

Bài báo khoa học

Nghiên cứu xác định bản đồ lớp phủ sử dụng công nghệ học máy phục vụ đánh giá tác động của biến động lớp phủ đến chế độ dòng chảy trên lưu vực sông Đà ngoài lãnh thổ vào Việt Nam

Trần Mạnh Cường^{1*}, Trần Anh Phương¹, Lê Mạnh An², Bùi Huyền Linh¹, Trần Văn Tú¹

¹ Viện Khoa học tài nguyên nước, Bộ Tài nguyên và Môi trường; man-hcuongkt11@gmail.com; phuongtran.monre@gmail.com; linhb.dctv@gmail.com; trantu.monre@gmail.com;

² Trung tâm Công nghệ Khí tượng thủy văn, Tổng Cục Khí tượng thủy văn, Bộ Tài nguyên và Môi trường; lemanhan82@gmail.com

*Tác giả liên hệ: manhcuongkt11@gmail.com; Tel.: +84-987986233

Ban Biên tập nhận bài: 17/9/2024; Ngày phản biện xong: 28/10/2024; Ngày đăng bài: 25/3/2025

Tóm tắt: Trong những năm qua, lớp phủ trên lưu vực sông Đà bên ngoài lãnh thổ Việt Nam có nhiều thay đổi, góp phần tác động đến dòng chảy vào Việt Nam. Tuy nhiên, do thiếu dữ liệu, việc đánh giá tác động này gặp nhiều thách thức. Bài báo này trình bày kết quả ứng dụng công nghệ học máy để xây dựng bản đồ lớp phủ từ ảnh viễn thám phục vụ đánh giá biến động của lớp phủ theo thời gian và tác động của nó đến dòng chảy trên lưu vực sông Đà ngoài lãnh thổ sử dụng mô hình SWAT. Dữ liệu lớp phủ các năm 2000 và 2020 làm đầu vào cho mô hình SWAT được xây dựng từ ảnh vệ tinh bằng thuật toán học máy trên nền tảng Google Earth Engine. Độ chính xác của lớp phủ được đánh giá bằng độ chính xác tổng thể và hệ số Kappa, cho thấy kết quả giải đoán ảnh tương đối tốt. Mô hình SWAT được thiết lập để so sánh sự thay đổi của lượng nước trên lưu vực tương ứng với lớp phủ năm 2000 (KB1) và năm 2020 (KB2). Lớp phủ xây dựng tại thời điểm năm 2000 và năm 2020 cho thấy diện tích khu vực xây dựng tăng 866,92%, rừng tăng 4,26%, đất trống tăng 3,39%, mặt nước tăng 63,89%, diện tích cây bụi - cỏ giảm 11,60%. Dòng chảy trong KB2 tăng lên so với KB1 trên phần lớn các tiểu lưu vực khi xét trong cả năm. Mùa khô dòng chảy trên toàn lưu vực giảm 0,15%, mùa mưa tăng 0,08% và cả năm tăng 0,04%.

Từ khóa: Sông Đà; Lớp phủ; Mô hình SWAT; Google Earth Engine.

1. Mở đầu

Dòng chảy từ thượng lưu sông Đà về Việt Nam đã có nhiều thay đổi trong những năm qua, tỷ lệ dòng chảy trong mùa mưa giảm xuống và tăng lên trong mùa khô [1]. Ngoài tác động từ quá trình vận hành các hồ thủy điện ở thượng nguồn thì sự thay đổi lớp phủ cũng tác động đến sự thay đổi của các quá trình thủy văn [2-4]. Trong bối cảnh Trung Quốc đang phát triển mạnh mẽ trong những thập kỷ qua, nhiều khu vực đã thay đổi hiện trạng như yêu cầu tất yếu để phù hợp với chiến lược phát triển. Nghiên cứu những tác động của thay đổi lớp phủ khu vực thượng nguồn sông Đà đến dòng chảy trên lưu vực sẽ cung cấp thông tin khoa học, đóng góp vào tổng thể nhiệm vụ quản lý tài nguyên nước trên hạ lưu sông Đà thuộc lãnh thổ Việt Nam. Việc thiếu thông tin, dữ liệu là một cản trở đối với trường hợp nghiên cứu khu vực nằm ngoài lãnh thổ, do thiếu các trạm khí tượng thủy văn hoặc hạn chế trong việc trao đổi thông tin giữa các nước [5, 6]. Ngày nay, các dữ liệu vệ tinh như lượng mưa, nhiệt độ, lớp phủ trở nên phổ biến hơn, với nhiều nguồn và có độ bao phủ gần như toàn cầu, giúp thực

Tạp chí Khí tượng Thủy văn **2025**, 771, 40-54; doi:10.36335/VNJHM.2025(771).40-54 <http://tapchikttv.vn/>

hiện những nghiên cứu ở khu vực thiếu hoặc không có dữ liệu [7]. Nghiên cứu [8] đánh giá dữ liệu vệ tinh có tiềm năng hỗ trợ cho các dự án nghiên cứu quản lý tài nguyên nước ở khu vực thiếu dữ liệu. Nghiên cứu [9] cho thấy sự hữu ích trong việc sử dụng dữ liệu vệ tinh trong nghiên cứu mô phỏng dòng chảy khi không có số liệu thực đo. Trong khi các nghiên cứu sử dụng dữ liệu vệ tinh về lượng mưa hay nhiệt độ để nghiên cứu quá trình thủy văn mưa - dòng chảy ngoài lãnh thổ đang được nhiều nhà khoa học nghiên cứu thì việc nghiên cứu ứng dụng dữ liệu viễn thám để đánh giá tác động của thảm phủ lên dòng chảy còn tương đối hạn chế. Trong những năm gần đây, với sự phát triển của công nghệ viễn thám và học máy, việc giải đoán bản đồ lớp phủ từ ảnh viễn thám sử dụng công nghệ học máy và đánh giá biến động của lớp phủ theo thời gian ngày càng trở nên phổ biến [10–12]. Nghiên cứu [13] đã xây dựng cơ sở dữ liệu bản đồ lớp phủ mặt đất từ dữ liệu viễn thám trên nền tảng điện toán đám mây Google Earth Engine (GEE) cho khu vực khó tiếp cận sử dụng thuật toán Random Forest (RF). Nghiên cứu [14] sử dụng nền tảng điện toán đám mây GEE cùng thuật toán học máy RF xây dựng dữ liệu lớp phủ dựa trên ảnh vệ tinh Sentinel 2 và Landsat 8 chính xác và nhanh chóng. Nghiên cứu [15] so sánh hiệu suất các thuật toán học máy CART (*Classification and Regression Tree*), RF, SVM (*Support Vector Machine*) trong việc lập bản đồ thảm phủ, trong nghiên cứu này RF đạt độ chính xác cao nhất.

Để phục vụ công tác quản lý tài nguyên nước trên lưu vực sông, việc đánh giá tác động của biến động lớp phủ đến chế độ dòng chảy là nhiệm vụ hết sức rất cần thiết do việc sử dụng đất luôn thay đổi theo thời gian. Nhiệm vụ này thường được thực hiện dựa vào việc sử dụng các mô hình thủy văn [16]. Trong các mô hình thủy văn, công cụ đánh giá đất và nước (SWAT) là một trong những mô hình thủy văn được sử dụng rộng rãi nhất do mô hình có khả năng tích hợp sự thay đổi của thảm phủ theo ô lưới đến chế độ dòng chảy [17–19]. Nghiên cứu [20] sử dụng mô hình SWAT để đánh giá tác động của việc thay đổi lớp phủ đến cân bằng nước trên một phần lưu vực sông Hoàng Hà của Trung Quốc. Nghiên cứu [16] sử dụng mô hình SWAT để đánh giá tính phù hợp của một số dữ liệu thảm phủ và tính toán dòng chảy trên lưu vực sông Yom của Thái Lan. Nghiên cứu [21] kiểm tra tác động của các dữ liệu thảm phủ khác nhau đối với thủy văn lưu vực Gilgel bằng mô hình thủy văn SWAT, nghiên cứu chỉ ra những thay đổi thủy văn ở quy mô tiểu lưu vực. Các nghiên cứu này đều cho thấy hiệu quả khi áp dụng mô hình SWAT cho các trường hợp sử dụng đất khác nhau để đánh giá những thay đổi của dòng chảy trên lưu vực sông.

