

ỨNG DỤNG VIỄN THÁM VÀ MÁY HỌC TRONG GIẢI ĐOÁN ĐỊA HÌNH VEN BỜ TẠI KHU VỰC CỬA TAM QUAN, BÌNH ĐỊNH

Vũ Văn Ngọc

Phòng Thí nghiệm trọng điểm Quốc gia về Động lực học sóng biển

Nguyễn Tiếp Tân, Trần Trung Dũng

Viện Khoa học Thủy lợi Việt Nam

Trần Thanh Tùng

Trường Đại học Thủy lợi

Tóm tắt: Dữ liệu địa hình ven biển đóng vai trò then chốt trong nghiên cứu kỹ thuật và quản lý vùng ven biển. Việc tiếp cận dữ liệu địa hình đa thời gian, kết hợp với thông tin khí tượng, thủy văn và hải văn, là yếu tố quan trọng để hiểu rõ động lực ven biển. Tuy nhiên, việc thu thập dữ liệu địa hình ven biển truyền thống thường gặp khó khăn do hạn chế về nguồn lực, tốn kém, đòi hỏi thiết bị chuyên dụng và chịu ảnh hưởng bởi điều kiện thời tiết. Công nghệ viễn thám nổi lên như một giải pháp thay thế hiệu quả. Ưu điểm của viễn thám bao gồm khả năng thu thập dữ liệu nhanh chóng, phạm vi bao phủ rộng lớn và khả năng tiếp cận kho lưu trữ ảnh lịch sử. Nghiên cứu này ứng dụng ảnh viễn thám để giải đoán địa hình khu vực cửa sông, ven biển và cồn cát tại cửa biển Tam Quan, tỉnh Bình Định. Hai phương pháp giải đoán được sử dụng và so sánh: công thức tỷ số của Stumpf và thuật toán máy học. Kết quả cho thấy cả hai phương pháp đều có khả năng trích xuất thông tin địa hình từ dữ liệu viễn thám. Phương pháp Stumpf đạt được hệ số tương quan $R^2 = 0.73$ so với dữ liệu khảo sát thực địa. Các thuật toán máy học cho thấy hiệu quả vượt trội hơn, trong đó thuật toán rừng ngẫu nhiên (RF) đạt độ chính xác cao nhất với $R^2 = 0.957$. Nghiên cứu này khẳng định tiềm năng của viễn thám, đặc biệt khi kết hợp với máy học, trong việc giải đoán địa hình đa thời gian ở vùng ven biển, góp phần hiểu rõ hơn về động lực ven biển và hỗ trợ quá trình ra quyết định.

Từ khóa: Viễn thám, Địa hình đáy biển, Giải đoán địa hình đáy, Máy học, cửa Tam Quan.

Summary: Coastal terrain data is crucial for coastal engineering and management research. Access to multi-temporal terrain data, combined with meteorological, hydrological, and oceanographic information, is essential for understanding coastal dynamics. However, traditional methods for acquiring coastal terrain data often face limitations due to resource constraints, high costs, specialized equipment requirements, and weather dependency. Remote sensing technology has emerged as an effective alternative, offering advantages such as rapid data acquisition, extensive spatial coverage, and access to historical imagery archives. This study utilizes remote sensing imagery to interpret the terrain of the estuary, coastal areas, and sand dunes at Tam Quan estuary in Binh Dinh province, Vietnam. Two interpretation methods are employed and compared: Stumpf's ratio formula and machine learning algorithms. Results indicate that both methods can effectively extract terrain information from remote sensing data. Stumpf's method achieves a correlation coefficient of $RMSE = 0.73$ compared to field survey data, while machine learning algorithms demonstrate superior performance, with the Random Forest (RF) algorithm achieving the highest accuracy ($RMSE = 0.957$). This research highlights the potential of remote sensing, particularly when integrated with machine learning, for multi-temporal terrain interpretation in coastal regions. This approach contributes to a deeper understanding of coastal dynamics and supports sustainable decision-making processes.

Keywords: Remote sensing, Coastal bathymetry, Bathymetry interpretation, Machine learning, Tam Quan estuary.

1. GIỚI THIỆU CHUNG

Vùng cửa sông và ven biển là khu vực chuyển

tiếp động lực giữa lục địa và đại dương, mang đặc điểm địa hình đa dạng và biến động phức tạp. Địa hình nơi đây bao gồm lớp phủ thực vật,

Ngày nhận bài: 17/10/2024

Ngày thông qua phản biện: 12/11/2024

Ngày duyệt đăng: 29/11/2024

núi đá, cồn cát và các công trình nhân tạo, chịu ảnh hưởng mạnh mẽ từ các quá trình tự nhiên như xói mòn, bồi lắng, sóng biển, thủy triều và dòng chảy ven bờ. Sự tương tác giữa các yếu tố này góp phần tạo nên tính biến động không ngừng của địa hình ven biển, cả trên cạn và dưới nước. Ví dụ, gió có thể gây biến dạng cồn cát, trong khi sóng và dòng chảy ven bờ tác động đến địa hình vùng ngập nước và chân cồn cát, gây xói lở bờ biển hoặc hình thành các bãi bồi mới.

Hiểu rõ biến động địa hình ven biển là yếu tố then chốt trong nghiên cứu kỹ thuật và quản lý vùng ven biển, phục vụ cho việc quy hoạch không gian, bảo vệ môi trường và phát triển kinh tế-xã hội bền vững. Thông tin về biến động địa hình giúp đánh giá tác động của biến đổi khí hậu, dự báo xói lở bờ biển, hỗ trợ thiết kế công trình ven biển và quản lý tài nguyên thiên nhiên.

