Tạp chí KHLN Số 3/2025: 133 - 143
©: Viện KHLNVN - VAFS
DOI: https://doi.org/10.70169/VJFS.1060
ISSN: 1859 - 0373 Đăng tải tại: https://vjfs.vafs.gov.vn

SO SÁNH PHƯƠNG PHÁP PHÂN LOẠI THẢM PHỦ SỬ DỤNG ĐẤT BẰNG THUẬT TOÁN RANDOM FOREST SỬ DỤNG CHỈ SỐ NDVI VỚI TỔ HỢP CHỈ SỐ NDVI, DEM, NDWI, NDSI TẠI HUYỆN ĐẮK GLONG VÀ HUYỆN KRÔNG NÔ - TỈNH ĐẮK NÔNG

Cao Thị Hoài, Nguyễn Thị Thanh Hương

Trường Đại học Tây Nguyên

TÓM TẮT

Nghiên cứu này tập trung đánh giá và so sánh hiệu quả phân loại lớp phủ sử dụng đất bằng thuật toán Random Forest (RF) với hai cách tiếp cận: (i) sử dụng riêng chỉ số thực vật chuẩn hóa (NDVI) và (ii) tích hợp tổ hợp chỉ số NDVI, mô hình số độ cao (DEM), chỉ số nước chuẩn hóa (NDWI) và chỉ số đất chuẩn hóa (NDSI) trên ảnh Landsat tại huyện Đắk Glong và Krông Nô, tỉnh Đắk Nông. Kết quả cho thấy, phương pháp sử dụng riêng NDVI cho độ chính xác tổng thể (OA) thấp, dao động từ 35,89% đến 59,35%, hệ số Kappa (K) chỉ đạt từ 0,24 đến 0,41. Trong khi đó, phương pháp tích hợp nhiều chỉ số đã cải thiện đáng kể hiệu suất mô hình, với OA tăng lên 81,75 - 86,04% và K đạt 0,75 - 0,78. Chênh lệch OA giữa hai phương pháp lên tới 45,86%, còn hệ số K tăng trung bình khoảng 0,40 đơn vị. Việc bổ sung DEM, NDWI và NDSI đã giúp khắc phục hiện tượng trộn phổ và tăng khả năng phân biệt giữa các lớp phủ có đặc tính tương tự. Trên cơ sở đó, nghiên cứu khuyến nghị áp dụng phương pháp tích hợp chỉ số NDVI, DEM, NDWI và NDSI để phân loại lớp phủ sử dụng đất sử dụng ảnh viễn thám, đặc biệt tại các vùng có địa hình phức tạp như Tây Nguyên, nhằm nâng cao độ chính xác và độ tin cậy của sản phẩm bản đồ phục vụ công tác quy hoạch và giám sát tài nguyên đất, rừng bền vững. Kết quả phân tích biến động lớp phủ đất tại huyện Đắk Glong và Krông Nô cho thấy xu hướng suy giảm mạnh diện tích rừng tự nhiên, trong khi rừng trồng có sự gia tăng rõ rệt, phản ánh quá trình chuyển đổi sử dụng đất theo hướng phát triển rừng sản xuất. Đồng thời, diện tích đất trống, đất nông nghiệp và khu dân cư cũng có nhiều biến động do tác động từ phát triển kinh tế - xã hội và nhu cầu sử dụng đất ngày càng tăng.

Từ khóa: Chỉ số đất chuẩn hóa, chỉ số nước chuẩn hóa, chỉ số thực vật chuẩn hóa, mô hình số độ cao, thảm phủ sử dụng đất.

COMPARISON OF LAND USE LAND COVER CLASSIFICATION METHODS USING THE RANDOM FOREST ALGORITHM WITH NDVI VERSUS A COMBINED INDEX APPROACH (NDVI, DEM, NDWI, AND NDSI) IN DAK GLONG AND KRONG NO DISTRICTS, DAK NONG PROVINCE

Cao Thi Hoai, Nguyen Thi Thanh Huong

Tay Nguyen University

ABSTRACT

This study focuses on evaluating and comparing the classification performance of Land use and land cover (LULC) using the Random Forest (RF) algorithm with two approaches: (i) employing only the Normalized difference vegetation index (NDVI), and (ii) integrating a combination of NDVI, Digital elevation model (DEM), Normalized difference water index (NDWI), and Normalized difference soil index (NDSI) derived from Landsat imagery in Dak Glong and Krong No districts, Dak Nong province. The results indicate that the NDVI-only approach yields relatively low overall accuracy (OA), ranging from 35.89% to 59.35%, with a Kappa coefficient (K) between 0.24 and 0.41. In contrast, the integrated index approach significantly enhances model performance, achieving OA between 81.75% and 86.04%, and K ranging from 0.75 to 0.78. The OA improvement reaches up to 45.86%, while the average K increase is approximately 0.40. The inclusion of DEM, NDWI, and NDSI effectively mitigates spectral confusion and improves the discrimination between spectrally similar land cover classes. Based on these findings, the study recommends applying the integrated NDVI, DEM, NDWI, and NDSI approach in LULC classification tasks using remote sensing data, especially in topographically complex regions such as the Central Highlands, to enhance classification accuracy and the reliability of land use maps for sustainable land and forest resource planning and monitoring. The analysis of

land cover changes in Dak Glong and Krong No districts reveals a significant decline in natural forest areas, while planted forests have increased markedly, reflecting a shift in land use towards the development of production forests. At the same time, the areas of bare land, agricultural land, and residential zones have also experienced considerable changes due to socio-economic development and the growing demand for land use.

Keywords: Normalized difference soil index, normalized difference water index, normalized difference vegetation index, digital elevation model, land use and land cover.

I. ĐẶT VẤN ĐỀ

Phân loại thảm phủ sử dụng đất là nền tảng quan trọng trong quy hoạch không gian, giám sát tài nguyên và quản lý môi trường bền vững. Tại các khu vực có địa hình phức tạp như Tây Nguyên, việc phân biệt ranh giới giữa các loại lớp phủ - rừng tự nhiên, đất nông nghiệp, đất trống hay mặt nước - trở nên đặc biệt khó khăn do hiện tượng trộn phổ phản xạ và sự thay đổi vi mô địa hình ảnh hưởng đến tín hiệu vệ tinh.