Nghiên cứu thực hiện đánh giá tác động từ thay đổi lớp phủ đến dòng chảy (thông qua lớp dòng chảy) trên lưu vực thượng nguồn sông Đà thông qua mô hình thủy văn SWAT. So với các nghiên cứu trước đây, nghiên cứu này kết hợp cả công nghệ giải đoán ảnh viễn thám sử dụng thuật toán học máy và mô hình thủy văn SWAT để đánh giá tác động của biến động lớp phủ đến chế độ dòng chảy. Dữ liệu lớp phủ được xây dựng từ dữ liệu vệ tinh Landsat vào thời điểm năm 2000 và 2020 bằng thuật toán học máy RF trên công cụ GEE. Dữ liệu lớp phủ sau đó được đưa vào mô hình SWAT để tính toán. Những thay đổi về dòng chảy giữa trường hợp lớp phủ năm 2000 (KB1) và năm 2020 (KB2) cung cấp thêm dữ liệu, làm cơ sở để nhận định, đánh giá xu hướng thay đổi trong tương lai gần, phục vụ nhiệm vụ quản lý tài nguyên nước bền vững trên hạ lưu sông Đà thuộc Việt Nam.

2. Phương pháp nghiên cứu và dữ liệu nghiên cứu

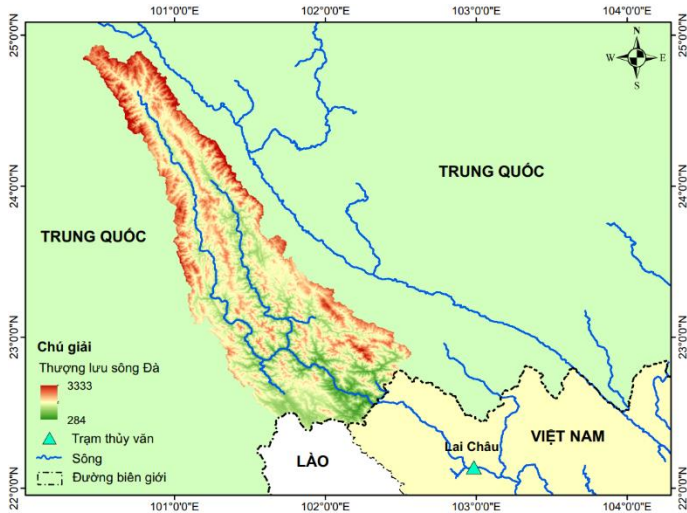
2.1. Khu vực nghiên cứu

Bài báo thực hiện nghiên cứu đối với lưu vực thượng lưu sông Đà bên ngoài lãnh thổ Việt Nam (Hình 1). Tổng diện tích lưu vực sông Đà là 52.900 km², trong đó phần thuộc lãnh thổ Trung Quốc chiếm khoảng 47%. Lưu vực sông nằm trong khu vực cao nguyên gió mùa cận nhiệt đới Nam Á, có độ che phủ thực vật cao, thảm thực vật thuộc vùng tiếp giáp giữa rừng lá rộng thường xanh cận nhiệt đới Nam Á, các khu rừng nhiệt đới gió mùa và rừng mưa nhiệt đới. Phía thượng lưu, thung lũng, đồi, và bồn địa đan xen nhau, địa hình đồi núi không

đồng đều và dòng sông tương đối hẹp. Phía hạ lưu, địa hình trở nên phẳng hơn, thung lũng sông tương đối rộng và dòng sông khá thẳng. Trên sông Đà, mùa lũ tập trung từ tháng 6 đến tháng 10, mùa kiệt từ tháng 2 đến tháng 5.

2.2. Dữ liệu nghiên cứu

Mô hình SWAT cần dữ liệu đầu vào bao gồm: dữ liệu địa hình (DEM), dữ liệu lớp phủ, dữ liệu thổ nhưỡng, lượng mưa, nhiệt độ. Vì khu vực nghiên cứu nằm ngoài lãnh thổ, không có số liệu thực đo vì vậy phần lớn các dữ liệu được thu thập từ nguồn vệ tinh, số liệu lưu lượng dòng chảy được thu thập tại trạm thủy văn Lai Châu. Thông tin về dữ liệu thu thập phục vụ nghiên cứu được tổng hợp trong Bảng 1.



Hình 1. Lưu vực sông Đà bên ngoài lãnh thổ Việt Nam.

Bảng 1. Dữ liệu sử dụng trong nghiên cứu làm đầu vào cho mô hình SWAT.

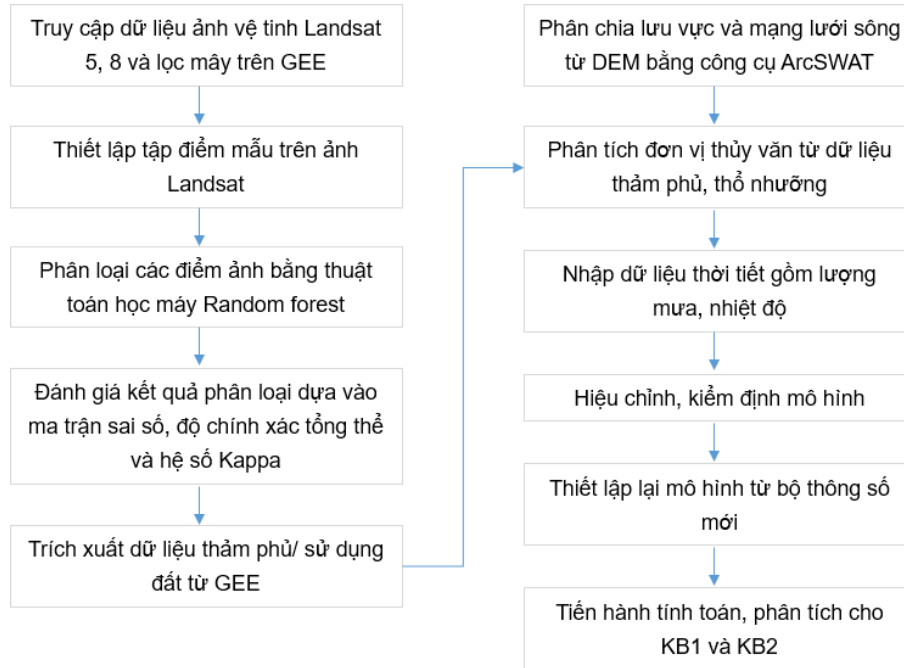
Dữ liệu	Tên dữ liệu	Độ phân giải/ Tỷ lệ	Thời gian thu thập	Mục đích sử dụng
DEM	NASADEM	30 m	2000	Thiết lập mô hình SWAT
Dữ liệu thảm phủ	Landsat 5,	30 m	2000,	Thiết lập mô hình SWAT và tính toán KB1
	Landsat 8	30m	2020	Tính toán KB2 bằng mô hình SWAT
Dữ liệu thổ nhưỡng	Soil Map of the World	1:5.000.000	1995	Thiết lập mô hình SWAT
Dữ liệu lượng mưa	CHIRPS	5.566 m	1986 - 1987,	Khởi động mô hình
			1988 - 1997,	Hiệu chỉnh mô hình
			1998 - 2002, 2020 - 2023	Kiểm định mô hình Tính toán kịch bản
Dữ liệu nhiệt độ	ERA5	11.132 m	1986 - 1987,	Khởi động mô hình
			1988 - 1997,	Hiệu chỉnh mô hình
			1998 - 2002, 2020 - 2023	Kiểm định mô hình Tính toán kịch bản
Dữ liệu thực đo trạm Lai Châu	Dữ liệu lưu lượng		1986 - 1987,	Khởi động mô hình
			1988 - 1997,	Hiệu chỉnh mô hình
			1998 - 2002, 2020 - 2023	Kiểm định mô hình Tính toán kịch bản

2.3. Phương pháp nghiên cứu

2.3.1. Sơ đồ nghiên cứu

Bài báo sử dụng mô hình thủy văn SWAT để đánh giá những thay đổi của dòng chảy do sự thay đổi của lớp phủ trên thượng nguồn lưu vực sông Đà bên ngoài lãnh thổ Việt Nam. Do khu vực biên giới Việt Nam - Trung Quốc trên sông Đà không có trạm đo thủy văn với chuỗi dữ liệu kéo dài nên bài báo xây dựng mô hình đến trạm thủy văn Lai Châu và sử dụng số liệu tại trạm để hiệu chỉnh, kiểm định mô hình. Dữ liệu lớp phủ đầu vào của mô hình được xây dựng từ ảnh vệ tinh Landsat 5 và Landsat 8 lần lượt cho thời điểm năm 2000 (KB1) và năm 2020 (KB2). Dữ liệu lớp phủ sau đó được sử dụng để thiết lập mô hình SWAT và tính

toán các kịch bản. Bộ thông số mô hình sau khi hiệu chỉnh, kiểm định được áp dụng để tính toán dòng chảy trên lưu vực trong khoảng thời gian 2020 - 2023 lần lượt với dữ liệu sử dụng đất KB1 và KB2, từ đó xác định được những biến động dòng chảy do thay đổi lớp phủ tạo ra trong cùng điều kiện thời tiết. Quá trình thực hiện nghiên cứu gồm các bước được thể hiện trong Hình 2.



Hình 2. Sơ đồ nghiên cứu.

