dụng dữ liệu SRTM DEM 30 m và dữ liệu độ dốc để loại bỏ chúng. Dữ liệu

SRTM DEM được cung cấp trên GEE cho phép loại bỏ một cách chính xác các khu vực này và không đưa chúng vào dữ liệu nghiên cứu.

Ngoài ra, sử dụng bản đồ kiểm kê rừng năm 2015 tỉnh Quảng Ninh (định dạng Shapefile, tỷ lệ 1:100.000) do Bộ Nông nghiệp và Phát triển nông thôn công bố để phân định ranh giới hành chính và ranh giới giữa rừng ngập mặn và các khu vực nội địa bởi vì thị xã Quảng Yên không có đường bờ biển. Hàng năm, rừng ngập mặn có thể phát triển thêm, chúng lấn ra biển hoặc có thể bị mất đi do nhiều nguyên nhân. Trên thực tế một số khu vực nằm trong ranh giới rừng ngập mặn (như đất nông nghiệp hoặc bãi bồi ven cửa sông...) cũng có thể được trồng mới thêm rừng ngập mặn. Các khu vực này chưa được phân loại để xác định là rừng ngập mặn. Bản đồ kiểm kê rừng được sử dụng để tạo tách khu vực ranh giới hành chính thị xã Quảng Yên, sau đó tạo các vùng tiềm năng có phân bố rừng ngập mặn bằng công cụ Buffer từ hộp công cụ trên phần mềm QGIS phiên bản 3.22 bằng cách tạo vùng đệm mở rộng 500 m tính từ ranh giới các lô kiểm kê rừng năm 2015. Các vùng tiềm năng có phân bố rừng ngập mặn (định dạng Shapefile,

hệ quy chiếu GSW84) sau khi trích xuất sẽ chuyển vào nền tảng GEE để làm dữ liệu tính toán.

3) Phân loại bằng Chỉ số CMRI và giá trị tán xạ ngược VH

Phân loại bằng Chỉ số CMRI: Chỉ số rừng ngập mặn kết hợp (CMRI - Combine Mangrove Recognition Index) được đề xuất bởi (Gupta và cộng sự, 2018). Biểu thức của CMRI như sau:

$$CMRI = NDVI - NDWI \quad (1)$$

Trong đó: NDVI là chỉ số thực vật khác biệt chuẩn hóa (Tucker, 1997); NDWI là chỉ số nước chênh lệch chuẩn hóa (Gao, 1996);

$$NDVI = (NIR - RED)/(NIR + RED) \quad (2)$$

$$NDWI = (GREEN - NIR)/(GREEN + NIR) \quad (3)$$

NIR, RED, GREEN lần lượt là giá trị phản xạ ở tầng trên khí quyển của các kênh ảnh: cận hồng ngoại, đỏ, green. Giá trị CMRI trung bình của các điểm mẫu trên toàn bộ các ảnh của năm 2020 sẽ được tính toán tự động trên GEE và phân tích để tìm ra các ngưỡng tối ưu phân loại với rừng ngập mặn và các lớp phủ khác.

Phân loại bằng giá trị tán xạ ngược VH:

Phân loại bằng ngưỡng giá trị hệ số tán xạ ngược VH được đo bằng cách tính giá trị trung bình của toàn bộ các cảnh ảnh đối với từng điểm mẫu nghiên cứu. Giá trị này được so sánh giữa các trạng thái lớp phủ với nhau để tìm ra ngưỡng tối ưu phân biệt các đối tượng.

4) Phân loại bằng phương pháp Random Forest

Trong nghiên cứu này, bộ phân loại RF dựa trên pixel đã được thực hiện trong GEE, tính khả thi cao của bộ phân loại RF trong việc lập bản đồ hệ sinh thái rừng ngập mặn (Zhao và Qin, 2020; Toosi và cộng sự, 2019; Nababa và cộng sự, 2020). Toàn bộ dữ liệu nghiên cứu bao gồm dữ liệu Sentinel-1 và Sentinel-2 của năm 2020 được đưa vào bộ phân loại RF. Trong khi đó, các mẫu tham chiếu được sử dụng để đào tạo bộ phân loại RF. Bộ phân loại RF có một số tham số điều chỉnh ảnh hưởng đến giai đoạn huấn

luyện của bước phân loại và ảnh hưởng trực tiếp đến kết quả phân loại. Trong số các tham số này, số lượng cây quyết định và biến tại mỗi nút là các tham số có ảnh hưởng lớn nhất (Belgiu và Drăguț, 2016). Kết quả phân loại bằng RF sẽ được thực hiện qua 3 thử sai dựa trên hiệu quả tính toán để chọn ra kết quả phù hợp nhất.

5) Đánh giá độ chính xác sau phân loại

Bất kỳ bản đồ chuyên đề nào có nguồn gốc từ dữ liệu viễn thám đều phải được đánh giá độ chính xác đáng tin cậy để đảm bảo chất lượng và độ tin cậy của nó (Ghorbanian và cộng sự, 2021). Do đó, cả hai phương pháp giải thích trực quan và phương pháp thống kê đã được thực hiện để đánh giá độ chính xác của bản đồ hệ sinh thái rừng ngập mặn cuối cùng.

Để đánh giá trực quan, bài báo sử dụng các hình ảnh vệ tinh có độ phân giải cao có sẵn trong Google Earth để đánh giá kết quả phân loại.

Về đánh giá độ chính xác thống kê, các mẫu thử nghiệm độc lập được kết hợp để tạo ra ma trận tính toán. Các ma trận phân loại và ma trận kiểm chứng được xây dựng để lập bảng chéo cho dữ liệu quan sát với dữ liệu tham chiếu bằng cách sử dụng hệ số Kappa (Long và cộng sự, 2014; Congalton, 1991). Hệ số Kappa là thước đo sự thống nhất giữa hai bản đồ, xem xét tất cả các yếu tố của ma trận sai số (Pham và Yoshino, 2016). Giá trị hệ số Kappa sau đó được phân loại thành bốn nhóm: Giá trị Kappa bằng 0 được coi là không có sự đồng nhất; từ 0,41 - 0,6 đồng nhất vừa phải; 0,61 - 0,8 đồng nhất đáng kể; và 0,81 - 1,0 coi như đồng nhất gần như hoàn hảo (Conchedda và cộng sự, 2008; Maryantika và Lin, 2017; Pham và Yoshino, 2016).