Trong bối cảnh đó, các phương pháp học máy hiện đại, đặc biệt là thuật toán Random Forest (RF), ngày càng được ứng dụng rộng rãi nhờ tính ổn định và khả năng xử lý dữ liệu lớn với độ chính xác cao (Eduardo *et al*., 2024; Yibai Xiong *et al*., 2022; Pelletier *et al.,* 2016; Lowe, Kulkarni, 2015). Tuy nhiên, hiệu quả của RF không chỉ phụ thuộc vào cấu trúc mô hình mà còn phụ thuộc đáng kể vào chất lượng và sự phù hợp của tập biến đầu vào (Dongping *et al*., 2016).

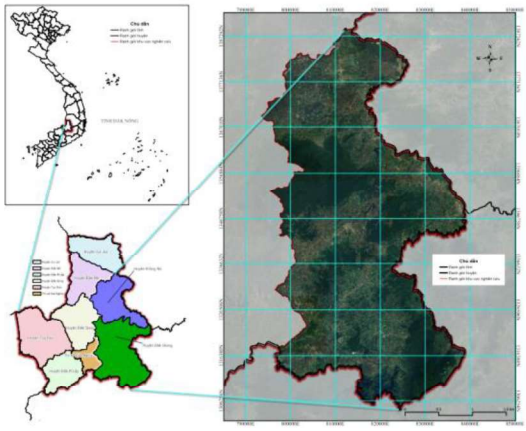
Chỉ số thực vật chuẩn hóa (Normalized difference vegetation index - NDVI) là một trong những chỉ số phổ biến nhất trong phân tích lớp phủ thực vật nhờ đơn giản và hiệu quả. Tuy nhiên, nhiều nghiên cứu đã chỉ ra rằng việc sử dụng đơn lẻ NDVI không đủ để phân biệt các lớp phủ có phổ phản xạ tương đồng, đặc biệt tại các vùng đồi núi có cấu trúc cảnh quan phức tạp (Ibrahim, 2023). Việc tích hợp thêm các chỉ số như mô hình số độ cao (Digital elevation model - DEM), chỉ số nước chuẩn hóa (Normalized difference water index - NDWI) và chỉ số đất chuẩn hóa (Normalized difference soil index - NDSI) có khả năng cải thiện đáng kể hiệu suất phân loại (Amoakoh *et al*, 2024; Eduardo *et al*., 2024; Ur Rehman, 2021).

Huyện Đắk Glong và Krông Nô, tỉnh Đắk Nông là minh chứng điển hình cho sự phức tạp trong phân loại thảm phủ sử dụng đất: địa hình chia cắt mạnh, thảm phủ đa dạng và tốc độ biến động sử dụng đất nhanh do phát triển nông lâm nghiệp và tác động của biến đổi khí hậu. Tuy nhiên, chưa có nhiều nghiên cứu định lượng nào so sánh trực tiếp hiệu quả của RF khi chỉ sử dụng NDVI và khi tích hợp với các chỉ số phổ - địa hình tại khu vực này.

Xuất phát từ đó, nghiên cứu này được thực hiện nhằm so sánh độ chính xác phân loại lớp phủ đất bằng thuật toán RF khi: (i) sử dụng chỉ số NDVI đơn lẻ và (ii) tích hợp tổ hợp chỉ số NDVI, DEM, NDWI và NDSI, trên ảnh vệ tinh Landsat tại khu vực Đắk Glong và Krông Nô trong các năm 2010, 2015 và 2023, nhằm đề xuất phương pháp phân loại phù hợp cho khu vực có tính chất địa hình và lớp phủ phức tạp.

II. KHU VỰC VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Khu vực nghiên cứu



Hình 1. Bản đồ khu vực nghiên cứu huyện Đắk Glong và Krông Nô thuộc tỉnh Đắk Nông, Việt Nam

Huyện Đắk Glong và Krông Nô thuộc tỉnh Đắk Nông, nằm ở khu vực Tây Nguyên của Việt Nam. Hai huyện này có tổng diện tích khoảng 1.880 km², với địa hình đa dạng gồm núi, thung lũng và vùng đất thấp dọc theo các con sông, dẫn đến sự khác biệt đáng kể về khí hậu và lớp phủ đất.

2.2. Phương pháp nghiên cứu

2.2.1. Dữ liệu sử dụng

Nghiên cứu sử dụng ảnh vệ tinh Landsat 5 TM (năm 2010) và Landsat 8 OLI (năm 2015 và 2023) do USGS cung cấp, với mức mây che phủ dưới 30%. Tổng cộng 55 cảnh ảnh được thu thập trên nền tảng Google Earth Engine (GEE), đảm bảo phủ kín khu vực huyện Đắk Glong và Krông Nô, tỉnh Đắk Nông. Ngoài ra, dữ liệu DEM từ SRTM và các dữ liệu thực địa (điểm GPS) được sử dụng để hỗ trợ phân loại và kiểm định độ chính xác.

Thông tin chi tiết về dữ liệu Landsat được trình bày ở bảng 1.

Bảng 1. Bảng thông tin thu thập ảnh vệ tinh Landsat

TT	Năm	Ảnh vệ tinh	Ngày bắt đầu	Ngày kết thúc	Tỷ lệ mây dưới (%)	Số cảnh ảnh
1	2010	LANDSAT/LT05/C02/T1_L2	01/01/2010	31/12/2010	30	10
2	2015	LANDSAT/LC08/C02/T1_L2	01/01/2015	31/12/2015	50	33
3	2023	LANDSAT/LC08/C02/T1_L2	01/01/2023	31/12/2023	26	12

2.2.2. Tiền xử lý ảnh trên GEE

Tiền xử lý ảnh vệ tinh là bước quan trọng nhằm nâng cao chất lượng dữ liệu và loại bỏ các yếu tố gây nhiễu trước khi tiến hành phân tích. Có nhiều biện pháp tiền xử lý ảnh để nâng cao chất lượng hình ảnh của ảnh vệ tinh như hiệu chỉnh hình học, hiệu chỉnh địa hình, tăng cương chất lượng ảnh, ... (Schowengerdt, 2017; Nguyễn Thị Thanh Hương, 2015). Tuy vậy, tùy theo mục đích nghiên cứu để lựa chọn các bước hiệu chỉnh cụ thể. Trong nghiên cứu này, các kỹ thuật tiền xử lý trên được triển khai trong nền tảng GEE thông qua mã lập trình, đảm bảo dữ liệu đầu vào chất lượng cho các bước phân tích tiếp theo.

