KHẢ NĂNG XÂY DỰNG BẢN ĐỒ LỚP PHỦ MẶT ĐẤT TỪ DỮ LIỆU VIỄN THÁM SENTINEL-2 THEO PHƯƠNG PHÁP PHÂN LOẠI RANDOM FOREST TRÊN NỀN TẢNG ĐIỆN TOÁN ĐÁM MÂY

VŨ THỊ PHƯƠNG THẢO, SOUKSAKONE SENGCHANH
Trường Đại học Mỏ - Địa chất

Tóm tắt:

Đây là nghiên cứu nhu cầu thực tế trong lĩnh vực viễn thám để xây dựng cơ sở dữ liệu lớp phủ mặt đất phục vụ theo dõi sự biến động các đối tượng trên bề mặt Trái đất trong các ứng dụng đa lĩnh vực.

Trong giám sát tài nguyên môi trường nói chung, tài nguyên nước trong các lưu vực sông nói riêng, sự kết hợp giữa thông tin, dữ liệu viễn thám và mô hình tính toán tài nguyên nước trong lưu vực để xây dựng các kịch bản tài nguyên nước phục vụ trực tiếp cho việc giám sát, hỗ trợ quyết định cho thích ứng thiên tai cũng như quy hoạch phát triển kinh tế - xã hội phụ thuộc nhiều vào yếu tố lớp phủ. Lớp phủ là sự kết hợp của nhiều thành phần như thực phủ, thổ nhưỡng, đá gốc và mặt nước chịu sự tác động của các nhân tố tự nhiên như nắng, gió, mưa bão và ảnh hưởng trực tiếp đến dòng chảy trong lưu vực. Dòng chảy trong lưu vực sông bên cạnh chịu ảnh hưởng của lớp phủ, còn chịu ảnh hưởng của địa hình. Thực tế, tồn tại mối quan hệ giữa phân bố lớp phủ và bề mặt địa hình, với các đặc trưng về dáng địa hình, thổ nhưỡng và điều kiện khí hậu.

Để tận dụng tối đa tính hữu dụng dùng phương pháp phân loại dữ liệu viễn thám quang học trong xây dựng bản đồ lớp phủ mặt đất, bài báo giới thiệu phương pháp xây dựng cơ sở dữ liệu, bản đồ lớp phủ mặt đất từ dữ liệu viễn thám Sentinel-2 theo phương pháp phân loại random forest trên nền tảng điện toán đám mây Google Earth Engine. Thực nghiệm sử dụng dữ liệu Sentinel-2 thu nhận năm 2020 tại khu vực tỉnh Louangphabang, là một tỉnh đầu nguồn lưu vực sông thuộc lãnh thổ của CHDCND Lào.

Từ khóa: Lớp phủ mặt đất; random forest; Sentinel-2.

1. Đặt vấn đề

Luật Tài nguyên nước [7] quy định rằng việc phòng, chống và khắc phục hậu quả tác hại do nước gây ra phải có kế hoạch và biện pháp chủ động phòng, tránh, giảm nhẹ, hạn chế tác hại do nước gây ra; bảo đảm kết hợp hài hòa giữa lợi ích của cả nước với các vùng, các ngành; giữa khoa học, công nghệ hiện đại

Ngày nhận bài: 1/5/2022, ngày chuyển phản biện: 5/5/2022, ngày chấp nhận phản biện: 9/5/2022, ngày chấp nhận đăng: 25/5/2022

với kinh nghiệm truyền thống của nhân dân và phù hợp với khả năng của nền kinh tế.

Việc xử lý, chiết tách thông tin lớp phủ từ dữ liệu viễn thám quang học làm đầu vào cho các mô hình thủy văn thủy lực trên cơ sở kết hợp số liệu thủy văn đã được khai thác khá phổ biến trong những năm gần đây. Mô hình hóa dòng chảy, ngập lũ, xâm nhập mặn và xói lỡ bồi lắng, cho đến bài toán tính toán nghiên cứu dòng chảy lưu vực vực sông cho thấy ứng dụng công nghệ viễn thám là cần thiết [4].

Hiện nay, các mô hình Mike 11, MikeFLOOD, MikeSHE, GeoSFM, SWAT … có giao diện thân thiện và có khả năng giải nhiều bài toán dòng chảy tích hợp nhiều loại dữ liệu. Mô hình thủy văn thủy lực thường mô phỏng dòng chảy từ mưa và các đặc trưng lớp phủ trên lưu vực [8], [9], [15].

Do đặc điểm tài nguyên nước trong lưu vực thường tại khu vực đồi núi trên vùng rộng, phương pháp đo đạc trực tiếp không hiệu quả, việc giám sát lưu vực sông có sự tham gia của dữ liệu viễn thám nói chung, thông tin lớp phủ mặt đất chiết tách từ dữ liệu viễn thám [4] nói riêng giúp ích nhiều cho khả năng dự báo và hạn chế các tác động của việc thay đổi dòng chảy thượng lưu [5], [14]. Hơn nữa, vấn đề chiết tách thông tin lớp phủ mặt đất phục vụ giám sát tài nguyên nước trong lưu vực là cần thiết.