Các phương pháp nghiên cứu biến động địa hình truyền thống thường dựa trên đo đạc thực địa. Tuy nhiên, phương pháp này tồn tại một số hạn chế. Việc đo đạc thực địa thường tốn kém, đòi hỏi nhiều nhân lực, thiết bị chuyên dụng và phụ thuộc vào điều kiện thời tiết. Hơn nữa, phương pháp này chỉ cung cấp thông tin địa hình tại một số điểm rời rạc và thời điểm khảo sát nhất định, khó khăn trong việc nắm bắt bức tranh toàn cảnh và diễn biến địa hình theo thời gian.

Sự phát triển của công nghệ viễn thám đã mang đến một giải pháp hiệu quả cho việc thu thập dữ liệu địa hình ven biển. Ảnh viễn thám, với khả năng bao quát diện rộng, thu thập dữ liệu nhanh chóng, đa thời gian và có tính đồng bộ cao, cho phép quan sát và phân tích biến động địa hình một cách chi tiết và toàn diện. Việc ứng dụng ảnh viễn thám giúp khắc phục những hạn chế của phương pháp đo đạc truyền thống, cung cấp dữ liệu phong phú phục vụ cho nghiên cứu, giám sát và quản lý vùng ven biển.

Bên cạnh đó, sự phát triển của trí tuệ nhân tạo (AI), đặc biệt là kỹ thuật máy học, đã mở ra hướng đi mới trong giải đoán thông tin từ ảnh viễn thám. Các thuật toán máy học có khả năng tự động phân tích, xử lý dữ liệu ảnh và trích xuất

thông tin địa hình với độ chính xác cao. Sự kết hợp giữa viễn thám và máy học hứa hẹn mang lại hiệu quả vượt trội trong nghiên cứu biến động địa hình ven biển, góp phần nâng cao hiệu quả quản lý và phát triển bền vững vùng ven biển.

2. SỐ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Số liệu sử dụng

Số liệu sử dụng được thu thập từ 02 Đề tài cấp Bộ NN& PTNT. Trong đó số liệu đo đạc địa hình dưới nước, khu vực bãi biển, cao độ đỉnh đê cửa Tam Quan được thu thập từ Đề tài “*Nghiên cứu ứng dụng giải pháp chuyên cát, chống bồi lấp cho các cửa sông miền Trung*” [1]. Số liệu đo đạc địa hình trên cạn và các khu vực cồn cát được thu thập từ Đề tài “*Nghiên cứu giải pháp quản lý, bảo vệ, phục hồi và phát triển cồn cát ven biển từ Quảng Bình đến Bình Thuận với vai trò là tuyến đê biển tự nhiên gắn với bảo vệ môi trường sinh thái*” [2].

2.2. Phương pháp xác định địa hình theo công thức tỉ số của Stumpf

Giải đoán địa hình đáy biển cho vùng nước nông ven bờ bằng ảnh vệ tinh đa phổ có độ phân giải cao đang được phát triển mạnh trong thời gian gần đây, khi nguồn ảnh vệ tinh đa phổ có độ phân giải cao miễn phí như LANDSAT, IKONOS, SENTINEL ngày càng nhiều. Stumpf và các cộng sự đã phát triển phương pháp giải đoán độ sâu đáy biển dựa trên tỉ lệ biến đổi tín hiệu theo hàm logarit của các dải tần ánh sáng của ảnh vệ tinh. Phương pháp của Stumpf đã áp dụng khá rộng rãi do có chi phí hợp lý và số lượng ảnh có chất lượng tốt ngày càng nhiều.

Mô hình tỉ số ước tính độ sâu thực tế của Stumpf (2003) [5] theo công thức sau:

$$Z = m_1 \frac{\ln(nR_w(\lambda_i))}{\ln(nR_w(\lambda_j))} - m_0 \quad (1)$$

Trong đó: Z là độ sâu ước tính

m_1 là hằng số có thể điều chỉnh được để chia tỉ lệ với độ sâu

n là hằng số cố định cho tất cả các khu vực

R_w là giá trị phản xạ của nước ở kênh i hoặc j

m_0 là độ lệch tại độ sâu $0m$

2.3. Phương pháp xác định địa hình bằng thuật toán máy học

Học máy (Machine Learning - ML) là một nhánh của trí tuệ nhân tạo, trao cho máy tính khả năng học từ dữ liệu mà không cần lập trình tường minh. Thay vì dựa trên các quy tắc được lập trình sẵn, máy tính sử dụng thuật toán để tự động nhận diện các mẫu, quy luật ẩn trong dữ liệu. Bằng cách này, máy tính có thể dự đoán kết quả cho dữ liệu mới dựa trên những gì đã học được.

Các mô hình trí tuệ nhân tạo trong nghiên cứu này gồm 07 mô hình được ứng dụng để giải đoán ảnh vệ tinh:

- RF (Random Forest): Mô hình rừng ngẫu nhiên là một phương pháp học tập hợp (ensemble learning), sử dụng nhiều cây quyết định để đưa ra dự đoán.
- LR (Linear Regression): Mô hình hồi quy tuyến tính thiết lập mối quan hệ tuyến tính giữa biến đầu vào và đầu ra.
- DT (Decision Tree): Mô hình cây quyết định sử dụng cấu trúc cây để phân loại hoặc dự đoán.
- SVR (Support Vector Regression): Mô hình hồi quy sử dụng máy vector hỗ trợ để dự đoán các biến liên tục.
- GB (Gradient Boosting): Mô hình tăng cường độ dốc là một kỹ thuật học tập hợp, xây dựng các mô hình theo chuỗi, mỗi mô hình cố gắng cải thiện độ chính xác bằng cách sửa chữa lỗi của mô hình trước đó.
- KNN (K-Nearest Neighbors): Mô hình K láng giềng gần nhất là một thuật toán đơn giản, dự đoán dựa trên các điểm dữ liệu huấn luyện gần nhất với điểm dữ liệu cần dự đoán.
- NN (Neural Network): Mô hình mạng nơ-ron nhân tạo mô phỏng cách hoạt động của não người, được thiết kế để nhận diện mẫu và đưa ra dự đoán.