+ Hiệu chỉnh hình học: Sử dụng các điểm khống chế mặt đất để điều chỉnh vị trí địa lý, đảm bảo ảnh trùng khớp với hệ tọa độ UTM và hệ quy chiếu WGS-84.

+ Tăng cường chất lượng ảnh:

Kỹ thuật trộn ảnh: Kết hợp các kênh phổ riêng biệt như Red, Green, Blue, NIR (Cận hồng ngoại) và SWIR (Hồng ngoại sóng ngắn) để tạo ra ảnh tổng hợp có thông tin phổ đa dạng, tăng khả năng phân biệt các đối tượng.

Biến đổi cấp độ xám: Áp dụng các hàm điều chỉnh độ sáng và độ tương phản (Contrast Stretching) nhằm làm rõ các chi tiết nhỏ trên ảnh.

Tổ hợp màu: Tích hợp các kênh phổ (Green, Red, NIR) - RGB để tạo ra hình ảnh giả màu, giúp phân biệt các loại lớp phủ dựa trên màu sắc đặc trưng.

+ Xử lý vùng mây che phủ: Sử dụng mặt nạ mây (Cloud Masking) tự động tích hợp sẵn trong GEE, dựa trên thuật toán FMask (Landsat).

+ Khử nhiễu đốm: Sử dụng bộ lọc không gian (Spatial Filters) để làm mịn ảnh, loại bỏ các điểm nhiễu rời rạc mà không làm mất chi tiết cần thiết (Haas, 2024).

+ Tính toán chỉ số phổ (NDVI, NDWI, NDSI): GEE hỗ trợ các phép toán raster trực tiếp trên các băng phổ của ảnh Landsat. NDVI, NDWI và NDSI được tính toán bằng các biểu thức đơn giản, sau đó được gộp thành ảnh tổ hợp đa lớp

(multi-band composite) làm đầu vào cho mô hình RF.

+ Tích hợp dữ liệu DEM: Mô hình số độ cao SRTM (30 m) được truy xuất từ nguồn USGS/SRTMGL1_003 trong GEE, đồng bộ với ảnh Landsat theo vùng nghiên cứu để bổ sung thông tin độ cao và độ dốc.

2.2.3. Thành lập tập mẫu phân loại

Tập mẫu phục vụ phân loại được xây dựng dựa trên sự kết hợp nhiều nguồn dữ liệu nhằm đảm bảo tính đại diện và độ chính xác. Cụ thể, các điểm GPS thu thập từ khảo sát thực địa được sử dụng làm cơ sở xác thực thông tin thực địa hiện tại. Đồng thời, ảnh vệ tinh độ phân giải cao từ Google Earth và kinh nghiệm giải đoán ảnh viễn thám được huy động để xác định đặc điểm lớp phủ trên toàn khu vực nghiên cứu.

Đối với các mốc thời gian trong quá khứ, việc giải đoán được thực hiện theo quy trình suy diễn ngược: (i) xác định trạng thái hiện tại thông qua ảnh mới nhất và điều tra thực địa; (ii) đối chiếu với ảnh Landsat các năm trước và bản đồ kiểm kê rừng gần nhất để phát hiện thay đổi; (iii) dựa trên chuỗi biến động không gian - thời gian để suy luận trạng thái lớp phủ tại thời điểm mục tiêu (Amani *et al*., 2020; Senf *et al*., 2013).

Dựa trên đặc điểm phổ phản xạ và khả năng phân biệt trên ảnh Landsat tổ hợp kênh Green, Red và NIR, nghiên cứu xác định 9 loại lớp phủ chính gồm: (1) rừng lá rộng thường xanh (RTX), (2) rừng bán thường xanh (RBTX), (3) rừng khộp (RK), (4) rừng trồng (RTG), (5) cao su trưởng thành (CAOSU), (6) đất nông nghiệp (gồm cây lâu năm và ngắn ngày) (NN), (7) khu dân cư (DC), (8) mặt nước (MN), và (9) các loại đất khác (trảng cỏ, cây bụi, đất trống, đất hoang, đất chưa thành rừng,...) (DK). Trong năm 2023, tổng cộng 650 điểm mẫu thực địa thuộc 9 lớp phủ chính được thu thập, kết hợp với khoảng 4.000 điểm mẫu giải đoán trên ảnh vệ tinh đảm bảo tính đại diện cho 9 lớp phủ. Dữ liệu thực địa được thu tại các khu vực khó phân biệt trên ảnh nhằm nâng cao độ tin cậy, với tọa độ ghi nhận bằng thiết bị GPS cầm tay, tham chiếu theo hệ tọa độ UTM và hệ quy chiếu WGS-84. Mỗi điểm được lấy tại trung tâm của đối tượng cần khảo sát để hạn chế sai số không gian, đồng thời đảm bảo phân bố đều trên toàn bộ vùng nghiên cứu.

Đối với các năm 2010 và 2015, lần lượt 4.900 và 2.500 điểm mẫu/năm được thiết lập thông qua phương pháp suy diễn từ hiện tại - tức là sử dụng hiện trạng lớp phủ để truy hồi thông tin quá khứ. Cách tiếp cận này dựa trên kinh nghiệm giải đoán ảnh, kết hợp đối chiếu ảnh Landsat đa thời gian và bản đồ kiểm kê rừng qua các giai đoạn tương ứng (Senf *et al*., 2013).

Tập dữ liệu mẫu được chia ngẫu nhiên thành hai phần độc lập: 70% để phân loại, 30% còn lại dùng để kiểm định độ chính xác. Ngoài các điểm mẫu dạng điểm (point), nghiên cứu còn khoanh vùng các đối tượng dưới dạng polygon để tăng dung lượng dữ liệu phân loại, trong khi đánh giá độ chính xác cuối cùng chỉ sử dụng tập dữ liệu dạng điểm.

2.2.4. Phân loại ảnh bằng thuật toán RF

Thuật toán RF được áp dụng để phân loại ảnh tổ hợp gồm các băng phổ RGB, NDVI, NDWI, NDSI và DEM. Mô hình được huấn luyện và thực thi hoàn toàn trên GEE. Hai siêu tham số quan trọng là: ntree: Số cây trong rừng (chọn từ 100-500) và mtry: Số biến ngẫu nhiên tại mỗi node (tối ưu tùy thuộc tập chỉ số đầu vào).