Ảnh viễn thấm quang học Sentinel-2 kể từ khi thu nhận được từ năm 2015 từ Chương trình Copernicus (của Cơ quan hàng không vũ trụ châu Âu) đã được đưa vào ứng dụng trong nhiều ngành, lĩnh vực. Trước đây, để xây dựng bản đồ lớp phủ mặt đất, thường sử dụng phương pháp giải đoán bằng mắt, những năm gần đáy,

Nghiên cứu phương pháp chiết tách thông tin lớp phủ mặt đất từ dữ liệu viễn thám quang học Sentinel-2 bằng cách sử dụng phương pháp random forrest (RF) góp phần giảm thiểu chi phí cũng như đẩy nhanh tốc độ thành lập bản đồ lớp phủ mặt đất làm đầu vào cho mô hình thủy văn thủy lực phục vụ bài toán giám sát tài nguyên nước nói riêng, giám sát tài nguyên thiên nhiên và môi trường nói chung.

2. Phương pháp nghiên cứu

2.1. Cơ sở lý luận sử dụng công nghệ viễn thám quang học trong chiết tách thông tin lớp phủ mặt đất

Để xây dựng lớp phủ mặt đất, đặc biệt là ở vùng núi thì công nghệ viễn thám là tối ưu. Công nghệ viễn thám có nhiều loại dữ liệu khác nhau, mỗi loại dữ liệu lại cung cấp những thông tin khác nhau với độ chính xác khác nhau. Do đó, công tác đánh giá phải gắn với những đối tượng cụ thể cũng như với từng loại ảnh cụ thể [12]. Nguyên tắc cơ bản để phân biệt các đối tượng lớp phủ mặt đất trên ảnh vệ tinh là dựa vào sự khác biệt về đặc tính phản xạ của chúng trên các kênh phổ. Việc chiết tách thông tin lớp phủ tại vùng núi sử dụng dữ liệu viễn thám quang học cần tập trung vào: (1) độ lớn của đối tượng có thể chiết tách trên ảnh, (2) đặc trưng yếu tố thủy văn trên ảnh, (3) đặc trưng yếu tố thực vật trên ảnh, (4) đặc trưng yếu tố dân cư trên ảnh, (5) đặc trưng yếu tố đất trống và các công trình xây dựng trên ảnh [2]. Trong đó, cần chú trọng sử dụng các mẫu giải đoán ảnh phục vụ bài toán phân loại có kiểm định dựa vào các dấu hiệu trên ảnh, bao gồm các dấu hiệu trực tiếp (Hình dạng, hình mẫu, kích thước, màu, bóng địa vật, cấu trúc và vị trí), các dấu hiệu gián tiếp (Vật thể có hình ảnh giống các vật xung quanh nhưng có thể nhìn thấy từ bóng, đường ống dưới lòng đất có thể nhìn thấy trên mặt đất do biến đổi của vùng đất các dấu vết đào bới để lại, các hiện tượng trong quá khứ hoặc các dấu vết của

chu kỳ sống để lại trên vùng đất…) và các dấu hiệu tổng hợp (Địa hình, thực vật, hiện trạng sử dụng đất, mạng lưới thủy văn, sông suối, hệ thống các khe nứt và các yếu tố dạng tuyến) [1], [4].

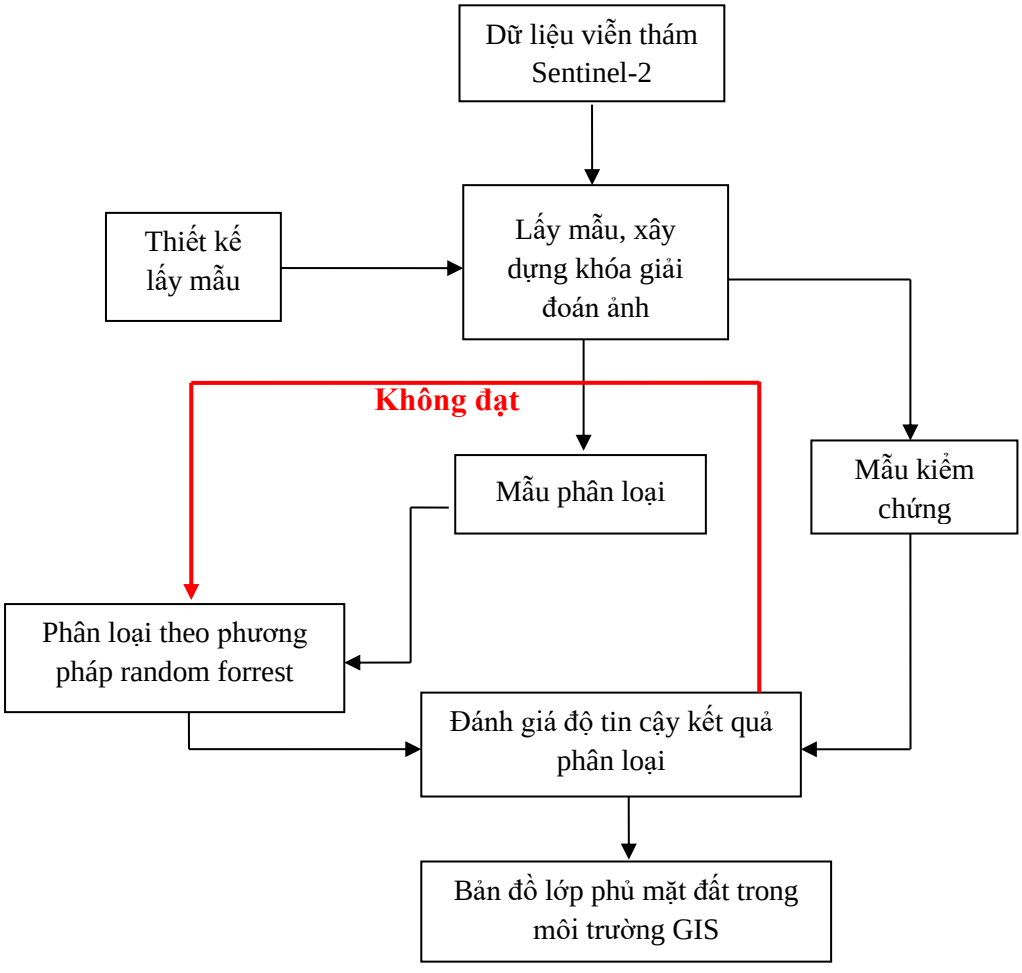
2.2. Phương pháp thành lập bộ dữ liệu lớp phủ mặt đất từ dữ liệu Sentinel-2