Giải đoán địa hình đáy biển bằng thuật toán học máy là một phương pháp tiên tiến, khai thác sức mạnh của trí tuệ nhân tạo để phân tích dữ liệu và trích xuất thông tin về độ sâu, hình dạng đáy biển. Thay vì dựa trên các phương pháp truyền thống như đo đạc thực địa tốn kém và hạn chế về mặt không gian, phương pháp này sử dụng

ảnh vệ tinh hoặc dữ liệu sonar làm đầu vào cho các thuật toán học máy. Các thuật toán này, sau khi được huấn luyện trên tập dữ liệu lớn gồm các điểm đo sâu thực tế và các đặc trưng tương ứng trên ảnh, có khả năng tự động nhận diện các mẫu và mối quan hệ phức tạp giữa dữ liệu ảnh và địa hình đáy biển.

Một số thuật toán học máy thường được sử dụng trong giải đoán địa hình đáy biển bao gồm mạng nơ-ron nhân tạo (NN), máy vector hỗ trợ (SVM), cây quyết định (DT) và rừng ngẫu nhiên (RF). Ưu điểm của phương pháp này là khả năng xử lý dữ liệu lớn, tự động hóa quy trình giải đoán, và tiềm năng đạt độ chính xác cao. Kết quả giải đoán địa hình đáy biển đóng vai trò quan trọng trong nhiều ứng dụng, từ lập bản đồ hải dương, quản lý tài nguyên biển, đến định vị tàu thuyền và khảo sát địa chất biển.

3. KẾT QUẢ NGHIÊN CỨU

3.1. Giải đoán địa hình đáy biển vùng ven bờ theo phương pháp Stumpf

Theo phương pháp mô hình tỉ số độ sâu của Stumpf (2003) [5], việc giải đoán độ sâu được thực hiện với ảnh tương ứng thời gian khảo sát địa hình đáy biển khu vực cửa Tam Quan ứng với đợt đo 7/2022 thuộc đề tài KHCN cấp Bộ NN&PTNT “Nghiên cứu ứng dụng giải pháp chuyển cát chống bồi lấp cho các cửa sông Miền Trung”

Hình 1a thể hiện quy trình thực hiện tính toán độ sâu theo công thức Stumpf ứng với 6 bước chính được mô tả. Trước đây việc thực hiện thông thường triển khai khá thủ công. Ở Việt Nam, tác giả Trần Thanh Tùng (2020) đã ứng dụng ảnh Sentinel 2 để giải đoán bãi ngập triều cho khu vực cửa Tiên Châu tỉnh Phú Yên. Trần Thanh Tùng (2023) [4], nghiên cứu và ứng dụng ảnh Sentinel 2 đánh giá diễn biến đường bờ và biến đổi địa hình đáy biển khu vực cửa Nhật Lệ, Quảng Bình. Trong hai nghiên cứu trên, việc thực hiện thủ công khá phức tạp và tốn nhiều thời gian cho mỗi cảnh ảnh. Đối với nghiên cứu trong bài báo này, nhóm tác giả thực hiện tự động hóa để giảm thiểu thời gian tính toán cho mỗi nghiên cứu. Việc tự động hóa được tiến hành trên nền tảng Google Earth

Engine (GEE), thể hiện như Hình 1b. Kết quả tính toán tự động từ GEE được triết xuất và biểu diễn theo Raster và thang màu tương ứng như Hình 2.

Bước 1. Tải ảnh vệ tinh, các ảnh vệ tinh được tích hợp sẵn trên GEE

Bước 2. Tách ranh giới đất và nước theo công thức của Gao 1996

Công thức chung: $NDWI = (NIR - SWIR) / (NIR + SWIR)(2)$

Đối với mỗi loại nguồn ảnh thì NDWI tính toán tùy theo nhãn hiệu các kênh ảnh sử dụng tương ứng như sau

$$NDWI (Landsat 8) = (B3 - B5) / (B3 + B5) \quad (3)$$

$$NDWI (Landsat 4 - 7) = (B2 - B4) / (B2 + B4) \quad (4)$$

$$NDWI (Sentinel 2) = (B3 - B8) / (B3 + B8) \quad (5)$$

Bước 3. Hiệu chỉnh ánh sáng theo Xu (2006)

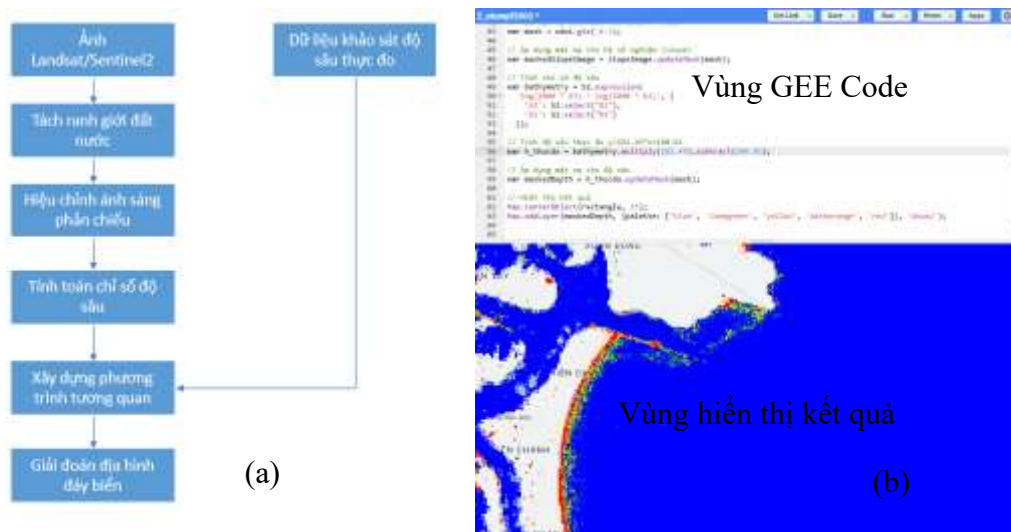
$$R'i = Ri - bi (RNIR - MinNIR) \quad (6)$$