3. KẾT QUẢ NGHIÊN CỨU

3.1. Phân loại rừng ngập mặn từ giá trị tán xạ ngược VH và chỉ số thực vật CMRI

Kết quả tính toán từ dữ liệu đào tạo tách từ 60% tổng số các điểm mẫu ngẫu nhiên và phân tích thống kê cho các ngưỡng phân loại đối với khu vực nghiên cứu thể hiện trong bảng 2.

Bảng 2. Ngưỡng phân loại các trạng thái của khu vực nghiên cứu

Dữ liệu ảnh/Chỉ số	Rừng ngập mặn	Đối tượng khác	Mặt nước
CMRI (Sentinel-2)	≥ 0,41	0,0-0,41	< 0,01
VH (Sentinel-1)	(-16,8) - (-10,0)	(-26,0) - (-16,8)	< (-26,0)

Dữ liệu tại bảng 2 cho thấy, các ngưỡng phân biệt rừng ngập mặn được xác định bằng CMRI từ Sentinel-2, lớn hơn từ 0,41, khu vực có sự hiện diện thường xuyên của nước có ngưỡng nhỏ hơn 0,01, khoảng từ 0,01 đến 0,41 được xác định là các đối tượng còn lại. Ngưỡng phân biệt rừng ngập mặn bằng cách sử dụng tán xạ ngược VH thu được từ Sentinel-1 được xác định từ (-16,8) đến (-10,0) trong khi các đối tượng khác có ngưỡng từ (-26,0) đến (-16,8) và nước có ngưỡng nhỏ hơn (-26,0). Kết quả này cho thấy rằng các lớp phủ nước được dễ dàng xác định bằng cách sử dụng ngưỡng của CMRI cũng như hệ số tán xạ ngược VH. Đối với ảnh Sentinel-1 vẫn còn một số điểm ảnh biệt lập của bãi bồi, đất bồi và đất trồng bị phân loại nhầm thành thủy vực vì chúng có ngưỡng hệ số tán xạ ngược VH < -26,0, tuy nhiên các điểm ảnh biệt lập được phát hiện với diện tích rất nhỏ khi sàng lọc trên toàn bộ địa điểm nghiên cứu.

Kết quả tính toán ma trận đánh giá độ chính xác của kết quả phân loại khi sử dụng các

ngưỡng đã tính toán trên cả 2 loại tư liệu cho khu vực nghiên cứu cho thấy, tư liệu Sentinel-1 có độ chính xác tổng thể đạt 96,1% và hệ số Kappa = 0,94, trong khi Sentinel-2 cũng có độ chính xác tổng thể tương đương với 95,9% và hệ số Kappa = 0,93. Ma trận phân loại còn cho thấy, nước là đối tượng dễ dàng phân loại hơn khi sử dụng cả hai tư liệu khi đó độ chính xác sản xuất cao nhất, sau đó là rừng ngập mặn với độ chính xác phân loại trên ảnh Sentinel-1 và Sentinel-2 lần lượt là 96,2% và 97,3%, giá trị này cao hơn các đối tượng khác (96,0% trên ảnh Sentinel-1 và 89,6% trên ảnh Sentinel-2).

Với các kết quả sau phân loại tự động từ các dữ liệu đào tạo, tiếp tục sử dụng dữ liệu kiểm chứng (từ 40% số điểm được lấy ngẫu nhiên từ các điểm đã được tạo) để tính toán độ chính xác tổng thể và hệ số Kappa sau phân loại với từng trạng thái lớp phủ, kết quả được thể hiện trong bảng 3.

Bảng 3. Ma trận độ chính xác kết quả kiểm chứng sau phân loại từ giá trị VH và Chỉ số CMRI trên hai loại ảnh

		Dữ liệu kiểm chứng				
VH_Sentinel-1		Rừng ngập mặn	Đối tượng khác	Mặt nước	Tổng	Độ chính xác của người dùng (%)
	Rừng ngập mặn	299	2	0	301	99,3
	Đối tượng khác	11	113	0	124	91,1
	Mặt nước	0	2	109	111	98,2
	Tổng	310	117	109	536	
Độ chính xác sản xuất (%)	96,5	96,6	100	Độ chính xác tổng thể: 96,2% Kappa = 0,95		
		Dữ liệu kiểm chứng				
CMRI_Sentinel-2		Rừng ngập mặn	Đối tượng khác	Mặt nước	Tổng	Độ chính xác của người dùng (%)
	Rừng ngập mặn	296	5	0	301	98,3
	Đối tượng khác	11	113	0	124	91,1
	Mặt nước	0	2	109	111	98,2
	Tổng	307	120	109	536	
Độ chính xác sản xuất (%)	96,4	94,2	100	Độ chính xác tổng thể: 95,9% Kappa = 0,94		

Kết quả đánh giá độ chính xác từ các dữ liệu kiểm chứng cho thấy, dữ liệu Radar Sentinel-1 có độ chính xác tổng thể cao (96,2%) và hệ số Kappa lớn (0,95) điều này cho thấy chúng cũng rất hữu ích cho việc phân biệt rừng ngập mặn vì nó có độ phân giải không gian tốt (10 m), không

bị ảnh hưởng bởi mây che phủ và các yếu tố thời tiết khác. Tương tự, đối với dữ liệu Sentinel-2, việc đánh giá kết quả phân loại cũng cho thấy Chỉ số CMRI là bộ phân loại tốt với độ chính xác phân loại tổng thể 95,9%, hệ số Kappa lớn (0,94). Các kết quả này chứng tỏ có sự thống

nhất rất cao giữa bản đồ sau phân loại và dữ liệu tham chiếu. Các độ chính xác này hàm ý rằng việc sử dụng giá trị tán xạ ngược VH từ dữ liệu Sentinel-1 và chỉ số CMRI từ Sentinel-2 có tiềm năng lớn cho việc giám sát và lập bản đồ rừng ngập mặn ở khu vực nghiên cứu.