Với tính đa dạng về độ phân giải, đa đầu thu, độ phủ rộng và khách quan, để tiết kiệm thời gian và chi phí, các đối tượng trên bề mặt Trái đất thường được chiết tách từ dữ liệu viễn thám [3], [6] [11]. Phương pháp chiết tách và thành lập bộ dữ liệu lớp phủ mặt đất từ dữ liệu viễn thám quang học thể hiện tại Hình 1.

Trong đo, các bước cần chú ý sau:

a) Thu thập dữ liệu viễn thám

Phụ thuộc phạm vị khu vực nghiên cứu và đảm bảo nguyên tắc đồng nhất, nhất quán về dữ liệu, do đó việc thu thập dữ liệu Sentinel-2 phải bao gồm các thông tin: (1) phản xạ phổ bề mặt (OLI/Sentinel), (2) nhiệt độ bề mặt (Sentinel), (3) đã được đánh giá chất lượng giá trị điểm ảnh (Sentinel), (4) phạm vi khu vực nghiên cứu (hệ tọa độ, vùng thực nghiệm) và (5) thông tin thời gian thu ảnh (ngày bắt đầu và kết thúc). Dữ liệu Sentinel-2 thu thập và xử lý sử dụng thuật toán của Zhu và Woodcock [17], đảm bảo độ che phủ mây dưới 25%.



Hình 1: Sơ đồ thành lập bộ dữ liệu lớp phủ mặt đất từ dữ liệu Sentinel-2

b) Phương pháp lấy mẫu

Lấy mẫu được thực hiện theo phương pháp chọn mẫu ngẫu nhiên phân tầng (Stratified Random Sampling - SRS(Eq)), là cách chọn ngẫu nhiên một số mẫu nhất định cho từng đối tượng lớp phủ (tầng). Số điểm lấy mẫu chọn đối tượng bằng tổng số điểm lấy mẫu chia cho số đối tượng lớp phủ.Đối với phương pháp lấy mẫu SRS(Eq), đánh giá ảnh hưởng của kích thước các điểm lấy mẫu đối với kết quả phân loại cũng được thực hiện. Để cố định kích thước điểm lấy mẫu ban đầu, công thức của Cochran [16] được sử dụng với giả định không biết trước tỉ lệ kích cỡ của mỗi đối tượng.

$$n_0 = \frac{Z^2 pq}{e^2} \quad (1)$$

Trong đó: n_0 – là kích cỡ cho từng lớp; Z– là giá trị cho một độ tin cậy nhất định; p – là tỷ lệ của lớp trong dân số; q – là bằng (1 - p) và e^2 – là biên độ lỗi.

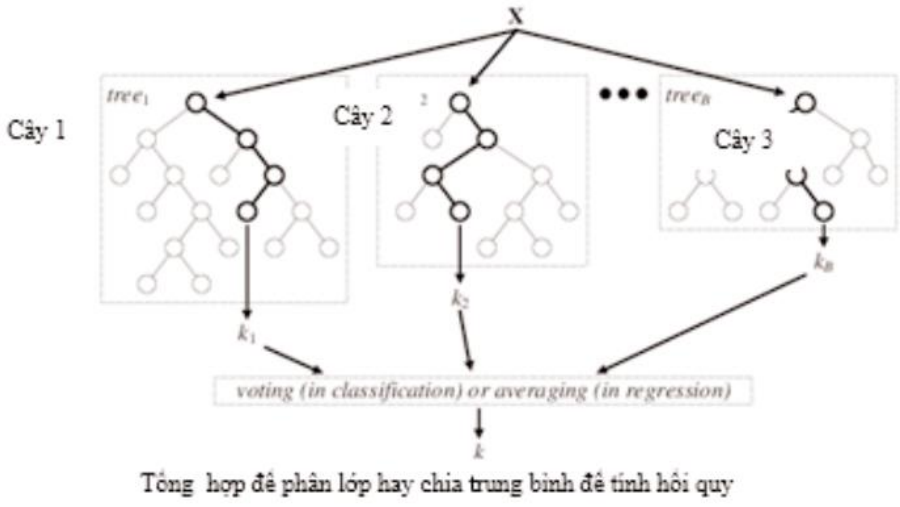
Với khả năng linh hoạt tối đa, giá trị của $p = 0,5$, giá trị của $e = 0,05$, và Z = $1,96$ cho độ tin cậy 95%. Lưu ý, tổng số điểm lấy mẫu bao gồm cả các điểm dành cho huấn luyện máy phục vụ phân loại (70%), và các điểm dùng để dánh giá độ tin cậy kết quả phân loại (30%).

