

TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT
KHOA TRẮC ĐỊA - BẢN ĐỒ VÀ QUẢN LÝ ĐẤT ĐAI
BỘ MÔN ĐO ẢNH VÀ VIỄN THĂM



HỘI NGHỊ KHOA HỌC QUỐC GIA
CÔNG NGHỆ ĐỊA KHÔNG GIAN VÌ SỰ PHÁT TRIỂN BỀN VỮNG
(NCGSD 2024)



NHÀ XUẤT BẢN KHOA HỌC TỰ NHIÊN VÀ CÔNG NGHỆ

**TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT
KHOA TRẮC ĐỊA BẢN ĐỒ VÀ QUẢN LÝ ĐẤT ĐAI
BỘ MÔN ĐO ẢNH VÀ VIỄN THĂM**



**HỘI NGHỊ KHOA HỌC QUỐC GIA
CÔNG NGHỆ ĐỊA KHÔNG GIAN VÌ SỰ PHÁT TRIỂN BỀN VỮNG
(NCGSD - 2024)**

Hà Nội - 2024

Nguồn ảnh bìa: Bộ môn Đo ảnh và Viễn thám, Khoa Trắc địa - Bản đồ và Quản lý Đất đai,
trường Đại học Mỏ - Địa chất

CÁC ĐƠN VỊ ĐỒNG TỔ CHỨC



Trường Đại học
Mỏ - Địa chất



Viện Khoa học Đo đạc và
Bản đồ



Học viện Kỹ thuật Quân sự



Trường Đại học Phenikaa



Hội Trắc địa - Bản đồ
- Viễn thám Việt Nam



Cục Viễn thám Quốc gia



Cục Bản đồ,
Bộ Tổng tham mưu



Trường Đại học Tài nguyên
và Môi trường Hà Nội

BAN TỔ CHỨC

- | | |
|--|------------------|
| 1. GS.TS Trần Thanh Hải, Trường Đại học Mỏ - Địa chất | - Trưởng ban |
| 2. TS Trần Trung Anh, Trường Đại học Mỏ - Địa chất | - Phó Trưởng ban |
| 3. PGS.TS Đỗ Ngọc Anh, Trường Đại học Mỏ - Địa chất | - Phó Trưởng ban |
| 4. TS Đỗ Thị Hoài, Viện Khoa học Đo đạc và Bản đồ | - Ủy viên |
| 5. PGS.TS Trịnh Lê Hùng, Học viện Kỹ thuật Quân sự | - Ủy viên |
| 6. GS.TS Trần Đức Tân, Trường Đại học Phenikaa | - Ủy viên |
| 7. GS.TS Võ Chí Mỹ, Hội Trắc địa - Bản đồ - Viễn thám Việt Nam | - Ủy viên |
| 8. PGS.TS Phạm Minh Hải, Cục Viễn thám Quốc gia | - Ủy viên |
| 9. TS Lê Đại Ngọc, Cục Bản đồ, Bộ tổng Tham mưu | - Ủy viên |
| 10. TS Vũ Danh Tuyên, Trường Đại học Tài nguyên và Môi trường Hà Nội | - Ủy viên |

BAN KHOA HỌC

- | | |
|--|------------------|
| 1. GS.TS Trần Thanh Hải, Trường Đại học Mỏ - Địa chất | - Trưởng ban |
| 2. TS Trần Trung Anh, Trường Đại học Mỏ - Địa chất | - Phó Trưởng ban |
| 3. PGS.TS Nguyễn Văn Trung, Trường Đại học Mỏ - Địa chất | - Phó Trưởng ban |
| 4. TS Dương Thành Trung, Trường Đại học Mỏ - Địa chất | - Ủy viên |
| 5. GS.TSKH Hoàng Ngọc Hà, Trường Đại học Mỏ - Địa chất | - Ủy viên |
| 6. GS.TS Võ Chí Mỹ, Hội Trắc địa - Bản đồ - Viễn thám Việt Nam | - Ủy viên |
| 7. PGS.TS Phạm Quốc Khánh, Trường Đại học Mỏ - Địa chất | - Ủy viên |
| 8. PGS.TS Nguyễn Quốc Long, Trường Đại học Mỏ - Địa chất | - Ủy viên |
| 9. TS Đỗ Thị Phương Thảo, Trường Đại học Mỏ - Địa chất | - Ủy viên |
| 10. TS Nguyễn Gia Trọng, Trường Đại học Mỏ - Địa chất | - Ủy viên |
| 11. TS Nguyễn Thị Kim Yến, Trường Đại học Mỏ - Địa chất | - Ủy viên |

BAN THƯ KÝ

- | | |
|---|------------------|
| 1. TS Trần Thanh Hà, Trường Đại học Mỏ - Địa chất | - Trưởng ban |
| 2. TS Đoàn Thị Nam Phương, Trường Đại học Mỏ - Địa chất | - Phó Trưởng ban |
| 3. ThS Lê Thanh Nghị, Trường Đại học Mỏ - Địa chất | - Ủy viên |
| 4. ThS Nguyễn Minh Hải, Trường Đại học Mỏ - Địa chất | - Ủy viên |
| 5. ThS Phạm Thị Thanh Hòa, Trường Đại học Mỏ - Địa chất | - Ủy viên |

BAN BIÊN TẬP

- | | |
|--|-------------------------|
| 1. PGS.TS Nguyễn Văn Trung, Trường Đại học Mở - Địa chất | - <i>Trưởng ban</i> |
| 2. ThS Nguyễn Minh Hải, Trường Đại học Mở - Địa chất | - <i>Phó Trưởng ban</i> |
| 3. TS Đoàn Thị Nam Phương, Trường Đại học Mở - Địa chất | - <i>Ủy viên</i> |
| 4. ThS Lê Thanh Nghị, Trường Đại học Mở - Địa chất | - <i>Ủy viên</i> |
| 5. ThS Phạm Thị Thanh Hòa, Trường Đại học Mở - Địa chất | - <i>Ủy viên</i> |

LỜI NÓI ĐẦU

“Hội nghị Khoa học Quốc gia Công nghệ địa không gian vì sự phát triển bền vững” (NCGSD - 2024) được tổ chức ngày 24 tháng 10 năm 2024 tại Trường Đại học Mỏ - Địa chất, thành phố Hà Nội. Hội nghị do Trường Đại học Mỏ - Địa chất phối hợp cùng Viện Khoa học Đo đạc và Bản đồ; Học viện Kỹ thuật Quân sự; Trường Đại học Phenikaa; Hội Trắc địa - Bản đồ - Viễn thám Việt Nam; Cục Viễn thám Quốc gia; Cục Đo đạc, Bản đồ và Thông tin địa lý Việt Nam; Cục Đo đạc Bản đồ, Bộ Tổng tham mưu; Trường Đại học Tài nguyên và Môi trường Hà Nội.

Ban Tổ chức Hội nghị đã nhận được 29 bài báo của các nhà khoa học trong cả nước gửi đến Hội nghị. Các bài báo khoa học đăng trong tuyển tập này trình bày những kết quả nghiên cứu về các lĩnh vực: Công nghệ viễn thám trong các Khoa học Trái đất; Định vị - dẫn đường và hệ thống định vị vệ tinh toàn cầu; Các mô hình xử lý và dự báo trong tai biến thiên nhiên và môi trường; Quan trắc và đánh giá môi trường chiến lược; Công nghệ quét LiDAR hàng không và mặt đất, công nghệ UAV; Ứng dụng nền tảng công nghệ 4.0 trong Địa không gian; Công nghệ trong quản lý và sử dụng tài nguyên đất bền vững; BIM - chuyển đổi số, tăng trưởng xanh, netzero,...; Các ứng dụng về dữ liệu lớn, học máy và trí tuệ nhân tạo; Công nghệ điện toán đám mây để xử lý, phân tích dữ liệu trong các lĩnh vực tài nguyên môi trường; Các vấn đề khoa học khác có liên quan đã được Ban Khoa học của Hội nghị thẩm định, biên tập.

Ban Tổ chức hy vọng rằng, Hội nghị không chỉ là diễn đàn để các nhà khoa học chia sẻ, thảo luận những kết quả nghiên cứu mới về các lĩnh vực khoa học liên quan mà còn là dịp để các tổ chức, cá nhân gặp gỡ, ký kết hợp tác nhằm nâng cao hiệu quả hoạt động khoa học và công nghệ trong các lĩnh vực chuyên môn.

Trong quá trình chuẩn bị tổ chức Hội nghị, Ban Tổ chức đã nhận được sự hỗ trợ cả về tinh thần và vật chất của các tổ chức và cá nhân. Ban Tổ chức Hội nghị xin gửi tới các tổ chức, cá nhân lời cảm ơn sâu sắc về sự ủng hộ, giúp đỡ quý báu đó.

Cuối cùng, thay mặt Ban Tổ chức, tôi chân thành cảm ơn Đảng ủy, Hội đồng Trường, Ban Giám hiệu Trường Đại học Mỏ - Địa chất và các đơn vị liên quan đã tạo mọi điều kiện thuận lợi, góp phần vào thành công của Hội nghị. Đặc biệt, trân trọng cảm ơn các tác giả bài báo, các phản biện, các nhà khoa học và các nhà tài trợ đã có đóng góp quan trọng vì sự thành công chung của Hội nghị.

Xin chân thành cảm ơn Nhà xuất bản Khoa học tự nhiên và Công nghệ đã giúp đỡ trong công tác xuất bản tuyển tập này.

