

Ứng dụng Viễn thám, GIS và Học máy để Xây dựng Bản đồ Dự đoán Nguy cơ Cháy rừng tại tỉnh Gia Lai, Việt Nam

Ninh Hải Đăng - MSSV: 21021411¹ and Lê Đức Lương - MSSV: 21021424¹

¹Viện Công nghệ Hàng không Vũ trụ, Trường Đại học Công nghệ - Đại học Quốc
gia Hà Nội, Hà Nội, Việt Nam

*Địa chỉ liên hệ: nindhaidangg@gmail.com

Tóm tắt nội dung

Cháy rừng là một hiểm họa nghiêm trọng gây thiệt hại lớn về kinh tế, xã hội và môi trường. Nghiên cứu này trình bày kết quả ứng dụng phương pháp tích hợp dữ liệu viễn thám đa nguồn và các mô hình học máy (RF và GTB) trên nền tảng GEE để xây dựng bản đồ phân cấp nguy cơ cháy rừng cho tỉnh Gia Lai. Dữ liệu cho nghiên cứu được thu thập và phân tích trong giai đoạn mùa khô từ tháng 12 năm 2024 đến tháng 4 năm 2025. Các kết quả của nghiên cứu cung cấp công cụ hữu ích cho các cơ quan chức năng trong công tác phòng chống cháy rừng.

1 Giới thiệu

Cháy rừng là một trong những thiên tai tàn khốc, gây ra những hậu quả nghiêm trọng đối với hệ sinh thái, kinh tế và đời sống con người. Việt Nam, với diện tích rừng đáng kể và điều kiện khí hậu nhiệt đới gió mùa, thường xuyên phải đối mặt với nguy cơ cháy rừng, đặc biệt trong mùa khô. Tỉnh Gia Lai, nằm ở khu vực Tây Nguyên, với đặc điểm thảm thực vật đa dạng và các yếu tố khí hậu khắc nghiệt, là một trong những điểm nóng về cháy rừng. Việc xây dựng các mô hình dự đoán sớm và chính xác nguy cơ cháy rừng đóng vai trò then chốt trong việc chủ động triển khai các biện pháp phòng ngừa, giảm thiệt hại.

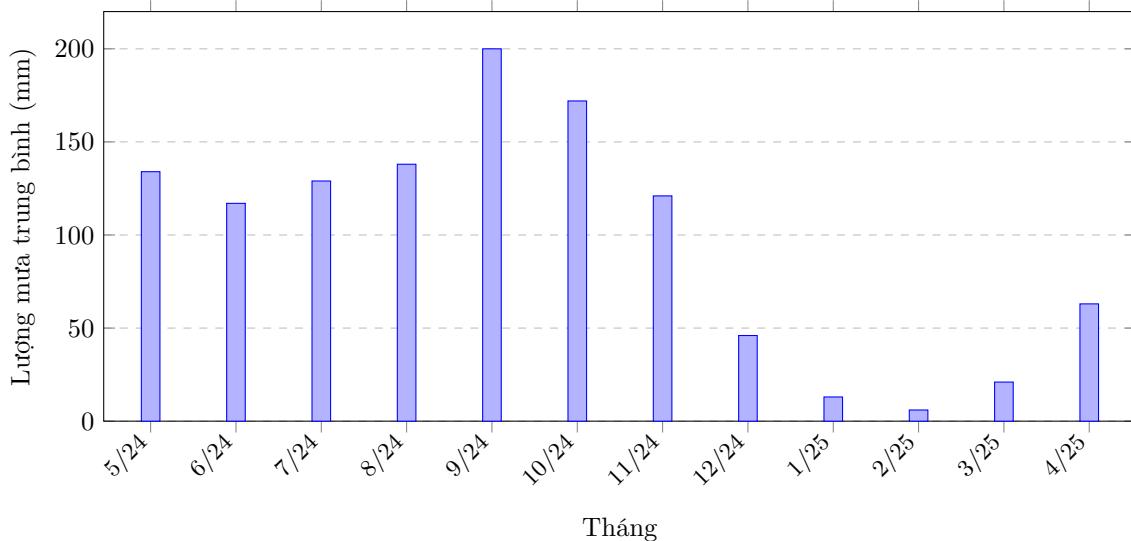
Trong những năm gần đây, sự phát triển của công nghệ viễn thám và hệ thống thông tin địa lý (GIS) đã mở ra những hướng tiếp cận mới và hiệu quả cho việc giám sát và mô hình hóa nguy cơ cháy rừng. Dữ liệu từ các vệ tinh như Sentinel-2, MODIS cung cấp thông tin phong phú về trạng thái thảm thực vật, độ ẩm, nhiệt độ bề mặt đất trên phạm vi rộng và có tính cập nhật cao. Đặc biệt, dữ liệu điểm cháy thực tế từ MODIS là nguồn thông tin quan trọng về sự xuất hiện của các vụ cháy, được sử dụng để tạo nhãn cháy cho việc huấn luyện và đánh giá mô hình. Bên cạnh đó, các thuật toán học máy (Machine Learning - ML) như Random Forest (RF) và Gradient Tree Boosting (GTB) đã chứng minh được hiệu quả vượt trội trong việc xử lý các bộ dữ liệu lớn, phức tạp và xây dựng các mô hình dự đoán với độ chính xác cao trong nhiều lĩnh vực, bao gồm cả khoa học môi trường và quản lý thiên tai [1–3]. Nền tảng điện toán đám mây Google Earth Engine (GEE) với khả

năng truy cập và xử lý lượng lớn dữ liệu không gian địa lý đã tạo điều kiện thuận lợi cho việc triển khai các nghiên cứu quy mô lớn [4].