2.3.2. Xây dựng dữ liệu lớp phủ

Dữ liệu ảnh vệ tinh Landsat 5 và Landsat 8 được truy cập trên nền tảng GEE mà không cần tải về máy tính, lần lượt theo các nguồn *LANDSAT/LT05/C02/T1_L2* và *LANDSAT/LC08/C02/T1_L2*. Các cảnh ảnh vệ tinh Landsat 5 được thu thập trong năm 2000, tiến hành lọc mây và chồng chập các cảnh ảnh. Tương tự các cảnh ảnh vệ tinh Landsat 8 được thu thập trong năm 2020, tiến hành lọc mây và chồng chập các cảnh ảnh.

GEE là một nền tảng điện toán đám mây cho phép người dùng có thể phân tích hình ảnh vệ tinh có sẵn bằng trình soạn thảo mã mà không cần tải dữ liệu này xuống máy cục bộ [22], giúp xử lý khối lượng dữ liệu lớn, phù hợp cho nhiệm vụ lập bản đồ lớp phủ ở các khu vực rộng lớn. GEE tích hợp nhiều thuật toán học máy, cung cấp công cụ để xây dựng dữ liệu lớp phủ bằng phương pháp phân loại có giám sát. Nhiều nghiên cứu đã sử dụng thuật toán RF để phân loại lớp phủ từ ảnh vệ tinh [13, 23, 24], thuật toán có ưu điểm khi chỉ cần tối ưu hóa ít tham số, giảm sai số từ tập dữ liệu nhiễu, tăng tốc độ xử lý bằng cách chọn các biến quan trọng [25]. Thuật toán RF là thuật toán phân loại tổng hợp tạo ra nhiều cây quyết định, sử dụng một tập hợp con các mẫu đào tạo được chọn ngẫu nhiên. RF tạo ra các phân loại tin cậy bằng cách sử dụng các dự đoán có được từ một tập hợp các cây quyết định. Thuật toán này có thể được sử dụng để lựa chọn và xếp hạng các biến (đặc điểm phổ của điểm ảnh) có khả năng phân loại tốt nhất [26]. Đây là một lợi thế quan trọng vì tính đa chiều cao của dữ liệu ảnh vệ tinh khiến việc lựa chọn các biến có liên quan nhất trở thành một nhiệm vụ tốn thời gian [27] và RF giúp giải quyết vấn đề này.

Các điểm mẫu dùng để đào tạo và kiểm định cho các điểm ảnh trên ảnh vệ tinh được thiết lập trên GEE, thuộc các nhóm lớp phủ: khu vực xây dựng, rừng, cây bụi - cỏ, đất trồng, mặt nước. Các điểm ảnh được phân loại theo các nhóm lớp phủ dựa vào đặc điểm phổ tương đồng với các điểm mẫu. Ngoài các kênh phổ trong ảnh Landsat, các chỉ số viễn thám được bổ sung vào thuộc tính đào tạo mô hình học máy, gồm NDVI (xác định khu vực thực vật), MNDWI (xác định khu vực mặt nước), NDBI (xác định khu vực xây dựng). Thuật toán học

máy RF thiết lập trên GEE giúp xác định mỗi điểm ảnh trên ảnh vệ tinh có đặc điểm phổ (thuộc tính đào tạo) gần tương đồng với nhóm điểm mẫu nào nhất. Sau quá trình phân loại toàn bộ các điểm ảnh được tách thành năm nhóm lớp phủ. Dữ liệu lớp phủ được trích xuất từ GEE để làm dữ liệu đầu vào cho mô hình SWAT.

Kết quả phân loại lớp phủ được đánh giá bằng độ chính xác tổng thể (OA) và hệ số Kappa (K). Độ chính xác tổng thể cho biết có bao nhiêu điểm ảnh được kiểm định là chính xác với điểm mẫu và được thể hiện thông qua phần trăm độ chính xác (công thức 1). Hệ số Kappa được sử dụng để đánh giá độ chính xác của các lớp đối tượng được phân loại (công thức 2).

$$OA = \frac{\text{Số điểm ảnh phân loại chính xác}}{\text{Số điểm ảnh được kiểm định}} \tag{1}$$

$$K = \frac{N \sum_{i=1}^n X_{ii} - \sum_{i=1}^n (X_{i+} - X_{+i})}{N^2 - \sum_{i=1}^n (X_{i+} - X_{+i})} \tag{2}$$

Trong đó n là số cột trong ma trận; X_{ii} là số lượng điểm ảnh trên hàng i và cột i trên đường chéo chính; X_{i+} là tổng điểm ảnh tại hàng i; X_{+i} tổng điểm ảnh tại cột i; N là tổng số điểm ảnh trong ma trận.

2.3.3. Mô hình SWAT

Mô hình SWAT là mô hình ở cấp độ lưu vực, được thiết kế để mô phỏng quá trình thủy văn trên lưu vực và đánh giá tác động của thảm phủ và các thay đổi [28]. Mô hình SWAT cho phép mô hình hóa nhiều quá trình vật lý trên cùng một lưu vực, với khả năng mô phỏng liên tục và kéo dài [29]. Các dữ liệu được đưa vào mô hình gồm địa hình, lớp phủ, thổ nhưỡng, lượng mưa, nhiệt độ. Chu trình thủy văn được mô tả trong mô hình SWAT dựa trên phương trình cân bằng nước (công thức 3).

$$SW_t = SW_0 + \sum_{i=1}^t (R_{\text{day}} - Q_{\text{surf}} - E_a - W_{\text{seep}} - Q_{\text{gw}}) \tag{3}$$

Trong đó SW_t là tổng lượng nước tại cuối thời đoạn tính toán (mm); SW_0 là tổng lượng nước ban đầu tại ngày thứ i (mm); t là thời gian (ngày); R_{day} là số tổng lượng mưa tại ngày thứ i (mm); Q_{surf} là tổng lượng nước mặt của ngày thứ i (mm); E_a là lượng bốc thoát hơi tại ngày thứ i (mm); W_{seep} là lượng nước đi vào tầng ngầm tại ngày thứ i (mm); Q_{gw} là số lượng nước hồi quy tại ngày thứ i (mm).

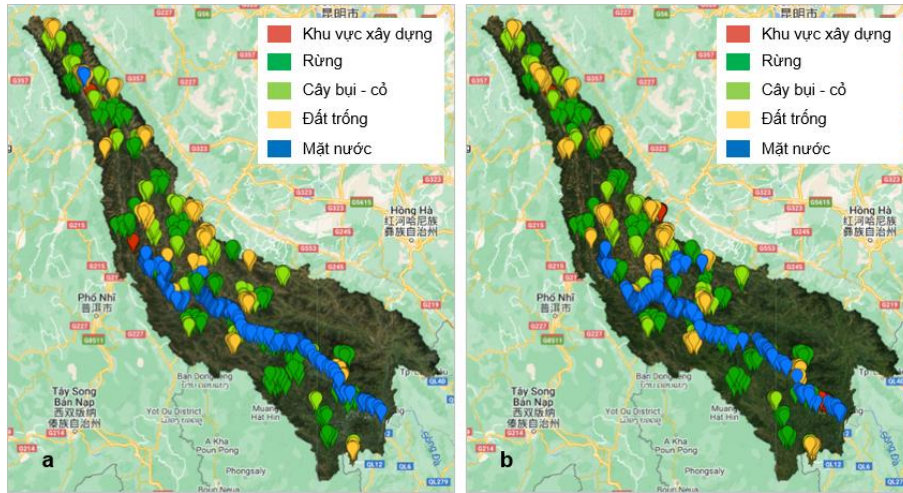
3. Kết quả

3.1. Dữ liệu lớp phủ

Dữ liệu vệ tinh Landsat 5, 8 được truy cập từ GEE, dựa vào các kênh QA_PIXEL, QA_RADSAT của ảnh vệ tinh để loại bỏ mây. Các cảnh ảnh sau khi loại bỏ mây được chồng chập bao phủ toàn bộ lưu vực nghiên cứu. Các điểm mẫu được thiết lập trên GEE để hiệu chỉnh, kiểm định kết quả phân loại bao gồm năm nhóm: khu vực xây dựng, rừng, cây bụi - cỏ, đất trống, mặt nước (Hình 3). Ảnh vệ tinh Landsat 5 được thiết lập 785 điểm mẫu, ảnh vệ tinh Landsat 8 được thiết lập 840 điểm mẫu (Bảng 2). Trong số các điểm mẫu này, 70% số điểm dùng để đào tạo dữ liệu, 30% số điểm dùng để kiểm định kết quả.

Bảng 2. Số lượng điểm mẫu của các nhóm lớp phủ.

Đối tượng thảm phủ	Landsat 5	Landsat 8
Khu vực xây dựng	45	100
Rừng	200	200
Cây bụi - cỏ	180	180
Đất trống	180	180
Mặt nước	180	180
Tổng	785	840



Hình 3. Ảnh vệ tinh được tiền xử lý và thiết lập các điểm mẫu: (a) Landsat 5, (b) Landsat 8.