3.2. Phân loại bằng phương pháp Random Forest

Kết quả đánh giá độ chính xác phân loại ba trạng thái lớp phủ từ các dữ liệu đào tạo bằng RF cả hai tư liệu đều có độ chính xác tổng thể rất tốt, trong đó phân loại RF từ ảnh Sentinel-2 gần như có độ chính xác tuyệt đối (độ chính xác tổng thể: 99,5%, Kappa: 0,99), ảnh Sentinel-1 có độ chính xác tổng thể thấp hơn (96,7%) và Kappa: 0,94. Kết quả này phản ánh rằng, các tư liệu sử dụng phù hợp và đáng tin để phân loại ngẫu nhiên rừng ngập mặn ở khu vực nghiên cứu, tuy vậy độ chính xác của tư liệu quang học Sentinel-2 cao hơn so với tư liệu SAR Sentinel-1

Mặc dù kết quả đánh giá độ chính xác từ dữ liệu đào tạo đối với ảnh Sentinel-1 thấp hơn, tuy nhiên, điều này có thể giải thích bởi sự tham gia của số lượng các kênh ảnh vào thuật toán RF trên nền tảng GEE. Rõ ràng, với số lượng các kênh ảnh tham gia lớn hơn vào quá trình tính toán, tư liệu quang học Sentinel-2 tạo ra nhiều cây quyết định và kết quả sẽ có độ chính xác cao hơn so với việc chỉ sử dụng hai giá trị VH và VV cho thuật toán này từ ảnh Sentinel-1.

Tương tự với kết quả phân loại, kết quả kiểm chứng độ chính xác độ chính xác sau phân loại cũng có sự khác biệt khá rõ giữa hai loại tư liệu. Kết quả đánh giá từ các dữ liệu tham chiếu cho thấy độ chính xác tổng thể đối với tư liệu quang học Sentinel-2 rất tốt (98%), hệ số Kappa: 0,97, trong khi độ chính xác tổng thể trên ảnh Sentinel-1 chỉ khoảng 91,4% và hệ số Kappa khoảng 0,85 (bảng 4).

Bảng 4. Ma trận độ chính xác kết quả kiểm chứng sau phân loại bằng phương pháp RF trên hai loại ảnh

		Dữ liệu kiểm chứng				Độ chính xác của người dùng (%)
		Rừng ngập mặn	Đối tượng khác	Mặt nước	Tổng	
RF_Sentinel-1	Rừng ngập mặn	285	29	1	301	90,4
	Đối tượng khác	13	97	1	124	87,4
	Mặt nước	2	0	108	111	98,2
	<i>Tổng</i>	300	126	110	536	
	Độ chính xác sản xuất (%)	95,0	76,9	98,2	Độ chính xác tổng thể: 91,4% Kappa = 0,85	
	Dữ liệu kiểm chứng					
		Rừng ngập mặn	Đối tượng khác	Mặt nước	Tổng	Độ chính xác của người dùng (%)
RF_Sentinel-2	Rừng ngập mặn	295	6	0	301	98,0
	Đối tượng khác	4	120	0	124	96,7
	Mặt nước	1	0	110	111	99,0
	<i>Tổng</i>	300	126	110	536	
	Độ chính xác sản xuất (%)	98,3	95,2	100	Độ chính xác tổng thể: 98,0% Kappa = 0,97	

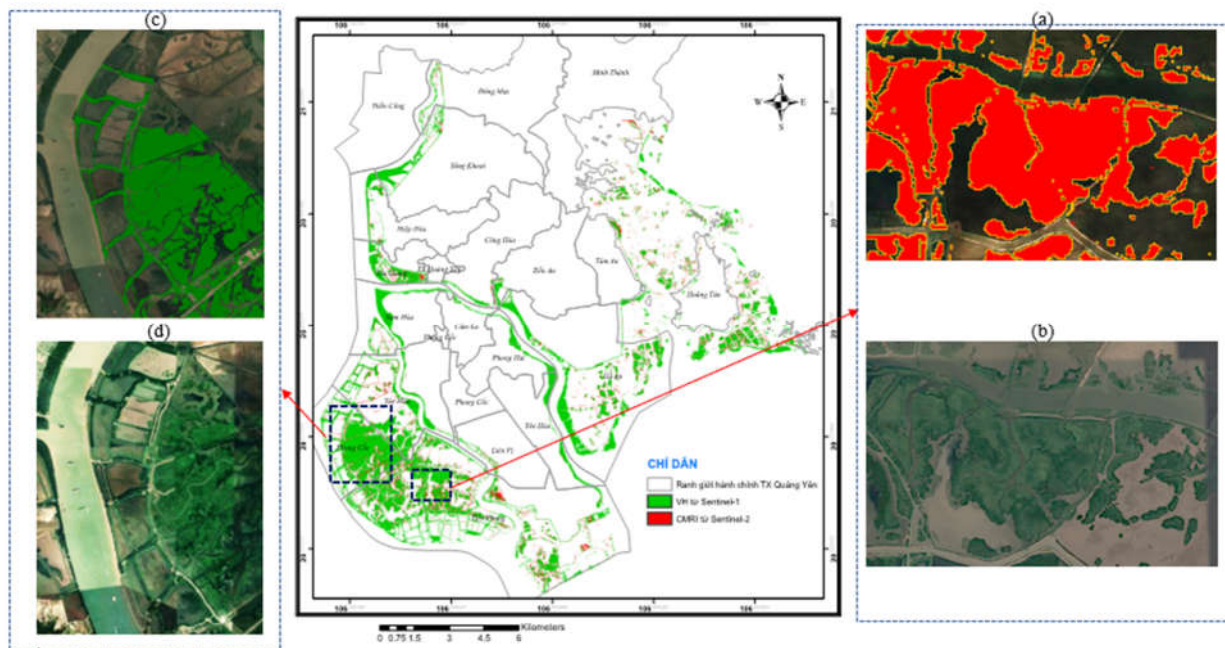
Trong khi độ chính xác giải đoán hiện trạng của trạng thái đối tượng khác trên ảnh Sentinel-2 đạt 95,2% thì trên ảnh Sentinel-1 chỉ đạt 76,9% mặc dù độ chính xác phân loại rừng ngập mặn và mặt nước ở cả hai tư liệu đều cao. Điều này chứng tỏ có một số điểm nhầm lẫn có thể xảy ra trong quá trình phân loại, như sự nhầm lẫn giữa rừng ngập mặn và một số khu vực có thể là đất ngập nước có thực vật sinh sống. Mặc dù vậy, tỉ lệ nhầm lẫn trên ảnh quang học lại ít

hơn so với ảnh Radar rất nhiều. Có thể thấy rằng, giá trị tán xạ ngược VH trên ảnh Sentinel-1 khá nhạy cảm với rừng ngập mặn nhưng không thích hợp để phân loại các trạng thái đối tượng khác. Trái ngược với kết quả này, có thể thấy tư liệu quang học Sentinel-2 lại thích hợp khi phân loại 3 trạng thái bởi vì sự chênh lệch về độ chính xác giữa 3 trạng thái lớp phủ gần như tương đương nhau (bảng 4).