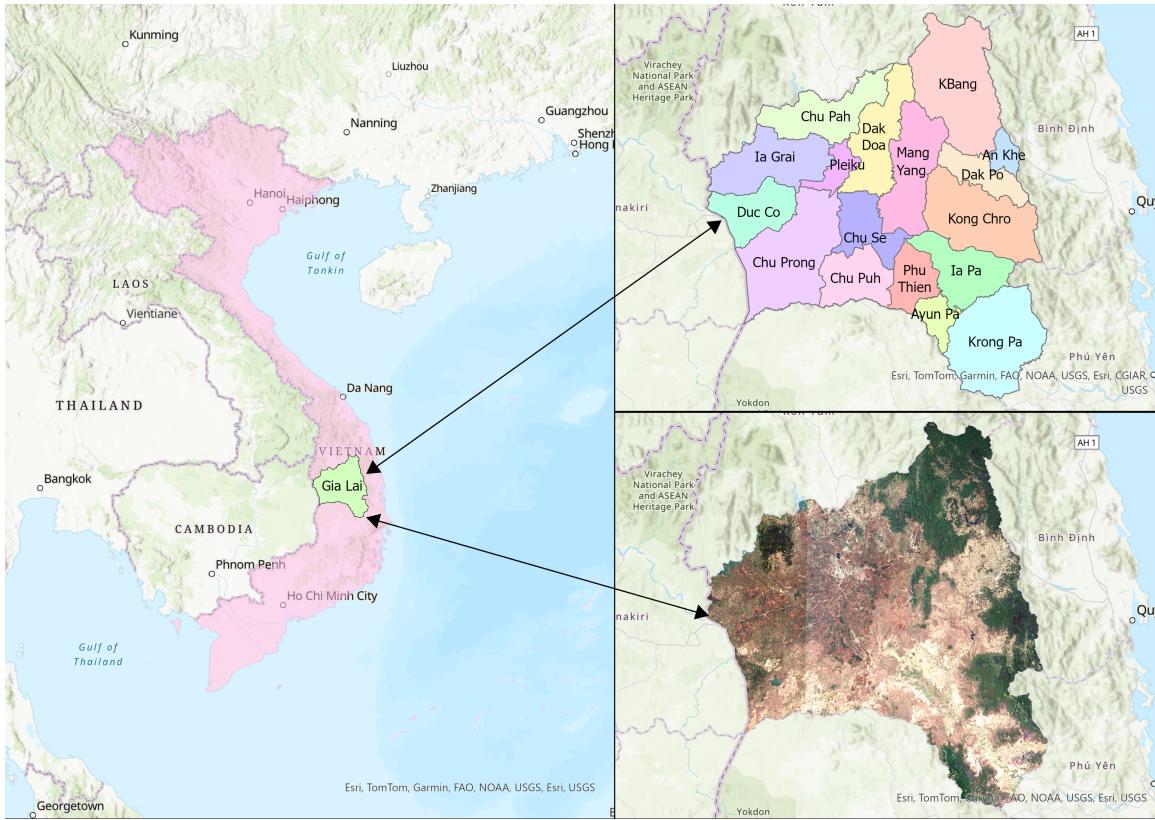
Nghiên cứu này nhằm mục tiêu ứng dụng phương pháp tích hợp dữ liệu viễn thám đa nguồn và các mô hình học máy (RF và GBT) trên nền tảng GEE để xây dựng bản đồ phân cấp nguy cơ cháy rừng cho tỉnh Gia Lai trong giai đoạn mùa khô dự kiến từ tháng 12 năm 2024 đến tháng 4 năm 2025. Các kết quả của nghiên cứu được kỳ vọng sẽ cung cấp công cụ hữu ích cho các cơ quan chức năng trong công tác phòng chống cháy rừng.

2 Khu vực nghiên cứu

Nghiên cứu được thực hiện trên toàn bộ địa bàn tỉnh Gia Lai, Việt Nam. Gia Lai có tọa độ địa lý từ $12^{\circ}58'40''$ đến $14^{\circ}37'00''$ vĩ độ Bắc và từ $107^{\circ}28'04''$ đến $108^{\circ}54'40''$ kinh độ Đông. Tỉnh có đặc điểm địa hình đa dạng, bao gồm chủ yếu là cao nguyên và núi thấp, nghiêng dần từ Bắc xuống Nam và từ Đông sang Tây. Phía Bắc là cao nguyên Kon Tum, phía Nam là cao nguyên Pleiku tương đối bằng phẳng, trong khi phía Đông và Đông Nam là vùng núi cao thuộc dãy Trường Sơn Nam, xen kẽ là các thung lũng. Thảm thực vật chủ yếu là rừng tự nhiên lá rộng thường xanh, rừng nửa rụng lá và rừng khộp đặc trưng của Tây Nguyên. Bên cạnh đó là các diện tích rừng trồi như thông, keo và các loại cây công nghiệp dài ngày. Khí hậu Gia Lai mang đặc trưng của khí hậu nhiệt đới gió mùa cao nguyên, chia thành hai mùa rõ rệt: mùa mưa và mùa khô. Mùa khô thường kéo dài từ tháng 11 đến tháng 4 năm sau, đây là thời kỳ có nguy cơ cháy rừng cao nhất.



Hình 1: Biểu đồ lượng mưa trung bình hàng tháng tại Gia Lai (giai đoạn 5/2024 - 4/2025).



Hình 2: Bản đồ vị trí khu vực nghiên cứu tỉnh Gia Lai.

3 Dữ liệu và Phương pháp

Phần này mô tả các bộ dữ liệu được sử dụng và quy trình phương pháp luận được áp dụng trong nghiên cứu.

3.1 Thu thập và Tiền xử lý Dữ liệu

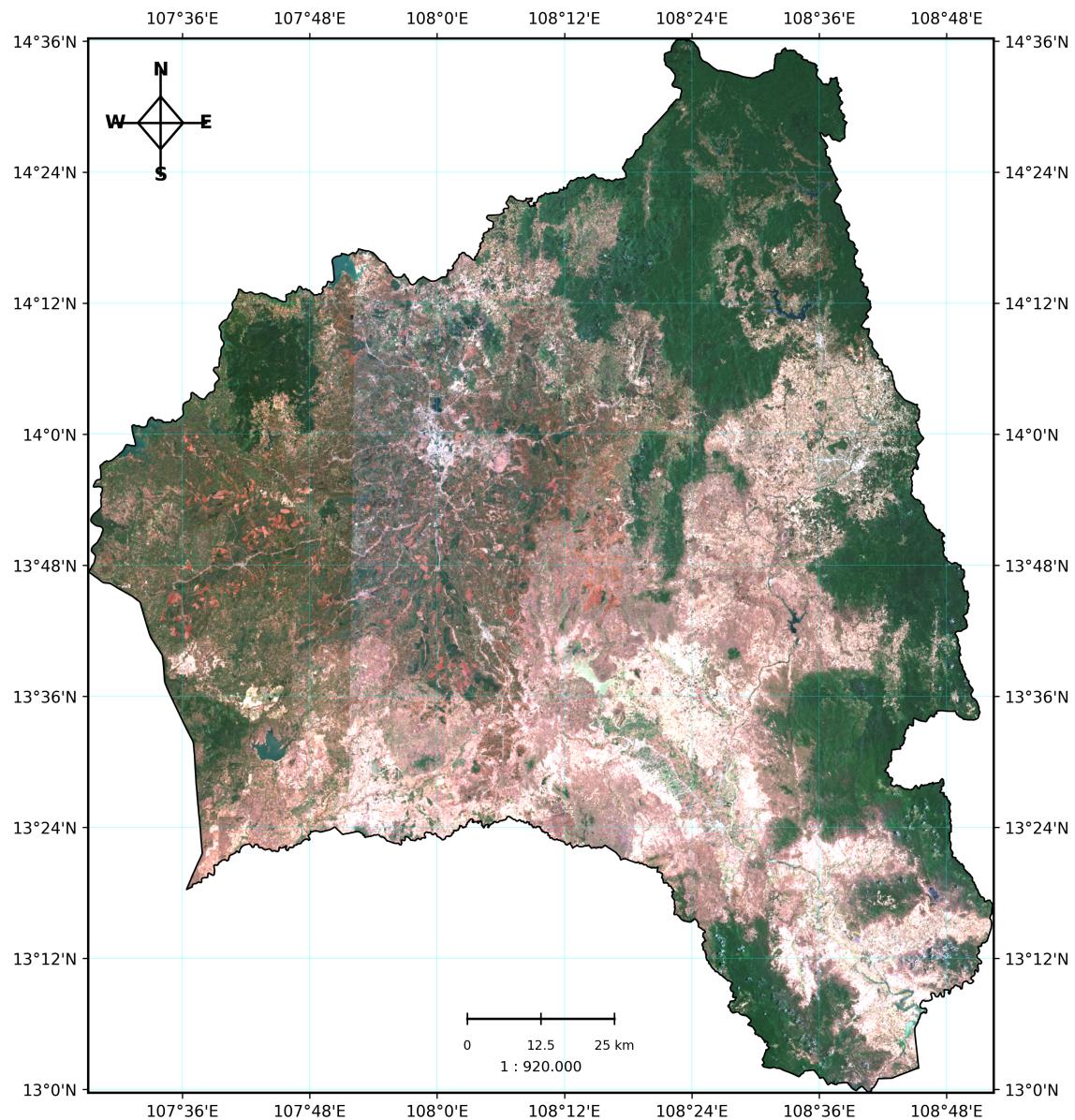
Các bộ dữ liệu viễn thám và phụ trợ sau đây đã được sử dụng trong nghiên cứu (Bảng 1). Các lớp dữ liệu raster đầu vào có độ phân giải không gian khác nhau; ví dụ, dữ liệu Sentinel-2 có các kênh ở 10m, 20m và 60m, trong khi dữ liệu Nhiệt độ Bề mặt Đất (LST) từ MODIS và chỉ số TCI được xử lý ở độ phân giải 1km. Dữ liệu Mô hình số độ cao (DEM) SRTM có độ phân giải 30m. Dữ liệu điểm cháy MODIS (FireMask) được sử dụng để tạo nhãn cháy cho mô hình được xử lý ở độ phân giải 500m.