Hai phương pháp phân loại được so sánh:

- Phương pháp 1: Chỉ sử dụng NDVI (và RGB).

- Phương pháp 2: Sử dụng tổ hợp NDVI + DEM + NDWI + NDSI (và RGB).

Các chỉ số được sử dụng

- NDVI = (NIR - RED)/(NIR + RED): phản ánh mức độ xanh của thảm thực vật (Nguyễn Thị

Thanh Hương *et al*., 2018; Đoàn Duy Hiếu, Nguyễn Thám, 2017)

- NDSI = (GREEN - SWIR)/(GREEN + SWIR): dùng để xác định đất trống và đất canh tác (Xu *et al*., 2006).

- NDWI = (GREEN - NIR)/(GREEN + NIR): phát hiện mặt nước (McFeeters, 1996).

- DEM: cung cấp thông tin độ cao địa hình, hỗ trợ phân biệt thảm phủ theo độ dốc và cao độ (Gan *et al*., 2023; Elumnoh, Shrestha, 2000).

2.2.5. Đánh giá độ chính xác kết quả phân loại

Độ chính xác kết quả phân loại được đánh giá thông qua ma trận sai số (confusion matrix), so sánh giữa các pixel phân loại và pixel thực địa. Từ đó, các chỉ số đánh giá bao gồm độ chính xác toàn bộ (OA), độ chính xác người sản xuất (PA), độ chính xác người sử dụng (UA) và hệ số Kappa (K) được tính toán theo công thức của Congalton và Green (1999).

Độ chính xác toàn bộ (OA) thể hiện tỷ lệ tổng số điểm phân loại đúng trên tổng số điểm kiểm tra.

$$OA=\frac{\sum_{i=1}^{r} X_{ii}}{N} \times 100 \quad [1]$$

Độ chính xác người sử dụng (User"s accurary - UA): phản ánh độ tin cậy của một lớp trong bản đồ phân loại.

$$U_a=\frac{X_{ii}}{X_{i+}} \times 100 \quad (2)$$

Độ chính xác người sản xuất (Producer"s accurary - PA): thể hiện khả năng của mô hình phân loại đúng một lớp cụ thể.

$$P_a=\frac{X_{ii}}{X_{+i}} \times 100 \quad (3)$$

Đồng thời, hệ số K được sử dụng để đo lường mức độ phù hợp giữa phân loại và thực tế, có hiệu chỉnh cho sai số do ngẫu nhiên. Giá trị của K nằm trong khoảng từ 0 đến 1, càng gần 1 thì mức độ phân loại càng chính xác (Navulur, 2006). Các mức đánh giá theo K được thể hiện trong bảng 2:

$$\kappa=\frac{N\sum_{i=1}^{r} x_{ii}-\sum_{i=1}^{r}(x_{i+}\times x_{+i})}{N^2-\sum_{i=1}^{r}(x_{i+}\times x_{+i})} \quad (4)$$

Trong đó:

Xii: số điểm được phân loại đúng ở lớp i;

Xi+: tổng số điểm thực tế ở lớp i (tổng theo hàng);

X+i: tổng số điểm được phân loại vào lớp i (tổng theo cột);

N: tổng số điểm kiểm tra;

r: số lớp phân loại.

Bảng 2. Bảng đánh giá độ chính xác phân loại dựa vào hệ số K

TT	Giá trị K	Mức độ phân biệt
1	<0	Thấp
2	0 - 0,20	Nhẹ
3	0,21 - 0,40	Vừa
4	0,41 - 0,60	Tương đối chặt
5	0,61 - 0,80	Chặt
6	0,81 - 1	Rất chặt

(Navulur, 2006)

2.2.6. So sánh lựa chọn phương pháp phân loại phù hợp với độ chính xác cao

Từ kết quả nghiên cứu độ chính xác OA, UA, PA và K, so sánh và lựa chọn phương pháp phân loại phù hợp với độ chính xác cao.

2.2.7. Thành lập bản đồ thay đổi lớp phủ

Sau khi hoàn tất phân loại, kết quả raster được xử lý lọc nhiễu bằng công cụ "Majority Filter" trong trình đơn Raster > Spatial Analyst Tools > Generalization trên phần mềm ArcGIS 10.4.1,

nhằm loại bỏ các điểm phân loại rời rạc và tăng tính liên kết không gian cho lớp phủ.

Tiếp theo, ảnh phân loại được chuyển sang định dạng vector bằng công cụ Raster to Polygon, giúp thuận tiện cho việc biên tập bản đồ lớp phủ và tính toán diện tích từng lớp. Diện tích các đối tượng lớp phủ được tính thông qua chức năng Calculate Geometry, với đơn vị chuẩn hóa là hecta.

III. KẾT QUẢ NGHIÊN CỨU

3.1. So sánh và lựa chọn phương pháp phân loại phù hợp với độ chính xác cao

Nghiên cứu tiến hành phân loại ảnh vệ tinh đa thời gian với phương pháp phân loại RF sử dụng chỉ số NDVI và sử dụng kết hợp nhiều chỉ số (NDVI, DEM, NDSI, và NDWI) năm 2010, 2015 và 2023. Từ đó so sánh và rút ra phương pháp phân loại tối ưu hóa độ chính xác. Kết quả so sánh phương pháp phân loại thảm phủ sử dụng đất bằng thuật toán RF sử dụng NDVI với sử dụng tích hợp NDVI, DEM, NDSI, NDWI qua các năm 2010, 2015 và 2023 tại huyện Đắk Glong và huyện Krông Nô, tỉnh Đắk Nông thể hiện trong bảng 3.