Thay mặt Ban Tổ chức
Trưởng Bộ môn Đo ảnh và Viễn thám



TS Trần Trung Anh

MỤC LỤC

Địa không gian thông minh trong hành trình thực hiện các mục tiêu phát triển bền vững	1
<i>Võ Chí Mỹ*</i>	
Giao dịch bất động sản - Thực trạng và biện pháp hạn chế rủi ro về mặt pháp lý	6
<i>Nguyễn Thị Dung*</i>	
Ảnh hưởng của mật độ điểm đo trực tiếp đến độ chính xác nội suy dị thường trọng lực bước sóng dài khi tính độ sâu từ dị thường trọng lực	12
<i>Nguyễn Văn Sáng*</i>	
Kết hợp đa chỉ số phổ trong chiết tách đất xây dựng ở khu vực Hà Nội từ ảnh vệ tinh Sentinel-2	18
<i>Nguyễn Văn Trung*, Lê Thị Thu Hà, Nguyễn Thành Len, Lê Thị Minh Phương</i>	
Phân tích tự tương quan không gian trong nghiên cứu kiểu mẫu phân bố bệnh sốt xuất huyết Dengue: Thử nghiệm tại Thành phố Hồ Chí Minh	28
<i>Vũ Danh Tuyên*, Vương Thị Hòe, Phạm Thị Thu Hương, Nguyễn Tiến Thành</i>	
Sử dụng dữ liệu vệ tinh Sentinel-2 thành lập bản đồ sinh vật đáy khu vực quần đảo Côn Đảo, tỉnh Bà Rịa - Vũng Tàu, Việt Nam	34
<i>Nguyễn Văn Trung*, Lê Thị Thu Hà, Đoàn Thị Nam Phương, Phạm Thị Hương, Nguyễn Quang Minh, Phạm Vọng Thành</i>	
Nghiên cứu giải pháp tự động chiết xuất đối tượng đường giao thông trên ảnh vệ tinh Sentinel- 2 MSI	44
<i>Lê Văn Phú, Trịnh Lê Hùng*, Tô Thị Phương</i>	
Phát hiện cháy rừng đang xảy ra sử dụng dữ liệu vệ tinh Sentinel-3 SLSTR khu vực xã Khánh Bình Tây, huyện Trần Văn Thời, tỉnh Cà Mau	51
<i>Đoàn Thị Nam Phương*, Nguyễn Văn Trung, Lê Thị Thu Hà, Nguyễn Quang Minh</i>	
Công nghệ SBAS xử lý dữ liệu lớn cho xác định biến dạng bề mặt địa hình vùng Thành phố Hồ Chí Minh	59
<i>Nguyễn Minh Hải*</i>	
Nghiên cứu phân bố không gian gió vùng biển vịnh Bắc Bộ giai đoạn 1981 - 2020 từ dữ liệu viễn thám	65
<i>Nguyễn Ngọc Tuấn, Nguyễn Hoàng Minh, Đỗ Thị Phương Thảo*</i>	
Nghiên cứu ứng dụng tư liệu ảnh Sentinel-5P trong giám sát chất lượng không khí tại các đô thị lớn	72
<i>Lê Minh Hằng, Đỗ Thị Hoài*</i>	
Thành lập mô hình lún để chẩn đoán Nhà máy lọc dầu Dung Quất	82
<i>Nguyễn Thị Kim Thanh*, Nguyễn Hà</i>	
Nghiên cứu ứng dụng phương pháp phân loại hướng đối tượng và mô hình học máy trong xây dựng bản đồ lớp phủ từ ảnh viễn thám quang học	90
<i>Khúc Thành Đông*, Trần Đình Trọng, Trần Văn Anh, Đỗ Quang Vinh, Hà Trung Khiên, Đào Duy Toàn, Trương Xuân Quang</i>	
Xác định lún đất khu vực Cà Mau bằng chuỗi ảnh radar trên nền tảng điện toán đám mây Google Colab	97
<i>Hà Trung Khiên, Trần Văn Anh*, Khúc Thành Đông, Nguyễn Đình Huy</i>	
Ứng dụng công nghệ GIS và viễn thám hỗ trợ phân tích, quản lý tài nguyên rừng trong điều kiện biến đổi khí hậu ở tỉnh Đắk Lắk	106
<i>Nguyễn Sách Thành*, Trịnh Lê Hùng, Lê Văn Phú, Nguyễn Văn Trung</i>	

Ứng dụng công nghệ địa tin học trong đánh giá hạn hán ở Quảng Nam	114
<i>Phạm Thị Thanh Hoà*, Trần Hồng Hạnh, Lưu Thị Diệu Chinh</i>	
Nghiên cứu ứng dụng mô hình LSTM (Long Short-Term Memory) trong dự báo độ mặn từ chuỗi dữ liệu đa biến tại trạm Đại Ngãi, Sóc Trăng	122
<i>Bùi Duy Quỳnh*, Hà Thị Hằng, Lưu Thị Diệu Chinh, Trần Xuân Trường</i>	
Kiểm định giả thuyết thống kê trong bình sai lưới quan trắc chuyển dịch ngang đập công trình thủy điện	131
<i>Trần Trung Anh*, Nguyễn Quang Hà</i>	
Phân tích, lọc và hiển thị dữ liệu đám mây điểm LiDAR với thư viện Pandas và Matplotlib của Python	138
<i>Nguyễn Thị Hữu Phương*, Phạm Thị Hải Vân, Đào Thị Hồng Thắm</i>	
Nghiên cứu xác định sự biến động độ sâu ven đảo xa bờ sử dụng dữ liệu viễn thám Sentinel-2 trên nền tảng Google Earth Engine	146
<i>Nguyễn Như Hùng*, Lê Văn Phú, Nguyễn Văn Sơn</i>	
Nghiên cứu một số thuật toán huấn luyện mạng nơron nhân tạo trong dự báo độ lún công trình thủy điện	154
<i>Phạm Quốc Khánh*</i>	
Công nghệ máy bay không người lái và thuật toán bán tự động trong việc xác định hệ thống khe nứt chính trên mô đá ốp lát	163
<i>Phạm Văn Việt*, Nguyễn Anh Tuấn, Trần Trung Anh, Phạm Văn Hoà</i>	
Phương trình LiDAR tính tọa độ điểm đo trong công nghệ bay quét laser hàng không	172
<i>Trần Trung Anh*, Trần Hồng Hạnh, Lưu Hải Âu, Quách Mạnh Tuấn</i>	
Lập bản đồ biến động hệ thực vật vùng đất ven biển khu vực Khánh Hòa bằng ảnh Landsat đa thời gian trên nền tảng Google Earth Engine	179
<i>Trần Thanh Hà*</i>	
Quan trắc trượt lở đất đá theo thời gian thực ở tầng khai thác mỏ lộ thiên dựa trên công nghệ GNSS/CORS	188
<i>Phạm Công Khải*, Nguyễn Văn Hải</i>	
Dự báo thay đổi lớp phủ bề mặt bằng thuật toán học máy	199
<i>Trần Thị Hoà*, Trần Thị Ngọc, Trần Đình Trí</i>	
Xây dựng cơ sở dữ liệu GIS phục vụ công tác tìm kiếm và cứu nạn khu vực vịnh Bắc Bộ	207
<i>Nguyễn Quang Minh*, Nguyễn Văn Trung, Bùi Tiến Diệu</i>	
Phân tích lượng mưa tích lũy và nguy cơ trượt lở đất: Nghiên cứu và mô phỏng số tại Km10+950, Quốc lộ 34, tỉnh Hà Giang	221
<i>Đỗ Văn Đăng*, Nguyễn Châu Lân, Hà Thị Hằng, Bùi Duy Quỳnh, Dương Công Hiếu, Lưu Thị Diệu Chinh</i>	
Nghiên cứu một số vấn đề chuyển đổi tọa độ từ WGS-84 về hệ tọa độ trắc địa sử dụng tại Việt Nam	228
<i>Hoàng Ngọc Hà*</i>	

Địa không gian thông minh trong hành trình thực hiện các mục tiêu phát triển bền vững

Võ Chí Mỹ*
Hội Trắc địa - Bản đồ - Viễn thám Việt Nam

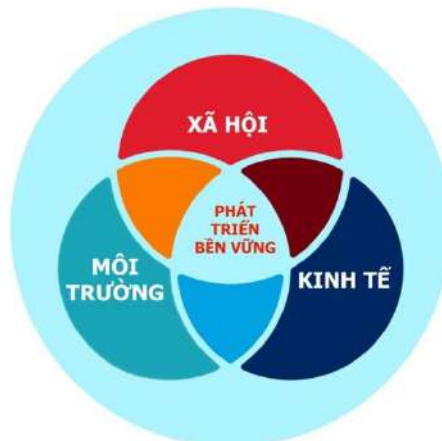
TÓM TẮT

Phát triển bền vững là xu thế tất yếu của thế giới. Trong hành trình thực hiện các mục tiêu phát triển bền vững (SDGs), công nghệ/dữ liệu địa không gian đóng vai trò quan trọng. Dữ liệu địa không gian đóng góp 80% trong các ngành kinh tế quốc dân và quốc phòng. Sự tích hợp các nền tảng công nghệ của cuộc Cách mạng 4.0 đã làm cho địa không gian thông minh hơn, hỗ trợ hiệu quả hơn các nội dung hoạt động kinh tế - xã hội và môi trường, thiết thực đẩy nhanh quá trình thực hiện các mục tiêu phát triển bền vững.

Từ khóa: Phát triển bền vững, công nghệ/dữ liệu địa không gian thông minh, GeoAI.

1. Phát triển bền vững - xu thế tất yếu

Từ thuở sơ khai của lịch sử, xã hội loài người luôn hướng tới sự phát triển vì phát triển là nhân tố đảm bảo sự tồn tại và nâng cao chất lượng cuộc sống. Phát triển là nhu cầu của mọi thời đại, là mục tiêu của mọi quốc gia và trách nhiệm của mọi Chính phủ. Phát triển là hoạt động tự nhiên của mỗi cá nhân và cộng đồng. Bốn cuộc cách mạng công nghiệp lần lượt ra đời bắt đầu từ giữa thế kỷ XVIII đã đưa nhân loại vào kỷ nguyên phát triển mới với các nền công nghiệp hiện đại. Bắt đầu từ cuối thế kỷ XX, loài người thức tỉnh và lo lắng cho cuộc sống của chính mình và đồng loại khi chứng kiến tình trạng cạn kiệt tài nguyên đồng hành với thảm họa môi trường kể cả ô nhiễm môi trường, suy thoái và tai biến môi trường. Sự phát triển thiếu kiểm soát gây ra biến đổi khí hậu và các hiện tượng thời tiết cực đoan là một thực tế hiện hữu và là mối hiểm họa lớn nhất trong thế kỷ XXI và xa hơn nữa. Đứng trước nguy cơ Trái đất bị huỷ diệt, đe dọa sự tồn tại và phát triển của xã hội loài người, thế giới quyết định chuyển đổi mô hình phát triển mới gọi là phát triển bền vững. Theo định nghĩa: “Phát triển bền vững là quá trình phát triển đáp ứng nhu cầu hiện tại mà không làm tổn hại đến khả năng của các thế hệ tương lai đáp ứng nhu cầu của họ”.



Hình 1. Phát triển bền vững - Sự kết hợp hài hoà các hợp phần kinh tế, xã hội và môi trường

Phát triển bền vững là sự kết hợp hài hoà của cả ba hợp phần: kinh tế, xã hội và môi trường (Hình 1). Thuật ngữ “phát triển bền vững” được nhắc đến lần đầu tiên năm 1992 tại Hội nghị Thượng đỉnh Trái đất tại Rio de Janeiro. Cùng với thời gian, lộ trình và mục tiêu phát triển bền vững được điều chỉnh, bổ sung tại các hội nghị thượng đỉnh New York (2000), Johannesburg (2002), Rio de Janeiro (2012). Năm 2015, Hội nghị Thượng đỉnh Liên hợp quốc tại New York đã thông qua Văn kiện “Chuyển đổi thế giới của chúng ta: Chương trình nghị sự 2030 vì sự phát triển bền vững”(Chương trình Nghị sự 2030) với 17 mục tiêu (SDGs) phải thực hiện được vào năm 2030 (Liên hợp quốc, 2015). Chương trình Nghị sự 2030 đã đề ra tầm nhìn chiến lược mới, thể hiện khát

* Tác giả liên hệ
Email: vochimy@humg.edu.vn

vọng chung của nhân loại được sống trong một thế giới hoà bình, công bằng, an toàn, xanh và sạch; tạo khung pháp lý và định hướng cho các quốc gia trong lộ trình thực hiện các mục tiêu phát triển bền vững cả về kinh tế - xã hội và môi trường.