Kết quả quá trình phân loại lớp phủ được đánh giá thông qua ma trận sai số, độ chính xác tổng thể và hệ số Kappa như Bảng 3, Bảng 4.

Bảng 3. Ma trận sai số của quá trình phân loại lớp phủ năm 2000.

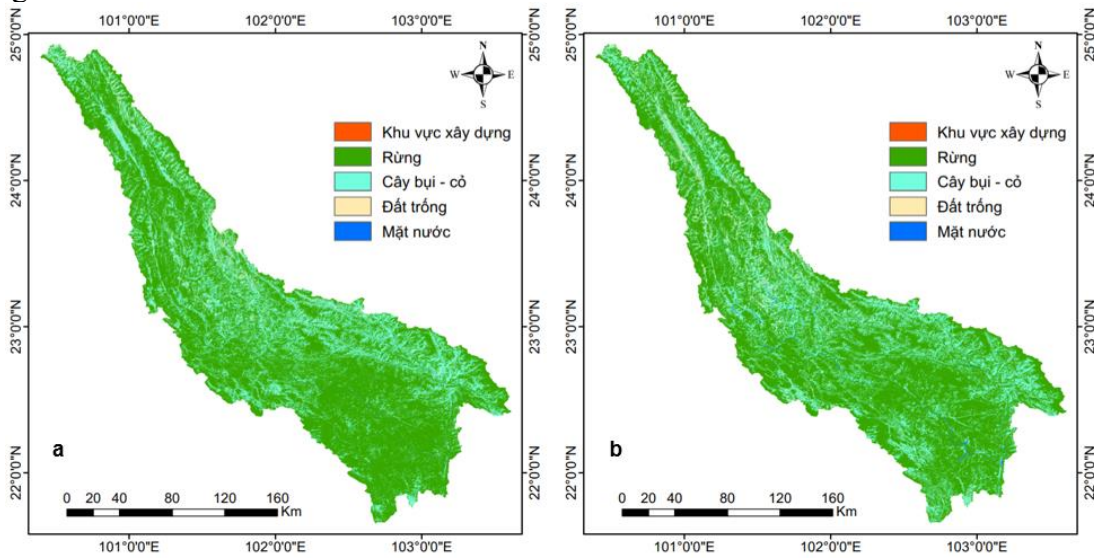
	Khu vực xây dựng	Rừng	Cây bụi - cỏ	Đất trống	Mặt nước	Tổng	Độ chính xác
Khu vực xây dựng	7	0	2	1	1	11	63,6%
Rừng	0	54	4	0	0	58	93,1%
Cây bụi - cỏ	0	2	48	3	0	53	90,6%
Đất trống	0	0	7	46	1	54	85,2%
Mặt nước	0	0	0	2	53	55	96,4%
Tổng	7	56	61	52	55	231	
Độ chính xác	100%	96,4%	78,7%	88,5%	96,4%		
Độ chính xác tổng thể				90%			
Hệ số Kappa				0,87			

Bảng 4. Ma trận sai số của quá trình phân loại lớp phủ năm 2020.

	Khu vực xây dựng	Rừng	Cây bụi - cỏ	Đất trống	Mặt nước	Tổng	Độ chính xác
Khu vực xây dựng	31	0	0	3	0	34	91,2%
Rừng	1	52	3	0	3	59	88,1%
Cây bụi - cỏ	0	0	48	3	0	51	94,1%
Đất trống	0	0	2	51	0	53	96,2%
Mặt nước	0	1	1	1	68	71	95,8%
Tổng	32	53	54	58	71	268	
Độ chính xác	96,9%	98,1%	88,9%	87,9%	95,8%		
Độ chính xác tổng thể				93%			
Hệ số Kappa				0,91			

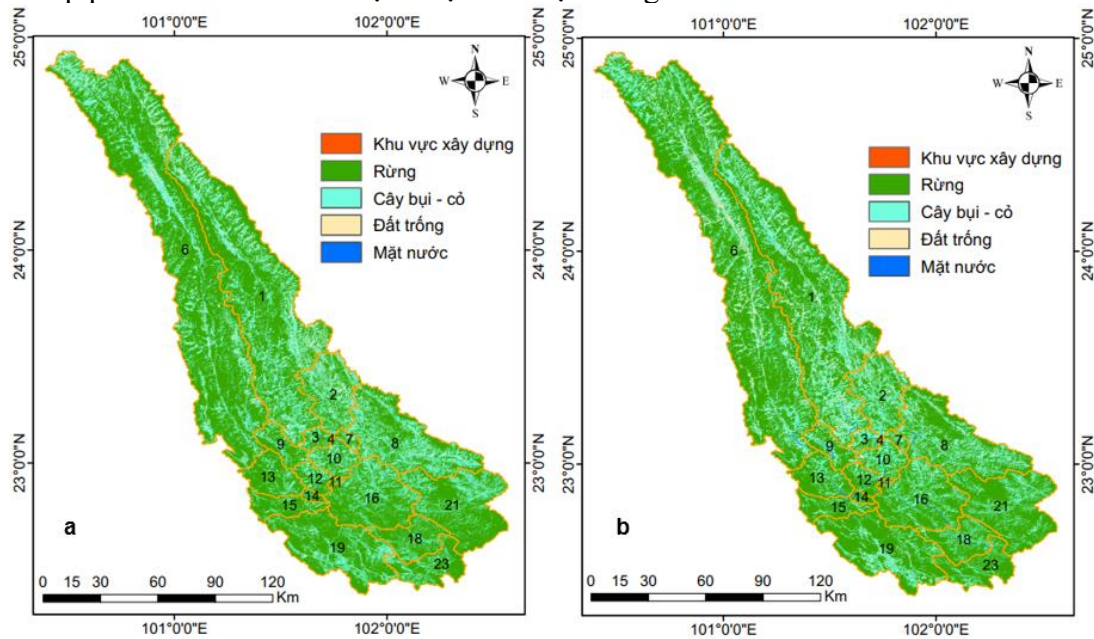
Hai trường hợp phân loại lớp phủ từ mô hình học máy bằng thuật toán RF đều đạt kết quả tốt với độ chính xác tổng thể cho dữ liệu năm 2000 và 2020 lần lượt là 90% và 93% (Hình 4). Hệ số Kappa cũng đạt giá trị cao với dữ liệu năm 2000 và năm 2020 lần lượt là 0,87 và 0,91. Dữ liệu lớp phủ năm 2000 được phân loại đạt độ chính xác cao với rừng, đất trống và mặt nước (trên 85%) và thấp nhất với khu vực xây dựng (63,64%). Dữ liệu lớp phủ năm 2020 được phân loại đạt độ chính xác cao nhất với khu vực xây dựng và mặt nước (trên 91%), các loại lớp phủ còn lại cũng đạt độ chính xác cao (trên 87%). Dựa vào độ chính xác tổng thể và hệ số Kappa có thể thấy kết quả phân loại lớp phủ từ ảnh vệ tinh Landsat 5 và

Landsat 8 bằng thuật toán RF trên GEE đảm bảo độ tin cậy. Dữ liệu lớp phủ năm 2000 được sử dụng để thiết lập mô hình SWAT và tính toán KB1, dữ liệu lớp phủ năm 2020 được sử dụng để tính toán KB2.



Hình 4. Dữ liệu lớp phủ năm 2000 (a) và năm 2020 (b).

Dữ liệu được xây dựng đến trạm thủy văn Lai Châu với 35 tiểu lưu vực, tuy vậy nghiên cứu này tập trung vào khu vực thượng lưu sông Đà bên ngoài lãnh thổ Việt Nam nên chỉ tính toán với 20/35 tiểu lưu vực trên (tên các tiểu lưu vực được đánh số theo mô hình SWAT). Vị trí và lớp phủ của các tiểu lưu vực được thể hiện trong Hình 5.



Hình 5. Dữ liệu lớp phủ theo tiểu lưu vực trên thượng lưu sông Đà.