c) Phân loại theo phương pháp random forrest

Phân loại lớp phủ sử dụng phương pháp RF trên nền tảng điện toán đám mây Google Earth Engine (GEE). RF cho thấy hiệu quả hơn so với các phương pháp phân loại thường được sử dụng vì có khả năng tìm ra thuộc tính nào quan trọng hơn so với những thuộc tính khác. RF theo Phạm Minh Hải, Nguyễn Ngọc Quang [10], là một phương pháp thống kê mô hình hóa bằng máy (machine learning statistic) dùng để phục vụ các mục đích phân loại, tính hồi quy và các nhiệm vụ khác bằng cách xây dựngnhiều cây quyết định (Decision tree). Cây quyết định bao gồm ba thành phần: nút quyết định, nút lá và nút gốc. Thuật toán cây quyết định chia tập dữ liệu huấn luyện thành các nhánh, tập dữ liệu này sẽ tách biệt thành các nhánh khác. Trình tự này tiếp tục cho đến khi đạt được một nút lá. Nút lá không thể được phân tách thêm. Các nút trong cây quyết định đại diện cho các thuộc tính được sử dụng để dự đoán kết quả. Các nút quyết định cung cấp một liên kết đến các lá Một khối lượng cây quyết định sản sinh ra bởi phương pháp RF, được huấn luyện thông qua kĩ thuật bagging. Trong đó, bagging là một thuật toán tổng hợp giúp cải thiện độ chính xác của các thuật toán học máy. Kỹ thuật bagging bao gồm việc sử dụng các mẫu khác nhau thay vì chỉ một mẫu dữ liệu. Tập dữ liệu huấn luyện bao gồm các đặc trưng được sử dụng để đưa ra dự đoán.

RF thiết lập kết quả dựa trên các dự đoán của cây quyết định, dự đoán bằng cách lấy giá trị trung bình hoặc giá trị trung bình của kết quả từ các cây khác nhau. Tăng số lượng cây làm tăng độ chính xác kết quả của phương pháp RF. RF làm giảm việc phải thu thập quá nhiều bộ dữ liệu, tạo ra các dự đoán mà không yêu cầu nhiều cấu hình máy.

Quá trình huấn luyện của RF cho các cây ra quyết định diễn tra trong máy phân loại RF. Mọi cây ra quyết định đều bao gồm các nút quyết định, nút nhánh và nút gốc. Nút nhánh của mỗi cây là đầu ra cuối cùng do cây ra quyết định cụ thể tạo ra. Việc lựa chọn đầu ra cuối cùng tuân theo đa số. Phương pháp hoạt động theo bốn bước: (1) chọn các mẫu ngẫu nhiên từ tập dữ liệu đã cho, (2) thiết lập cây quyết định cho từng đối tượng và nhận kết quả dự đoán từ mỗi quyết định cây, (3) xác lập kết quả dự đoán và (4) chọn kết quả được dự đoán nhiều nhất là kết quả cuối cùng.



Hình 2: Sơ đồ biểu diễn các cây quyết định trong phương pháp random forest

Đối với nghiên cứu này, các đối tượng lớp phủ cần cung cấp trong môi trường GEE để phân loại theo phương pháp RF bao gồm: rừng, đất cây trồng, cỏ, mặt nước, dân cư và cơ sở hạ tầng, đất khác.

c) Đánh giá độ tin cậy của kết quả phân loại

Sử dụng 30% mẫu giải đoán làm mẫu kiểm chứng (validation data) để đánh giá độ chính xác kết quả phân loại. Độ chính xác tổng thể từ 60% trở lên có thể chấp nhận được kết quả phân loại. Các bước thực hiện như sau:

- Xác định số lượng mẫu cần để đánh giá, số lượng mẫu phụ thuộc vào độ chính xác giả định cần đạt được khi đánh giá. Độ chính xác giả định càng cao thì số lượng mẫu cần phải lớn và ngược lai. Dung lượng mẫu đánh giá được xác đinh bằng công thức sau [16] :

$$n = \frac{Z^2 O(1-O)}{d^2} \quad (2)$$

Trong đó: O là độ chính xác tổng thể giả định cần đạt được, z: là phần trăm so với phân bố chuẩn chuẩn (z = 1,96 cho khoảng tin cậy 95%, z = 1,645 cho khoảng tin cậy 90%); d là nửa giá trị của O.

- Thiết kế lấy mẫu: sử dụng phương pháp lấy mẫu ngẫu nhiên phân tầng để đánh giá được độ chính xác của tất cả các lớp phân loại.

- Thiết kế phản hồi sử dụng.

- Phân tích, đánh giá độ chính xác, độ tin cậy.

Quá trình đánh giá độ độ tin cậy trước tiên phải thực hiện xây dựng ma trận sai số (Confusion matrix) giữa kết quả phân loại với mẫu kiểm chứng [8]. Mỗi trạng thái lớp phủ cần đảm bảo tối thiểu có 50 điểm kiểm tra.

d) Thành lập bản đồ lớp phủ sử dụng đất

Các đối tượng lớp phủ sau khi đã được phân loại, đánh giá độ tin cậy được đưa sang môi trường GIS để khái quát hóa, gộp các polygon nhỏ, độ lớn theo quy phạm tương ứng tỷ lệ bản đồ giám sát và được biên tập thành bản đồ chuyên đề lớp phủ mặt đất.