2. Lộ trình phát triển bền vững ở Việt Nam

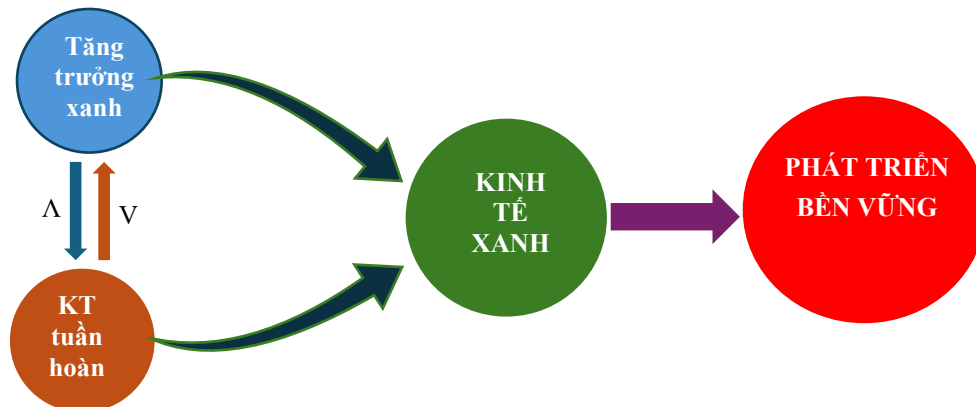
Việt Nam là một trong những quốc gia đầu tiên cam kết thực hiện các mục tiêu phát triển bền vững từ Chương trình Nghị sự 21 đến Chương trình Nghị sự 2030. Những nỗ lực của Việt Nam được cụ thể hoá trong nhiều văn kiện của Đảng, Nhà nước và Chính phủ. Hai văn bản pháp lý gần đây nhất bao gồm:

- Nghị quyết 136/2020 NQ-CP ngày 29/9/2020 của Chính phủ: Về phát triển bền vững;
- Quyết định số 841/2023 QĐ-TTg ngày 14/7/2023 của Thủ tướng Chính phủ: Ban hành lộ trình thực hiện các mục tiêu phát triển bền vững Việt Nam đến năm 2030.

Từ nội hàm các nghị quyết và quyết định của Chính phủ (Chính phủ, 2020) (Thủ tướng Chính phủ, 2023) có thể khẳng định rằng: Để hướng tới phát triển bền vững, Việt Nam sẽ chuyển đổi sang nền kinh tế xanh theo lộ trình tăng trưởng xanh và kinh tế tuần hoàn. Nền kinh tế xanh được Liên hợp quốc chính thức định danh và xác định là nền kinh tế tất yếu hướng tới các mục tiêu phát triển bền vững của các quốc gia trên thế giới. Chỉ có kinh tế xanh mới đồng thời kết hợp hài hoà hạnh phúc cho con người, công bằng xã hội, sử dụng hiệu quả tài nguyên. Nghị quyết Đại hội lần thứ 13 của Đảng Cộng sản Việt Nam đã khẳng định chủ trương xây dựng và phát triển nền kinh tế xanh. Với Việt Nam, việc chuyển hướng từ kinh tế nâu sang kinh tế xanh không chỉ là lựa chọn tất yếu mà còn là cơ hội để hội nhập với xu thế phát triển bền vững của thế giới.

Để hướng tới nền kinh tế xanh, các mô hình tăng trưởng xanh đã và đang được thực hiện ở Việt Nam. Là một nước đang phát triển, tăng trưởng xanh được coi là giai đoạn chuyển tiếp, cần thiết trong hành trình chuyển đổi toàn diện từ kinh tế nâu sang kinh tế xanh. Tháng 7 năm 2022, Thủ tướng Chính phủ Việt Nam đã ký quyết định số 822/QĐ-TTg “Phê duyệt kế hoạch hành động quốc gia về tăng trưởng xanh giai đoạn 2021-2030”. Mục tiêu chiến lược tăng trưởng xanh của Việt Nam là góp phần thúc đẩy cơ cấu lại nền kinh tế, gắn với đổi mới mô hình tăng trưởng nhằm đạt thịnh vượng về kinh tế, bền vững môi trường và công bằng xã hội, hướng tới nền kinh tế xanh.

Kinh tế tuần hoàn là một trong những cấu phần hướng tới nền kinh tế xanh và phát triển bền vững. Kinh tế tuần hoàn đang được thúc đẩy ở Việt Nam nhằm thay thế nền kinh tế tuyến tính. Trong bối cảnh hoà nhập quốc tế, kinh tế tuần hoàn đã và đang trở thành xu thế tất yếu trong các nước đang phát triển, trong đó có Việt Nam. Năm 2022, thủ tướng Chính phủ đã ban hành Quyết định số 687/QĐ-TTg Quyết định phê duyệt đề án phát triển kinh tế tuần hoàn ở Việt Nam (Hình 2).



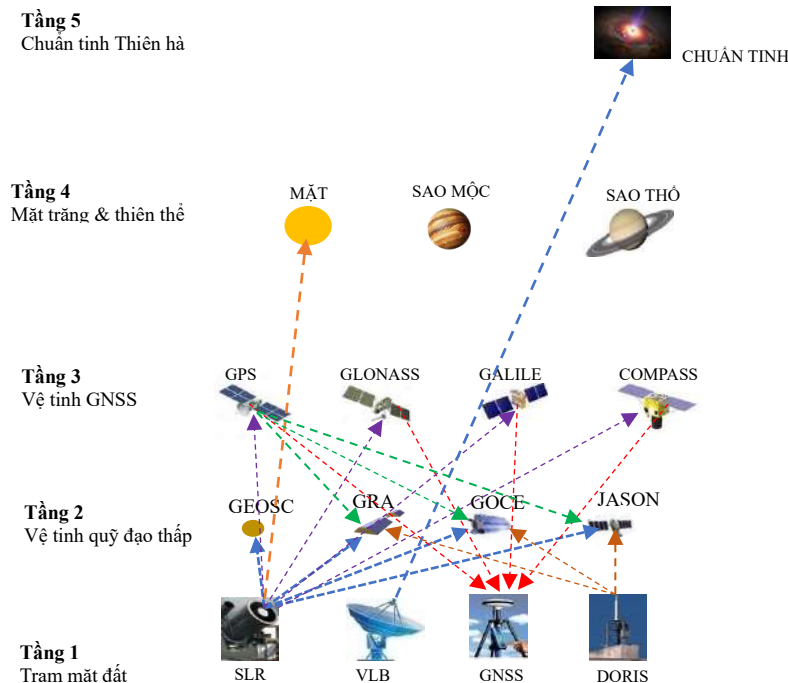
Hình 2. Tăng trưởng xanh và kinh tế tuần hoàn – lộ trình tất yếu hướng tới phát triển bền vững

3. Sự phát triển của công nghiệp địa không gian

Từ thời tiền sử, con người đã biết sử dụng dữ liệu địa không gian để phục vụ cho cuộc sống, duy trì sự tồn tại và phát triển. Mức độ khai thác và sử dụng dữ liệu địa không gian ngày càng gia tăng mạnh mẽ khi thế giới trải qua các cuộc cách mạng công nghiệp. Từ những phép đo đạc thô sơ đến các thiết bị quang cơ và hôm nay, trong bối cảnh cách mạng công nghiệp lần 4.0, các thiết bị và công nghệ hiện đại thông minh như GNSS, viễn thám, UAV, Lidar,... đang sản xuất, khai thác hàng loạt các loại hình dữ liệu địa không gian phong phú, đa dạng. Các thiết bị và công nghệ địa không gian ngày càng được hiện đại hoá, phân bố trên nhiều tầng quan trắc cho phép thu thập, khai thác thông tin ngày càng phong phú và đa dạng, từ các thiết bị trên mặt đất đến vệ tinh quỹ đạo thấp (LEO), hệ thống vệ tinh dẫn đường toàn cầu (GNSS) đến hệ thống nghiên cứu Mặt trăng và các thiết bị giám sát chuẩn tinh và thiên thể ngoài Thiên hà. Đánh giá cao vai trò của dữ liệu và thông tin địa không gian trong lộ trình thực hiện các mục tiêu phát triển bền vững, Ủy ban Quản lý Thông tin địa không gian toàn cầu của Liên hợp quốc (UN-GGIM) đã ban hành hướng dẫn khung thông tin địa không gian tích hợp (UN-IGIF) (United Nation,

2023). Mục đích của UN-IGIF là hướng dẫn cho các quốc gia tăng cường cơ sở hạ tầng và mô hình quản lý, bao gồm một bộ nguyên tắc và các giải pháp khai thác, quản lý, sử dụng, tránh xung đột, dư thừa và dễ dàng chia sẻ thông tin địa không gian. UN-IGIF được thiết kế để hỗ trợ các quốc gia tích hợp thông tin địa không gian vào quá trình ra quyết định, thúc đẩy việc sử dụng thông tin địa không gian trong các lĩnh vực khác nhau, kể cả quản lý, giám sát môi trường; ứng phó, chống chịu với thảm họa thiên nhiên và biến đổi khí hậu,... Trong kỷ nguyên kỹ thuật số, vai trò của dữ liệu địa không gian bao gồm dữ liệu không gian và phi không gian đóng vai trò quan trọng trong sự thúc đẩy sự phát triển. Thế giới đang ở trong “nền kinh tế dữ liệu”, các đổi mới dựa trên dữ liệu DDI (Data Documentation Initiative) ngày càng trở nên quan trọng hơn trong mọi hoạt động vì một xã hội bền vững. Cốt lõi của cuộc “cách mạng dữ liệu” là hạ tầng dữ liệu địa không gian, là nền tảng gồm nhiều thành phần để chia sẻ, khai thác giữa các đối tượng cung cấp và sử dụng. Nền kinh tế dữ liệu dựa trên nguyên tắc “dữ liệu làm ra thông tin, thông tin làm ra giá trị sản phẩm”. Cơ sở hạ tầng dữ liệu địa không gian được coi là tài nguyên có giá trị nhất để kích thích mọi mặt tăng trưởng của xã hội, nhân tố then chốt của cuộc cách mạng công nghiệp 4.0.

Trong những năm gần đây, hệ thống trắc địa quan sát toàn cầu GGOS được Hội trắc địa quốc tế IAG thiết lập nhằm đóng góp công nghệ, khung tham chiếu, cung cấp cơ sở dữ liệu và thông tin địa không gian cho các nội dung quan sát và nghiên cứu Trái đất và sự biến động các thành phần tài nguyên - môi trường. Để thực hiện sứ mệnh quan sát Trái đất với nhiều nhiệm vụ khác nhau, GGOS đã được sử dụng một hạ tầng công nghệ hiện đại và đầy đủ, kể cả các hạ tầng mặt đất và không gian. Mạng lưới trên mặt đất, bên cạnh các thiết bị hiện đại như các máy đo trọng lực siêu dẫn, máy đo trọng lực tuyệt đối, các phương pháp công nghệ trắc địa không gian đa năng như: giao thoa cạnh đáy dài VLBI, đo khoảng cách laser SLR, đo khoảng cách laser Mặt trăng LLR và kỹ thuật Doppler DORIS. Trong những năm gần đây, các hệ thống định vị toàn cầu không ngừng được hoàn thiện và nâng cấp. Vệ tinh thế hệ mới GPS III SV01 đã được đưa lên quỹ đạo năm 2019 với độ chính xác định vị gấp ba lần và khả năng chống nhiễu gấp tám lần so với vệ tinh thế hệ cũ. Vệ tinh GPS III có kết cấu vững chắc, có tuổi thọ 15 năm nghĩa là 25% dài hơn so với vệ tinh thế hệ cũ. Các hệ thống định vị GLONASS, Galileo, COMPASS cũng đang được nâng cấp, hoàn thiện kể cả số lượng vệ tinh và công suất tín hiệu. Ở quỹ đạo tầm thấp LEO, ngoài các vệ tinh trọng lực CAMP, GRACE, GOCE còn có LAGEOS là vệ tinh laser của NASA nghiên cứu sự chuyển động của các mảng kiến tạo và JASON-2 là vệ tinh đo độ cao để xác định mực nước biển và đại dương.