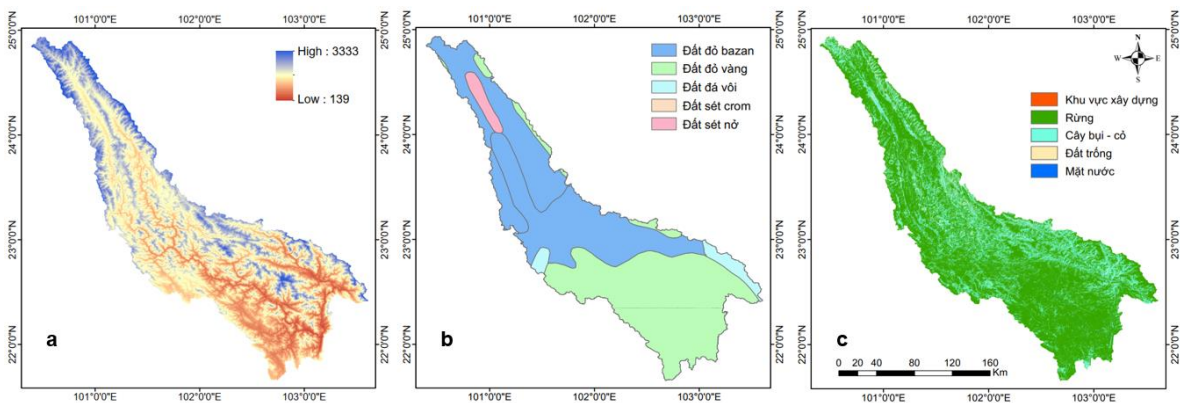
Dữ liệu sử dụng đất được xây dựng cho thấy có sự thay đổi giữa thời điểm năm 2000 (KB1) và năm 2020 (KB2) trên 20 tiểu lưu vực thượng lưu sông Đà bên ngoài lãnh thổ (Bảng 5). Diện tích khu vực xây dựng, rừng, đất trống và mặt nước tăng lên trong khi diện tích cây bụi - cỏ giảm. Diện tích khu vực xây dựng có tỷ lệ tăng lớn nhất với 866,92% (toàn bộ 20 tiểu lưu vực đều tăng), rừng tăng 4,26% (17 tiểu lưu vực tăng, 3 tiểu lưu vực giảm), đất trống tăng 3,39% (5 tiểu lưu vực tăng, 15 tiểu lưu vực giảm), mặt nước tăng 63,89% (14 tiểu lưu vực tăng, 6 tiểu lưu vực giảm), diện tích cây bụi - cỏ giảm 11,60% (3 tiểu lưu vực tăng, 17 tiểu lưu vực giảm).

Bảng 5. So sánh diện tích sử dụng đất lưu vực thượng lưu sông Đà KB2 với KB1.

STT	Tên lưu vực	Khu vực xây dựng		Rừng		Cây bụi - cỏ		Đất trồng		Mặt nước	
		Diện tích	%	Diện tích	%	Diện tích	%	Diện tích	%	Diện tích	%
		tăng/ giảm (+/-) (km ²)	tăng/ giảm (+/-)	tăng/ giảm (+/-) (km ²)	tăng/ giảm (+/-)	tăng/ giảm (+/-) (km ²)	tăng/ giảm (+/-)	tăng/ giảm (+/-) (km ²)	tăng/ giảm (+/-)	tăng/ giảm (+/-) (km ²)	tăng/ giảm (+/-)
1	1	13,66	3473,4	-45,77	-1,8	-39,96	-3,1	59,62	27,6	12,45	271,1
2	2	2,64	597,9	21,38	5,3	-25,14	-7,3	0,09	0,2	1,03	68,5
3	3	0,49	5190,9	16,60	33,4	-11,80	-16,2	-7,21	-56,2	1,93	556,4
4	4	0,02	300,0	-0,01	-0,6	0,19	14,3	-0,30	-85,6	0,09	100,0
5	5	0,0009	100,0	0,06	34,3	0,01	3,1	-0,08	-98,9	0,01	100,0
6	6	23,12	876,0	-176,10	-4,0	1,63	0,1	144,14	78,6	7,21	49,3
7	7	0,08	370,8	15,75	51,2	-10,24	-30,0	-5,57	-80,6	-0,01	-13,0
8	8	3,85	470,1	181,08	20,8	-118,69	-20,5	-70,06	-54,7	3,82	56,6
9	9	0,93	6347,1	26,53	15,1	-28,32	-25,2	-3,77	-25,8	4,64	1958,0
10	10	0,49	3000,0	55,72	44,1	-39,66	-23,5	-16,64	-62,2	0,09	3,9
11	11	0,0017	15,4	3,27	23,9	-2,65	-34,9	-0,70	-76,9	0,08	7,6
12	12	0,21	24700	39,00	34,0	-36,79	-44,3	-2,46	-41,3	0,05	2,3
13	13	0,43	3340,0	39,07	12,2	-38,33	-28,6	-1,04	-15,1	-0,13	-3,0
14	14	0,0043	100,0	9,72	27,6	-9,18	-53,2	-0,39	-43,3	-0,16	-6,0
15	15	0,37	14500	24,72	9,4	-26,47	-32,8	1,49	35,3	-0,12	-7,5
16	16	1,07	202,6	113,52	15,5	-83,13	-22,4	-33,08	-64,4	1,62	23,5
17	18	0,77	178,2	35,93	10,3	-30,52	-20,5	-6,17	-72,3	-0,02	-0,5
18	19	2,35	601,5	47,42	4,4	-56,00	-15,6	4,32	13,2	1,91	402,9
19	21	3,61	865,0	80,19	7,8	-56,24	-16,9	-29,21	-79,1	1,65	76,0
20	23	0,24	220,8	65,56	17,2	-59,80	-40,4	-5,71	-62,8	-0,29	-32,2

3.2. Hiệu chỉnh, kiểm định mô hình SWAT

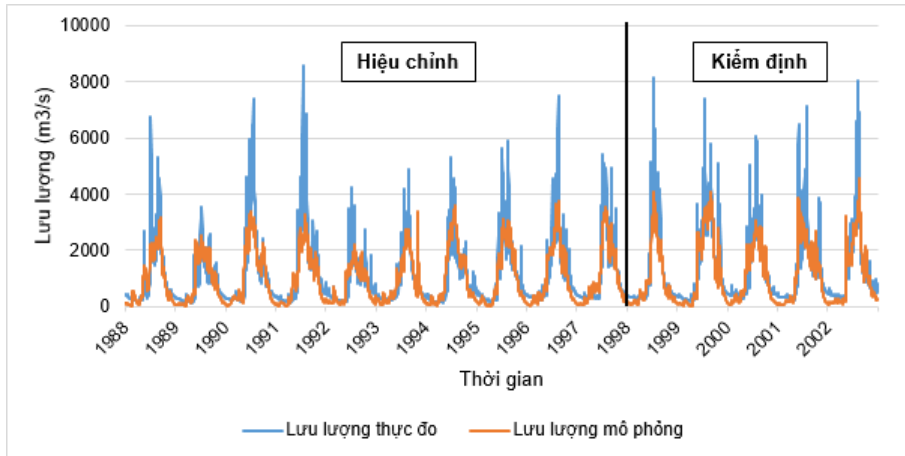
Do không có số liệu đo đạc trên sông Đà tại biên giới Việt Nam nên bài báo thực hiện mô phỏng dòng chảy đến trạm thủy văn Lai Châu để sử dụng số liệu tại trạm hiệu chỉnh, kiểm định mô hình. Bộ thông số từ mô hình đã hiệu chỉnh, kiểm định được áp dụng để tính toán cho các tiểu lưu vực ở thượng nguồn sông Đà bên ngoài lãnh thổ. Dữ liệu phục vụ hiệu chỉnh và kiểm định mô hình được trình bày ở Bảng 1. Mô hình sử dụng số liệu trong giai đoạn 1986-2002 trước khi các hồ chứa trên lưu vực sông Đà xây dựng. Quá trình chạy và hiệu chỉnh, kiểm định được chia thành ba giai đoạn, trong đó giai đoạn 1986-1987 được dùng để khởi động mô hình, giai đoạn 1988 - 1997 được dùng để hiệu chỉnh các tham số, giai đoạn 1998-2002 được dùng để kiểm định mô hình. Chất lượng mô phỏng của mô hình được đánh giá bằng việc so sánh kết quả dòng chảy với số liệu thực đo tại trạm thủy văn Lai Châu thông qua các chỉ tiêu: hệ số hiệu quả Nash & Sutcliffe (NSE), hệ số xác định (R²) và độ thiên lệch (PBIAS).



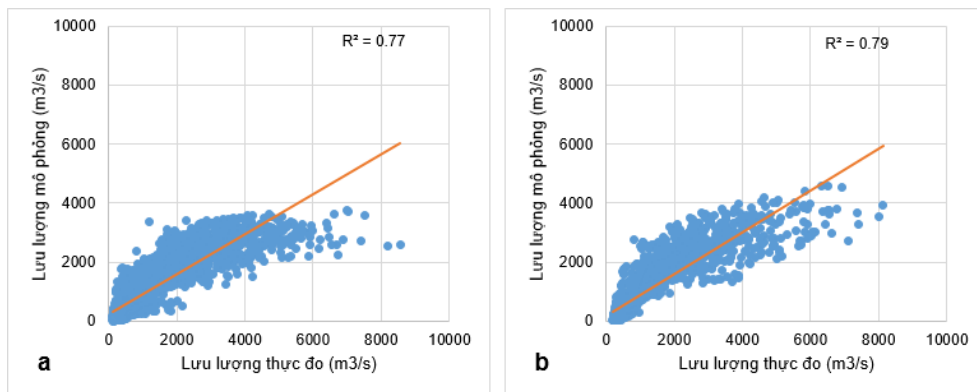
Hình 6. Dữ liệu địa hình (a), dữ liệu thổ nhưỡng (b), dữ liệu lớp phủ năm 2000 (c).