3.3. Bản đồ phân loại và phân bố rừng ngập mặn

Đối với phương pháp phân loại bằng chỉ số CMRI và giá trị VH: Hình 3 trình bày bản đồ phân loại hệ sinh thái rừng ngập mặn khu vực nghiên cứu từ hai loại tư liệu gồm kết quả phân loại rừng ngập mặn từ VH trên ảnh Sentinel-1 và phân loại từ chỉ số CMRI. Bên cạnh kết quả kiểm chứng bằng dữ liệu thống kê, các kết quả kiểm chứng trực quan bằng ảnh vệ tinh độ phân giải cao cung cấp bởi Google Earth, các hình từ 3a đến 3d minh họa phóng to một khu vực cụ

thể đã được phân loại là rừng ngập mặn, trong đó các khu vực rừng ngập mặn được xác định thành công. Cụ thể, kiểm chứng giữa kết quả phân loại bằng CMRI trên Sentinel-2 (hình 3a) với ảnh độ phân giải cao Google Earth (hình 3b) và kết quả phân loại bằng VH trên Sentinel-1 (hình 3c) với ảnh độ phân giải cao Google Earth (hình 3d) đối với hai khu vực cụ thể đều cho kết quả trùng khớp lớn. Thậm chí, các khu vực ngập nước thường xuyên như bãi cỏ, bãi bồi, khu thực vật trong các đầm trũng... cũng đã được loại bỏ dễ dàng.



Hình 3. Bản đồ phân loại rừng ngập mặn được phân loại từ hai giá trị VH, CMRI và hình ảnh kiểm chứng trực quan từ ảnh độ phân giải cao năm 2020 trên Google Earth

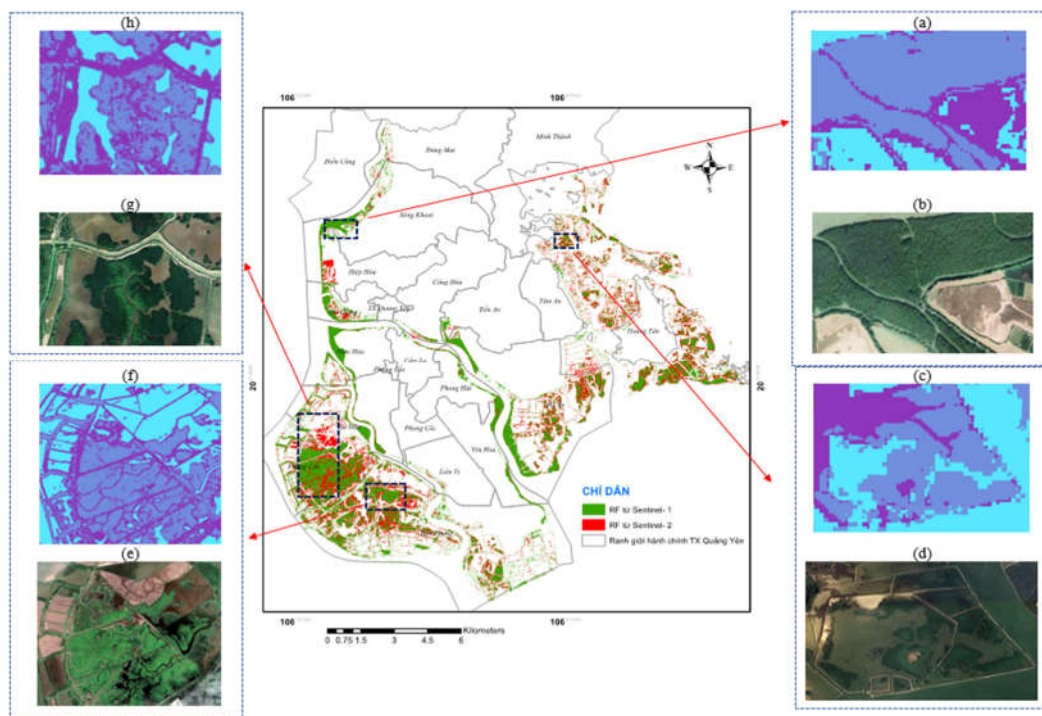
Kết quả này một lần nữa chứng minh rằng, sử dụng ngưỡng phân loại rừng ngập mặn của giá trị tán xạ ngược VH (trên ảnh Sentinel-1) và Chỉ số CMRI (trên ảnh Sentinel-2) là hoàn toàn phù hợp với mục tiêu phân loại rừng ngập mặn ở khu vực nghiên cứu.

Đối với phương pháp phân loại bằng thuật toán Random Forest: Hình 4 trình bày bản đồ hệ sinh thái rừng ngập mặn thu được (tức là chỉ các lớp trong hệ sinh thái rừng ngập mặn) và các kết quả kiểm chứng trực quan (từ hình 4a đến 4d đối với dữ liệu Sentinel-2 và hình 4e đến 4h đối với dữ liệu Sentinel-1). Về mặt trực quan, bản đồ chuyên đề có độ chính xác chấp nhận được, cho thấy tiềm năng cao của phương pháp RF được sử dụng để phân định các lớp khác nhau.

Kết quả phân loại bằng phương pháp rừng ngẫu nhiên RF từ tư liệu Sentinel-2 cho thấy, các khu vực rừng ngập mặn được mô tả chính xác và môi trường xung quanh chúng cũng được phân loại là mặt nước và các đối tượng khác. Ngoài ra, cùng với các hình ảnh vệ tinh có độ phân giải cao tương ứng trên Google Earth và hai khu vực phóng to cũng được cung cấp trong hình 4 để giải thích trực quan hơn. Ví dụ, hình 4a, 4b minh họa một khu vực phóng to, trong đó các khu vực rừng ngập mặn được xác định rất rõ ràng trên ảnh Sentinel-2. Hơn nữa, hình 4c và 4d cho thấy một số đối tượng mặc dù rất khó phân biệt cũng được tách ra khỏi hệ sinh thái rừng ngập mặn, có thể thấy các khu vực rừng ngập mặn và môi trường xung quanh (như các đầm nước có cây mọc nhưng không phải rừng

ngập mặn) đã được phân biệt rõ ràng. Ngoài ra, các khu vực khác không có lớp phủ cụ thể, ẩm ướt (do sự tồn tại của các kênh nước hẹp), cũng được phân loại chính xác (hình 4e và 4f). Tuy

nhiên, ở hình 4g, 4h cho thấy có một số khu vực là rừng ngập mặn nhưng đã bị phân loại là đối tượng khác trên ảnh Sentinel-1.