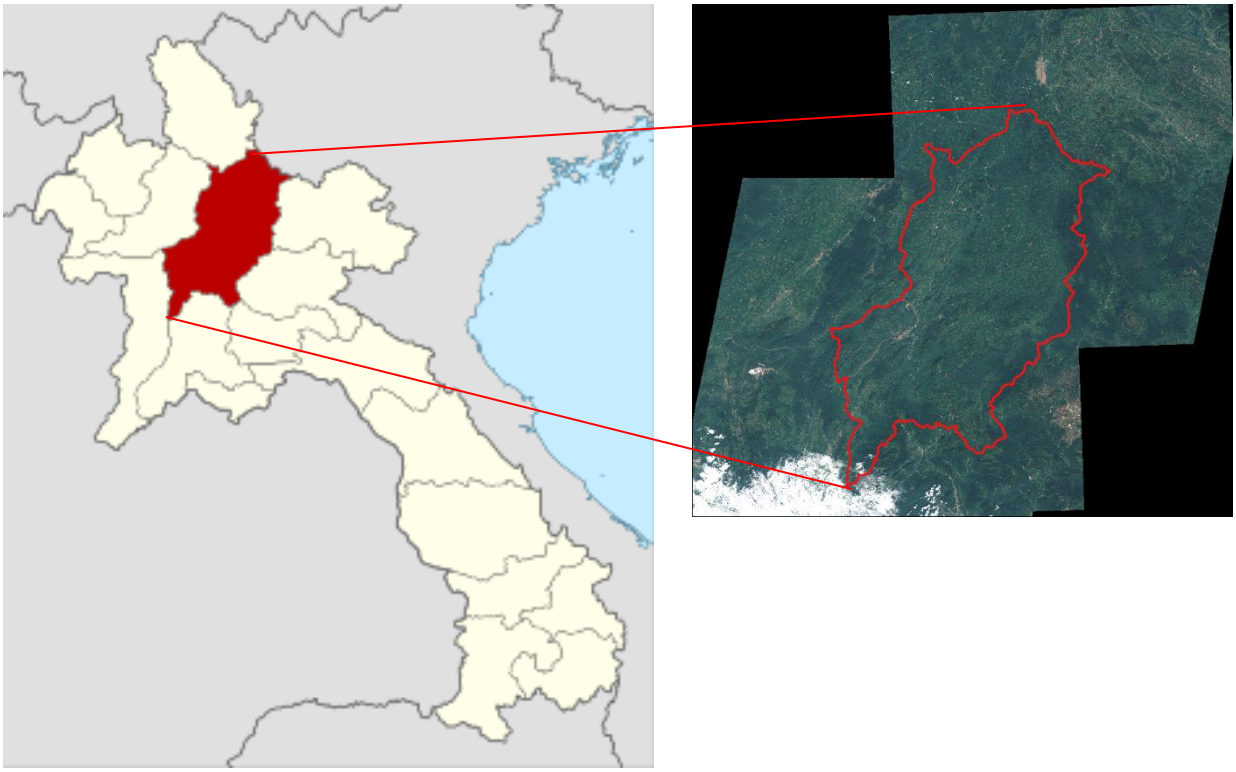
3. Kết quả và đánh giá kết quả

3.1. Khái quát khu vực nghiên cứu

a) Địa lý

Vị trí địa lý khu vực nghiên cứu từ $101^0$41'50" đến $103^0$25'02" độ kinh Đông, từ $19^0$02'31" đến $21^0$08'36" độ vĩ Bắc, diện tích 16.875 km^2, giáp tỉnh Phongsaly (Lào) về phía

Bắc, tỉnh Điện Biên và Sơn La của Việt Nam phía đông bắc, tỉnh Houaphan (Lào) về phía đông, tỉnh Xiangkhouang (Lào) về phía đông nam, tỉnh Vientiane (Viên Chăn, Lào) về phía nam, tỉnh Xaignabouli (Lào) về phía tây nam, và tỉnh Oudomxay (Lào) ở phía tây.



Hình 3: Vị trí khu vực nghiên cứu

b) Địa hình

Khu vực nghiên cứu có địa hình đa dạng, bao gồm núi đất và núi đá có độ dốc lớn, xen kẽ với thung lũng và núi non hiểm trở. Núi cao nằm ở phía bắc thoải dần xuống phía nam. Nổi bật về địa hình trên khu vực nghiên cứu là dãy núi Louangprabang chạy gần như theo hướng Bắc - Nam. Dãy núi Louangprabang với nhiều đỉnh cao như Phu Soi Dao (2 120 m), Phu Khe (2 079 m), Doi Phu Kha (1 980 m).

c) Thủy hệ

Trên khu vực nghiên cứu có hệ thống sông suối dày đặc và đều thuộc lưu vực sông Mê Kông; giới hạn đồng thời bởi sông Mekong chảy theo hướng nam dọc theo phía tây và sông Mae Kok, một nhánh của sông Mekong chảy từ phía đông nhưng ở góc phía bắc ngay trước chỗ hợp lưu giữa hai con sông.