Hình 3. Hệ thống địa không gian đa tầng

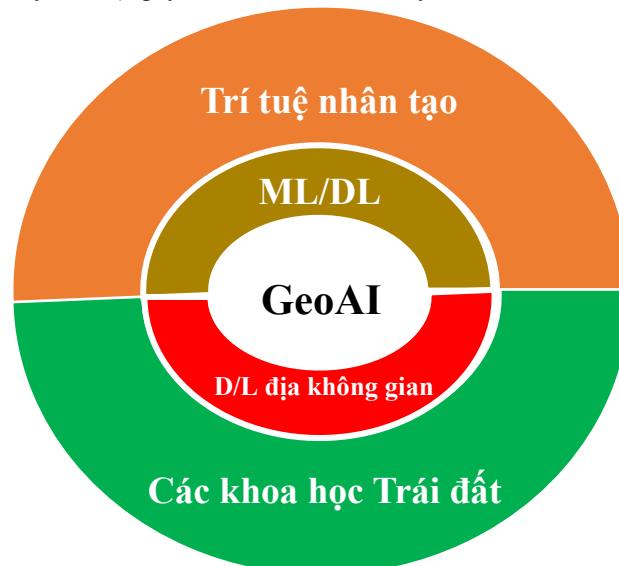
Sau 30 năm số hoá và quy trình số, từ năm 2016 đến nay, thế giới đang ở trong làn sóng chuyển đổi số là quá trình chuyển các hoạt động quản lý, mô hình tổ chức, sản xuất, kinh doanh,... từ xã hội thực lên không gian số. Từ kinh nghiệm và thành tựu của chuyển đổi số trong tất cả các lĩnh vực từ chính phủ số, kinh tế số và xã hội số, càng ngày, thế giới càng nhận thức được chuyển đổi số là xu thế tất yếu của xã hội tương lai (Thomas M. Siebel, 2019). Sự phát triển nhanh chóng của công nghệ đã biến chuyển đổi số trở thành động lực thiết yếu cho quá trình hiện thực hoá các mục tiêu phát triển bền vững, đóng góp tích cực và hiệu quả trong cả ba hợp phần phát triển

kinh tế - xã hội và môi trường. Chương trình chuyển đổi số phải dựa trên hạ tầng số trong đó, hạ tầng dữ liệu là then chốt quyết định sự thành công của chuyển đổi số. Với 80% dữ liệu trong các hoạt động xã hội của một đất nước, hạ tầng dữ liệu địa không gian (GDI) là cấu phần quan trọng là nền tảng trong hạ tầng dữ liệu. GDI cho phép tích hợp các bộ dữ liệu địa không gian đa dạng từ nhiều nguồn khác nhau, tạo điều kiện cho các hệ thống khác nhau làm việc với một bộ thông tin không gian thống nhất và toàn diện. GDI cho phép phân tích tối ưu, trực quan hoá thông tin hỗ trợ ra quyết định trong nhiều lĩnh vực kinh tế quốc dân và quốc phòng; quy hoạch và quản lý đô thị, quản lý môi trường và ứng phó thiên tai,... GDI tạo môi trường tương tác với công dân tạo điều kiện cho công chúng tham gia các chương trình tham vấn cộng đồng đóng góp vào các quy trình ra quyết định trong các hoạt động kinh tế - xã hội và môi trường. GDI là cơ sở tạo ra bản sao số (digital twin) sử dụng dữ liệu địa không gian thời gian thực để mô phỏng các đối tượng thực thể cho phép quản lý và giám sát các đối tượng kịp thời. Các mục tiêu phát triển bền vững không thể thực hiện dễ dàng nếu không có chuyển đổi số. Chương trình chuyển đổi số quốc gia không dễ dàng thực hiện nếu không có dữ liệu và thông tin địa không gian.

Trong kỷ nguyên số của cách mạng công nghiệp 4.0, công nghệ/dữ liệu và thông tin địa không gian trực tiếp hoặc gián tiếp thúc đẩy hình thành và đổi mới các ngành công nghệ và ngược lại. Các giải pháp GeoAI, in 3D, thực tế tăng cường và thực tế ảo (AR/VR), xe tự lái,... là những đổi mới sáng tạo đã góp phần nâng cao hiệu quả kinh tế, chuyển đổi phương thức làm việc, mang lại sản phẩm và dịch vụ tốt hơn cho con người, nâng cao mức sống; giám sát hiệu quả sự biến động các thành phần tài nguyên và môi trường (Võ Chí Mỹ, 2024; UN-GGIM, 2021).

4. Địa không gian thông minh (GeoAI)

Geospatial Artificial Intelligence - GeoAI (tạm dịch là địa không gian thông minh) được thảo luận đầu tiên tại Hội thảo của Hiệp hội Máy tính (ACM - Association for Computing Machinery) của Nhóm quan tâm đặc biệt về thông tin không gian SIGSPATIAL tổ chức tháng 11/2017 tại Redondo Beach, California (Hoa Kỳ). Kể từ đó, GeoAI không ngừng được nghiên cứu và phát triển dựa vào sự phát triển của trí tuệ nhân tạo, khoa học dữ liệu, phân tích không gian và điện toán. Trí tuệ nhân tạo đã được ứng dụng trước đó trong các ngành khoa học Trái đất nói chung và địa lý nói riêng. Nhưng GeoAI được hình thành và phát triển kể từ khi xuất hiện các mô hình học máy và học sâu với khả năng xử lý hiệu quả dữ liệu lớn, đa dạng, phức tạp có cấu trúc, bán cấu trúc và phi cấu trúc; tự động phân loại, phân tích không gian và chiết xuất các đối tượng từ dữ liệu thô, khả năng khái quát hoá và chuyển giao huấn luyện,... (Nguyễn Phi Sơn, 2024; Krzysztof Janowicz, 2019) (Hình 4).



Hình 4. Học máy/học sâu - công cụ then chốt của GeoAI

Ngoài trí tuệ nhân tạo, hầu hết các nền tảng công nghệ CM 4.0 đều tham gia vào GeoAI:

- Big data cung cấp khối lượng dữ liệu lớn, đa dạng theo thời gian thực, mô hình hoá không gian và thời gian,... Cũng lưu ý rằng: công cụ then chốt của GeoAI là mô hình học sâu chỉ có thể phát huy khả năng nổi trội này khi sử dụng khối lượng dữ liệu lớn đa nguồn với các định dạng khác nhau.
- Điện toán đám mây cung cấp cơ sở hạ tầng có thể lưu trữ mở rộng, quản lý, phân tích, cho phép các ứng dụng GeoAI xử lý hiệu quả các khối lượng dữ liệu lớn từ nhiều nguồn khác nhau với các yêu cầu tính toán khác nhau.
- Với vai trò thu thập, lưu trữ, quản lý, phân tích và trực quan hoá thông tin, GIS là công cụ then chốt của GeoAI. Sự tích hợp với các thuật toán học máy cho phép GIS giải các bài toán địa không gian phức tạp nhằm nâng cao độ tin cậy quy trình ra quyết định các vấn đề kinh tế - xã hội và môi trường.
- Cảm biến IoT là nguồn cung cấp dữ liệu theo không gian và thời gian thực. Tích hợp IoT với GeoAI sẽ nâng

cao độ chính xác thông tin, hỗ trợ hiệu quả ra quyết định trong nhiều lĩnh vực khác nhau.

- Viễn thám là nền tảng của GeoAI. Viễn thám cung cấp dữ liệu địa không gian từ nhiều nguồn khác nhau như ảnh vệ tinh, ảnh hàng không, máy bay không người lái, Lidar,... Hệ thống dữ liệu raster này chính là đối tượng xử lý của các mô hình GeoAI hiện nay. Được kết hợp phân tích bằng các thuật toán AI và học sâu, dữ liệu này cung cấp các thông tin về hiện trạng bề mặt Trái đất và sự biến động các thành phần tài nguyên - môi trường (Qiangqiang Yuan, 2020) - những thông tin đặc biệt quan trọng để nhân loại tìm kiếm các giải pháp ứng phó và chống chịu với thảm họa do biến đổi khí hậu toàn cầu.

- Chuỗi khối (blockchain) bảo đảm dữ liệu địa không gian sử dụng trong các mô hình GeoAI được bảo đảm về tính toàn vẹn, tính bảo mật, nguồn gốc dữ liệu, chia sẻ dữ liệu phi tập trung,... Nhờ thế, blockchain có thể tích hợp an toàn dữ liệu địa không gian từ nhiều nguồn khác nhau, cho phép GeoAI đưa ra quyết định chính xác và kịp thời.

GeoAI được coi là cuộc cách mạng trong công nghệ/dữ liệu địa không gian. Sự tích hợp các nền tảng công nghệ 4.0 trong mô hình GeoAI sẽ nâng cao độ chính xác, độ tin cậy, giải quyết được nhiều bài toán đa dạng phức tạp trên nhiều lĩnh vực khác nhau theo không gian và thời gian thực (Nguyễn Phi Sơn, 2024; ESCAP, 2022; Pengyuan Liu, 2022). Nghiên cứu triển khai GeoAI ở Việt Nam là xu thế tất yếu vì sự phát triển của ngành công nghiệp địa không gian phù hợp với xu hướng phát triển của thế giới. Nhờ những khả năng và ưu điểm nổi trội, địa không gian thông minh sẽ là động lực thúc đẩy đất nước hoàn thành các mục tiêu phát triển bền vững theo cam kết quốc tế của Chính phủ.

5. Kết luận

Khi được kết hợp với các nền tảng công nghệ 4.0, dữ liệu địa không gian trở thành công cụ thông minh cung cấp thông tin tin cậy cho các nhà hoạch định chính sách lập kế hoạch và ra quyết định. Các chỉ số về sự phát triển kinh tế - xã hội, sự biến động các thành phần tài nguyên và môi trường được nhanh chóng xác định là cơ sở để giám sát và đánh giá tiến độ thực hiện các mục tiêu phát triển bền vững. Việc ứng dụng dữ liệu địa không gian thông minh sẽ nâng cao tính minh bạch và công bằng của các dự án phát triển bền vững. Dữ liệu địa không gian giúp dự báo, cảnh báo và giảm thiểu tác động của biến đổi khí hậu và thời tiết cực đoan - một thách thức to lớn đối với sự phát triển bền vững. Trong bối cảnh đô thị hóa nhanh chóng, việc xây dựng các thành phố thông minh và bền vững là một trong những mục tiêu quan trọng của Việt Nam. Dữ liệu địa không gian giúp quy hoạch đô thị một cách hợp lý, tối ưu hóa việc sử dụng đất đai, phát triển hạ tầng giao thông, logistic và quản lý tài nguyên môi trường đô thị. Dữ liệu địa không gian thông minh đóng vai trò then chốt thúc đẩy nhanh hơn hành trình thực hiện 17 mục tiêu phát triển bền vững (SDGs) ở Việt Nam.