Hình 4. Bản đồ phân loại rừng ngập mặn bằng phương pháp RF và hình ảnh kiểm chứng trực quan từ ảnh độ phân giải cao năm 2020 trên Google Earth

Kết quả phân loại rừng ngập mặn từ hai phương pháp đã cho phép tạo các vùng phân bố rừng ngập mặn đồng thời tính toán được diện tích của chúng trên GEE (bảng 5).

Bảng 5. Kết quả phân loại rừng ngập mặn từ các phương pháp trên hai loại tư liệu như sau

TT	Tư liệu ảnh	Phương pháp phân loại			Chênh lệch giữa hai phương pháp
		Chỉ số CMRI/	Giá trị tán xạ ngược VH	Random Forest	
1	Sentinel-1	-	2790,4	2503,1	287,3
2	Sentinel-2	2802,2	-	2945,6	143,4

Có thể thấy, xét trên cùng một phương pháp phân loại, hầu như không có sự khác biệt giữa diện tích rừng ngập mặn phân loại từ CMRI lấy từ Sentinel-2 và diện tích rừng ngập mặn từ tán xạ ngược VH được chiết xuất từ Sentinel-1. Kết quả phân loại cho thấy, diện tích rừng ngập mặn phân loại từ CMRI trên ảnh Sentinel-2 và VH trên Sentinel-1 lần lượt là 2802,2 ha và 2790,4 ha (chênh lệch khoảng 11,8 ha). Trong khi đó, đối với phương pháp Random Forest quan sát thấy có sự chênh lệch rất lớn về diện tích từ hai loại ảnh (diện tích rừng ngập mặn phân loại từ

ảnh Sentinel-1 là 2503,1 ha và từ ảnh Sentinel-2 là 2945,6 ha, chênh lệch khoảng 442,5 ha).

Ngoài ra, xét trên cùng một tư liệu ảnh thì diện tích quan sát được giữa phương pháp sử dụng ngưỡng chỉ số và phương pháp RF có sự không thống nhất trên cùng một tư liệu. Rõ ràng, diện tích rừng ngập mặn quan sát được từ hai phương pháp trên tư liệu Sentinel-1 có sự chênh lệch lên đến 287,3 ha, trong khi đó trên tư liệu Sentinel-2 thấp hơn khoảng một nửa (143,4 ha). Điều này cho thấy, dữ liệu quang học Sentinel-2 có sự ổn định hơn so với dữ liệu Sentinel-1 khi phân loại rừng ngập mặn đối với khu vực nghiên cứu.

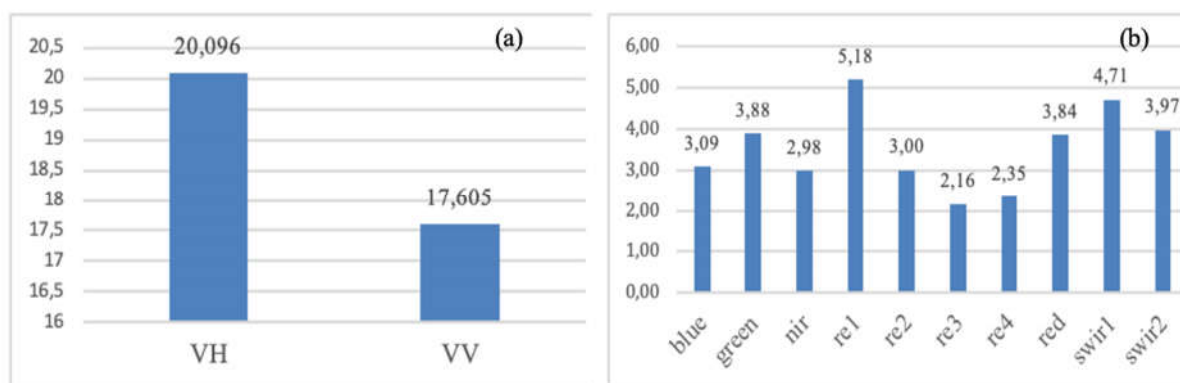
4. THẢO LUẬN

Các hệ sinh thái rừng ngập mặn thường xuyên bị ngập úng và nằm trong vùng thủy triều liên tục (Zhu và cộng sự, 2019). Do đó, việc khảo sát thực địa để thu thập mẫu đối chứng bằng thiết bị hệ thống định vị toàn cầu (GPS) gặp nhiều khó khăn (Zhao và Qin, 2020). Sự sẵn có của các mẫu tham chiếu có độ chính xác cao thông qua khảo sát thực địa là rất quan trọng để lập bản đồ đáng tin cậy. Tuy nhiên, việc giải thích trực quan chính xác các hình ảnh vệ tinh có độ phân giải cao có thể bù đắp cho hạn chế này. Các phương pháp phân loại từ các ngưỡng chỉ số thực vật cũng như sử dụng hệ thống phân loại ngẫu nhiên có thể đạt được các giá trị độ chính xác tổng thể cũng như hệ số Kappa khá cao, điều này cho thấy sự phù hợp của việc áp dụng các phương pháp này khi việc điều tra thực địa gặp nhiều khó khăn (Ghorbanian và cộng sự, 2021).

Có thể thấy, các phương pháp phân loại khác nhau cho kết quả không đồng nhất. Các kết quả sử dụng ngưỡng chỉ số theo cách chủ động cho phép phân loại tốt rừng ngập mặn, tuy vậy đôi khi chúng có thể phân loại các loại lớp phủ có phản xạ tương tự rừng ngập mặn (cây ngập nước quanh năm) là rừng ngập mặn (Hu và cộng sự, 2020), phương pháp này cho phép phân loại một cách chính xác nếu dữ liệu mặt đất phục vụ phân loại được thu thập lớn và đảm bảo độ chính xác cao (Jia và cộng sự, 2019). Ngoài ra, do phương pháp sử dụng ngưỡng các chỉ số có sự chọn lọc các kênh ảnh kỹ lưỡng bởi các chỉ số được lựa chọn chỉ sử dụng một số kênh ảnh vào biểu thức toán học cho nên tạo ra kết quả chủ quan theo ngưỡng cho trước. Không chỉ vậy, phương pháp này còn cho thấy giá trị tán xạ ngược VH trên ảnh Sentinel-1 có kết quả phân loại tốt hơn so với ảnh quang học Sentinel-2. Bên cạnh đó, phương pháp sử dụng ngưỡng chỉ số CMRI và giá trị tán xạ ngược VH cũng có nhiều ưu điểm bởi chúng đã được nghiên cứu phát triển trước đây và được chứng minh là thích hợp để phân loại rừng ngập mặn ra khỏi các lớp phủ khác (Ali và Nayyar, 2020; Alvin B và cộng sự, 2020; Chen và cộng sự, 2017; Gupta và cộng sự, 2018). Ngoài ra, việc phân loại theo ngưỡng là phương pháp mang tính chủ động cao, phù hợp

với nhiều khu vực bởi vì có thể điều chỉnh ngưỡng với từng tiểu vùng cụ thể. Tuy nhiên, phương pháp này phụ thuộc lớn vào chủ quan của người sử dụng và mục đích sử dụng bởi vì mỗi chỉ số có một biểu thức toán học khác nhau và các kênh ảnh sử dụng trong biểu thức cũng không giống nhau.