Bảng 3. So sánh phương pháp phân loại thảm phủ sử dụng đất bằng thuật toán RF sử dụng NDVI với sử dụng tích hợp NDVI, DEM, NDSI, NDWI qua các năm 2010, 2015 và 2023 tại huyện Đắk Glong và huyện Krông Nô, tỉnh Đắk Nông

Đơn vị: %

Trạng thái	NDVI						Tích hợp NDVI, DEM, NDSI, và NDWI					
	2010		2015		2023		2010		2015		2023	
	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA
RTX	47,95	43,75	78,59	81,57	45,54	40,00	86,76	80,17	94,26	93,28	87,95	81,74
RBTX	58,59	47,54	56,00	91,26	8,00	6,67	88,89	92,63	94,00	91,26	56,00	48,28
RK	0,00	0,00	12,50	9,09	13,64	19,35	66,67	70,00	62,50	76,92	65,91	78,38
RTG	40,25	39,51	4,17	2,94	22,92	23,74	88,05	89,74	54,17	59,09	88,19	90,71
CAOSU	47,01	49,22	12,90	14,81	32,10	36,36	87,31	84,78	51,61	47,06	80,86	83,44
NN	40,23	45,26	16,67	15,91	22,61	22,61	87,36	83,52	59,52	62,50	81,74	81,74
DKH	26,19	33,33	19,61	24,39	43,58	39,39	71,43	76,92	52,94	58,70	84,92	77,55
DC	59,09	53,53	46,67	26,92	50,00	60,76	83,12	87,07	46,67	41,18	64,58	83,78
MN	73,21	77,27	100,0	100,0	92,86	100,0	89,95	95,92	100,0	100,0	100,00	100,0
OA%	49,33		59,35		35,89		86,04		83,83		81,75	
K	0,41		0,38		0,24		0,84		0,75		0,78	

Trong đó: (TX) rừng gỗ tự nhiên lá rộng thường xanh; (RBTX) rừng gỗ tự nhiên lá rộng bán thường xanh; (RK) rừng Khộp; (RTG) rừng trồng; (CAOSU) cao su; (NN) đất nông nghiệp; (DC) dân cư, (MN) mặt nước, (DKH) đất khác.

Kết quả phân loại lớp phủ sử dụng đất cho khu vực huyện Đắk Glong và Krông Nô trong các năm 2010, 2015 và 2023 cho thấy sự khác biệt rõ rệt về độ chính xác giữa hai phương pháp phân loại: sử dụng riêng NDVI và tích hợp nhiều chỉ số (NDVI, DEM, NDWI, NDSI).

Cụ thể, phương pháp sử dụng riêng NDVI đạt độ chính xác OA thấp, dao động từ 35,89% đến 59,35% và hệ số K chỉ đạt từ 0,24 đến 0,41, tương ứng với mức phân biệt "nhẹ đến trung bình" theo thang đánh giá của Navulur (2006). Độ chính xác thấp này chủ yếu do hạn chế của NDVI trong việc phân biệt các lớp phủ có phổ phản xạ tương tự, đặc biệt là khu vực đồi núi hoặc có sự pha trộn giữa đất trống, nước và thực vật.

Trong khi đó, phương pháp tích hợp nhiều chỉ số NDVI, DEM, NDWI và NDSI đã cải thiện đáng kể kết quả phân loại. Độ chính xác OA đạt

từ 81,75% đến 86,04%, hệ số K đạt từ 0,75 đến 0,78, đạt mức "chặt" và tiệm cận "rất chặt" theo tiêu chuẩn đánh giá của Navulur (2006). Chênh lệch OA giữa hai phương pháp lên đến 45,86%, trong khi hệ số K tăng trung bình khoảng 0,40 đơn vị, cho thấy ảnh hưởng rõ rệt của tổ hợp chỉ số đầu vào đến hiệu suất mô hình. Việc bổ sung DEM giúp nâng cao khả năng phân biệt các lớp phủ theo đặc điểm địa hình; NDWI và NDSI tăng cường khả năng phát hiện vùng nước và đất trống, khắc phục nhược điểm khi chỉ sử dụng NDVI.

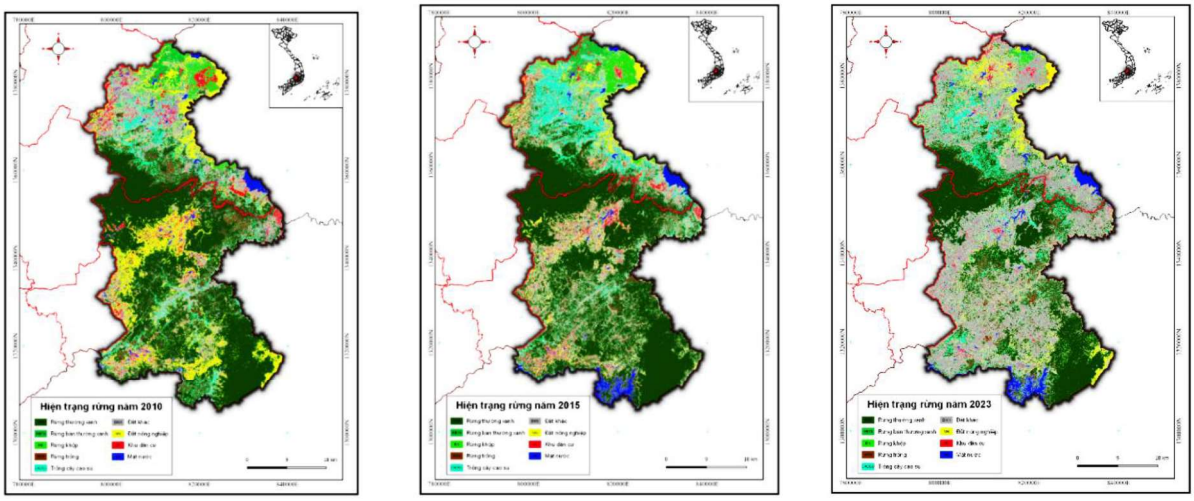
Kết quả nghiên cứu phù hợp với xu hướng và kết luận của nhiều nghiên cứu trước đó. Eduardo và đồng tác giả (2024) và Ur Rehman (2021) cũng chỉ ra rằng việc tích hợp chỉ số NDVI với các biến địa hình và độ ẩm như DEM, NDWI và NDSI đã giúp tăng độ chính xác phân loại lớp phủ sử dụng đất từ dưới 65% lên trên 80% trong các khu vực có địa hình phức tạp như miền núi Trung Quốc và Pakistan. Đồng thời, nghiên cứu của Ibrahim (2023) khi kết hợp dữ liệu Sentinel với DEM cho thấy chỉ số NDVI đơn lẻ dễ bị nhầm lẫn ở các vùng đất canh tác theo mùa hoặc vùng đất trống xen kẽ cây bụi.

Khác biệt của nghiên cứu này là đã kiểm chứng rõ ràng trên địa hình đồi núi đặc trưng của Tây Nguyên, với kết quả phân loại cho ba mốc thời gian cách biệt (năm 2010, năm 2015 và năm 2023), giúp đánh giá hiệu quả mô hình một cách ổn định theo chuỗi thời gian.