Tài liệu tham khảo

- Chính phủ (2020). Nghị quyết 136/2020 NQ-CP ngày 29/9/2020 của Chính phủ: Về phát triển bền vững.
- ESCAP (2022). Strategic Foresight to Applications of Geospatial Artificial Intelligence (GeoAI) to Achieve Disaster-related Sustainable Development Goals. Space Applications Sections Information and Communications Technology and Disaster Risk Reduction Division.
- Krzysztof Janowicz, S. G. (2019). GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, 625-636.
- Liên hợp quốc (2015). Chuyển đổi thế giới của chúng ta: Chương trình nghị sự 2030 vì sự phát triển bền vững. New York: Hội nghị Thượng đỉnh Liên hợp quốc.
- Nguyễn Phi Sơn, N. Q. (2024). GeoAI: Công nghệ và ứng dụng trong địa không gian. Hội thảo khoa học và công nghệ với chủ đề "Công nghệ địa không gian thông minh trong quản lý TN&MT". Hà Nội: Viện Khoa học Đo đạc và Bản đồ.
- Pengyuan Liu, F. B. (2022). A review of spatially-explicit GeoAI applications in Urban Geography. *International Journal of Applied Earth Observation and Geoinformation-Elsevier*.
- Qiangqiang Yuan, H. S. (2020). Deep Learning in Environmental Remote Sensing: Achievements and Challenges. *Remote sensing of Environment*.
- Thomas M. Siebel. (2019). Digital Transformation: Survive and Thrive in an Era of Mass Extinction. RosettaBooks.
- Thủ tướng Chính phủ (2023). Quyết định số 841/2023 QĐ-TTg ngày 14/7/2023 của Thủ tướng Chính phủ: Ban hành lộ trình thực hiện các mục tiêu phát triển bền vững Việt Nam đến năm 2030.
- UN-GGIM (2021). Geospatial industry advancing sustainable development goals. *Geospatial World*.
- United Nation (2023). Integrated Geospatial Information Framework (UN-IGIF). Geomares Publishing, GIM Issue 1-2023.
- Võ Chí Mỹ, V. N. (2024). Vai trò của công nghệ và dữ liệu địa không gian trong hành trình thực hiện mục tiêu các mô hình tăng trưởng mới của đất nước. Hội thảo khoa học và công nghệ với chủ đề "Công nghệ địa không gian thông minh trong quản lý TN&MT", Hà Nội, Viện Khoa học Đo đạc và Bản đồ.

Giao dịch bất động sản - Thực trạng và biện pháp hạn chế rủi ro về mặt pháp lý

Nguyễn Thị Dung^{1,2*}

¹Khoa Trắc địa - Bản đồ và Quản lý đất đai, Trường Đại học Mô - Địa chất,

²Nhóm nghiên cứu Quản lý đất đai và Công nghệ địa chính phục vụ phát triển bền vững

TÓM TẮT

Giao dịch bất động sản là nội dung quan trọng của thị trường bất động sản, giao dịch có giá trị lớn và ảnh hưởng tới nhiều bên liên quan nên yêu cầu thủ tục pháp lý chặt chẽ. Thực tế có nhiều rủi ro và nhiều tranh chấp xảy ra khi thực hiện giao dịch bất động sản trong thời gian qua. Mục đích bài viết là làm rõ tầm quan trọng của vấn đề pháp lý bất động sản; thông tin pháp lý đầy đủ của bất động sản, một số rủi ro khi giao dịch bất động sản không đủ điều kiện và các giải pháp. Phương pháp được sử dụng gồm phương pháp thu thập tài liệu, số liệu và phương pháp phân tích, đánh giá, tổng hợp; bài viết cũng đề xuất một số biện pháp hạn chế rủi ro pháp lý bất động sản khi các bên tham gia giao dịch. Giao dịch bất động sản là một trong những lĩnh vực tồn tại nhiều rủi ro đối với cả người mua lẫn người bán. Bên cạnh những khía cạnh kinh tế phức tạp, việc tham gia vào thị trường này còn đòi hỏi sự nắm vững trong việc hiểu biết về pháp lý. Pháp lý bất động sản cũng là yếu tố có ảnh hưởng trực tiếp tới quyền lợi, nghĩa vụ của cả chủ đầu tư lẫn các đối tượng liên quan, là nền tảng cốt lõi xác định quyền và nghĩa vụ của người mua, người bán và các bên liên quan khác. Việc rõ ràng cũng như đảm bảo các vấn đề pháp lý là một trong những cách để tránh việc xảy ra tranh chấp sau này. Xây dựng các quy trình đúng và tuân thủ quy định luật pháp sẽ mang lại một môi trường giao dịch bất động sản tốt hơn, thúc đẩy sự phát triển bền vững và tạo dựng niềm tin cho tất cả chủ thể tham gia vào thị trường. Vì vậy, việc tìm hiểu kỹ tính pháp lý bất động sản trước khi tiến hành ký hợp đồng giao dịch là một trong những lưu ý quan trọng đối với các nhà đầu tư.

Từ khóa: Pháp lý bất động sản, luật đất đai, giao dịch bất động sản.

1. Đặt vấn đề

Pháp lý đóng vai trò vô cùng quan trọng trong đầu tư bất động sản (BDS) để bảo đảm quyền lợi giữa các bên khi thực hiện giao dịch mua bán bất động sản. Đáng chú ý là những thủ tục và giấy tờ pháp lý liên quan phải theo đúng quy định của Nhà nước, bảo đảm phù hợp với từng loại sản phẩm bất động sản. Pháp lý trong bất động sản càng rõ ràng càng bảo đảm được vốn đầu tư an toàn và tăng khả năng sinh lời. Đặc biệt, bất động sản là một tài sản lớn nên với nhiều chủ đầu tư, họ sẽ quan tâm đến tính pháp lý đầu tiên. Nhà đầu tư cần phải chắc chắn đây là tài sản thuộc quyền sở hữu, quyền sử dụng của mình. Nếu có tranh chấp và chuyển nhượng sẽ được pháp luật bảo vệ.

Hiện nay, có rất nhiều vụ lùm xùm về mặt pháp lý bất động sản, dự án không được cấp phép, dự án không có thực khiến nhà đầu tư gặp rủi ro và gây ra nhiều thiệt hại. Do đó, việc tìm hiểu kỹ pháp lý bất động sản trước khi tiến hành ký hợp đồng mua bán là điều nhà đầu tư cần phải đặc biệt quan tâm.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Dữ liệu và phương pháp điều tra thu thập tài liệu, số liệu

Thu thập các tài liệu, số liệu về tình hình ban hành văn bản quy phạm pháp luật đất đai; về lĩnh vực kinh tế đất đai; dịch vụ hành chính công, thị trường bất động sản. Các tài liệu được thu thập tại các cơ quan như Bộ Tài nguyên và Môi trường, các địa phương (thông qua các cổng thông tin điện tử).

Các văn bản về giao dịch bất động sản, phát triển thị trường bất động sản từ khi có Luật Đất đai 2013 đến nay sử dụng trong nghiên cứu này như: (i) Nghị quyết số 18-NQ/TW của Ban Chấp hành Trung ương Đảng khóa XIII; (ii) Luật Đất đai năm 2013, Luật Nhà ở 2014, Luật Kinh doanh bất động sản 2014; (iii) Các văn bản của Chính phủ như: Chỉ thị số 01/CT-TTg ngày 03/01/2018 của Thủ tướng Chính phủ, Nghị định số 44/2014/NĐ-CP; Nghị định số 117/2015/NĐ-CP, Nghị định số 25/2020/NĐ-CP, Nghị định số 69/2021/NĐ-CP, Nghị định số 49/2021/NĐ-CP, Nghị định số 44/2022/NĐ-CP, Nghị định số 16/2022/NĐ-CP.

Ngoài ra, nghiên cứu này còn sử dụng các báo cáo, đánh giá của Bộ Tài nguyên và Môi trường và của một số một địa phương, cơ quan khác.

* Tác giả liên hệ

Email: nguyenthidung@humg.edu.vn

2.2. Phương pháp phân tích, đánh giá tổng hợp

Từ các số liệu, tài liệu thực trạng tiến hành tổng hợp, lựa chọn các cơ sở lý luận liên quan đến vấn đề nghiên cứu: quản lý đất đai, giao dịch bất động sản, đồng thời nắm bắt được các chủ trương, chính sách của Đảng, định hướng của Nhà nước liên quan đến quản lý đất đai nói chung và thị trường bất động sản nói riêng trong thời gian tới.

3. Nội dung nghiên cứu

3.1. Tầm quan trọng của vấn đề pháp lý trong giao dịch bất động sản

Pháp lý bất động sản được hiểu là các hồ sơ giấy tờ pháp lý của các dự án bất động sản đó và đặc biệt những hồ sơ pháp lý đó phải đúng theo quy định của pháp luật.

Pháp lý bất động sản có vai trò quan trọng, nhằm đảm bảo quyền lợi giữa các bên khi thực hiện các giao dịch mua bán sản phẩm bất động sản nhà nước quy định các thủ tục, giấy tờ pháp lý liên quan đảm bảo phù hợp với từng loại sản phẩm. Pháp lý bất động sản là yếu tố quan trọng, được đặt lên hàng đầu trong việc mua bán và đầu tư. Các nhà đầu tư không những quan tâm đến vị trí, kết nối của dự án, mà còn quan tâm hơn nữa về an toàn pháp lý dự án. Nhằm đảm bảo cho quyền lợi giữa các bên khi thực hiện những giao dịch mua bán sản phẩm bất động sản, nhà nước quy định thủ tục và các loại giấy tờ pháp lý liên quan. Đặc biệt, hồ sơ pháp lý đó phải đúng theo quy định của pháp luật, đảm bảo phù hợp với từng loại sản phẩm bất động sản.

Theo Rich Nguyen (2023), những bất động sản có pháp lý rõ ràng sẽ đem lại những lợi ích như sau:

(i) Thu hút được sự quan tâm của nhiều nhà đầu tư bất động sản

Nhà đầu tư càng ngày càng chú ý đến những bất động sản có tính pháp lý minh bạch, rõ ràng. Khách hàng sẵn sàng trả mức phí cao hơn cho những bất động sản có tính pháp lý đầy đủ, rõ ràng. Trong bối cảnh xã hội có nhiều dự án chậm trễ, những bất động sản có pháp lý rõ ràng sẽ được ưu tiên hơn.

(ii) Khả năng an toàn của bất động sản có pháp lý đầy đủ, rõ ràng cao hơn

Pháp lý trong bất động sản càng rõ ràng càng bảo đảm được vốn đầu tư an toàn và tăng khả năng sinh lời. Đặc biệt, bất động sản là một tài sản lớn nên với nhiều nhà đầu tư sẽ quan tâm đến tính pháp lý đầu tiên.

3.2. Pháp lý đầy đủ của một dự án bất động sản

Pháp lý bất động sản có tầm ảnh hưởng rất lớn và là một trong những yếu tố mà người mua, nhà đầu tư (Rich Nguyen, 2023) cần phải đặc biệt lưu ý để tránh những tranh chấp về sau, bởi trên thực tế, rất nhiều trường hợp đã gặp phải những vấn đề liên quan tới tranh chấp và kiện tụng kéo dài vì bất động sản không có giấy tờ, hồ sơ pháp lý rõ ràng.

Trước khi giao dịch bất cứ loại hình bất động sản nào cần phải tìm hiểu kỹ tính pháp lý bất động sản một cách đầy đủ và đúng nội dung. Pháp lý dự án là toàn bộ những giấy tờ, hồ sơ mà một dự án đầu tư BĐS cần phải có theo quy định của pháp luật (Luật Đất đai 2013, Luật Đầu tư 2020, Luật Kinh doanh bất động sản 2014, Luật Nhà ở 2014, Nghị định số 148/2020/NĐ-CP,...).

Các loại giấy tờ pháp lý cơ bản (Admin, 2021) bao gồm:

(1) Giấy phép kinh doanh bất động sản, quyết định chủ trương đầu tư, giấy chứng nhận đầu tư, quyết định phê duyệt dự án của chủ đầu tư

Các loại giấy phép này chứng nhận cho việc chủ đầu tư dự án có quyền được kinh doanh và đầu tư bất động sản hay không. Một chủ thể thông thường khi có quyền sử dụng đất mà không có các loại giấy tờ trên thì sẽ không có quyền thương mại mở bán công khai các sản phẩm của dự án.