Bước 4. Tính toán chỉ số độ sâu theo Stumpf (2003) [5]

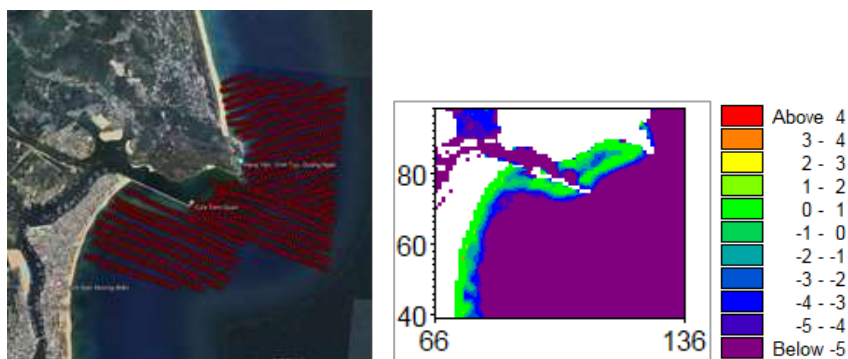
Bước 5. Xây dựng phương trình tương quan chỉ số độ sâu và độ sâu thực tế

Bước 6. Từ phương trình tương quan, tính toán qui đổi từ chỉ số độ sâu Stumpf sang độ sâu thực tế

Ghi chú: **NIR**: cận hồng ngoại, **SWIR** hồng ngoại sóng ngắn, **Ri** là giá trị pixel phản xạ từ kênh nhìn thấy, **bi** là độ dốc hồi quy, **RNIR** là giá trị pixel tại bước sóng cận hồng ngoại, **MinNIR** là giá trị min của NIR tại tập hợp mẫu.



Hình 1: Sơ đồ quy trình tính toán (a), triển khai tính toán trên GEE (b)



Hình 2: Các điểm đo thực tế (a), kết quả tính theo công thức Stumpf biểu diễn raster

Ứng với việc sử dụng công thức tỉ số độ sâu của Stumpf (2003)[3] [5], kết quả đạt được là

phương trình tương quan giữa độ sâu giải đoán và độ sâu đo đạc thực tế đạt được là: $y = 181.45x$

-180.81; trong đó x là độ sâu giải đoán từ ảnh vệ tinh và y là độ sâu đo đạc), tương quan giữa số liệu giải đoán và số liệu đo đạc có hệ số $RMSE = 0.73$

Qua kết quả giải đoán cho thấy khả năng ứng dụng phương pháp tỉ số của Stumpf (2003) [3][5], đối với khu vực nghiên cứu đã đạt được là tương đối tốt. Tuy nhiên phương pháp này không mô tả được địa hình khu vực ở phần không ngập nước.

3.2. Giải đoán địa hình đáy biển vùng ven bờ bằng thuật toán máy học

Trong bối cảnh Cách mạng Công nghiệp 4.0, trí tuệ nhân tạo (AI) đang trở thành công nghệ chủ chốt, được ứng dụng rộng rãi trong nghiên cứu trên toàn cầu. Nghiên cứu gần đây đã chỉ ra rằng, áp dụng các phương pháp học sâu có thể mở ra những khả năng mới trong việc ước lượng độ sâu từ ảnh vệ tinh, cho phép cải thiện độ chính xác và giảm thiểu sự phụ thuộc vào dữ liệu thực địa (Najar et al., 2021; Gupta et al., 2023) [6] [7]. Sự phát triển của công nghệ và các thuật toán mới giúp tăng cường khả năng từ ảnh vệ tinh, làm cho nó trở thành công cụ hữu ích trong kỹ thuật và quản lý (Pacheco et al., 2015; Sukmono et al., 2022; Apicella et al., 2023)[8] [9] [10].

Tác giả Rifqi Muhammad Harrys [11] xây dựng công cụ SDB_GUI trên nền tảng Python và cung cấp mã nguồn mở trên Github. Công cụ có giao diện đồ họa và 4 mô hình máy học KNN, MLR, RF, SVM để dự đoán độ sâu đáy biển. Công cụ có hạn chế ở số lượng giới hạn 4 mô hình, không đồng thời được 4 mô hình cùng lúc, không thực hiện tính toán hàng loạt ảnh vệ tinh để cho kết quả độ sâu. Mã nguồn chương trình SDB_GUI [11] được tác giả kế thừa và cải tiến bổ sung lượng lớn mô hình máy học đồng thời,

bổ sung khả năng sử dụng mô hình tốt nhất để dự đoán với hàng loạt ảnh vệ tinh từ một thư mục đầu vào không cần phân loại.

Tại Việt Nam, việc ứng dụng AI vẫn còn khá mới mẻ nhưng đầy tiềm năng. Nghiên cứu này tận dụng các thuật toán AI để giải đoán thông tin từ ảnh vệ tinh, cụ thể là trích xuất thông tin địa hình từ giá trị các kênh phổ của ảnh kết hợp với dữ liệu khảo sát thực địa. Hình 3 thể hiện các dữ liệu đầu vào gồm năm loại giá trị tương ứng với các băng ảnh từ Band1 đến Band5, trong đó có giá trị thực đo về cao độ địa hình ở phần trên cạn, và giá trị thực đo về độ sâu đáy biển ở phần dưới nước. Giá trị các băng ảnh là giá trị nhận được từ ảnh viễn thám. Giá trị tại các điểm đo dưới nước là giá trị độ sâu đáy biển được đo đạc từ đề tài “Nghiên cứu ứng dụng giải pháp chuyển cát chống bồi lấp cho các cửa sông Miền Trung” [1]. Giá trị tại các điểm đo trên cạn là giá trị độ cao được thu thập và đo đạc từ đề tài “Nghiên cứu giải pháp quản lý, bảo vệ, phục hồi và phát triển cồn cát ven biển từ Quảng bình đến Bình thuận với vai trò là tuyến đê biển tự nhiên gắn với bảo vệ môi trường sinh thái” [2]. Dữ liệu được đồng bộ cùng hệ tọa độ, bao phủ được phần trên cạn và dưới nước.