- Ảnh vệ tinh Sentinel-2 MSI (COPERNICUS/S2_SR) [5] được sử dụng để tính toán các chỉ số thực vật và độ ẩm. Các ảnh trong khoảng thời gian từ `startDate` ("2024-12-01") đến `endDate` ("2025-04-30") có độ che phủ mây dưới 10% được lựa chọn. Mây được che phủ dựa trên kênh xác suất mây ($MSK_CLDPRB < 30$). Ảnh tổng hợp trung vị (median composite) được tạo ra và cắt theo ranh giới tỉnh Gia Lai.

- Dữ liệu Nhiệt độ Bề mặt Đất (LST) ban ngày từ MODIS (MODIS/061/MOD11A2) [6] được sử dụng để phản ánh điều kiện nhiệt.
- Dữ liệu Mô hình số độ cao (DEM) từ SRTM (USGS/SRTMGL1_003) [7] được sử dụng để tính toán các yếu tố địa hình.
- Dữ liệu Lượng mưa hàng ngày từ CHIRPS (UCSB-CHG/CHIRPS/DAILY) [8].
- Dữ liệu Gió hàng ngày từ ERA5 Land (ECMWF/ERA5_LAND/DAILY_AGGR) [9].
- Dữ liệu điểm cháy lịch sử từ MODIS (MODIS/061/MOD14A1) [10] được sử dụng để tạo nhãn cháy cho việc huấn luyện mô hình. Các pixel có giá trị **FireMask** > 6 được coi là có cháy.

Bảng 1: Tổng hợp các bộ dữ liệu đã sử dụng.

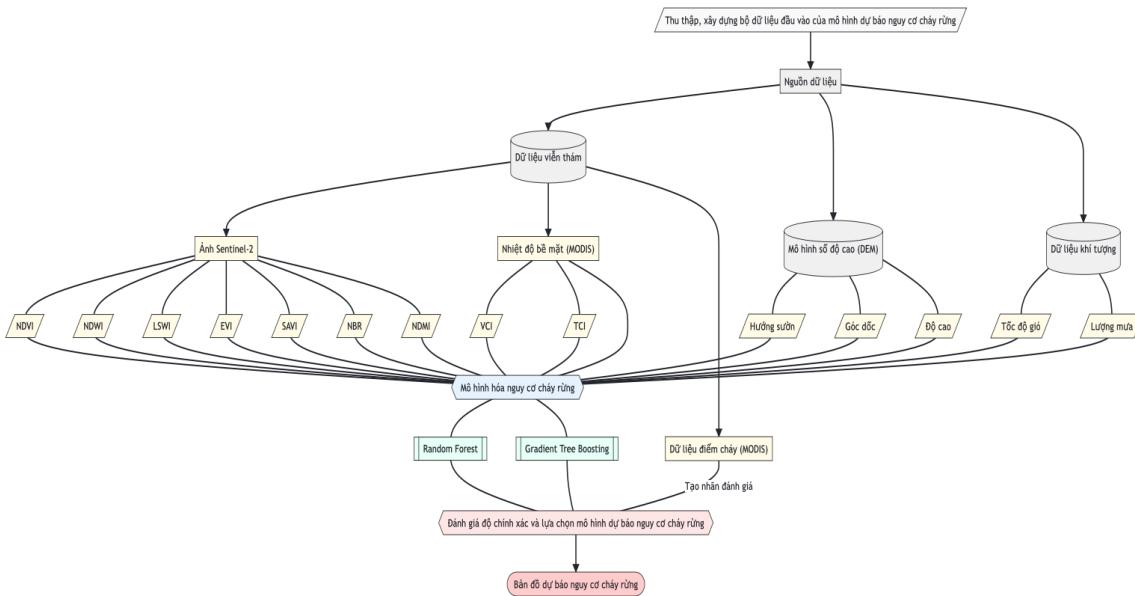
Tên dữ liệu	Nguồn gốc (Mã sản phẩm GEE)	Độ phân giải
Ảnh Sentinel-2 MSI	COPERNICUS/S2_SR	10m, 20m, 60m
Nhiệt độ bề mặt đất MODIS	MODIS/061/MOD11A2	1km
Mô hình số độ cao SRTM	USGS/SRTMGL1_003	30m
Lượng mưa CHIRPS	UCSB-CHG/CHIRPS	~5.5km
Gió ERA5 Land	ECMWF/ERA5_LAND	~9km
Điểm cháy MODIS	MODIS/061/MOD14A1	500m



Hình 3: Ảnh màu thực Sentinel-2 tỉnh Gia Lai.