Khi nhà đầu tư xem xét giấy phép kinh doanh của chủ đầu tư dự án, cần lưu ý:

(i) Đúng lĩnh vực bất động sản và phải có chức năng xây dựng, đầu tư kinh doanh;

(ii) Các loại giấy tờ trên được ký và đóng dấu đỏ của cơ quan có thẩm quyền, còn hiệu lực.

(2) Các giấy tờ về quyền sử dụng đất

- Quyết định giao đất

Chủ thể có nhu cầu sử dụng đất sẽ được Nhà nước trao quyền sử dụng đất thông qua quyết định giao đất. Người sử dụng đất được nhà nước giao đất gồm: giao đất có thu tiền và giao đất không thu tiền sử dụng đất. Trong trường hợp thực hiện các dự án bất động sản thì được Nhà nước giao đất có thu tiền.

Bên cạnh đó, Quyết định giao đất là cơ sở để cơ quan nhà nước có thẩm quyền ra quyết định cấp Giấy chứng nhận quyền sử dụng đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất cho dự án.

Điều 59, Luật Đất đai 2013 quy định về thẩm quyền giao đất như sau:

(i) Ủy ban nhân dân cấp tỉnh quyết định giao đất trong các trường hợp: Giao đất đối với tổ chức; cơ sở tôn giáo; người Việt Nam định cư ở nước ngoài, doanh nghiệp có vốn đầu tư nước ngoài.

(ii) Ủy ban nhân dân cấp huyện quyết định giao đất trong các trường hợp: Giao đất đối với hộ gia đình, cá nhân; cộng đồng dân cư.

- Giấy chứng nhận quyền sử dụng đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất

Giấy chứng nhận quyền sử dụng đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất như là một minh chứng tốt nhất về tính minh bạch và tính pháp lý đối với đất đai. Một dự án khi có Giấy chứng nhận quyền sử dụng đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất, tức là diện tích đất đó đã được nhà nước cấp phép sử dụng với các quyền:

- (i) Chứng minh quyền sở hữu, quyền sử dụng hợp pháp đối với bất động sản.
- (ii) Quy định rõ ràng, cụ thể về thời gian sử dụng đất.
- (iii) Là cơ sở để được bồi thường, hỗ trợ về đất và tài sản có trên đất khi nhà nước thu hồi đất.
- (iv) Là cơ sở thực hiện các quyền về chuyển nhượng, thế chấp, tặng cho,...

(3) Bản đồ quy hoạch chi tiết xây dựng (tỷ lệ 1/500) hoặc bản vẽ tổng mặt bằng và hồ sơ thiết kế bản vẽ thi công

Bản đồ quy hoạch chi tiết xây dựng (tỷ lệ 1/500) là quy hoạch tổng mặt bằng của các dự án bất động sản, cơ sở để định vị công trình, thiết kế cơ sở, thiết kế hạ tầng kỹ thuật và thực hiện xây dựng.

Một dự án bất động sản có bản đồ quy hoạch chi tiết xây dựng (tỷ lệ 1/500) được phê duyệt mới có đầy đủ pháp lý.

Bản đồ quy hoạch chi tiết xây dựng (tỷ lệ 1/500) cung cấp các thông tin sau:

(i) Tổng diện tích đất dự án, trong đó bao gồm diện tích xây dựng nhà ở và diện tích các công trình tiện ích, hạ tầng giao thông.

(ii) Theo quy chế kỹ thuật quốc gia: Mật độ xây dựng (%) = $\frac{\text{Diện tích chiếm đất của công trình kiến trúc (m}^2\text{)}}{\text{Tổng diện tích lô đất xây dựng (m}^2\text{)}} \times 100\%$.

(iii) Xác định mạng lưới giao thông: Bản đồ quy hoạch chi tiết xây dựng (tỷ lệ 1/500) sẽ thể hiện rõ mặt cắt các tuyến đường, chỉ giới đường.

(iv) Cách bố trí tiện ích: Trong bản đồ quy hoạch chi tiết xây dựng (tỷ lệ 1/500) sẽ thể hiện vị trí nào xây dựng tiện ích gì và được thiết kế để sắp xếp khoa học giúp khách hàng có được cuộc sống tốt nhất.

(v) Ranh giới hành chính: Bản đồ quy hoạch chi tiết xây dựng (tỷ lệ 1/500) thể hiện rõ ranh giới hành chính đối với các khu vực tiếp giáp dự án, tránh tình trạng chồng chéo và tranh chấp trong quy hoạch.

(vi) Hình thái kiến trúc của bản đồ quy hoạch chi tiết xây dựng (tỷ lệ 1/500) thông thường sẽ bao gồm cả bản vẽ kiến trúc cảnh quan, cung cấp thông tin trực quan về hình thái kiến trúc toàn bộ khu vực dự án.

(4) Giấy phép xây dựng

Giấy phép xây dựng là văn bản pháp lý do cơ quan Nhà nước có thẩm quyền cấp, cho phép cá nhân, tổ chức được xây dựng, cải tạo nhà cửa, công trình,...

Đây là văn bản pháp lý bất động sản bắt buộc các dự án phải có trước khi khởi công công trình.

Để được cấp phép, dự án cần đáp ứng các điều kiện sau:

- (i) Phù hợp với quy hoạch xây dựng chi tiết được duyệt.
- (ii) Bảo đảm các quy định về ranh giới, hành lang bảo vệ các công trình công cộng, di sản văn hóa,...
- (iii) Công trình sửa chữa, cải tạo không làm ảnh hưởng đến các công trình khác về kết cấu, khoảng cách đến các công trình xung quanh, cấp, thoát nước, thông gió, ánh sáng, vệ sinh cảnh quan môi trường, phòng, chống cháy, nổ.

(iv) Bảo đảm khoảng cách theo quy định đối với công trình là kho chứa hóa chất độc hại, các công trình khác có khả năng gây ô nhiễm môi trường. Không làm ảnh hưởng đến dân cư sinh sống liền kề xung quanh.

(v) Đối với công trình nhà cao tầng cấp đặc biệt và cấp I phải có thiết kế tầng hầm, trừ các trường hợp khác có yêu cầu riêng về thiết kế tầng hầm.

(5) Văn bản nghiệm thu việc hoàn thành xây dựng cơ sở hạ tầng kỹ thuật của dự án

Theo quy định tại Điều 55 Luật Kinh doanh bất động sản (2014), điều kiện của bất động sản hình thành trong tương lai được đưa vào kinh doanh:

(i) Có giấy tờ về quyền sử dụng đất, hồ sơ dự án, thiết kế bản vẽ thi công đã được cấp có thẩm quyền phê duyệt, Giấy phép xây dựng đối với trường hợp phải có giấy phép xây dựng, giấy tờ về nghiệm thu việc hoàn thành xây dựng cơ sở hạ tầng kỹ thuật tương ứng theo tiến độ dự án; trường hợp là nhà chung cư, tòa nhà hỗn hợp có mục đích để ở hình thành trong tương lai thì phải có biên bản nghiệm thu đã hoàn thành xong phần móng của tòa nhà đó.

(ii) Trước khi bán, cho thuê mua nhà ở hình thành trong tương lai, chủ đầu tư phải có văn bản thông báo cho cơ quan quản lý nhà ở cấp tỉnh về việc nhà ở đủ điều kiện được bán, cho thuê mua.

Trong thời hạn 15 ngày kể từ ngày nhận được thông báo, cơ quan quản lý nhà ở cấp tỉnh có trách nhiệm trả lời bằng văn bản cho chủ đầu tư về nhà ở đủ điều kiện được bán, cho thuê mua; trường hợp không đủ điều kiện phải nêu rõ lý do.

(6) Giấy tờ bảo lãnh của ngân hàng

Quy định bảo lãnh ngân hàng nhằm đảm bảo nghĩa vụ tài chính của chủ đầu tư với dự án hình thành trong tương lai, đảm bảo sự an tâm cho nhà đầu tư vào bất động sản. Nếu có bảo lãnh ngân hàng thì khi chủ đầu tư dự án không có khả năng triển khai dự án xây dựng thì ngân hàng sẽ hoàn tiền lại cho người mua.

(7) Giấy phép phòng cháy chữa cháy

Giấy phép phòng cháy chữa cháy (PCCC) là một phần pháp lý của dự án. Sau nhiều vụ hỏa hoạn làm nhiều tài sản và tính mạng con người bị thiệt hại, Nhà nước đã siết chặt hơn công tác thanh kiểm tra hoạt động PCCC. Các dự án bất động sản không có giấy phép PCCC sẽ không được nghiệm thu toàn phần và không có cơ sở để cấp Giấy chứng nhận quyền sử dụng đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất cho các căn hộ trong dự án.

3.3. Một số rủi ro khi tham gia giao dịch bất động sản không đủ điều kiện pháp lý

Các dự án đầu tư bất động sản luôn tồn tại rất nhiều rủi ro đối với cả người mua lẫn người bán. Việc nắm rõ các rủi ro pháp lý khi đầu tư bất động sản luôn được ưu tiên hàng đầu để tránh mất tiền cũng như hạn chế những tranh chấp không đáng có.

Những rủi ro có thể gặp phải khi đầu tư BĐS ở các dự án không đủ điều kiện pháp lý:

- Không được hoặc chậm được cấp Giấy chứng nhận quyền sở hữu đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất.

- Bị lừa đảo chiếm đoạt tiền do mua phải dự án ma hoặc chủ đầu tư bán cùng 1 sản phẩm cho nhiều người mua để thu lợi bất chính.

- Dự án chậm bàn giao nhà theo cam kết.

Tại thị trường bất động sản (BDS) Việt Nam, xu hướng đầu tư BDS đặc biệt là các loại BDS hình thành trong tương lai ngày càng gia tăng. Những loại BDS này thường chưa hoàn thiện về mặt pháp lý, có khả năng gây ra rủi ro cho nhà đầu tư và gây ảnh hưởng xấu đến lĩnh vực kinh doanh BDS nói chung. Do vậy, để tránh những hậu quả không đáng có, nhà đầu tư cần phải xác minh và kiểm tra tính pháp lý của bất động sản đó (Hoài Thu, 2022) theo các thông tin sau:

- + Chủ đầu tư dự án bất động sản hình thành trong tương lai có “lách luật” hay không.

Việc “lách luật” ở đây có thể là chủ đầu tư dự án bất động sản hình thành trong tương lai chưa xây xong phần móng công trình, thậm chí chưa được cấp phép xây dựng nhưng đã rao bán công khai và nhận tiền thông qua hợp đồng góp vốn từ người mua (hình thức trá hình của đặt cọc mua sản phẩm). Thực tế đã có nhiều dự án chưa được cấp phép xây dựng nhưng chủ đầu tư dự án đã rao bán công khai và nhận đặt cọc của người đầu tư với số tiền đặt cọc từ 5-10% tổng giá trị bất động sản.

Đối với các dự án có vị trí tốt có rất nhiều nhà đầu tư sẵn sàng bỏ tiền ra để đặt cọc, tiềm ẩn rủi ro rất cao với khách hàng. Trong trường hợp chủ đầu tư không có khả năng triển khai dự án hoặc không nhận được giấy phép xây dựng, nhà đầu tư rất khó được hoàn lại tiền đặt cọc. Thậm chí, nếu khởi kiện ra tòa án thì quyền lợi của nhà đầu tư cũng khó đảm bảo (An Vũ, 2022).