Bước 1: Tiền xử lý

Dữ liệu đầu vào bao gồm 2 loại tệp: 1) Dữ liệu ảnh vệ tinh dưới lưu trữ dưới dạng tệp raster có đuôi mở rộng tif chứa 5 band ảnh; 2) Dữ liệu các điểm đo đạc có đuôi mở rộng shp chứa các thông tin tọa độ x,y và độ sâu z . Việc xử lý ban đầu nhằm đồng bộ điểm đo đạc và raster ảnh ở cùng một hệ tọa độ. Hình 3 thể hiện kết quả công đoạn tiền xử lý.



Hình 3: Các điểm đo được đồng bộ hệ tọa độ theo ảnh vệ tinh chứa các band ảnh

Bước 2: Làm sạch dữ liệu

Dữ liệu được làm sạch để chuẩn bị cho huấn luyện và kiểm định mô hình máy học. Các dữ liệu band ảnh và dữ liệu điểm đo được đồng bộ hóa theo tiêu chí lấy tọa độ điểm đo x,y. Như vậy tại vị trí từng điểm đo sẽ có dữ liệu độ sâu z, dữ liệu giá trị band 1 đến band 5. Những vị trí không có điểm đo được làm sạch bằng cách loại bỏ

Bước 3: Chia đoạn dữ liệu để huấn luyện và kiểm định mô hình máy học

Việc chia dữ liệu được thực hiện với tỉ lệ 75% để huấn luyện, 25% để kiểm định. Cú pháp `train_size = 0.75` được đặt để thực thi.

Bước 4: Thiết lập mô hình máy học

Các mô hình được thành lập dựa trên thư viện `sklearn` trong phần mềm Python [11]. Trong nghiên cứu này, bộ công cụ được cải tiến để thực hiện 7 mô hình máy học với các thiết lập cụ thể như Hình 4.

```
# Khởi tạo 7 mô hình máy RF- LR- DT- SVR- GB- KNN- NN
models = {
    'Random Forest': RandomForestRegressor(n_estimators=300, random_state=random_state, n_jobs=-1),
    'Linear Regression': LinearRegression(),
    'Decision Tree': DecisionTreeRegressor(random_state=random_state),
    'SVR': SVR(),
    'Gradient Boosting': GradientBoostingRegressor(n_estimators=300, random_state=random_state),
    'KNN': KNeighborsRegressor(n_neighbors=5),
    'Neural Network': MLPRegressor(random_state=random_state, max_iter=10000)
}
```

Hình 4: Bộ thông số mô hình máy học

Bước 5: Huấn luyện và kiểm định mô hình máy học

Số liệu đã chuẩn bị ở các bước trước được sử dụng trong bước này như Hình 5 để huấn luyện và kiểm thử mô hình theo tỉ lệ. Trong đó 75% số liệu được dùng để huấn luyện với các tham số đầu vào là `features_train` là 75% số liệu của 5 band ảnh, `z_train` là 75% số liệu của điểm đo đặc. Lượng còn lại 25% để kiểm định được thể hiện bởi `features_test` và `z_test`. Mô hình học tập thông qua số liệu huấn luyện sau đó thực hiện tính toán với dữ liệu kiểm thử, kết quả đạt được là `z_pred`.

Bước 6: Đánh giá hiệu suất mô hình

Sau khi thực hiện tính toán kiểm thử và đạt được kết quả đầu ra là `z_pred`, tiến hành đánh giá hiệu suất mô hình thông qua đánh giá các chỉ số MAE (sai số tuyệt đối trung bình), MSE (sai số bình phương trung bình), RMSE (sai số bình phương trung bình) giữa số liệu `z_pred` và `z_test`. Quá trình này được điều khiển như Hình 5.

```
# Huấn luyện và đánh giá các mô hình
results = {}
for name, model in models.items():
    model.fit(features_train, z_train)
    z_pred = model.predict(features_test)
    mae = metrics.mean_absolute_error(z_test, z_pred)
    mse = metrics.mean_squared_error(z_test, z_pred)
    r2 = metrics.r2_score(z_test, z_pred)
    results[name] = {'MAE': mae, 'MSE': mse, 'R^2': r2}
    print(f'{name}:')
    print(f'MAE: {mae}')
    print(f'MSE: {mse}')
    print(f'R^2: {r2}\n')
```