Bài báo sử dụng 11 thông số để hiệu chỉnh mô hình, các thông số này được điều chỉnh giá trị và phân tích độ nhạy dựa trên thuật toán SUFI-2 (có sẵn trên công cụ SWAT-CUP) cho đến khi đạt được mô phỏng tối ưu. Trong các thông số này, EPCO, OV_N, GWQMN, CN2 là các thông số có ảnh hưởng nhất đến quá trình hiệu chỉnh dòng chảy, cho thấy yếu tố thấm phủ bề mặt ảnh hưởng nhiều đến mô hình, tiếp đến là đặc điểm tầng chứa nước nông.

Dòng chảy mô phỏng được so sánh với số liệu thực đo theo ngày tại trạm Lai Châu như Hình 7, Hình 8. Nhìn chung, đường quá trình lưu lượng mô phỏng từ mô hình tương đồng với đường quá trình lưu lượng thực tế. Tuy nhiên, mô hình chưa mô phỏng chính xác đỉnh lũ, quá trình lũ mô phỏng thấp hơn so với thực tế, điều này có thể do dữ liệu mưa vệ tinh trong thời kỳ lũ chưa tốt.



Hình 7. Đường quá trình lưu lượng thực đo và mô phỏng từ mô hình SWAT giai đoạn hiệu chỉnh và kiểm định.



Hình 8. Biểu đồ phân tán so sánh lưu lượng thực đo và mô phỏng giai đoạn hiệu chỉnh (a) và kiểm định (b).

Hiệu suất mô phỏng của mô hình đánh giá thông qua các chỉ tiêu NSE, R^2 , PBIAS thể hiện trong Bảng 6.

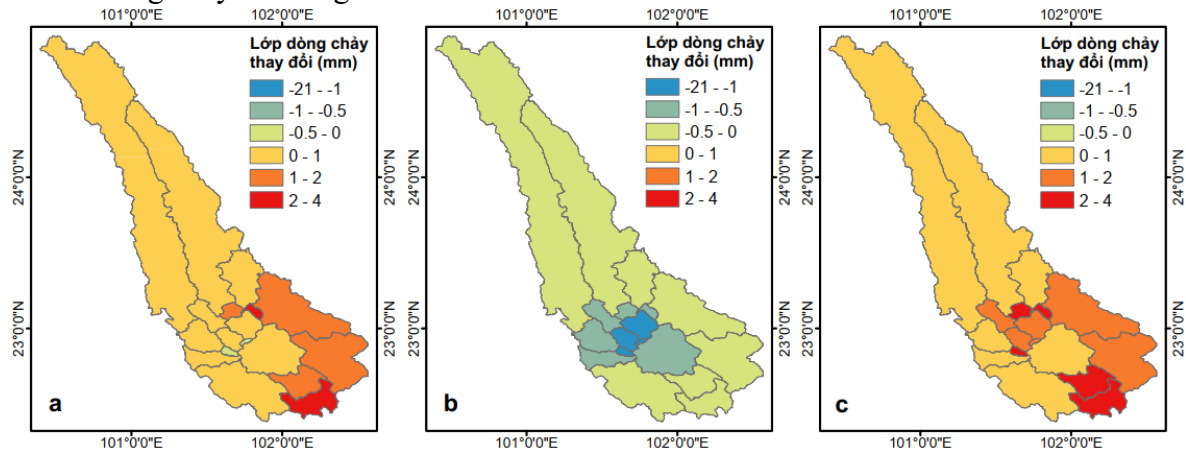
Bảng 6. Các chỉ số đánh giá khả năng mô phỏng của mô hình.

Giai đoạn	NSE	R^2	PBIAS
Hiệu chỉnh	0,75	0,77	12,71
Kiểm định	0,76	0,79	14,09

Các hệ số NSE, R^2 , PBIAS của mô hình đạt kết quả tốt trong giai đoạn hiệu chỉnh và giai đoạn kiểm định. Qua đó cho thấy mô hình đã xây dựng có khả năng mô phỏng dòng chảy trên lưu vực. Bộ thông số mô hình sau khi hiệu chỉnh, kiểm định được áp dụng để đánh giá tác động của các trường hợp lớp phủ khác nhau đến dòng chảy trên lưu vực thượng lưu sông Đà.

3.3. Sự thay đổi dòng chảy theo không gian

Dòng chảy mà mỗi tiểu lưu vực đóng góp vào sông chính được đánh giá thông qua lớp dòng chảy trên mỗi tiểu lưu vực. Sự thay đổi giữa lớp dòng chảy KB2 và lớp dòng chảy KB1 trên 20 tiểu lưu vực được thể hiện trong Hình 9. Lớp phủ trong KB2 thay đổi so với KB1, diện tích khu vực xây dựng, rừng, đất trồng và mặt nước tăng lên còn diện tích cây bụi - cỏ giảm làm cho dòng chảy mặt có sự thay đổi. Tuy các tiểu lưu vực có sự thay đổi dòng chảy không nhiều nhưng cũng đã cho thấy tác động của lớp phủ chủ yếu làm tăng dòng chảy mặt trên lưu vực. Trong mùa khô, hầu hết các tiểu lưu vực giảm dòng chảy trong KB2 so với KB1, trong khi vào mùa mưa hầu hết các tiểu lưu vực tăng dòng chảy. Mùa khô lượng mưa ít, trong khi diện tích khu vực xây dựng, rừng và mặt nước tăng làm cho lượng bốc thoát hơi lớn hơn, lượng nước đóng góp cho dòng chảy giảm. Mùa mưa lượng mưa nhiều, diện tích cây bụi - cỏ giảm nhiều cùng với diện tích khu vực xây dựng tăng làm cho lớp nước bề mặt tăng lên. Trong cả năm, dòng chảy tăng lên ở hầu hết các tiểu lưu vực trong KB2 so với KB1, ngoại trừ các tiểu lưu vực 4, 5, 11, 14; các tiểu lưu vực 3, 7, 18, 21, 23 có mức tăng lớn nhất. Dòng chảy giảm ở các tiểu lưu vực có diện tích nhỏ nhất, đối với tiểu lưu vực có diện tích lớn thì dòng chảy đều tăng.



Hình 9. Sự thay đổi lớp dòng chảy trên các tiểu lưu vực KB2 so với KB1: (a) cả năm, (b) mùa khô, (c) mùa mưa.

Trong các tiểu lưu vực có mức tăng dòng chảy nhiều nhất, tiểu lưu vực 3 tăng diện tích khu vực xây dựng 5190,9%, mặt nước 556,4% và giảm diện tích cây bụi - cỏ 16,2%. Tiểu lưu vực 7 tăng diện tích khu vực xây dựng 370,8% và giảm diện tích cây bụi - cỏ 30,0%. Tiểu lưu vực 18 tăng diện tích khu vực xây dựng 178,2% và giảm diện tích cây bụi - cỏ 20,5%. Tiểu lưu vực 21 tăng diện tích khu vực xây dựng 865,0%, mặt nước 76,0% và giảm diện tích cây bụi - cỏ 16,9%. Tiểu lưu vực 23 tăng diện tích khu vực xây dựng 220,8% và giảm diện tích cây bụi - cỏ 40,4%. Các tiểu lưu vực giảm dòng chảy gồm 4, 5, 11, 14 đều giảm diện tích đất trồng với tỷ lệ giảm khá lớn so với các tiểu lưu vực còn lại (đặc biệt là tiểu lưu vực 4, 5), lần lượt giảm 85,6%, 98,9%, 76,9% và 43,3%. Dòng chảy trên các tiểu lưu vực này đều ảnh hưởng bởi khu vực xây dựng, mặt nước, đất trồng, cây bụi - cỏ. Diện tích khu vực xây dựng, mặt nước tăng, trong khi diện tích cây bụi - cỏ giảm có xu hướng làm tăng dòng chảy và diện tích đất trồng giảm có xu hướng làm giảm dòng chảy.