3.2. Dữ liệu đầu vào

- Dữ liệu viễn thám Sentinel-2 của Cơ quan hàng không vũ trụ châu Âu (ESA)được thu thập để tiến hành phần thực nghiệm. Việc lựa chọn dữ liệu viễn thám Sentinel-2 cho phần thực nghiệm vì dữ liệu Sentinel-2 có độ phân giải 10 m, phù hợp cho việc thành lập bản đồ lớp phủ mặt đất ở tỷ lệ 1:50.000, tần suất chụp lặp trung bình là 4 ngày (chụp xen kẽ vệ tinh Sentinel 2A và 2B là 5 ngày; một số vùng giao 2 dải có thể là 3 ngày) [13]. Dữ liệu Sentinel-2 với độ mây che phủ dưới 20%, ngày 12/12//2020 [18] thu nhận được tận dụng thế mạnh của vệ tinh theo thiết kế trong giám

sát các hoạt độ canh tác nông nghiệp, rừng, sử dụng đất, thay đổi thực phủ/ sử dụng đất.

- Bản đồ địa hình khu vực tỉnh Louangphabang, CHDCND Lào, tỷ lệ 1:50.000 thu thập được từ kết quả Chương trình Việt Nam giúp đỡ chính phủ Lào hiện chỉnh bản đồ địa hình UTM tỷ lệ 1:50.000 bằng dữ liệu viễn thám SPOT 4.

3.3. Kết quả nghiên cứu

a) Kết quả xử lý ảnh Sentinel – 2 năm 2020:

Đánh giá độ tin cậy kết quả phân loại ảnh Sentinel-2 các đối tượng trong nghiên cứu sử dụng phương pháp ma trận sai số Kappa Khat. Bước đầu tiên trong quá trình kiểm định là xác định các khu vực kiểm tra trên Google Earth năm 2020. Điểm kiểm tra được tạo ngẫu nhiên trong phần mềm GIS, sau đó được tạo thành các vùng đệm kích thước 1ha. Các vùng đệm này được trích xuất và nạp vào Google Earth. Qua các phân tích trực quan trên ảnh, tính chất của lớp phủ được gán vào các điểm kiểm tra ngẫu nhiên. Số lượng điểm kiểm tra này sau đó được kiểm định thuộc tính một lần nữa với dữ liệu bản đồ địa hình thu thập được. Quá trình kiểm định được thực hiện cho toàn khu vực trên tổng số 191 điểm lấy mẫu cho năm tại thời kỳ 2020. Kết quả đạt đạt độ tin cậy đến 69%, cho thấy phương pháp này đạt giới hạn độ chính xác khá cao. Kết quả xây dựng bộ dữ liệu lớp phủ thể hiện đầy đủ toàn bộ các đối tượng phủ mặt đất chính, các đối tượng này đều dễ dàng thu nhận được từ dữ liệu viễn thám quang học nói chung, ảnh Sentinel-2 nói riêng.

Bảng 1: Kết quả đánh giá độ tin cậy phân loại lớp phủ khu vực nghiên cứu

STT	**Mã phân loại**	**1**	**2**	**3**	**4**	**5**	**6**	**Tổng**	**PA**	**UA**	**OA**	**Kappa**
		Rừng	Cây trồng	Cỏ	Mặt nước	Dân cư và cơ sở hạ tầng	Đất khác					
1	Rừng	**38**	1	5	3	5	5	***57***	**88%**	**92%**	**85%**	**0.69**
2	Cây trồng	4	**12**			6		***22***	**82%**	**88%**		
3	Cỏ	2		**18**		5	5	***30***	**86%**	**81%**		
4	Mặt nước	3	1	7	**13**	2	3	***29***	**74%**	**87%**		
5	Dân cư và cơ sở hạ tầng		3		3	**15**	8	***29***	**90%**	**85%**		
6	Đất khác	3	1	5	1	4	**10**	***24***	**76%**	**68%**		
Tổng		***50***	***18***	***35***	***20***	***37***	***31***	**191**				
		92%	88%	81%	87%	85%	68%					

b) Kết quả thành lập bản đồ lớp phủ khu vực nghiên cứu

Kết quả thành lập bản đồ lớp phủ khu vực tỉnh Louangphabang năm 2020 sử dụng dữ liệu Seltinel 2 thể hiện tại Hình 4, bao gồm:

- Hệ thống giao thông và ghi chú, được lấy từ bản đồ địa hình đã thu thập, hiện chỉnh theo ảnh Sentinel-2 đã xử lý, tương ứng tỷ lệ 1:50.000;