Hình 5: Huấn luyện và đánh giá các mô hình

Bước 7: Tìm mô hình tốt nhất và dự tính từ ảnh tif mới

Nghiên cứu tiến hành đồng thời 7 mô hình, do đó cần xem xét mô hình nào đạt hiệu suất tốt nhất. Từ đó lưu các tham số mô hình để thực hiện tính toán từ ảnh tif mới và tính toán hàng loạt địa hình từ các ảnh vệ tinh tương ứng. Mô hình tốt nhất là mô hình có các chỉ số MAE, MSE, RMSE đạt giá trị nhỏ nhất. Tiến trình này tự động hóa trên bộ công cụ, đồng thời chỉ ra tên của mô hình tốt nhất và tự động lưu các thông tin của mô hình tốt nhất, xuất thông tin ra dạng tệp `txt` để người dùng kiểm soát thông tin.

```
# Tìm mô hình tốt nhất
best_model = min(results, key=lambda x: (results[x]['MAE'], results[x]['MSE'], -results[x]['R^2']))

# Xuất siêu tham số của mô hình tốt nhất ra tệp .txt
with open('best_model_params.txt', 'w', encoding='utf-8') as f:
    f.write(f"Siêu tham số của mô hình tốt nhất:\n")
    f.write(f"{best_model.get_params()}\n") # Lấy và ghi siêu tham số của mô hình
else:

# Tải mô hình tốt nhất
best_model_loaded = joblib.load('best_model.pkl')

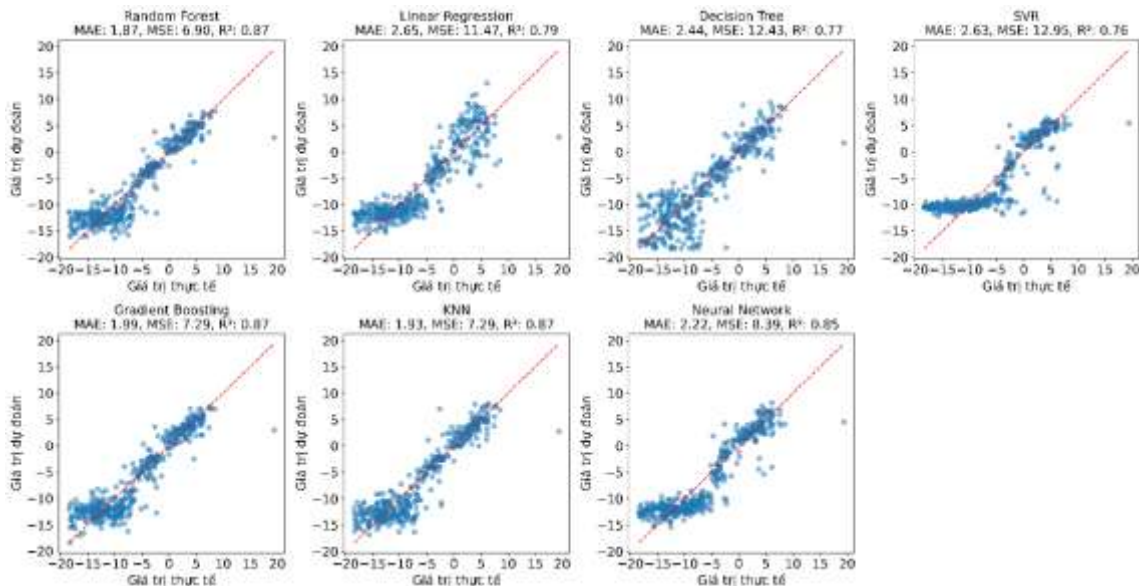
# Lấy danh sách các ảnh từ thư mục
new_images = [os.path.join(new_images_folder, f) for f in os.listdir(new_images_folder) if f.endswith('.tif')]

# Hàm dự đoán và Lưu ảnh dự đoán
def predict_and_save(image_path, model, save_folder):
```

Hình 6: Các tiến trình chính để xác định hiệu suất mô hình, bộ tham số và dự tính hàng loạt địa hình từ ảnh tif

Nghiên cứu đã thể hiện các bước nghiên cứu chủ yếu đối với phương pháp máy học được triển khai trên bộ công cụ được lập trình trên Python [11]. Bảng 1 trình bày kết quả đánh giá các mô hình máy học. Trong 7 mô hình máy học đã huấn luyện thì mô hình rừng ngẫu nhiên (RF) dự báo đạt độ chính xác cao nhất (Bảng

1). Mô hình được huấn luyện và kiểm định đạt độ chính xác được dùng để giải đoán hàng loạt địa hình theo thời gian từ các ảnh tif. Kết quả cho thấy phân bố địa hình theo không gian có sự tương đồng khá rõ khi đối chiếu trực quan với ảnh Google Earth như Hình 9.



Hình 7: Tương quan giá trị khảo sát và giá trị giải đoán từ ảnh L5

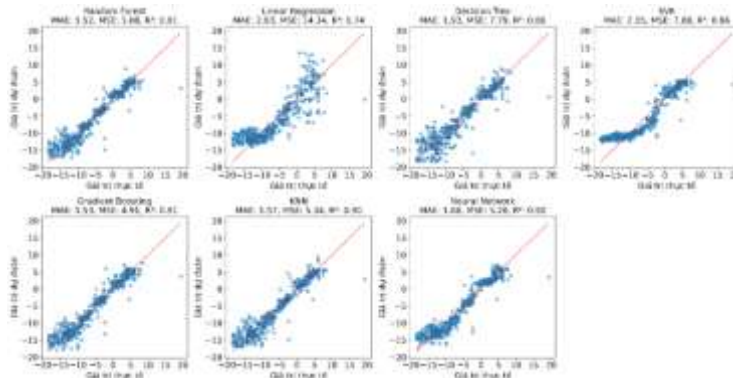
Bảng 1: Kết quả đánh giá mô hình dự đoán

Vệ tinh	Thông số	Mô hình trí tuệ nhân tạo						
		RF	LR	DT	SVR	GB	KNN	NN
L5	MAE	1.869	2.651	2.436	2.631	1.993	1.929	2.221
	MSE	6.901	11.467	12.425	12.947	7.295	7.295	8.394
	RMSE	0.874	0.790	0.773	0.763	0.867	0.867	0.846
L9	MAE	0.845	2.219	1.097	1.697	1.051	0.919	1.341
	MSE	2.326	9.108	4.730	5.504	3.412	2.497	3.856
	RMSE	0.957	0.833	0.913	0.899	0.938	0.954	0.929

Ghi chú: Trong nghiên cứu sử dụng kỹ thuật Cross-validation, không có over fitting.

Hình 7 thể hiện kết quả tính toán ứng với 7 mô hình máy học với nguồn dữ liệu ảnh Landsat 5. Trục tung thể hiện giá trị dự đoán, trục hoành thể hiện giá trị thực tế. Ứng với nguồn ảnh Landsat 5 thì việc tính toán cho kết quả tốt nhất ứng với mô hình Rừng ngẫu nhiên (RF), hệ số RMSE = 0.874 (Bảng 1).