3.2. Thành lập bản đồ lớp phủ khu vực nghiên cứu

Dựa trên cơ sở kết quả phân loại thảm phủ bằng thuật toán RF với phương pháp tích hợp nhiều chỉ số (NDVI, DEM, NDWI, NDSI), các bản đồ thảm phủ sử dụng đất cho ba mốc thời gian năm 2010, 2015 và 2023 đã được xây dựng thành công cho khu vực huyện Đắk Glong và Krông Nô, tỉnh Đắk Nông. Đây là kết quả quan trọng nhằm phục vụ cho công tác giám sát tài nguyên đất, rừng và định hướng quản lý, quy hoạch sử dụng đất bền vững tại địa phương.

Kết quả thành lập bản đồ chuyên đề thảm phủ sử dụng đất các năm 2010, 2015 và 2023 sau khi được sàng lọc nhiễu ảnh và vector hóa được thể hiện ở hình 2. Các bản đồ cho thấy sự phân bố và xu thế biến động rõ rệt của từng loại lớp phủ theo không gian và thời gian. Diện tích và tỷ lệ giữa các nhóm lớp phủ qua các năm được tổng hợp ở bảng 4.

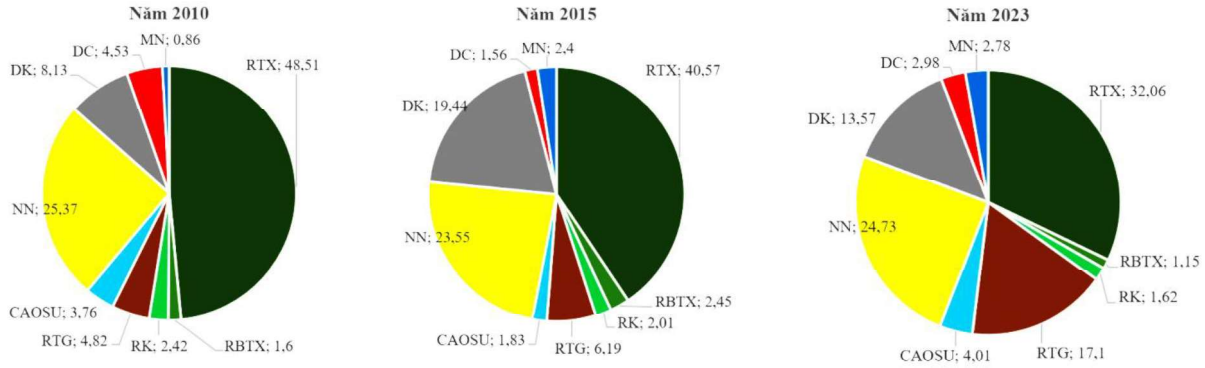


Bản đồ phân loại năm 2010 Bản đồ phân loại năm 2015 Bản đồ phân loại năm 2023

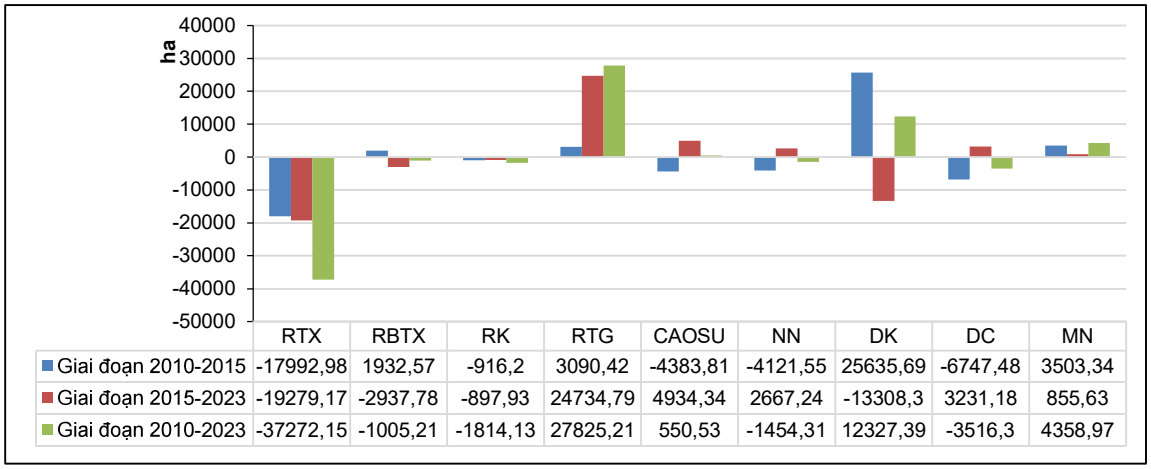
Hình 2. Bản đồ phân loại thảm phủ sử dụng đất huyện Đắk Glong và huyện Krông Nô, tỉnh Đắk Nông qua các năm 2010, 2015 và 2023

Bảng 4. Kết quả thống kê diện tích của các loại thảm phủ sử dụng đất huyện Đắk Glong và huyện Krông Nô, tỉnh Đắk Nông qua các năm 2010, 2015 và 2023

Lớp phủ sử dụng đất	Năm 2010		Năm 2015		Năm 2023	
	Diện tích (ha)	Tỷ lệ (%)	Diện tích (ha)	Tỷ lệ (%)	Diện tích (ha)	Tỷ lệ (%)
RTX	109.906,11	48,51	91.913,13	40,57	72.633,96	32,06
RBTX	3.617,01	1,6	5.549,58	2,45	2.611,80	1,15
RK	5.473,53	2,42	4.557,33	2,01	3.659,40	1,62
RTG	10.926,45	4,82	14.016,87	6,19	38.751,66	17,10
CAOSU	8.524,80	3,76	4.140,99	1,83	9.075,33	4,01
NN	57.485,52	25,37	53.363,97	23,55	56.031,21	24,73
DK	18.412,65	8,13	44.048,34	19,44	30.740,04	13,57
DC	10.270,71	4,53	3.523,23	1,56	6.754,41	2,98
MN	1.940,58	0,86	5.443,92	2,40	6.299,55	2,78



Hình 3. Tỷ lệ phần trăm của các loại thảm phủ sử dụng đất huyện Đắk Glong và huyện Krông Nô, tỉnh Đắk Nông qua các năm 2010, 2015 và 2023 (%)



Hình 4. Thay đổi thảm phủ sử dụng đất huyện Đắk Glong và huyện Krông Nô, tỉnh Đắk Nông qua các năm 2010, 2015 và 2023

Kết quả phân tích biến động lớp phủ đất qua ba giai đoạn cho thấy xu hướng giảm diện tích rừng tự nhiên (RTX và RBTX) diễn ra liên tục và đáng kể. Cụ thể, diện tích RTX từ 109.906,11 ha năm 2010 đã giảm xuống còn 72.633,96 ha năm 2023, tương ứng giảm 33,91%. RBTX cũng có xu hướng giảm mạnh, từ 3.617,01 ha (năm 2010) còn 2.611,80 ha (năm 2023). Đây là tín hiệu cảnh báo về áp lực suy thoái tài nguyên rừng, đòi hỏi các biện pháp bảo vệ và phục hồi rừng hiệu quả hơn trong giai đoạn tới.