Trái ngược với việc sử dụng ngưỡng các chỉ số, phương pháp sử dụng thuật toán Random Forest thể hiện ưu thế và hoạt động rất tốt khi phân loại tự động các trạng thái ở khu vực nghiên cứu, đặc biệt là với tư liệu quang học, khi chủ động sử dụng toàn bộ kênh ảnh của tư liệu này và cho kết quả phân loại rất cao độ chính xác tổng thể chỉ đạt 98%, Kappa: 0,97). Tuy nhiên, với tư liệu SAR Sentinel-1, sử dụng giá trị phản xạ VH và VV có kết quả đánh giá độ chính xác sau phân loại thấp hơn (độ chính xác tổng thể chỉ đạt 91,4%, Kappa: 0,85) điều này nhận định rằng đối với rừng ngập mặn ở khu vực nghiên cứu, sử dụng phương pháp Random Forest trên ảnh quang học Sentinel-2 thích hợp hơn so với ảnh Sentinel-1. Điều này có thể được quan sát cụ thể tại hình 5 kết quả tính toán thể hiện sự ảnh hưởng của các biến là các kênh ảnh đến kết quả của Random Forest. Có thể thấy, các kênh ảnh GREEN, RED1, SWIR1 và SWIR2 (tương ứng với các band 3, 5, 11 và 12) trên tư liệu Sentinel-2 đã chi phối rất lớn đến kết quả nghiên cứu. Cụ thể, mức độ tham gia của các biến kênh 3 (3,88), kênh 5 (5,18), kênh 11 (4,71) và kênh 12 (3,97) chúng là những dải có khả năng làm nổi bật các giá trị phản xạ của rừng ngập mặn vốn là một hệ sinh thái đặc trưng, bên cạnh đó các dải SWIR chủ yếu được sử dụng để đặc trưng cho các tính năng liên quan đến độ ẩm hoặc nước, do các đặc tính hút phản xạ của nước xung quanh các dải này và được cho là tốt hơn trong việc mô tả các đặc điểm và chất lượng của rừng ngập mặn (Manna và Raychaudhuri, 2018). So sánh với tư liệu Sentinel-1, sự tham gia của hai giá trị phản xạ là VH và VV tạo nên kết quả phân loại thấp hơn. Mặc dù vậy, tần suất tham gia của giá trị VH cao (20,096) là giá trị quyết định đến kết quả phân loại rừng ngập mặn ở khu vực nghiên cứu.



Hình 5. Biểu đồ tần suất tham gia của các band ảnh trên tư liệu Sentinel-1 (a) và Sentinel-2 (b) đối với kết quả phân loại bằng thuật toán rừng ngẫu nhiên RF

Trong nghiên cứu này, nền tảng điện toán đám mây GEE được sử dụng để lập bản đồ hệ sinh thái rừng ngập mặn. GEE cung cấp cơ hội rất lớn để sử dụng dữ liệu chuỗi thời gian dày đặc từ các hình ảnh vệ tinh miễn phí và nó cũng chứa nhiều thuật toán máy học và xử lý hình ảnh tích hợp cho các thao tác dữ liệu vệ tinh (Amani và cộng sự, 2020; Gorelick và cộng sự, 2017). Hơn nữa, sự tồn tại của dữ liệu sẵn sàng sử dụng cùng với các đặc điểm khác, góp phần hướng tới việc phát triển các phương pháp hiệu quả với tỷ lệ tự động hóa cao hơn và do đó cho phép giảm thời gian dành riêng cho việc thu thập, hiệu chỉnh và tiền xử lý dữ liệu (Cárdenas và cộng sự, 2017).

Có thể thấy, dữ liệu quang học với sự đa dạng về kênh phổ đang tỏ ra là tài nguyên có khả năng đáp ứng tốt yêu cầu về giám sát, thành lập bản đồ rừng ngập mặn, đặc biệt nền tảng GEE đã cho phép xử lý dữ liệu đối với các khu vực rộng lớn trong khoảng thời bất kỳ điều mà trước đây, sử dụng các phần mềm chuyên ngành vốn tiêu tốn nhiều thời gian và công sức. Bên cạnh đó, dữ liệu SAR Sentinel-1 cũng có nhiều lợi thế bởi không ảnh hưởng của thời tiết và phù hợp với khu vực có địa hình bằng phẳng, tuy nhiên nghiên cứu này không đề cập đến dữ liệu đơn đơn thời gian mà sử dụng ảnh trung bình năm 2020 để giảm sự nhiễu động về chất lượng ảnh, điều này ít ảnh hưởng đến kết quả phân loại rừng ngập mặn bởi các loài cây rừng ngập mặn chủ yếu là loài sinh trưởng chậm, bên cạnh đó dữ liệu cảm biến từ xa đa chiều có thể cải thiện khả năng phân loại rừng ngập mặn tốt hơn so với dữ liệu ảnh đơn chiều (Zhang và cộng sự, 2018).