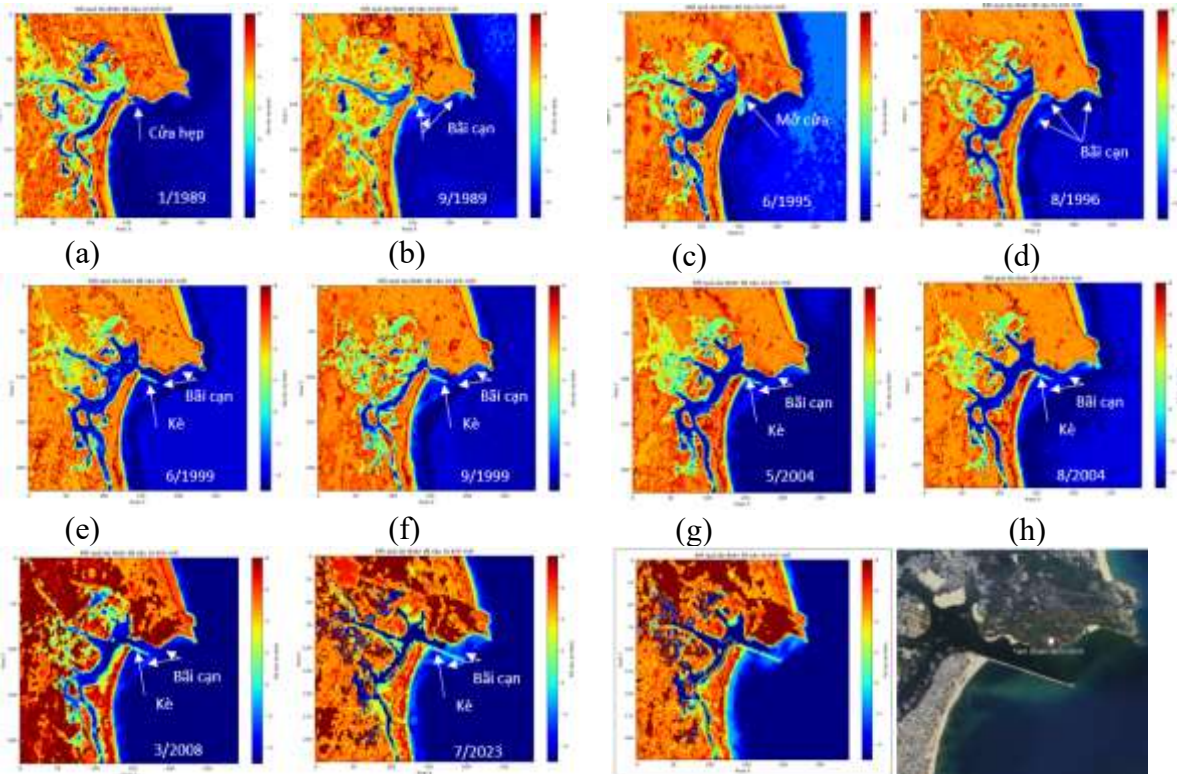
Hình 8 thể hiện kết quả tính toán với nguồn dữ liệu ảnh Landsat 9. Ứng với nguồn ảnh Landsat 5 thì việc tính toán cho kết quả tốt nhất ứng với mô hình Rừng ngẫu nhiên (RF), hệ số RMSE = 0.957 (Bảng 1).



Hình 8: Tương quan giá trị khảo sát và giá trị giải đoán từ ảnh L9

Qua nghiên cứu này cho thấy việc ứng dụng các thuật toán trí tuệ nhân tạo để giải địa hình từ ảnh vệ tinh đạt được kết quả khá tốt. Hệ số tương quan đạt được theo các mô hình học máy đều đạt

RMSE > 0.76 cao hơn so với phương pháp Stumpf hiện chỉ đạt được RMSE = 0.73. Ứng dụng AI có ưu điểm giải đoán đồng thời được phân trên cạn và phân dưới nước.



(i)

(j)

(k)

(l)

Hình 9: Xu thế biến động địa hình khu vực cửa Tam Quan

Hình 9a đến Hình 9f thể hiện xu thế bồi lấp khu vực cửa biển từ năm 1988- 2000 cho thấy những đặc điểm chính là:

Khu vực ranh giới trong và ngoài cửa diễn ra quá trình thu hẹp và mở rộng, doi cát phía Nam có xu thế kéo dài theo hướng của đê NCGS hiện nay. Ngay ranh giới phía ngoài cửa biển có xu thế bồi lắng tạo thành cồn ngầm có dạng vòng cung. Trong một số thời điểm vùng này tách hẳn bờ và dịch chuyển về phía Nam như ở tháng 6/1995.

Khu vực cách cửa biển khoảng 600-650m tồn tại cồn ngầm phân bố liên tục, cồn ngầm này có hình dạng tương đối thẳng, chắn ngang cửa, nối thẳng từ mũi đá đến vị trí bờ Nam nơi cách cửa biển khoảng 650m.

Hình 9g đến Hình 9l thể hiện xu thế bồi lấp khu vực cửa biển từ năm 2000- 2023 cho thấy những đặc điểm chính là:

Toàn bộ khu vực có xu thế bồi lắng, hình dạng các khu vực bồi lắng không thay đổi rõ rệt sau khi có đê NCGS. Toàn bộ khu vực bồi lắng vẫn chủ đạo nằm trong một khu vực tam giác giới hạn giữa cạnh bờ 650m và mũi đá phía bờ Bắc

Khi đê NCGS có độ dài ngắn hơn so với hiện nay thì hình dạng vùng bồi lắng hầu như tương đương với dạng bồi lắng tự nhiên trước khi có đê NCGS. Thời gian đê NCGS kéo dài như hiện nay, khu vực bồi càng trở lên rõ rệt hơn trên bức tranh tổng thể địa hình khu vực.