Ngược lại, diện tích RTG có xu hướng tăng nhanh và rõ rệt, đặc biệt từ năm 2015 đến 2023, tăng từ 14.016,87 ha lên 38.751,66 ha, tương ứng tăng 176,52%. Sự gia tăng này phản ánh nỗ lực trồng rừng và phục hồi đất trống của địa phương, đồng thời cũng thể hiện quá trình chuyển đổi cơ cấu sử dụng đất theo hướng trồng rừng sản xuất để thay thế rừng tự nhiên bị suy giảm.

Đáng chú ý, nhóm đất NN có sự dao động nhẹ về diện tích qua các năm, dao động quanh mức 24 - 25% tổng diện tích tự nhiên toàn vùng nghiên cứu. Tuy nhiên, nhóm DK bao gồm đất trống, đất chưa sử dụng, trảng cỏ và cây bụi lại có xu hướng tăng mạnh giai đoạn 2010 - 2015 (tăng hơn 25.000 ha), sau đó giảm nhẹ vào năm 2023. Điều này phản ánh tác động của quá trình chuyển đổi đất đai và biến động kinh tế - xã hội đến sử dụng đất tại địa phương.

Diện tích khu DC có xu hướng giảm nhẹ giai đoạn 2010 - 2015 nhưng sau đó tăng trở lại vào năm 2023, điều này phản ánh sự phát triển và mở rộng khu vực dân cư trong điều kiện gia tăng dân số và nhu cầu nhà ở. Các vùng MN cũng có sự gia tăng nhẹ diện tích qua các năm, phù hợp với điều kiện phát triển thủy lợi và các hồ chứa nước phục vụ sản xuất nông nghiệp.

Nhìn chung, kết quả thành lập bản đồ lớp phủ đất qua ba giai đoạn tại huyện Đắk Glong và Krông Nô đã phản ánh rõ nét xu thế biến động sử dụng đất theo hai hướng: (i) suy giảm diện tích rừng tự nhiên và gia tăng diện tích đất sản xuất, đất trống, đất rừng trồng; (ii) tái cơ cấu lại sử dụng đất phù hợp với điều kiện tự nhiên và phát triển kinh tế - xã hội địa phương. Kết quả này hoàn toàn phù hợp với thực trạng sử dụng đất của tỉnh Đắk Nông và là cơ sở khoa học quan trọng phục vụ cho công tác quy hoạch, quản lý tài nguyên, cũng như định hướng phát triển nông lâm nghiệp bền vững trong bối cảnh chịu nhiều tác động từ hoạt động con người và biến đổi khí hậu.

IV. KẾT LUẬN

Kết quả nghiên cứu đã cho thấy sự khác biệt rõ rệt về hiệu quả phân loại giữa hai phương pháp áp dụng thuật toán RF. Khi chỉ sử dụng chỉ số NDVI đơn lẻ, độ chính xác OA dao động trong khoảng từ 35,89 - 59,35%, với hệ số K chỉ đạt 0,24 - 0,41, phản ánh mức độ phân loại còn thấp và chưa đủ độ tin cậy trong khu vực có đặc điểm địa hình và thảm phủ phức tạp.

Ngược lại, phương pháp tích hợp tổ hợp chỉ số NDVI, DEM, NDWI và NDSI đã giúp nâng cao đáng kể hiệu quả phân loại, với OA tăng lên mức 81,75 - 86,04% và hệ số K đạt từ 0,75 - 0,78. Mức chênh lệch về OA giữa hai phương pháp lên tới 45,86%, trong khi hệ số K tăng trung bình khoảng 0,40 đơn vị, cho thấy hiệu quả vượt trội của phương pháp tích hợp trong việc nhận diện chính xác các lớp phủ khác nhau. Việc kết hợp thông tin DEM, NDWI, và NDSI với NDVI đã giúp khắc phục rõ rệt tình trạng trộn lẫn phổ phản xạ, đồng thời cải thiện khả năng phân biệt tại các khu vực rừng, đất trống, nước mặt và nông nghiệp.

Từ các kết quả trên, nghiên cứu khuyến nghị ứng dụng phương pháp tích hợp các chỉ số NDVI, DEM, NDWI, NDSI trong các bài toán phân loại lớp phủ sử dụng đất sử dụng ảnh viễn thám, đặc biệt tại những khu vực có địa hình phức tạp như vùng Tây Nguyên. Việc lựa chọn

tổ hợp chỉ số phù hợp sẽ không chỉ giúp nâng cao độ chính xác phân loại mà còn đảm bảo độ tin cậy của sản phẩm bản đồ phục vụ cho quy hoạch, giám sát và phát triển sử dụng đất một cách bền vững.

Kết quả phân tích cũng cho thấy diện tích RTX giảm mạnh từ 109.906,11 ha năm 2010 xuống còn 72.633,96 ha năm 2023, tương ứng giảm 33,91%; RBTX cũng giảm từ 3.617,01 ha xuống 2.611,80 ha. Ngược lại, RTG tăng nhanh từ 14.016,87 ha lên 38.751,66 ha, tăng 176,52%. DK tăng hơn 25.000 ha trong giai đoạn 2010 - 2015, sau đó giảm nhẹ. Diện tích NN dao động ổn định quanh mức 24 - 25% tổng diện tích tự nhiên, trong khi khu dân cư và mặt nước có xu hướng tăng nhẹ. Những biến động này phản ánh rõ xu hướng suy giảm rừng tự nhiên, gia tăng đất sản xuất và rừng trồng, phù hợp với quá trình phát triển kinh tế - xã hội tại địa phương, là cơ sở quan trọng phục vụ công tác quy hoạch, quản lý tài nguyên và phát triển bền vững.