3.2 Các Biến Dự báo (Predictor Variables)

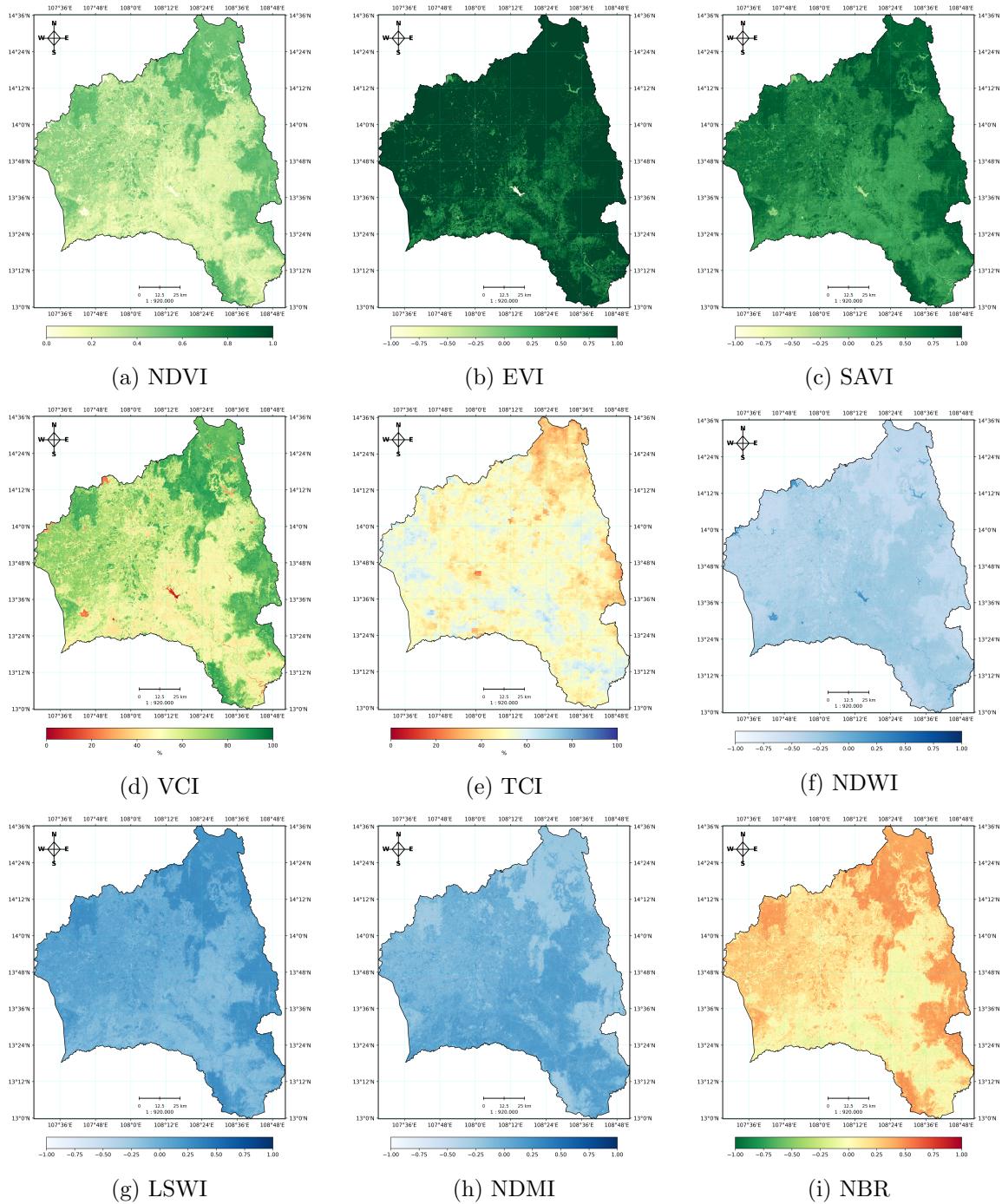
Dựa trên các dữ liệu đã thu thập và tổng quan tài liệu, 15 biến dự báo sau đây đã được tính toán và sử dụng để huấn luyện mô hình (tham khảo quy trình trong Hình 4):

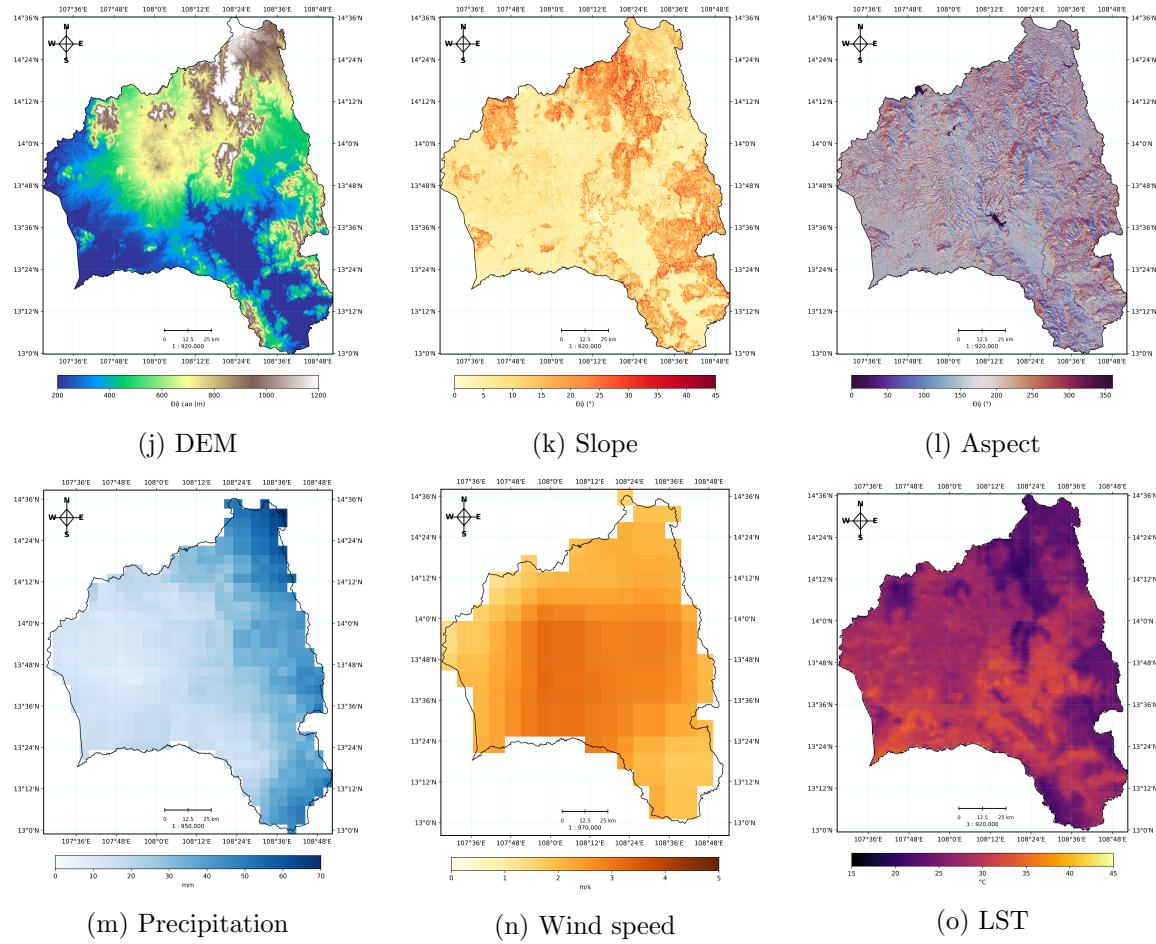


Hình 4: Lưu đồ quy trình xử lý dữ liệu và xây dựng mô hình dự báo nguy cơ cháy rừng với 15 biến đầu vào, trong đó dữ liệu điểm cháy thực tế được sử dụng để tạo nhãn đánh giá mô hình.