4. KẾT LUẬN VÀ KIẾN NGHỊ

Nghiên cứu đã trình bày phương pháp giải đoán địa hình đa thời gian khu vực cửa sông Tam Quan, tỉnh Bình Định bằng ảnh viễn thám. Hai phương pháp được sử dụng bao gồm giải đoán

theo công thức tỷ số của Stumpf và giải đoán bằng thuật toán máy học. Kết quả cho thấy cả hai phương pháp đều đạt được độ chính xác chấp nhận được. Đối với phương pháp Stumpf, việc giải đoán chỉ thực hiện phù hợp đối với phần địa hình dưới nước, kết quả cho thấy sự phân bố địa hình khá phù hợp với thực tế. Độ chính xác thể hiện qua tương quan giữa kết quả giải đoán và kết quả khảo sát đạt được $RMSE = 0.73$. Tuy nhiên, phương pháp ứng dụng thuật toán máy học cho thấy hiệu quả vượt trội với hệ số tương quan cao hơn đáng kể so với phương pháp Stumpf. Đặc biệt, thuật toán rừng ngẫu nhiên (RF) cho kết quả có độ chính xác cao nhất, $RMSE = 0.874 \div 0.957$ ứng với ảnh vệ tinh Landsat 5, 7.

Nghiên cứu này đã khẳng định tiềm năng của kỹ thuật viễn thám, đặc biệt khi kết hợp với trí tuệ nhân tạo, trong việc giải đoán địa hình đa thời gian vùng ven biển. Phương pháp ứng dụng công nghệ mới này cung cấp một giải pháp hiệu quả về chi phí và thời gian, đồng thời cho phép thu thập dữ liệu trên diện rộng và ở những khu vực khó tiếp cận. Kết quả nghiên cứu có ý nghĩa quan trọng trong việc quản lý và khai thác bền vững vùng ven biển, góp phần vào sự phát triển kinh tế- xã hội của địa phương.

LỜI CẢM ƠN:

Nghiên cứu này được thực hiện trong khuôn khổ các đề tài cấp Bộ NN và PTNT “Nghiên cứu giải pháp quản lý, bảo vệ, phục hồi và phát triển cồn cát ven biển từ Quảng Bình đến Bình Thuận với vai trò là tuyến đê biển tự nhiên gắn với bảo vệ môi trường sinh thái” và đề tài “Nghiên cứu ứng dụng giải pháp chuyển cát, chống bồi lấp cho các cửa sông miền Trung”.

TÀI LIỆU THAM KHẢO

- [1] Trần Thanh Tùng, 2022. Đề tài KHCN cấp Bộ Nông nghiệp và Phát triển Nông thôn “Nghiên cứu ứng dụng giải pháp chuyển cát, chống bồi lấp cho các cửa sông miền Trung”. Số liệu và báo cáo khảo sát địa hình khu vực biển Tam Quan tỉnh Bình Định.
- [2] Nguyễn Tiếp Tân, 2022. Đề tài KHCN cấp Bộ Nông nghiệp và Phát triển Nông thôn “Nghiên cứu giải pháp quản lý, bảo vệ, phục hồi và phát triển cồn cát ven biển từ quảng bình đến bình

- thuận với vai trò là tuyến đê biển tự nhiên gắn với bảo vệ môi trường sinh thái”. Số liệu và báo cáo khảo sát, thu thập địa hình các cồn cát ven biển Miền Trung.
- [3] Vũ Văn Ngọc, Trần Thanh Tùng, 2023. Phân tích diễn biến địa hình đáy biển khu vực cửa tam quan, tỉnh bình định. Tuyển tập hội nghị khoa học thường niên năm 2023. Trường ĐH Thủy lợi
 - [4] Trần Thanh Tùng, Dương Công Điền, 2020. Giải đoán bãi ngập triều bằng ảnh Sentinel 2, áp dụng cho khu vực cửa Tiên Châu tỉnh Phú Yên. Tuyển tập hội nghị khoa học thường niên năm 2020. Trường ĐH Thủy lợi
 - [5] R. P. Stumpf, K. Holderied and M. Sinclair. 2003."Determination of water depth with high-resolution satellite imagery over variable bottom types," *Limnology and Oceanography* Vol. 48, pp. 547-556.
 - [6] Najar, M., Thoumyre, G., Bergsma, E., Almar, R., Benshila, R., & Wilson, D. (2021). Satellite derived bathymetry using deep learning. *Machine Learning*, 112(4), 1107-1130. <https://doi.org/10.1007/s10994-021-05977-w>
 - [7] Gupta, G. K., Bhat, R. V., & Balan, M. S. (2023). Satellite-derived bathymetry of an inland reservoir in india.. <https://doi.org/10.36227/techrxiv.22210603.v1>
 - [8] Pacheco, A., Horta, J., Loureiro, C., & Ferreira, Ó. (2015). Retrieval of nearshore bathymetry from landsat 8 images: a tool for coastal monitoring in shallow waters. *Remote Sensing of Environment*, 159, 102-116. <https://doi.org/10.1016/j.rse.2014.12.004>
 - [9] Sukmono, A., Aji, S., Amarrohman, F. J., Bashit, N., & Saputra, L. R. (2022). The extraction of near-shore bathymetry using sentinel-2a satellite imagery: algorithms and their modifications. *TEM Journal*, 150-158. <https://doi.org/10.18421/tem111-17>
 - [10] Apicella, L., Martino, M. D., Ferrando, I., Quarati, A., & Federici, B. (2023). Deriving coastal shallow bathymetry from sentinel 2-, aircraft- and uav-derived orthophotos: a case study in ligurian marinas. *Journal of Marine Science and Engineering*, 11(3), 671. <https://doi.org/10.3390/jmse11030671>
 - [11] GitHub - rifqiharrys/sdb_gui: Python based GUI for Satellite Derived Bathymetry processing