TÀI LIỆU THAM KHẢO

1. Amani M., Ghorbanian A., Ahmadi S.A., Kakooei M., Moghimi A. & Others, 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. https://doi.org/10.1109/JSTARS.2020.3021052.
2. Amoakoh A.O., Aplin P., Rodríguez-Veiga P., Moses C., Alonso C.P., Cortés J.A., Delgado-Fernandez I., Kankam S., Mensah J.C., Nortey D.D.N., 2024. Predictive Modelling of Land Cover Changes in the Greater Amanzule Peatlands Using Multi-Source Remote Sensing and Machine Learning Techniques. Remote Sens. 2024, *16*, 4013. https://doi.org/10.3390/rs16214013.
3. Congalton R.G. & Green K., 1999. Assessing the accuracy of remotely sensed data: Principles and practices. Lewis Publishers. Boca Raton FL.
4. Dongping Ming, Tianning Zhou, Min Wang, Tian Tan, 2016. Land cover classification using random forest with genetic algorithm-based parameter optimization. Journal of Applied Remote Sensing, 10(3). doi:10.1117/1.JRS.10.035021.
5. Eduardo S. N., Fernanda Y. W., Maria L. B. T., Erivaldo A., 2024. Assessing Land Use and Cover Changes arising from the 2022 water crisis in Southeast China: A comparative analysis of Remote Sensing Imagery classifications and Machine Learning algorithms. ISPRS TC III Mid-term Symposium “Beyond the canopy: technologies and applications of remote sensing”, 4-8 November 2024, Belém, Brazil.
6. Elumnoh, A., & Shrestha, R. P., 2000. Application of DEM data to Landsat image classification: Evaluation in a tropical wet-dry landscape of Thailand. *Photogrammetric Engineering & Remote Sensing*, 66(3), 297-304. Retrieved from https://www.asprs.org/wp-content/uploads/pers/2000journal/march/2000_mar_297-304.pdf.
7. Gan Y., Wang Q., Lio A., 2023. Tree Crown Detection and Delineation in a Temperate Deciduous Forest from UAV RGB Imagery Using Deep Learning Approaches: Effects of Spatial Resolution and Species Characteristics. Remote Sens. 2023, 15(3), 778; https://doi.org/10.3390/rs15030778
8. Haas J., 2024. Image processing of Sentinel-1 SAR (S-1) and Sentinel-2 multispectral (MSI) (S-2) data for applications in landscape archaeology. Zenodo. doi.org/10.5281/zenodo.10687538
9. Đoàn Duy Hiếu và Nguyễn Thám, 2017. Đánh giá biến động rừng huyện Ia Pa, tỉnh Gia Lai trên tư liệu viễn thám đa thời gian và GIS, Tạp chí Khoa học và Giáo dục, Trường Đại học Sư phạm Huế, 2(42)/2017, tr. 116 - 126.
10. Nguyễn Thị Thanh Hương, 2015. Ứng dụng công nghệ viễn thám trong nghiên cứu và quản lý lâm nghiệp. Tạp chí Khoa học và Công nghệ Việt Nam.
11. Nguyễn Thị Thanh Hương, Nguyễn Đức Định, Phan Thị Hằng, Cao Thị Hoài, Nguyễn Thế Hiển, Đặng Ngọc Trí, Ngô Thế Sơn, Đoàn Minh Trung, Châu Thị Như Quỳnh, Hồ Đình Bảo và Ngô Thị Thùy Phương, 2018. Báo cáo thay đổi sử dụng đất tại tỉnh tỉnh Đắk Nông, giai đoạn 2005 - 2017, Tài liệu Hội thảo tập huấn tại tỉnh Đắk Nông. Đại học Tây Nguyên và Sở NN & PTNT Đắk Nông.

12. Ibrahim S., 2023. Improving Land Use/Cover Classification Accuracy from Random Forest Feature Importance Selection Based on Synergistic Use of Sentinel Data and Digital Elevation Model in Agriculturally Dominated Landscape. Agriculture, *13*(1), 98. https://doi.org/10.3390/agriculture13010098.
13. Lowe B. and Kulkarni A., 2015. *Multispectral Image Analysis Using Random Forest*". International Journal on Soft Computing (IJSC), 6(1), February 2015.
14. McFEETERS S. K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425-1432. https://doi.org/10.1080/01431169608948714.
15. Navulur K., 2006, Multispectral Image Analysis Using the Object-Oriented Paradigm, *CRC Press.*Senf C., Pflugmacher D., van der Linden S., & Hostert P., 2013. Mapping Rubber Plantations and Natural Forests in Xishuangbanna (Southwest China) Using Multi-Spectral Phenological Metrics from MODIS Time Series. Remote Sensing, 5(6), 2795-2812. https://doi.org/10.3390/rs5062795.
16. Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016. *Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas,* Remote Sensing of Environment, 187(2016): p.156-168.
17. Schowengerdt, R.A., 2017. Remote Sensing: Models and Methods for Image Processing. Academic Press.
18. Senf, C., Pflugmacher, D., van der Linden, S., & Hostert, P. (2013). Mapping Rubber Plantations and Natural Forests in Xishuangbanna (Southwest China) Using Multi-Spectral Phenological Metrics from MODIS Time Series. *Remote Sensing*, 5(6), 2795-2812. https://doi.org/10.3390/rs5062795.
19. Ur Rehman A., Ullah S., Shafique M., 2021. Combining Landsat-8 spectral bands with ancillary variables for land cover classification in mountainous terrains of northern Pakistan. J. Mt. Sci. 18, 2388-2401 (2021). https://doi.org/10.1007/s11629-020-6548-7.
20. Xu H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing, *27*(14), 3025-3033. https://doi.org/10.1080/01431160600589179.
21. Yibai Xiong, Yan Ma, Lianguo Ruan, Dan Li, Cheng Lu, Luqi Huang, 2022. The National traditinal chinese medicine medical team (*Comparing different machine learning techniques for predicting covid 19 severity*, Infectious diseases of poverty, munber 19 (2022).

Email tác giả liên hệ: cthoai@ttn.edu.vn
Ngày nhận bài: 11/04/2025
Ngày phản biện đánh giá và sửa chữa: 14/04/2025; 20/04/2025
Ngày duyệt đăng: 21/04/2025