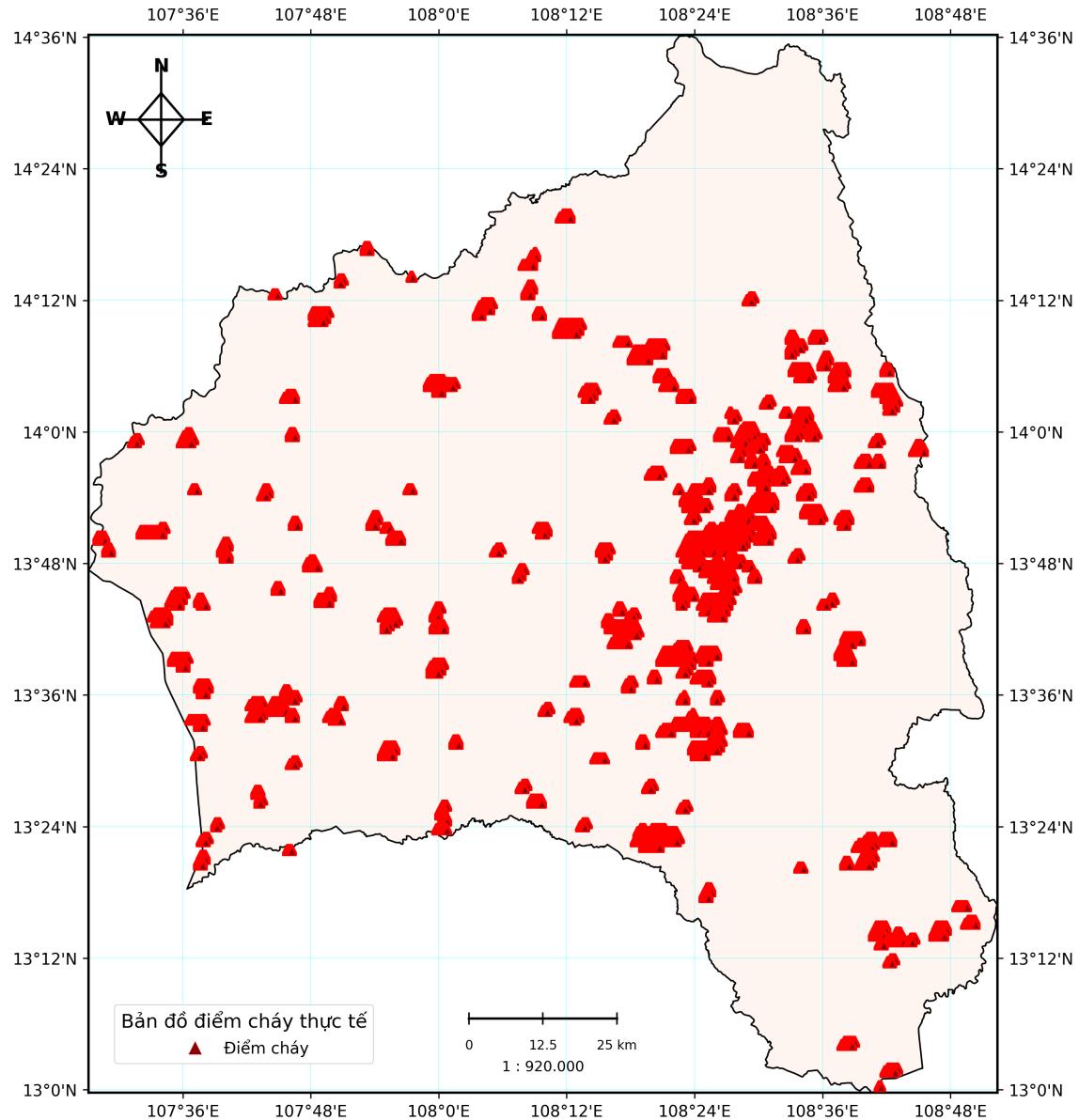
Các nhóm chỉ số dự báo được sử dụng trong nghiên cứu:

- **Chỉ số thực vật (Vegetation Indices):** NDVI (Normalized Difference Vegetation Index), EVI (Enhanced Vegetation Index), SAVI (Soil-Adjusted Vegetation Index), VCI (Vegetation Condition Index), TCI (Temperature Condition Index).
- **Chỉ số nước và ẩm (Water & Moisture Indices):** NDWI (Normalized Difference Water Index), LSWI (Land Surface Water Index), NDMI (Normalized Difference Moisture Index).
- **Chỉ số cháy/phục hồi (Burn or Recovery Indices):** NBR (Normalized Burn Ratio).
- **Dữ liệu địa hình (Topographic Variables):** DEM (Digital Elevation Model), Slope (Độ dốc), Aspect (Hướng dốc).
- **Dữ liệu khí tượng (Meteorological Variables):** Precipitation (Lượng mưa), Wind speed (Tốc độ gió).





Hình 5: Các bản đồ phân bố không gian của 15 biến đầu vào chính sử dụng trong mô hình dự báo nguy cơ cháy rừng.



Hình 6: Bản đồ các điểm cháy thực tế từ MODIS được sử dụng làm nhãn để huấn luyện và đánh giá mô hình.

3.3 Xây dựng Mô hình Học máy

Quy trình xây dựng mô hình bao gồm các bước sau:

- Chuẩn bị dữ liệu:** Tất cả 15 biến đầu vào bao gồm các chỉ số thực vật, độ ẩm, địa hình và khí hậu được chuẩn hóa và kết hợp thành một bộ dữ liệu đa chiều. Dữ liệu điểm cháy từ MODIS FireMask được sử dụng để tạo lớp nhãn cháy (Fire_Label) cho việc huấn luyện và đánh giá mô hình. Các pixel có giá trị FireMask > 6 được gán nhãn là cháy (1), các pixel khác

là không cháy (0).

2. **Lấy mẫu dữ liệu:** Để đảm bảo tính nhất quán, tất cả các biến dự báo (sau khi được xử lý ở các độ phân giải gốc hoặc độ phân giải phân tích phù hợp như mô tả ở Mục 3.1, ví dụ: các chỉ số Sentinel ở 10-30m, LST ở 1km, nhãn cháy MODIS ở 500m) và lớp nhãn cháy đã được lấy mẫu thống nhất tại quy mô không gian 500m sử dụng phương pháp `stratifiedSample` trong Google Earth Engine. Tổng cộng 1000 điểm mẫu đã được thu thập theo cách này, phân tầng theo lớp nhãn cháy (`Fire_Label`). Các điểm mẫu này sau đó được chia ngẫu nhiên thành 70% cho tập huấn luyện (khoảng 700 điểm) và 30% cho tập kiểm tra (khoảng 300 điểm).
3. **Huấn luyện mô hình:** Hai thuật toán học máy là Random Forest (RF) và Gradient Tree Boosting (GTB) được huấn luyện. Đối với Random Forest (RF), các tham số chính được sử dụng trong Google Earth Engine bao gồm: `numberOfTrees` = 100 (Số lượng cây trong rừng), `minLeafPopulation` = 5 (Số lượng mẫu tối thiểu tại một nút lá), `bagFraction` = 0.7 (Tỷ lệ mẫu con cho mỗi cây), và `seed` = 42 (Giá trị khởi tạo cho bộ sinh số ngẫu nhiên). Đối với Gradient Tree Boosting (GTB), các tham số chính được sử dụng bao gồm: `numberOfTrees` = 100 (Số lượng cây), `shrinkage` = 0.05 (Tốc độ học), `samplingRate` = 0.7 (Tỷ lệ mẫu con của dữ liệu đầu vào), `maxNodes` = 10 (Số lượng nút tối đa cho mỗi cây), và `seed` = 42. Mô hình được huấn luyện ở chế độ `PROBABILITY` để tạo bản đồ xác suất và `CLASSIFICATION` để đánh giá độ chính xác trên tập kiểm tra.
4. **Tạo bản đồ dự đoán:** Các mô hình đã huấn luyện sau đó được áp dụng lên một chồng dữ liệu các biến dự báo đầu vào. Chồng dữ liệu này đã được đồng nhất về độ phân giải không gian 30m (sử dụng phương pháp nội suy song tuyến tính - bilinear interpolation) để tạo ra các bản đồ nguy cơ cháy rừng cuối cùng với độ phân giải 30m.