Vì vậy, để đảm bảo quyền lợi của mình, nhà đầu tư yêu cầu chủ đầu tư dự án cho xem nghiệm thu phần móng công trình và giấy phép xây dựng được cấp bởi cơ quan có thẩm quyền.

- + Dự án đã được ngân hàng bảo lãnh hay chưa

Theo quy định của pháp luật về điều kiện bán nhà ở hình thành trong tương lai, một dự án bất động sản bắt buộc phải được một ngân hàng thương mại có đủ năng lực thực hiện bảo lãnh nghĩa vụ tài chính của chủ đầu tư đối với khách hàng. Nhà đầu tư có quyền yêu cầu chủ đầu tư xuất trình cam kết bảo lãnh của ngân hàng.

Tuy nhiên, thực tế hiện nay cho thấy nhà đầu tư bất động sản không mấy quan tâm đến quy định này nên khả năng cao gặp phải những rủi ro pháp lý khi dự án bị chậm hoặc dừng triển khai do chủ đầu tư dự án thiếu vốn.

Khi chủ đầu tư dự án không có khả năng để triển khai dự án thì cơ hội hủy hợp đồng và đòi lại tiền mua nhà không hề dễ dàng. Nhà đầu tư trong trường hợp này rơi vào thế yếu, ít có cơ sở để bảo đảm quyền lợi của mình.

- + Trốn thuế

Khi chuyển nhượng BĐS, người bán có trách nhiệm nộp thuế thu nhập từ chuyển quyền sử dụng đất đầy đủ theo quy định của pháp luật, giá chuyển nhượng ghi trong hợp đồng là căn cứ để tính thuế thu nhập cá nhân và các phí, lệ phí liên quan. Nếu các bên trong giao dịch cố tình kê khai giá bán không chính xác trong hợp đồng chuyển nhượng để chịu thuế, phí thấp thì được coi là hành vi vi phạm pháp luật. Những hành động này tiềm ẩn nhiều rủi ro thậm chí có thể bị truy cứu trách nhiệm hình sự. Tùy thuộc vào mức độ vi phạm của cá nhân đó mà cơ quan chức năng sẽ tiến hành truy thu, xử phạt hành chính, thậm chí có thể bị xử lý hình sự về tội trốn thuế (Điều 200 Bộ Luật Hình sự năm 2015).

Hiện nay, các quy định mới đã chặt chẽ hơn trong vấn đề này, cũng là điểm thay đổi giúp thị trường minh bạch và hạn chế rủi ro cho các bên giao dịch bất động sản. Trong giao dịch chuyển nhượng nhà đất, hợp đồng chỉ đảm bảo tính pháp lý trước pháp luật khi có công chứng. Điều này đồng nghĩa với việc nếu công chứng viên biết các bên kê khai không đúng giá trị mua bán thực tế nhưng vẫn công chứng hợp đồng mua bán thì được xem là hành vi phạm pháp và sẽ phải liên đới chịu trách nhiệm, xử lý theo quy định pháp luật (Phương Uyên, 2022).

3.4. Phương pháp kiểm tra hồ sơ pháp lý dự án bất động sản

Tìm hiểu về hồ sơ pháp lý BĐS là việc quan trọng nhà đầu tư nên thực hiện khi tìm hiểu về bất cứ dự án BĐS

nào bên cạnh giá bán, thiết kế, tiện ích,... Tuy nhiên, nhiều nhà đầu BĐS đang chủ quan, chưa kiểm tra kỹ lưỡng dẫn đến hậu quả bất lợi, rắc rối về lâu dài tốn kém không ít thời gian và tiền bạc.

Điều quan trọng trước khi quyết định đầu tư bất động sản và tránh các rủi ro, rắc rối phát sinh cần nắm rõ cách kiểm tra pháp lý dự án BĐS để lựa chọn một BĐS tốt:

- Làm việc trực tiếp với đại diện chủ đầu tư dự án: Nhà đầu tư cần yêu cầu chủ đầu tư dự án xuất trình các loại giấy tờ pháp lý bắt buộc liên quan đến dự án.

Kiểm tra kỹ hợp đồng mua bán, chuyển nhượng BĐS. Các thông tin cần lưu ý trên hợp đồng đó là:

+ Đại diện ký hợp đồng là người có thẩm quyền hoặc ủy quyền từ phía chủ đầu tư dự án hay không.

+ Về thanh toán: tiến độ, phương thức, thỏa thuận phạt nếu có sai phạm của cả 2 phía.

+ Thông tin về sản phẩm như số tầng, mã căn, phí bảo trì chung cư, tiện ích,...

+ Trường hợp mua nhà ở riêng lẻ thì cần kiểm tra các giấy tờ như: vị trí thửa đất (số tờ bản đồ, số thửa,...), quy hoạch của thửa đất phân lô bán nền,...

- Kiểm tra tại cơ quan nhà nước có thẩm quyền. Nhà đầu tư cần chủ động tới các cơ quan nhà nước có thẩm quyền ở địa phương có dự án để tìm hiểu về quy hoạch sử dụng đất, nghĩa vụ tài chính của chủ đầu tư đối với dự án đó.

Bên cạnh đó, nhà đầu tư cần hỏi chủ đầu tư dự án BĐS về thời gian bàn giao Giấy chứng nhận quyền sử dụng đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất và các phí cần nộp khi nhận sổ. Theo quy định của Nhà nước, chủ đầu tư dự án BĐS phải cung cấp hồ sơ đầy đủ để làm thủ tục cấp Giấy chứng nhận quyền sử dụng đất, quyền sở hữu nhà ở và tài sản khác gắn liền với đất cho người mua trong vòng 50 ngày sau khi bàn giao nhà và người mua không mất bất cứ khoản phí nào (khoản 4 Điều 13 Luật Kinh doanh bất động sản 2014).

- Kiểm tra thông qua hình thức vay trả góp

Nhà đầu tư BĐS sử dụng hình thức này để được ngân hàng thẩm định tính hợp lệ của dự án BĐS. Với các dự án có vấn đề về hồ sơ pháp lý sẽ không được cho vay và ngược lại.

- Tư vấn từ các công ty môi giới bất động sản uy tín

Ở các công ty môi giới uy tín, các chuyên gia sẽ tìm hiểu các thông tin và tư vấn cho khách hàng về hồ sơ pháp lý dự án BĐS. Nhờ đó, khách hàng lựa chọn được dự án có chủ đầu tư uy tín.

3.5. Một số biện pháp hạn chế rủi ro pháp lý bất động sản

Bất động sản là một lĩnh vực đòi hỏi sự quan tâm đặc biệt về pháp lý, để hạn chế những rủi ro pháp lý khi đầu tư bất động sản, nhà đầu tư nên chủ động thực hiện các biện pháp sau:

Thứ nhất, nắm vững quy định pháp luật: Hiểu rõ và nắm vững các quy định pháp luật liên quan đến giao dịch bất động sản là cách quan trọng để tránh rủi ro. Điều này bao gồm Luật Đất đai, Luật Nhà ở, Luật Kinh doanh bất động sản, Luật Giao dịch tài sản và các quy định về thuế, phí.

Thứ hai, kiểm tra thông tin về bất động sản: Trước khi đầu tư, cần kiểm tra và xác minh thông tin về bất động sản, bao gồm quyền sở hữu, quyền sử dụng, giấy tờ pháp lý, lịch sử giao dịch trước đây, nghĩa vụ thuế và phí liên quan đến bất động sản. Điều này giúp nhà đầu tư tránh việc mua vào tài sản có vấn đề pháp lý.

Thứ ba, lựa chọn đối tác tin cậy: Lựa chọn chủ đầu tư, đơn vị môi giới và các đối tác liên quan có danh tiếng tốt và uy tín trong ngành. Điều này giúp tăng khả năng nhà đầu tư tham gia vào các giao dịch an toàn và minh bạch.

Thứ tư, theo dõi thay đổi chính sách pháp luật đất đai, thị trường bất động sản: Chính sách trong lĩnh vực đất đai luôn được sự quan tâm của Đảng, Nhà nước và được điều chỉnh, sửa đổi, bổ sung cho phù hợp với tình hình phát triển của đất nước qua mỗi giai đoạn phát triển. Nhà đầu tư cần theo dõi các thay đổi này để nâng cao ý thức tôn trọng, tuân thủ, chấp hành pháp luật. Tuân thủ đúng các thủ tục pháp lý liên quan đến việc mua bán, chuyển nhượng và sở hữu bất động sản là điều cần thiết để tránh bất kỳ tranh chấp nào trong tương lai.

Thứ năm, thận trọng trong việc tìm hiểu dự án: Nếu đầu tư vào dự án mới, cần thận trọng trong việc tìm hiểu về chủ đầu tư, tiến độ dự án, quy hoạch và các vấn đề liên quan khác. Điều này giúp nhà đầu tư tránh việc đầu tư vào dự án vi phạm quy định quy hoạch hoặc gặp vấn đề sau này khi quy hoạch thay đổi.

Trước khi ký kết bất kỳ hợp đồng đảm bảo xác minh toàn bộ giấy tờ pháp lý liên quan đến bất động sản, bao gồm Giấy chứng nhận quyền sử dụng đất, giấy phép xây dựng và các văn bản khác,...

Thứ sáu, kiểm tra lịch sử của chủ đầu tư: Nên kiểm tra lịch sử và danh tiếng của chủ đầu tư dự án trong các dự án trước đây. Điều này giúp nhà đầu tư đánh giá khả năng hoàn thành dự án và tuân thủ quy định pháp luật (Bùi Tuấn An, 2023).

4. Kết luận

Trong một xã hội thượng tôn pháp luật, giao dịch bất động sản không chỉ đơn thuần là việc trao đổi vật chất mà còn là sự giao thiệp phức tạp về quyền lợi và trách nhiệm pháp lý. Hiểu rõ và tuân thủ quy định pháp luật không chỉ là biện pháp để tránh những khúc mắc, mâu thuẫn trong tương lai mà còn đảm bảo sự ổn định và an toàn cho các bên trong giao dịch.

Để hạn chế rủi ro pháp lý khi đầu tư vào bất động sản tại Việt Nam, cần kết hợp kiến thức pháp luật, tư vấn từ chuyên gia, và thực hiện các biện pháp kiểm tra và đánh giá cẩn thận trước khi tham gia vào bất kỳ giao dịch nào.

Tài liệu tham khảo

Ban Chấp hành Trung ương Đảng khóa XIII (2022). Nghị quyết số 18-NQ/TW ngày 16/6/2022 về “tiếp tục đổi mới, hoàn thiện thể chế, chính sách, nâng cao hiệu lực, hiệu quả quản lý và sử dụng đất, tạo động lực đưa nước ta trở thành nước phát triển có thu nhập cao”.

Bộ Chính trị (2018). Kết luận số 36-KL/TW ngày 06/9/2018 về việc đẩy mạnh thực hiện Nghị quyết Trung ương 6 (khóa XI) về tiếp tục đổi mới chính sách pháp luật đất đai trong thời kỳ đẩy mạnh toàn diện công cuộc đổi mới, tạo nền tảng để đến năm 2020 nước ta cơ bản trở thành nước công nghiệp theo hướng hiện đại.

Luật số 45/2013/QH13, Luật Đất đai 2013.

Luật số 65/2014/QH13, Luật Nhà ở 2014.

Luật số 66/2014/QH13, Luật Kinh doanh bất động sản 2014.

Bộ Tài nguyên và Môi trường (2022). Báo cáo đánh giá tình hình thi hành Luật Đất đai 2013 và định hướng sửa đổi Luật Đất đai.