3.4. Sự thay đổi dòng chảy theo thời gian

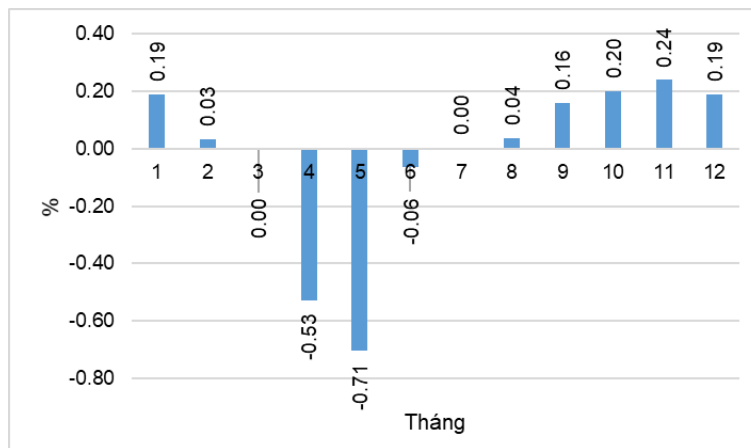
Dòng chảy các tiểu lưu vực đóng góp có sự thay đổi theo các thời điểm trong năm. Bảng 7 và Hình 10 cho thấy sự thay đổi theo tháng và theo mùa trong năm trên lưu vực nghiên cứu.

Lớp dòng chảy trên lưu vực có sự thay đổi ở KB2 so với KB1, tăng lên ở hầu hết các tháng nhưng giảm vào các tháng 4, 5, 6. Mùa khô lớp dòng chảy trên toàn lưu vực giảm 0,15%, mùa mưa tăng 0,08%. Sự chênh lệch dòng chảy ở KB2 so với KB1 tăng giảm theo

mùa, dòng chảy tăng lên từ mùa khô sang mùa mưa. Trong mùa mưa, dòng chảy chênh lệch giữa KB2 so với KB1 tăng dần từ tháng 6 (-0,06%) đến lớn nhất vào tháng 11 (0,24%). Trong mùa khô, dòng chảy chênh lệch giữa KB2 so với KB1 giảm từ tháng 12 (0,19%) và thấp nhất vào tháng 5 (-0,71%). Về tổng thể, lượng nước trung bình năm lưu vực đóng góp trong KB2 lớn hơn KB1 (tăng 0,04%).

Bảng 7. Lớp dòng chảy theo tháng, mùa trên lưu vực thượng lưu sông Đà.

Tháng/ mùa	Lớp dòng chảy KB1 (mm)	Lớp dòng chảy KB2 (mm)	KB2 so với KB1 (%)
1	25,33	25,38	0,19
2	25,29	25,30	0,03
3	21,52	21,52	0,00
4	19,71	19,61	-0,53
5	46,72	46,39	-0,71
6	115,76	115,69	-0,06
7	169,98	169,98	0,00
8	231,41	231,50	0,04
9	180,97	181,26	0,16
10	130,76	131,02	0,20
11	74,14	74,32	0,24
12	47,79	47,88	0,19
Mùa khô	186,37	186,08	-0,15
Mùa mưa	903,03	903,78	0,08



Hình 10. Sự thay đổi lớp dòng chảy trên lưu vực theo tháng KB2 so với KB1.

3.5. Thảo luận

Dữ liệu lớp phủ được xây dựng trên dữ liệu ảnh vệ tinh Landsat 5 và Landsat 8 với độ phân giải 30 m, tuy chưa phải là độ phân giải cao nhưng đây là dữ liệu miễn phí sẵn có, phù hợp với yêu cầu các thời điểm nghiên cứu khác nhau. Một số dữ liệu lớp phủ được xây dựng bao phủ gần như toàn cầu nhưng độ phân giải quá lớn hoặc không có dữ liệu cho các năm quá khứ nên không phù hợp với nghiên cứu. Mặc dù ảnh vệ tinh đã được tiền xử lý nhưng do chất lượng một số cảnh ảnh vệ tinh bị ảnh hưởng nhiều bởi khí quyển, mây khiến quá trình thu nhận quang phổ có những sai khác nhất định so với thực tế, điều này gây ra việc phân loại một số điểm ảnh không đúng.

Mô hình SWAT được thiết lập mô phỏng xu thế của dòng chảy phù hợp với số liệu thực đo, tuy nhiên chưa mô phỏng tốt thời kỳ lũ. Trong thời kỳ lũ, mô phỏng của mô hình luôn thấp hơn so với thực tế, dẫn đến giảm tổng lượng dòng chảy. Một số nguyên nhân có thể ảnh hưởng đến sự mô phỏng chưa tốt của mô hình gồm dữ liệu thời tiết từ vệ tinh, dữ liệu lớp phủ có sai khác so với thực tế, hay do đường cong dòng chảy SCS được thiết lập trong mô hình. Do nghiên cứu thực hiện việc so sánh dòng chảy tương ứng với hai trường hợp sử dụng

đất trong điều kiện thời tiết như nhau nên việc không mô phỏng tốt thời kỳ lũ không ảnh hưởng nhiều đến kết quả nghiên cứu.

Dòng chảy trên các tiểu lưu vực tuy thay đổi không nhiều với trung bình toàn lưu vực là 0,04% nhưng nghiên cứu này chứng tỏ rằng phương pháp nghiên cứu có khả năng đánh giá được tác động của sự thay đổi lớp phủ đến dòng chảy mặt trên lưu vực theo thời gian. Các loại sử dụng đất trên mỗi tiểu lưu vực có sự tăng, giảm diện tích khác nhau, dòng chảy thay đổi tùy thuộc vào mức độ tác động tổng thể của các loại sử dụng đất trên mỗi tiểu lưu vực. Điều này là một nguyên nhân dẫn đến việc diện tích các loại lớp phủ có thay đổi nhưng dòng chảy không thay đổi nhiều. Việc lớp phủ thay đổi trên các loại thổ nhưỡng khác nhau cũng dẫn đến mức tăng, giảm dòng chảy khác nhau, chẳng hạn như trên lưu vực nghiên cứu thì đất đỏ bazan thường thấm nước tốt hơn so với đất đỏ vàng.

4. Kết luận

Bài báo xây dựng dữ liệu lớp phủ cho lưu vực thượng nguồn sông Đà bên ngoài lãnh thổ Việt Nam dựa vào thuật toán phân loại học máy RF trên nền tảng GEE. Dữ liệu lớp phủ được xây dựng đảm bảo độ tin cậy dựa vào việc kiểm định với các điểm mẫu, qua đó cho thấy phương pháp này có thể được áp dụng cho khu vực không có sẵn dữ liệu lớp phủ. Do không có số liệu thực đo của khu vực nghiên cứu, mô hình SWAT được thiết lập từ đầu vào là các dữ liệu vệ tinh để mô phỏng thủy văn lưu vực. Quá trình hiệu chỉnh, kiểm định mô hình cho dòng chảy đạt kết quả cao, trong đó hệ số NSE và R^2 từ 0,75 trở lên. Kết quả hiệu chỉnh và kiểm định chứng tỏ hiệu suất mô phỏng của mô hình tốt, phù hợp để áp dụng đánh giá tác động của lớp phủ đến dòng chảy trên lưu vực nghiên cứu. Dòng chảy trên lưu vực trong trường hợp lớp phủ năm 2020 thay đổi so với trường hợp lớp phủ năm 2000 cả về không gian và thời gian. Dòng chảy trên phần lớn các tiểu lưu vực tăng lên với mức tăng trung bình của cả lưu vực là 0,04%. Mùa khô dòng chảy trên toàn lưu vực giảm 0,15%, mùa mưa tăng 0,08%. Kết quả này có thể đóng góp vào những nghiên cứu tiếp theo với mục tiêu đánh giá những thay đổi thủy văn trên lưu vực thượng nguồn sông Đà bên ngoài lãnh thổ và hỗ trợ công tác quản lý tài nguyên nước trên sông Đà thuộc Việt Nam. Việc sử dụng dữ liệu vệ tinh có độ phân giải chưa cao dẫn đến sẽ có sai số trong tính toán, tuy vậy chất lượng có thể cải thiện nhiều hơn với các nguồn vệ tinh thương mại. Phương pháp nghiên cứu có thể được áp dụng ở các khu vực khác để đánh giá sự thay đổi dòng chảy trên lưu vực do thay đổi sử dụng đất, đặc biệt là đối với các thời điểm nghiên cứu trong quá khứ, không có sẵn dữ liệu sử dụng đất.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: T.M.C., T.A.P.; Lựa chọn phương pháp nghiên cứu: T.M.C., T.A.P., T.V.T.; Xử lý số liệu: T.M.C., L.M.A., B.H.L.; Viết bản thảo bài báo: T.M.C., L.M.A., B.H.L., T.V.T.; Chỉnh sửa bài báo: T.M.C., T.A.P.

Lời cảm ơn: Chúng tôi xin cảm ơn đề tài “Nghiên cứu đánh giá tác động của hệ thống hồ chứa và sử dụng đất ngoài lãnh thổ Việt Nam trên lưu vực sông Đà ảnh hưởng đến biến đổi dòng chảy và hàm lượng phù sa bùn cát vào Việt Nam”, Mã số: TNMT.2023.02.30 đã hỗ trợ để thực hiện nghiên cứu này.