3.4 Dánh giá Mô hình

Hiệu suất của các mô hình được đánh giá trên tập dữ liệu kiểm tra (30% dữ liệu mẫu) bằng cách sử dụng ma trận nhầm lẫn (Confusion Matrix) và các chỉ số thống kê bao gồm: Độ chính xác tổng thể (Overall Accuracy), Độ chính xác cho lớp cháy (Precision), Độ nhạy cho lớp cháy (Recall), và F1-Score cho lớp cháy. Ngoài ra, để đánh giá thêm khả năng tổng quát hóa của mô hình, kiểm định chéo 5-fold (5-fold cross-validation) cũng đã được thực hiện trên tập dữ liệu huấn luyện (70% dữ liệu mẫu, có khả năng sử dụng dữ liệu đã được xuất ra và xử lý bằng một quy trình bên ngoài Google Earth Engine). Kết quả của kiểm định chéo này cũng được sử dụng để so sánh hiệu suất giữa hai mô hình.

4 Kết quả

4.1 Hiệu suất Mô hình

Kết quả đánh giá hiệu suất của hai mô hình được trình bày trong Bảng 2, Bảng 3 và Bảng 4.

Bảng 2: Ma trận nhầm lẩn (Confusion Matrix) cho mô hình Random Forest (RF) trên tập kiểm tra.

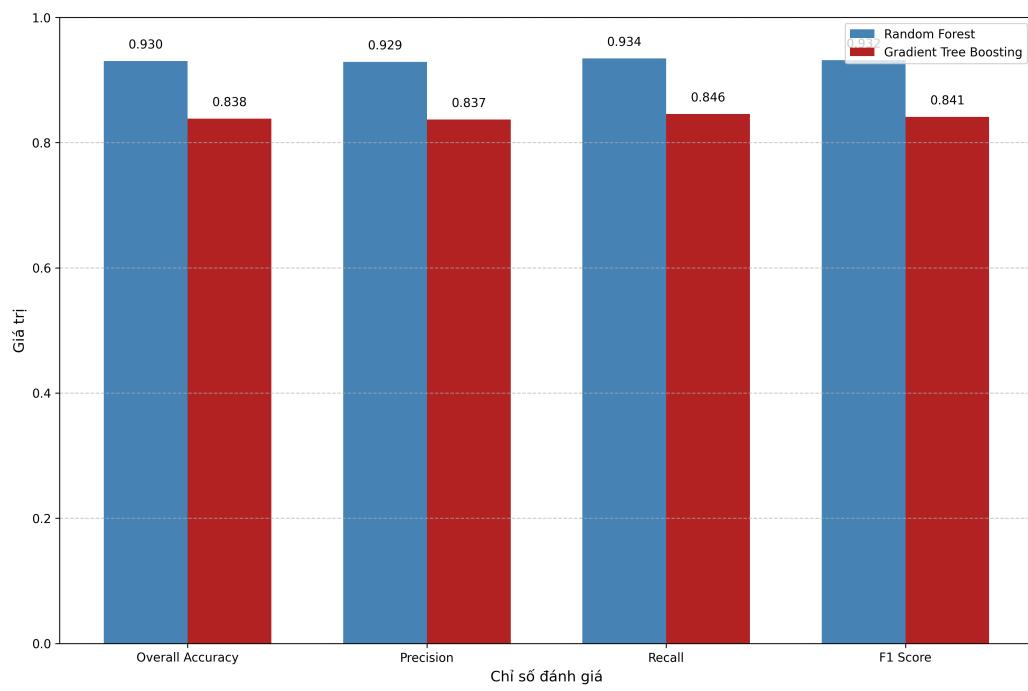
	Dự đoán Không cháy	Dự đoán Cháy
Thực tế Không cháy	254 (TN)	79 (FP)
Thực tế Cháy	64 (FN)	250 (TP)

Bảng 3: Ma trận nhầm lẩn (Confusion Matrix) cho mô hình Gradient Tree Boosting (GTB) trên tập kiểm tra.

	Dự đoán Không cháy	Dự đoán Cháy
Thực tế Không cháy	258 (TN)	75 (FP)
Thực tế Cháy	72 (FN)	242 (TP)

Bảng 4: So sánh các chỉ số đánh giá hiệu suất của mô hình RF và GTB.

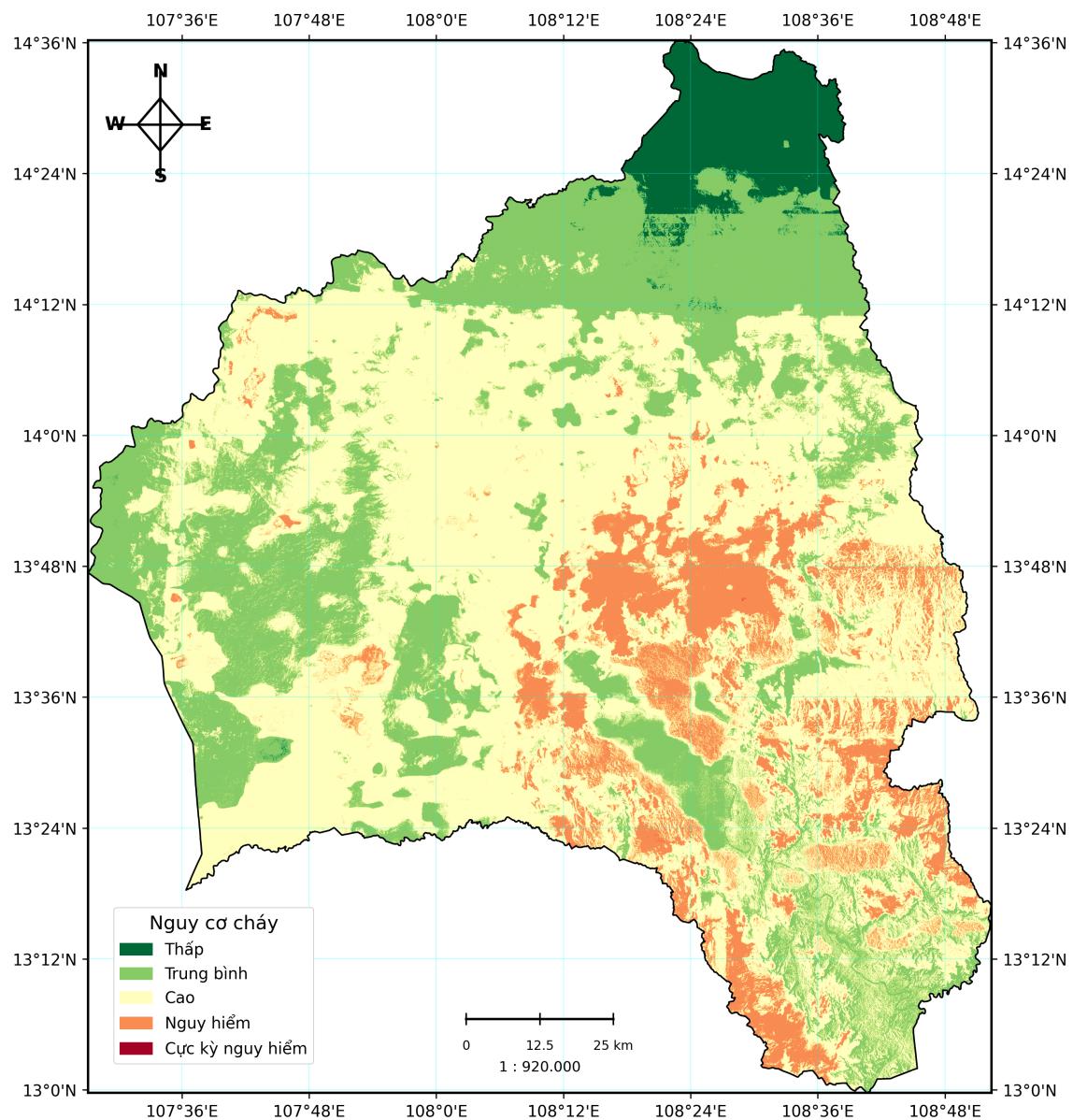
Chỉ số đánh giá	Random Forest	Gradient Tree Boosting
Overall Accuracy	93.05%	83.81%
Recall (Lớp Cháy)	0.9344	0.8455
Precision (Lớp Cháy)	0.9290	0.8369
F1-Score (Lớp Cháy)	0.9317	0.8412
Cross-validation (5-fold)	77.90%	77.28%