5. KẾT LUẬN

Lập bản đồ chi tiết về hệ sinh thái rừng ngập mặn là cần thiết cho việc theo dõi các mục tiêu phát triển bền vững và bảo tồn tài nguyên thiên nhiên. Trong bài báo này, hai phương pháp phân loại rừng ngập mặn (sử dụng các ngưỡng chỉ thực vật và sử dụng phương pháp Random Forest) đã được sử dụng trên cả hai dữ liệu viễn thám quang học (Sentinel-2) và SAR (Sentinel-1) trên nền tảng điện toán đám mây GEE. Nhìn chung, xét đến cả tiêu chí đánh giá định tính (kiểm chứng bằng mắt) và định lượng (đánh giá độ chính xác thống kê), cả hai phương pháp được sử dụng đã khẳng định khả năng ứng dụng của nó trong việc lập bản đồ chi tiết hệ sinh thái rừng ngập mặn với sự thích hợp giữa kết quả phân loại và kiểm chứng trên ảnh độ phân giải cao cũng như các kết quả đánh giá độ chính xác đều chứng minh là phù hợp. Đối với dữ liệu quang học Sentinel-2, kết quả đánh giá độ chính xác tổng thể của hai phương pháp sử dụng ngưỡng Chỉ số CMRI và phương pháp Random Forest lần lượt là 95,9%, (Kappa = 0,95) và 98% (Kappa = 0,97). Trong khi đó, dữ liệu SAR Sentinel-1 kém ổn định và có sự chênh lệch khá lớn về độ chính xác tổng thể với 96,2% (Kappa = 0,95) từ phương pháp dùng ngưỡng giá trị tán xạ ngược VH và 91,4% (Kappa = 0,85) từ phương pháp sử dụng thuật toán Random Forest. Kết quả tính diện tích khu vực nghiên cứu cho biết, dữ liệu Sentinel-2 có sự chênh lệch về diện tích quan sát được giữa hai phương pháp khoảng 143 ha, trong khi dữ liệu SAR Sentinel-1 có sự chênh lệch lớn hơn với khoảng 287 ha. Các kết quả này cho thấy, dữ liệu quang học

Sentinel 2 có sự ổn định hơn khi phân loại rừng ngập mặn đối với khu vực thị xã Quảng Yên, tỉnh Quảng Ninh.

TÀI LIỆU THAM KHẢO

1. Ammar Abdul Aziz, Stuart Phinn and Paul Dargusch (2015). 'Investigating the decline of ecosystem services in a production mangrove forest using Landsat and object-based image analysis', *Estuarine, Coastal and Shelf Science*, 164: 353-366.
2. Ahmed Ali and Zeeshan Alam Nayyar (2020). 'Extraction of mangrove forest through Landsat 8 Mangrove Index (L8MI)', *Arabian Journal of Geosciences*, 13.
3. Baloloy Alvin B, Blanco Ariel C, Sta. Ana Raymund Rhommel C and Nadaoka Kazuo (2020). 'Development and application of a new mangrove vegetation index (MVI) for rapid and accurate mangrove mapping', *ISPRS Journal of Photogrammetry and Remote Sensing*, 166: 95-117.
4. Meisam Amani, Arsalan Ghorbanian, Seyed Ali Ahmadi, Mohammad Kakooei, Armin Moghimi, S. Mohammad Mirmazloumi, Sayyed Hamed Alizadeh Moghaddam, Sahel Mahdavi, Masoud Ghahremanloo, Saeid Parsian, Qiusheng Wu and Brian Brisco (2020). 'Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 5326-5350.
5. Mariana Belgiu and Lucian Drăguț (2016). 'Random forest in remote sensing: A review of applications and future directions', *ISPRS Journal of Photogrammetry and Remote Sensing*, 114: 24-31.
6. Leo Breiman (2001). 'Random Forests', *Machine Learning*, 45: 5-32.
7. Younes Nicolás Cárdenas, Karen E Joyce and Stefan W Maier (2017). *Monitoring mangrove forests: Are we taking full advantage of technology?*
8. Bangqian Chen, Xiangming Xiao, Xiangping Li, Lianghao Pan, Russell Doughty, Jun Ma, Jinwei Dong, Yuanwei Qin, Bin Zhao, Zhixiang Wu, Rui Sun, Guoyu Lan, Guishui Xie, Nicholas Clinton and Chandra Giri (2017). 'A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform', *ISPRS Journal of Photogrammetry and Remote Sensing*, 131: 104-120.
9. Giulia Conchedda, Laurent Durieux and Philippe Mayaux (2008). 'An object-based method for mapping and change analysis in mangrove ecosystems', *ISPRS Journal of Photogrammetry and Remote Sensing*, 63: 578-589.
10. R.G Congalton (1991). 'A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment', *Remote Sens*, 37: 35-46.
11. Nguyễn Trọng Cường, Trần Quang Bảo, Nguyễn Hải Hòa, Phạm Văn Duẩn và Phạm Ngọc Hải (2021). 'Tổng quan sử dụng tư liệu viễn thám để lập bản đồ rừng ngập mặn', *KH&CN Lâm Nghiệp*, 03: 65-76.
12. Bo-Cai Gao (1996). 'NDWI-A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water From Space', *Elsevier*, 96.
13. Anang D.Purwanto Gathot Winarso, Doddy M.Yuwono (2014). 'New Mangrove Index As Degradation Health Indicator Using Remote Sensing Data: Segara Anakan and Alas Purwo Case Study'.
14. Arsalan Ghorbanian, Mohammad Kakooei, Meisam Amani, Sahel Mahdavi, Ali Mohammadzadeh and Mahdi Hasanlou (2020). 'Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples', *ISPRS Journal of Photogrammetry and Remote Sensing*, 167: 276-288.
15. Arsalan Ghorbanian, Soheil Zaghian, Reza Mohammadi Asiyabi, Meisam Amani, Ali Mohammadzadeh and Sadegh Jamali (2021). 'Mangrove Ecosystem Mapping Using Sentinel-1 and Sentinel-2 Satellite Images and Random Forest Algorithm in Google Earth Engine', *Remote Sensing*, 13.
16. C. Giri, J. Long, S. Abbas, R. M. Murali, F. M. Qamer, B. Pengra and D. Thau (2015). 'Distribution and dynamics of mangrove forests of South Asia', *J Environ Manage*, 148: 101-111.
17. Chandra Giri, Jordan Long and Larry Tieszen (2011). 'Mapping and Monitoring Louisiana's Mangroves in the Aftermath of the 2010 Gulf of Mexico Oil Spill', *Journal of Coastal Research*, 277: 1059-1064.
18. Noel Gorelick, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau and Rebecca Moore (2017). 'Google Earth Engine: Planetary-scale geospatial analysis for everyone', *Remote Sensing of Environment*, 202: 18-27.
19. K Gupta, A Mukhopadhyay, S Giri, A Chanda, S Datta Majumdar, S Samanta, D Mitra, R. N Samal, A. K Pattnaik and S Hazra (2018). 'An index for discrimination of mangroves from non-mangroves using LANDSAT 8 OLI imagery', *MethodsX*, 5: 1129-1139.
20. Luojia Hu, Nan Xu, Jian Liang, Zhichao Li, Luzhen Chen and Feng Zhao (2020). 'Advancing the Mapping of Mangrove Forests at National-Scale Using Sentinel-1 and Sentinel-2 Time-Series Data with Google Earth Engine: A Case Study in China', *Remote Sensing*, 12.
21. Mingming Jia, Zongming Wang, Chao Wang, Dehua Mao and Yuanzhi Zhang (2019). 'A New Vegetation Index to Detect Periodically Submerged Mangrove Forest Using Single-Tide Sentinel-2 Imagery', *Remote Sensing*, 11.
22. Tsai-Ming Lee and Hui-Chung Yeh (2009). 'Applying remote sensing techniques to monitor shifting wetland vegetation: A case study of Danshui River estuary mangrove communities, Taiwan', *Ecological Engineering*, 35: 487-496.
23. Jordan Long, Darrell Napton, Chandra Giri and Jordan Graesser (2014). 'A Mapping and Monitoring Assessment of the Philippines' Mangrove Forests from 1990 to 2010', *Journal of Coastal Research*, 294: 260-271.
24. Sudip Manna and Barun Raychaudhuri (2018). 'Mapping distribution of Sundarban mangroves using Sentinel-2 data and new spectral metric for detecting their health condition', *Geocarto International*, 35: 434-452.