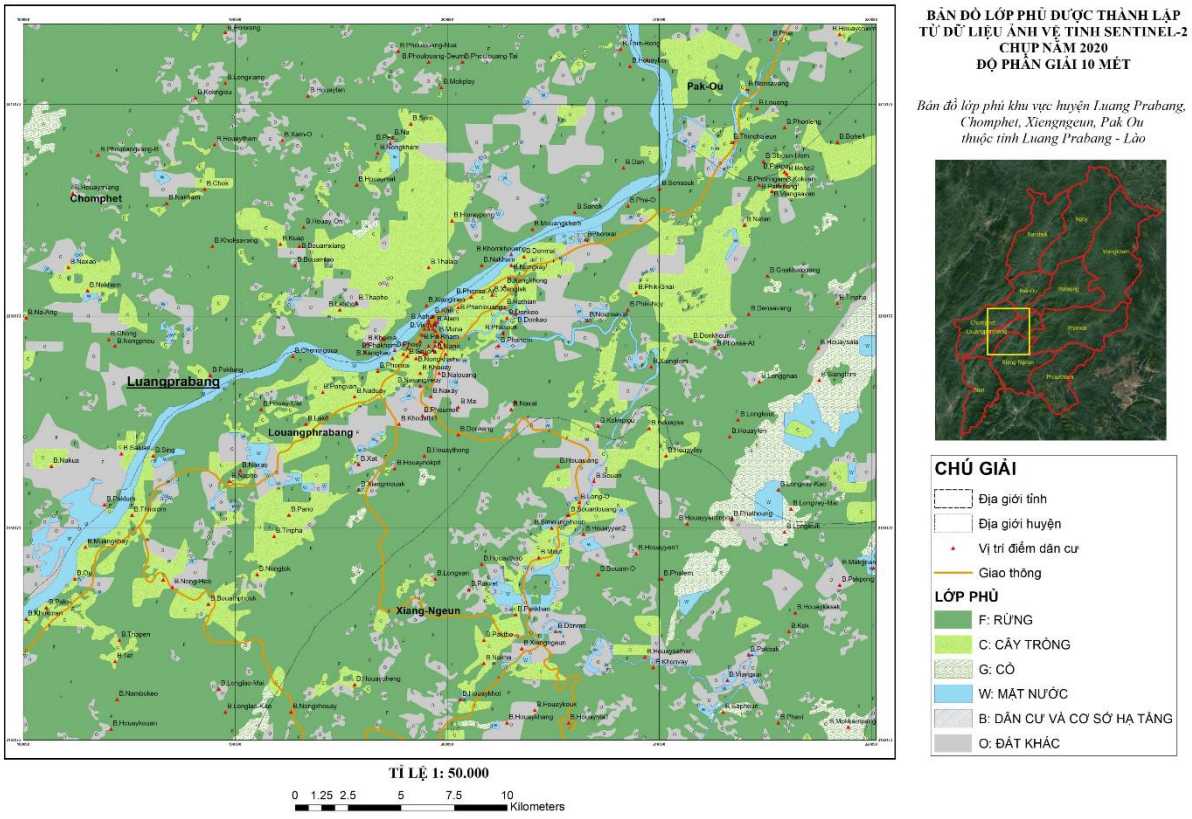
- Hệ thống thủy hệ được chiết tách nội dung chuyên đề lớp phủ trên cơ sở phân loại ảnh Sentinel-2, hệ thống ghi chú được lấy trên bản đồ địa hình tỷ lệ 1:50.000 thu thập được;

- Hệ thống địa hình và ghi chú, hệ thống này được lấy từ bản đồ địa hình tỷ lệ 1:50.000 thu thập được;

- Ghi chú dân cư được lấy từ bản đồ địa hình tỷ lệ 1:50.000 thu thập được;

- Đường ranh giới các cấp được cập nhật mới ranh giới mới nhất đã được công bố;

- Hệ thống các đối tượng chuyên đề lớp phủ mặt đất được chiết tách từ ảnh Sentinel-2 đã xử lý.



Hình 4: Bản đồ lớp phủ mặt đất khu vực Louangphabang năm 2020 (Sử dụng dữ liệu viễn thám Sentinel - 2)

4. Kết luận

Với tần suất chụp lặp là 04 ngày, dữ liệu Sentinel-2 cho phép chiết tách thông tin lớp phủ mặt đất nói chung, khu vực có đặc trưng vùng núi phục vụ giám sát tài nguyên nước thượng nguồn lưu vực sông.

Kết quả phân loại các đối tượng lớp phủ bằng phương pháp RF, sau khi đánh giá độ tin cậy đạt được là 68%, cho phép xây dựng được

bộ dữ liệu lớp phủ đạt yêu cầu thành lập cơ sở dữ liệu/bản đồ lớp phủ tỷ lệ 1:50.000, đây lá bộ thông số đầu vào quan trong cho mô hình thủy văn thủy lực đã qua quá trình hiệu chỉnh và kiểm định mô hình tại trạm thủy văn khu vực nghiên cứu, theo số liệu dòng chảy lịch sử.

Như vậy, phân loại tự động theo phương pháp Random forest đã trở thành công cụ hữu dụng thành lập bộ dữ liệu lơp phủ phục vụ bài toán giám sát tài nguyên nước, dựa vào tính hữu dụng của dữ liệu viễn thám quang học về độ chính xác, đa thời gian, độ phủ rộng phù hợp với các khu vực khó tiếp cận như thượng nguồn các lưu vực sông.

Việc đề xuất và thực nghiệm quy trình nêu trên cho thấy thông tin lớp phủ mặt đất chiết tách từ dữ liệu viễn thám quang học Sentinel-2 nói riêng, dữ liệu viễn thám quang học nói chung đạt độ chính xác cao, hữu hiệu xác định biến động của các đối tượng lớp phủ trên bề mặt trái đất nói chung, tại các khu vực thượng nguồn luwu vực sông nói riêng. Việc áp dụng công nghệ viễn thám còn giúp thực hiện các hoạt động giám sát, theo dõi biến động lớp phủ và tài nguyên nước trong quá trình hoạch định sử dụng hợp lý tài nguyên nước.❍

Tài liệu tham khảo

[1]. Áp dụng quy định kỹ thuật của ipcc cho bộ dữ liệu lớp phủ phục vụ tính toán phát thải khí nhà kính/các bon. Lê Quốc Hưng và nnk. 2020. Tạp chí Khoa học Tài nguyên và Môi trường, số 29, trang 29-41.

[2]. Nguyễn Xuân Lâm. 2006. Nghiên cứu ứng dụng phương pháp viễn thám và hệ thống thông tin địa lý phục vụ mục đích giám sát một số thành phần tài nguyên, môi trường tại các khu vực xây dựng công trình thủy điện. Báo cáo đề tài nghiên cứu khoa học cấp Bộ(Bộ Tài nguyên và Môi trường).