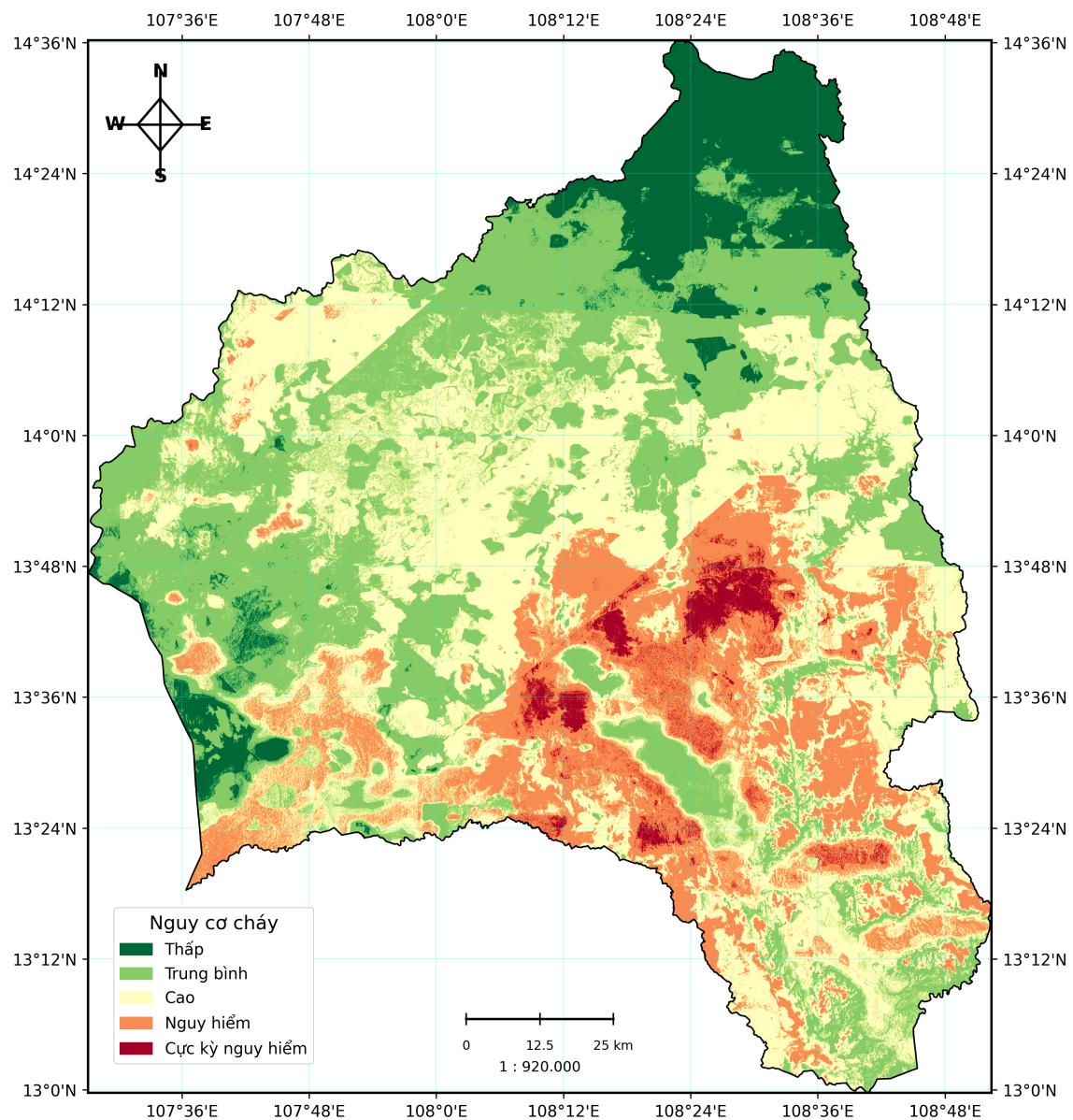
Hình 7: So sánh các chỉ số đánh giá hiệu suất giữa mô hình Random Forest và Gradient Tree Boosting.

4.2 Bản đồ Nguy cơ Cháy rừng

Bản đồ dự đoán nguy cơ cháy rừng cho tỉnh Gia Lai, được tạo ra từ mô hình RF và GTB, được trình bày trong Hình 8 và Hình 9.



Hình 8: Bản đồ dự đoán nguy cơ cháy rừng tỉnh Gia Lai sử dụng mô hình Random Forest.



Hình 9: Bản đồ dự đoán nguy cơ cháy rừng tỉnh Gia Lai sử dụng mô hình Gradient Tree Boosting.

4.3 Thông kê Diện tích theo Cấp độ Nguy cơ

Thông kê diện tích và tỷ lệ phần trăm của từng cấp độ nguy cơ cháy (dựa trên mô hình Random Forest và Gradient Tree Boosting) được trình bày trong Bảng 5 và Bảng 6.

Bảng 5: Thống kê diện tích và tỷ lệ phần trăm theo từng cấp độ nguy cơ cháy (Mô hình RF).

Cấp độ nguy cơ	Diện tích (ha)	Tỷ lệ (%)
1: Thấp	75,830.850	4.9%
2: Trung bình	471,075.570	30.3%
3: Cao	867,867.390	55.8%
4: Nguy hiểm	192,581.460	12.4%
5: Cực kỳ nguy hiểm	823.320	0.1%
Tổng cộng	1,608,178.590	100%

Bảng 6: Thống kê diện tích và tỷ lệ phần trăm theo từng cấp độ nguy cơ cháy (Mô hình GTB).

Cấp độ nguy cơ	Diện tích (ha)	Tỷ lệ (%)
1: Thấp	145,292.310	9.3%
2: Trung bình	535,017.060	34.4%
3: Cao	605,722.050	38.9%
4: Nguy hiểm	289,231.560	18.6%
5: Cực kỳ nguy hiểm	32,915.610	2.1%
Tổng cộng	1,608,178.590	100%

5 Thảo luận

Nghiên cứu này đã thành công trong việc xây dựng và đánh giá các mô hình học máy để dự đoán nguy cơ cháy rừng tại tỉnh Gia Lai bằng cách sử dụng dữ liệu viễn thám đa nguồn trên nền tảng Google Earth Engine. Kết quả đánh giá trên tập kiểm tra cho thấy mô hình RF thể hiện hiệu suất vượt trội hơn trong việc xác định các khu vực có nguy cơ cháy, với độ chính xác tổng thể 93.05% và F1-score cho lớp cháy 0.9317, cao hơn đáng kể so với mô hình GTB (83.81% và 0.8412). Tuy nhiên, kết quả kiểm định chéo 5-fold trên tập huấn luyện cho thấy cả hai mô hình có độ chính xác trung bình khá tương đồng (RF: 77.90%, GTB: 77.28).