25. Norida Maryantika and Chinsu Lin (2017). 'Exploring changes of land use and mangrove distribution in the economic area of Sidoarjo District, East Java using multi-temporal Landsat images', *Information Processing in Agriculture*, 4: 321-332.
26. Iliya Nababa, Elias Symeonakis, Sotirios Koukoulas, Thomas Higginbottom, Gina Cavan and Stuart Marsden (2020). 'Land Cover Dynamics and Mangrove Degradation in the Niger Delta Region', *Remote Sensing*, 12.
27. Tien Dat Pham and Kunihiko Yoshino (2016). 'Impacts of mangrove management systems on mangrove changes in the Northern Coast of Vietnam', *Tropics*, 24: 141-151.
28. Abdullah F. Rahman, Danilo Dragoni, Kamel Didan, Armando Barreto-Munoz and Joseph A. Hutabarat (2013). 'Detecting large scale conversion of mangroves to aquaculture with change point and mixed-pixel analyses of high-fidelity MODIS data', *Remote Sensing of Environment*, 130: 96-107.
29. Neda Bihamta Toosi, Ali Reza Soffianian, Sima Fakheran, Saeid Pourmanafi, Christian Ginzler and Lars T. Waser (2019). 'Comparing different classification algorithms for monitoring mangrove cover changes in southern Iran', *Global Ecology and Conservation*, 19.
30. C.J Tucker (1997). 'Red and Photographic Infrared 1,linear Combinations for Monitoring Vegetation', *REMOTE SENSING OF ENVIRONMENT* 8: 127-150.
31. Ivan Valiela, Jennifer L Bowen and Joanna K York (2001). 'Mangrove Forests: One of the World's Threatened Major Tropical Environments', *BioScience*, 51.
32. Bijeesh Kozhikkodan Veetil, Sebastian Felipe Ruiz Pereira and Ngo Xuan Quang (2018). 'Rapidly diminishing mangrove forests in Myanmar (Burma): a review', *Hydrobiologia*, 822: 19-35.
33. Dezhi Wang, Bo Wan, Penghua Qiu, Yanjun Su, Qinghua Guo, Run Wang, Fei Sun and Xincai Wu (2018). 'Evaluating the Performance of Sentinel-2, Landsat 8 and Pléiades-1 in Mapping Mangrove Extent and Species', *Remote Sensing*, 10.
34. Hongsheng Zhang, Ting Wang, Mingfeng Liu, Mingming Jia, Hui Lin, LM Chu and Adam Thomas Devlin (2018). 'Potential of Combining Optical and Dual Polarimetric SAR Data for Improving Mangrove Species Discrimination Using Rotation Forest', *Remote Sensing*, 10: 467.
35. Chuanpeng Zhao and Cheng-Zhi Qin. (2020). '10-m-resolution mangrove maps of China derived from multi-source and multi-temporal satellite observations', *ISPRS Journal of Photogrammetry and Remote Sensing*, 169: 389-405.
36. Xudong Zhu, Yuwen Hou, Qihao Weng and Luzhen Chen (2019). 'Integrating UAV optical imagery and LiDAR data for assessing the spatial relationship between mangrove and inundation across a subtropical estuarine wetland', *ISPRS Journal of Photogrammetry and Remote Sensing*, 149: 146-156.

USE GOOGLE EARTH ENGINE PLATFORM TO AUTOMATICALLY CLASSIFY MANGROVE FORESTS FROM SENTINEL-1 AND SENTINEL-2 FOR QUANG YEN AREA, QUANG NINH PROVINCE

Nguyen Trong Cuong¹, Tran Quang Bao², Nguyen Hai Hoa¹

¹Vietnam National University of Forestry

²Vietnam Administration of Forestry

SUMMARY

This study used Google Earth Engine (GEE) platform to test and automatically classify mangrove forests for Quang Yen area, Quang Ninh province. All Sentinel-1 Radar satellite images and Sentinel-2 optics with a cloud ratio of less than 20% in 2020 of the study area were used. Applying two methods of mangrove classification including (1) using the threshold to classify mangroves from the backscatter value of VH on Radar Sentinel-1 images and the threshold of mangrove index CMRI (Combine Mangrove Recognition Index) on the Sentinel-2 optical image; (2) random forest method (Random Forest - RF) on both data types. The results of mangrove classification of both methods have high accuracy. Sentinel-2 optical data have more stability when the overall accuracy of the two methods using the threshold of CMRI Index and Random Forest method is 95.9% (Kappa = 0.95) and 98% (Kappa = 0.97), respectively. The Sentinel-1 SAR data were less stable with an overall accuracy of 96.2% (Kappa = 0.95) using the VH backscatter value threshold method and 91.4% (Kappa = 0.85) using the Random Forest method. The results of calculating the area of mangroves in the study area show that the Sentinel-2 data has less area difference between the two methods (about 143 ha), while the Sentinel-1 SAR data has a larger area difference (about 287 ha).

Keywords: google earth engine, mangrove forest, mangrove mapping, randomforest, sentinel-1, sentinel-2.

Ngày nhận bài : 23/12/2021

Ngày phản biện : 18/02/2021

Ngày quyết định đăng : 25/02/2022