[3]. Hoàng Minh Tuyển. 2006. Nghiên cứu xây dựng khung hỗ trợ ra quyết định trong quản lý tài nguyên nước lưu vực sông Cả. Báo cáo đề tài nghiên cứu khoa học cấp Bộ (Bộ Tài nguyên và Môi trường).

[4]. Lê Quốc Hưng, Vũ Thị Phương Thảo, Trần Thu Huyền. Khả năng tính toán phát thải khí các bon trong lĩnh vực sử dụng đất, thay đổi sử dụng đất và rừng sử dụng dữ liệu ảnh viễn thám. Tạp chí Khoa học Đo đạc và Bản đồ, số 42, trang 44-50.

[5]. Lê Quốc Hưng, Đặng Trường Giang, Kiều Thị Thanh Tâm 2014. Khả năng sử dụng một số phương pháp thành lập DEM vùng núi có độ dốc lớn từ ảnh vệ tinh phục vụ giám sát nước thượng lưu.. Tuyển tập báo cáo Hội nghị Khoa học và Công nghệ Trắc địa và Bản đồ, 07, trang 193-202.

[6]. Tạ Thị Vân Anh, Vũ Thị Phương Thảo. 2021. Kết hợp công nghệ viễn thám và mô hình số trị tính toán dòng chảy lũ về hồ chứa - Thực nghiệm tại lưu vực thủy điện Bản Chát. Tạp chí Khoa học Tài nguyên và Môi trường, số 38, trang 49-62.

[7]. Luật tài nguyên nước số 17/2012/QH13 năm 2012.

[8]. Lê Quốc Hưng, Vũ Thị Phương Thảo, Lê Thị Mai Vân, Nguyễn Ngọc Quang 2014. Kết hợp công nghệ viễn thám và mô hình thủy văn, thủy lực để xây dựng kịch bản tài nguyên nước hồ chứa và sơ bộ đánh giá thiệt hại một số đối tượng trong trường hợp vỡ đập.. Tạp chí Khoa học Tài nguyên và Môi trường, số 03, trang 82-88.

[9]. Nghiên cứu một số mô hình dự báo dòng chảy ở Việt Nam. Trung tâm Quy hoạch và Điều tra tài nguyên nước quốc gia, http://nawapi.gov.vn/index.php?option=com_content&view=article&id=3489%3Anghien-cu-mt-s-mo-hinh-d-bao-dong-chy--vit-

nam&catid=70%3Anhim-v-chuyen-mon-ang-thc-hin&Itemid=135&lang=vi

[10]. Phạm Minh Hải, Nguyễn Ngọc Quang. 2019. Tạp chí Khoa học Đo đạc và bản đồ, 39. 15-19.

[11]. Tô Trung Nghĩa và Lê Hùng Nam. 2004. Xây dựng quy trình vận hành hệ thống lien hồ chứa Hòa Bình, Thác Bà, Tuyên Quang phục vụ cấp nước trong mùa khô cho hạ du lưu vực sông Hồng – Thái Bình, Viện Quy hoạch Thủy lợi.

[12]. Viễn thám và Hệ thông tin địa lý ứng dụng. Nguyễn Ngọc Thạch và nnk. 2003.

[13]. Carrasco, L; O'Neil, A.W; Morton, R. D; Rowland, CS. 2019. Evaluating Combinations of Temporally Aggregated Sentinel-1, Sentinel-2 and Landsat 8 for Land Cover Mapping with Google Earth Engine. Remote Sens., 11(3), 288.

[14]. Kääb A. 2002. Monitoring high-mountaing terrain deformation from repeated air- and spaceborne optical data: examples using digital aerial imagery and ASTER data, ISPRS Journal of photogrammetry & Remote Sensing, 57, 39-52.

[15]. McCuen, R. H. 1998. Hydrologic analysis and design, 2 Ed., Prentice-Hall, Inc., Upper Saddle River, N.J., 814.

[16]. Sampling Techniques. Cochran, W.G. 1977 3rd Edition, John Wiley and Sons, New York.

[17]. Zhu, Z., & Woodcock, C. E. 2014a. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. Remote Sensing of Environment, 152, 217-234.

[18]. https://earthexplorer.usgs.gov/.❍

Summary

Ability to build the landcover map from remote sensing data sentinel-2 based on random forest classification method on cloud computing platform

Vu Thi Phuong Thao, Souksakone Sengchanh - Hanoi University of Mining and Geology

The study focused on actual demands in the field of remote sensing for building the landcover map from remote sensing data Sentinel-2 based on the random forest classification method. In the orientation of water resource monitoring, the combination of remote sensing data and water resource computational model builds water resource scenarios for direct service of monitoring and decision support for climate change adaptation as well as socio-economic development planning. Land cover is the combination of many components such as vegetation cover, soil and water surface under advantages of natural factors such as the sun, wind, rainstorm, etc. Flows in the river basin are not only impacted by the land cover but also by topographic impacts. In fact, there existed relationships between the distribution of land cover objects and topographic surface with the characteristics of topography, soil and climatic conditions. The article introduced the method to build the landcover database from remote sensing data Sentinel-2 based on the random forest classification method on the cloud computing platform Google Earth Engine. A case study using Sentinel-2 data collected in 2020 in Louangphabang province in northern part of Lao PDR.❍

Keyword: Landcover; random forest; Sentinel-2.