Phân tích độ quan trọng của các biến (dựa trên kết quả từ cả hai mô hình) đã chỉ ra rằng nhiệt độ bề mặt đất (Temperature), độ cao địa hình (DEM), lượng mưa (Precipitation) và tốc độ gió (WindSpeed) đóng vai trò quyết định trong việc hình thành nguy cơ cháy tại khu vực nghiên cứu. Đây là những yếu tố liên quan trực tiếp đến điều kiện khí hậu và địa hình, điều này phù hợp với lý thuyết về cơ chế hình thành cháy rừng. Điều đáng chú ý là các chỉ số thực vật truyền thống như NDVI không nằm trong top 5 biến quan trọng nhất đối với cả hai mô hình, mặc dù chúng thường được coi là chỉ báo quan trọng của trạng thái thực vật và nguy cơ cháy trong nhiều nghiên cứu khác. Ví dụ, đối với mô hình RF, 5 biến quan trọng nhất là Temperature, DEM, Precipitation, WindSpeed, và NBR; trong khi đối với GTB là Temperature, WindSpeed, Precipitation, DEM, và EVI.

Bản đồ nguy cơ cháy được tạo ra từ hai mô hình (Hình 8 và 9) cho thấy sự phân bố không gian của các vùng có mức độ rủi ro khác nhau. Nhìn chung, cả hai bản đồ đều xác định các khu vực có nguy cơ cao tập trung ở phía trung tâm và phía đông của tỉnh, tuy nhiên bản đồ từ mô hình RF có vẻ mịn hơn và ít phân mảnh hơn so với GTB. Điều này có thể giải thích do đặc tính khác nhau của hai thuật toán: RF xây dựng nhiều cây quyết định độc lập và lấy trung bình kết quả, trong khi GTB xây dựng các cây tuần tự để khắc phục lỗi của các cây trước đó.

Khi so sánh với bản đồ các điểm cháy thực tế (Hình 6), có thể thấy rằng phần lớn các điểm cháy ghi nhận được nằm trong những vùng được dự đoán có nguy cơ từ cao đến rất cao, đặc biệt là theo bản đồ từ mô hình RF. Điều này góp phần xác nhận tính hiệu quả của mô hình trong việc xác định các khu vực có nguy cơ cháy cao.

6 Kết luận

Nghiên cứu này đã chứng minh tiềm năng của việc tích hợp dữ liệu viễn thám và các thuật toán học máy tiên tiến trên nền tảng Google Earth Engine để xây dựng bản đồ dự đoán nguy cơ cháy rừng hiệu quả cho tỉnh Gia Lai. Cả hai mô hình RF và GTB đều cho thấy hiệu suất tốt, với mô hình RF đạt độ chính xác tổng thể cao hơn (93.05%) nhưng mô hình GTB có khả năng tổng quát hóa tốt hơn trên dữ liệu mới theo kết quả cross-validation.

Phân tích độ quan trọng của các biến đã xác định bốn yếu tố có ảnh hưởng lớn nhất đến nguy cơ cháy rừng tại Gia Lai, bao gồm: nhiệt độ bề mặt đất (Temperature), độ cao địa hình (DEM), lượng mưa (Precipitation) và tốc độ gió (WindSpeed). Việc sử dụng dữ liệu điểm cháy thực tế từ MODIS làm nhãn đánh giá đã đóng vai trò quan trọng trong việc đánh giá hiệu suất của mô hình. Kết quả này nhấn mạnh vai trò quan trọng của các yếu tố khí hậu và địa hình trong việc xác định nguy cơ cháy, và gợi ý rằng các nỗ lực giám sát cháy rừng nên tập trung đặc biệt vào những khu vực có nhiệt độ cao, độ cao trung bình, lượng mưa thấp và tốc độ gió mạnh.

Bản đồ nguy cơ cháy rừng được tạo ra từ mô hình RF là một công cụ có giá trị, có thể hỗ trợ đắc lực cho cơ quan quản lý rừng và phòng chống cháy rừng tại Gia Lai trong việc phân bổ nguồn lực, lập kế hoạch tuần tra và triển khai các biện pháp phòng ngừa hiệu quả, đặc biệt tại các khu vực có nguy cơ cao và cực kỳ nguy hiểm đã được xác định trên bản đồ.

Các hướng nghiên cứu tiếp theo có thể tập trung vào: (1) Tích hợp các yếu tố xã hội-kinh tế và hoạt động của con người vào mô hình dự báo; (2) Phát triển mô hình dự báo theo thời gian thực hoặc ngắn hạn để hỗ trợ cảnh báo sớm; (3) Khảo sát thực địa để kiểm chứng và hiệu chỉnh kết quả mô hình; và (4) Mở rộng phạm vi áp dụng cho các khu vực khác có nguy cơ cháy rừng cao ở Việt Nam.

Lời cảm ơn

Nghiên cứu này được thực hiện với sự hỗ trợ của TS. Hà Minh Cường và ThS. Hoàng Tích Phúc của Viện Công nghệ Hàng không Vũ trụ, Trường Đại học Công nghệ - Đại học Quốc gia Hà Nội. Chúng tôi xin chân thành cảm ơn.

Mã nguồn và Dữ liệu

Các bộ dữ liệu gốc được sử dụng trong nghiên cứu này có nguồn gốc từ các kho lưu trữ công cộng đã được mô tả chi tiết trong phần Phương pháp. Toàn bộ mã nguồn xử lý dữ liệu, huấn luyện mô hình và tạo ra các kết quả được trình bày trong bài báo này được công bố công khai tại GitHub: <https://github.com/ninhaidang/GiaLai-forest-fire.git>. Các sản phẩm trung gian, bao gồm dữ liệu đã xử lý và bản đồ nguy cơ chi tiết, sẵn sàng được cung cấp theo yêu cầu hợp lý gửi đến tác giả liên hệ, nhằm mục đích hỗ trợ việc kiểm tra và tái tạo kết quả nghiên cứu.

Tài liệu

1. Breiman L. Random Forests. *Machine Learning* 2001;45:5–32.
2. Friedman JH. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics* 2001;29:1189–232.
3. Adnan M, Ahmad T, Akhter J, Pourghasemi HR **and** Naghibi S. GIS-based forest fire susceptibility modeling using advanced machine learning techniques: A case study in Northern Pakistan. *Geocarto International* 2022;37:6585–609.
4. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D **and** Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 2017;202:18–27.
5. Drusch M, Del Bello U, Carlier S **and others**. Sentinel-2: ESA’s Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment* 2012;120:25–36.
6. Wan Z. New Refinements and Validation of the MODIS Land-Surface Temperature/Emissivity Products. *Remote Sensing of Environment* 2014;140:132–44.
7. Farr TG, Rosen PA, Caro E **and others**. The Shuttle Radar Topography Mission. *Reviews of Geophysics* 2007;45, RG2004.
8. Funk C, Peterson P, Landsfeld M **and others**. The climate hazards infrared precipitation with stations. *Scientific Data* 2015;2:150066.
9. Muñoz-Sabater J, Dutra E, Agustí-Panareda A **and others**. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data* 2021;13:4349–83.
10. Justice CO, Giglio L, Korontzi S **and others**. The MODIS fire products. *Remote Sensing of Environment* 2002;83:244–62.