Lời cam đoan: Tập thể tác giả cam đoan bài báo này là công trình nghiên cứu của tập thể tác giả, chưa được công bố ở đâu, không được sao chép từ những nghiên cứu trước đây, không có sự tranh chấp lợi ích trong nhóm tác giả.

Tài liệu tham khảo

1. Linh, B.H.; Phương, T.A. Đánh giá ảnh hưởng của các hồ chứa đến dòng chảy trên sông Đà. *Tạp chí Khí tượng Thủy văn* 2021, 731, 97–107.
2. Mensah, J.K.; Ofosu, E.A.; Yidana, S.M.; Akpoti, K.; Kabo-bah, A.T. Integrated modeling of hydrological processes and groundwater recharge based on land use land cover, and climate changes: A systematic review. *Environ. Adv.* 2022, 8, 100224.

3. de Oliveira Serrão, E.A.; Silva, M.T.; Ferreira, T.R.; de Ataíde, L.C.P.; dos Santos, C.A.; de Lima, A.M.M.; ...Gomes, D.J.C. Impacts of land use and land cover changes on hydrological processes and sediment yield determined using the SWAT model. *Int. J. Sediment Res.* **2022**, *37*(1), 54–69.
4. Alawi, S.A.; Özkul, S. Evaluation of land use/land cover datasets in hydrological modelling using the SWAT model. *H2Open J.* **2023**, *6*(1), 63–74.
5. Komori, D.; Rangsiwanichpong, P.; Inoue, N.; Ono, K.; Watanabe, S.; Kazama, S. Distributed probability of slope failure in Thailand under climate change. *Clim. Risk Manage.* **2018**, *20*, 126–137.
6. Gumindoga, W.; Rwasoka, D.T.; Nhapi, I.; Dube, T. Ungauged runoff simulation in Upper Manyame Catchment, Zimbabwe: Application of the HEC-HMS model. *Phys. Chem. Earth. A/B/C/* **2017**, *100*, 371–382.
7. Dinku, T.; Ceccato, P.; Connor, S.J. Challenges of satellite rainfall estimation over mountainous and arid parts of east Africa. *Int. J. Remote Sens.* **2011**, *32*(21), 5965–5979.
8. Sheffield, J.; Wood, E.F.; Pan, M.; Beck, H.; Coccia, G.; Serrat-Capdevila, A.; Verbist, K.J.W.R.R. Satellite remote sensing for water resources management: Potential for supporting sustainable development in data-poor regions. *Water Resour. Res.* **2018**, *54*(12), 9724–9758.
9. Jiang, D.; Wang, K. The role of satellite-based remote sensing in improving simulated streamflow: A review. *Water J.* **2019**, *11*(8), 1615.
10. Steinhausen, M.J.; Wagner, P.D.; Narasimhan, B.; Waske, B. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 595–604.
11. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; ... Zhu, Z. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172.
12. Wulder, M.A.; White, J.C.; Loveland, T.R.; Woodcock, C.E.; Belward, A.S.; Cohen, W.B.; ... Roy, D.P. The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* **2016**, *185*, 271–283.
13. Thảo, V.T.P.; Sengchanh, S. 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. *Tap chí Khoa học Đo đạc và Bản đồ* **2022**, *52*, 26–35.
14. Nasiri, V.; Deljouei, A.; Moradi, F.; Sadeghi, S.M.M.; Borz, S.A. Land use and land cover mapping using Sentinel-2, Landsat-8 Satellite Images, and Google Earth Engine: A comparison of two composition methods. *Remote Sens.* **2022**, *14*(9), 1977.
15. Zafar, Z.; Zubair, M.; Zha, Y.; Fahd, S.; Nadeem, A.A. Performance assessment of machine learning algorithms for mapping of land use/land cover using remote sensing data. *Egypt. J. Remote Sens. Space Sci.* **2024**, *27*(2), 216–226.
16. Chirachawala, C.; Shrestha, S.; Babel, M.S.; Viridis, S.G.; Wichakul, S. Evaluation of global land use/land cover products for hydrologic simulation in the Upper Yom River Basin, Thailand. *Sci. Total Environ.* **2020**, *708*, 135148.
17. Lee, J.; Kim, J.; Jang, W.S.; Lim, K.J.; Engel, B.A. Assessment of baseflow estimates considering recession characteristics in SWAT. *Water J.* **2018**, *10*(4), 371.
18. Asante, K.; Leh, M.D.; Cothren, J.D.; Luzio, M.D.; Brahana, J.V. Effects of land-use land-cover data resolution and classification methods on SWAT model flow predictive reliability. *Int. J. Hydrol. Sci. Technol.* **2017**, *7*(1), 39–62.
19. Aloui, S.; Mazzoni, A.; Elomri, A.; Aouissi, J.; Boufekane, A.; Zghibi, A. A review of Soil and Water Assessment Tool (SWAT) studies of Mediterranean catchments: Applications, feasibility, and future directions. *J. Environ. Manage.* **2023**, *326*, 116799.

20. Liu, Z.; Rong, L.; Wei, W. Impacts of land use/cover change on water balance by using the SWAT model in a typical loess hilly watershed of China. *Geogr. Sustainability* **2023**, 4(1), 19–28.
21. Abuhay, W.; Gashaw, T.; Tsegaye, L. Assessing impacts of land use/land cover changes on the hydrology of Upper Gilgel Abbay watershed using the SWAT model. *J. Agric. Food Res.* **2023**, 12, 100535.
22. Tamiminia, H.; Salehi, B.; Mahdianpari, M.; Quackenbush, L.; Adeli, S.; Brisco, B. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS J. Photogramm. Remote Sens.* **2020**, 164, 152–170.
23. Maxwell, A.E.; Strager, M.P.; Warner, T.A.; Ramezan, C.A.; Morgan, A.N.; Pauley, C.E. Large-area, high spatial resolution land cover mapping using random forests, GEOBIA, and NAIP orthophotography: Findings and recommendations. *Remote Sens.* **2019**, 11(12), 1409.
24. Maxwell, A.E.; Strager, M.P.; Warner, T.A.; Ramezan, C.A.; Morgan, A.N.; Pauley, C.E. Large-area, high spatial resolution land cover mapping using random forests, GEOBIA, and NAIP orthophotography: Findings and recommendations. *Remote Sens.* **2019**, 11(12), 1409.
25. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, 39(9), 2784–2817.
26. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, 114, 24–31.
27. Körting, T.S.; Fonseca, L.M.G.; Câmara, G. GeoDMA-Geographic data mining analyst. *Comput. Geosci.* **2013**, 57, 133–145.
28. Arnold, J.G.; Srinivasan, R.; Muttiah, R.S.; Williams, J.R. Large area hydrologic modeling and assessment part I: model development 1. *JAWRA*, **1998**, 34(1), 73–89.
29. Tuppad, P.; Douglas-Mankin, K.R.; Lee, T.; Srinivasan, R.; Arnold, J.G. Soil and water assessment tool (SWAT) hydrologic/water quality model: Extended capability and wider adoption. *Trans. ASABE* **2011**, 54(5), 1677–1684.

Research on Mapping Land Cover Using Machine Learning Technology for Evaluating the Impact of Land Cover Changes on Flow Regime in the Da River Basin Outside Vietnam

Tran Manh Cuong^{1*}, Tran Anh Phuong¹, Le Manh An², Bui Huyen Linh¹, Tran Van Tu¹

¹ Water Resources Institute, Ministry of Natural Resources Environment; manhcuongkt11@gmail.com; phuongtran.monre@gmail.com; linhb.dctv@gmail.com; trantu.monre@gmail.com

² Hydrometeorological Technology Center, Viet Nam Meteorological and Hydrological Administration, Ministry of Natural Resources Environment; lemanhan82@gmail.com

Abstract: In recent years, land cover in the Da River basin outside Vietnam has undergone significant changes, impacting on water flow into Vietnam. However, due to a lack of data, assessing this impact has faced numerous challenges. This article presents the results of applying machine learning technology to create land cover maps from remote sensing imagery to evaluate land cover changes over time and their effects on flow in the Da River basin using the SWAT model. Land cover data for the years 2000 and 2020 were used as input for the SWAT model, created from satellite images using machine learning algorithms on the Google Earth Engine platform. The accuracy of the land cover was evaluated through overall accuracy and Kappa coefficient, achieving relatively good results. The SWAT model

was set up to compare changes in water quantity in the basin corresponding to land cover in 2000 (KB1) and 2020 (KB2). The land cover in 2000 and 2020 indicates that the built-up area increased by 866,92%, the forest increased by 4,26%, the bare land increased by 3,39%, the water increased by 63,89%, and the area of shrubs - grasses decreased by 11,60%. Flow in KB2 increased compared to KB1 in most sub-basins when considered over the entire year. During the dry season, flow across the basin decreased by 0,15%, while it increased by 0,08% during the rainy season, resulting in an overall annual increase of 0,04%.

Keywords: Da river; Land cover; SWAT model; Google Earth Engine.