An Vũ (2022). Rủi ro pháp lý là nỗi lo thường trực của thị trường bất động sản. <https://Reatimes.vn>.

Rich Nguyen (2023). Pháp lý bất động sản và những kiến thức nhà đầu tư cần biết. <https://richnguyen.vn/>.

Bùi Tuấn An (2023). Pháp lý - yếu tố các nhà đầu tư bất động sản cần đặc biệt quan tâm. <https://luatminhkhue.vn/>.

Phương Uyên (2022). Kinh doanh bất động sản - Không đầy rủi ro pháp lý cho người dân <https://thuvienphapluat.vn/>.

Admin (2021): Pháp lý bất động sản là gì? Có những giấy tờ nào? <https://bacnamland.com>.

Hoài Thu (2022). Những vấn đề pháp lý cần lưu ý khi giao dịch dự án bất động sản chưa đủ điều kiện đưa vào kinh doanh. <https://htlegalvn.com>.

ABSTRACT

Real estate transactions - Current status and measures to limit legal risks

Dung Thi Nguyen^{1,2*}

¹*Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology*

²*Land Management and Cadastral Technology for Sustainable Development Research Group*

Real estate transactions are an important content of the real estate market, transactions have great value and affect many stakeholders, so strict legal procedures are required. In fact, there are many risks and many disputes that have occurred when conducting real estate transactions in recent times. The purpose of the article is to clarify the importance of real estate legal issues; Full legal information of real estate, some risks when transacting unqualified real estate and solutions. Methods used include methods of collecting documents, data and methods of analysis, evaluation and synthesis; The article also proposes a number of measures to limit real estate legal risks when parties participate in transactions. Real estate transactions are one of the areas with many risks for both buyers and sellers. Besides the complex economic aspects, participating in this market also requires a solid grasp of legal knowledge. Real estate legality is also a factor that directly affects the rights and obligations of both investors and related entities, and is the core foundation for determining the rights and obligations of buyers, sellers and stakeholders other relevant parties. Being clear and ensuring legal issues is one of the ways to avoid future disputes. Building correct processes and complying with legal regulations will bring a better real estate transaction environment, promote sustainable development and build trust for all market participants. Therefore, carefully understanding the legality of real estate before signing a transaction contract is one of the important notes for investors.

Keywords: Real estate legal, land law, real estate transactions.

Ảnh hưởng của mật độ điểm đo trực tiếp đến độ chính xác nội suy dị thường trọng lực bước sóng dài khi tính độ sâu từ dị thường trọng lực

Nguyễn Văn Sáng^{1*}

¹Trường Đại học Mở - Địa chất

TÓM TẮT

Mục đích của nghiên cứu này là xác định xem mật độ của điểm đo trực tiếp dị thường trọng lực và độ sâu ảnh hưởng thế nào đến độ chính xác nội suy dị thường trọng lực bước sóng dài trong bài toán xác định độ sâu đáy biển từ số liệu dị thường trọng lực. Phương pháp Collocation đã được sử dụng để nội suy dị thường trọng lực bước sóng dài. Khu vực thực nghiệm có độ vĩ Bắc từ 14,5°-16,0°; độ kinh Đông từ 115,0°-116,5°, nằm ở phía Đông - Nam quần đảo Hoàng Sa của Việt Nam. Các tính toán được thực hiện ở 8 phương án có mật độ điểm đo khác nhau. Kết quả thực nghiệm cho thấy: Khi khoảng cách tối thiểu giữa các điểm đo tăng từ 183 m (0,1 x 0,1') đến 3667 m (2' x 2') thì độ lệch trung phương nội suy dị thường trọng lực bước sóng dài tăng từ $\pm 1,014$ mGal đến $\pm 3,945$ mGal. Kết quả nghiên cứu này rất có ý nghĩa khi xác định mật độ điểm đo trực tiếp cần thiết khi tính toán độ sâu từ số liệu dị thường trọng lực.

Từ khóa: Mật độ điểm đo trực tiếp, độ chính xác, dị thường trọng lực bước sóng dài, độ sâu trọng lực.

1. Đặt vấn đề

Trong bài toán xác định độ sâu đáy biển từ trọng lực, dị thường trọng lực được chia làm 2 phần: Phần thứ nhất liên quan đến mật độ đất đá dưới sâu và rộng của đáy biển, được gọi là dị thường trọng lực bước sóng dài, ký hiệu là $\Delta g^{dài}$; phần thứ hai liên quan đến địa hình đáy biển, được gọi là dị thường trọng lực bước sóng ngắn, ký hiệu là $\Delta g^{ngắn}$. Một trong những công việc quan trọng trong bài toán này là nội suy $\Delta g^{dài}$ từ các điểm đã biết độ sâu và dị thường trọng lực (Nguyễn Văn Sáng và nnk., 2023).

Những năm gần đây, trên thế giới, nhiều nhà khoa học đã có nghiên cứu về vấn đề xác định độ sâu đáy biển từ dị thường trọng lực như: Năm 2010, các tác giả đã sử dụng phương pháp địa chất - trọng lực (GGM) để xác định độ sâu của vùng biển phía Đông Nhật Bản từ dị thường trọng lực có nguồn gốc từ đo cao vệ tinh kết hợp độ sâu đo bằng tàu (Kwang Bae Kim và nnk., 2010); năm 2011, các tác giả đã nghiên cứu phương pháp “downward continuation - DWC” để xác định độ sâu đáy biển từ số liệu dị thường trọng lực có nguồn gốc từ đo cao vệ tinh và số liệu đo sâu bằng tàu ở hai khu vực nghiên cứu, nằm ở phía nam Greenland (Yu-Shen Hsiao và nnk., 2011); năm 2017, các tác giả đã nghiên cứu xây dựng mô hình dị thường trọng lực bước sóng dài để xác định độ sâu từ số liệu dị thường trọng lực. Các kết quả nghiên cứu chỉ ra rằng mô hình dị thường trọng lực bước sóng dài là rất quan trọng và ảnh hưởng đến độ chính xác xác định độ sâu từ số liệu trọng lực (Xueshuang Xiang và nnk., 2017). Năm 2018, Yeon Yeu và các đồng nghiệp đã nghiên cứu xác định được độ sâu từ số liệu dị thường trọng lực có nguồn gốc từ đo cao vệ tinh kết hợp với các số liệu đo sâu hồi âm đơn tia và đa tia, số liệu đo bằng Lidar ở vùng biển phía Tây Hàn Quốc,... Tuy nhiên, chưa có công trình nào nghiên cứu về ảnh hưởng của mật độ điểm đo trực tiếp trọng lực và độ sâu đến độ chính xác nội suy $\Delta g^{dài}$.

Công trình nghiên cứu này sẽ trình bày nội suy dị thường trọng lực bước sóng dài theo phương pháp Collocation và khảo sát ảnh hưởng của mật độ điểm đo trực tiếp trọng lực và độ sâu đến độ chính xác nội suy $\Delta g^{dài}$ bằng cách thực hiện các phương án nội suy với mật độ điểm đo trực tiếp khác nhau, các kết quả nội suy này được so sánh với dị thường trọng lực đã biết để đánh giá độ chính xác.

2. Số liệu và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu và số liệu nghiên cứu

Khu vực nghiên cứu

Khu vực nghiên cứu nằm ở phía Đông - Nam quần đảo Hoàng Sa (Hình 1), thuộc vùng trũng sâu của Biển Đông: độ sâu nhỏ nhất là -626,0 m; độ sâu lớn nhất là -4413,0 m; độ sâu trung bình là -4100,7 m. Kích thước của khu vực nghiên cứu mỗi chiều 1,5°, được giới hạn bởi: Độ vĩ Bắc từ 14,5°-16,0°; độ kinh Đông từ 115,0°-116,5°, diện tích khoảng 27 225 km².

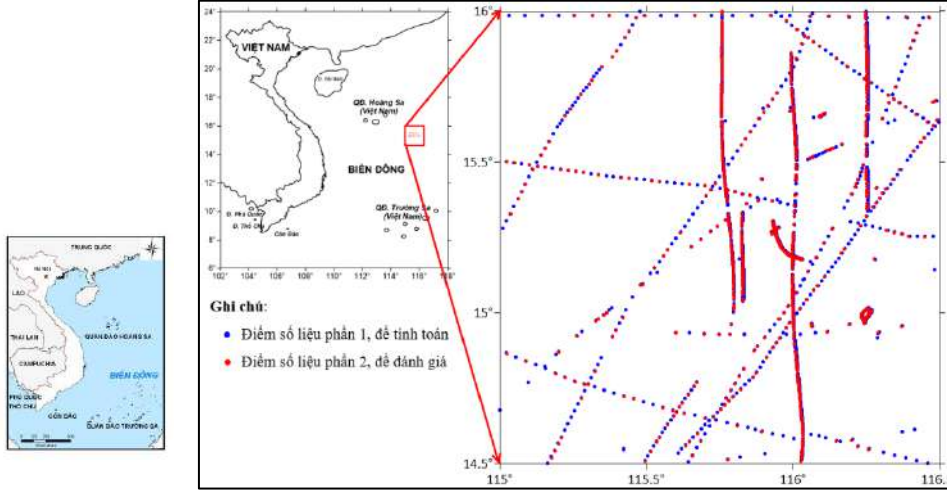
Số liệu nghiên cứu

Số liệu đo sâu và đo trọng lực trực tiếp trên khu vực nghiên cứu được cung cấp bởi Tổ chức Trọng lực quốc tế

* Tác giả liên hệ

Email: nguyenvansang@humg.edu.vn

(BGI). Đây là số liệu đo trọng lực trực tiếp bằng tàu biển. Tổng số có 4298 điểm đo (các điểm màu **đỏ** và **xanh** trên Hình 1). Tọa độ của các điểm trong hệ WGS-84. Số liệu này được chia làm 2 phần: phần 1 dùng để tính toán, bao gồm 3224 điểm màu xanh trên Hình 1; phần 2 dùng để đánh giá độ chính xác, bao gồm 1074 điểm màu đỏ trên Hình 1.



Hình 1. Khu vực nghiên cứu và số liệu nghiên cứu

2.2. Nội suy dị thường trọng lực bước sóng dài theo phương pháp Collocation

2.2.1. Tính dị thường trọng lực bước sóng dài

Giả sử có n điểm đã biết dị thường trọng lực bước sóng dài $\Delta g_1^{dài}, \Delta g_2^{dài}, \dots, \Delta g_n^{dài}$ trên khu vực nghiên cứu. Khi đó, theo phương pháp Collocation, dị thường trọng lực bước sóng dài của điểm P được tính bằng công thức (Neiman, 2010):

$$\Delta g_P^{dài} = C^T(\Delta g^{dài}, \Delta g_P^{dài}) \cdot [C(\Delta g^{dài}, \Delta g^{dài}) + C_{\Delta g^{dài}}]^{-1} \cdot \Delta g^{dài} \quad (1)$$

trong đó: $C^T(\Delta g^{dài}, \Delta g_P^{dài})$ là ma trận hiệp phương sai dị thường trọng lực bước sóng dài giữa điểm P và các điểm đã biết dị thường trọng lực:

$$C^T(\Delta g^{dài}, \Delta g_P^{dài}) = [C(\Delta g_1^{dài}, \Delta g_P^{dài}) \ C(\Delta g_2^{dài}, \Delta g_P^{dài}) \ \dots \ C(\Delta g_n^{dài}, \Delta g_P^{dài})] \quad (2)$$

$C(\Delta g^{dài}, \Delta g^{dài})$ là ma trận hiệp phương sai dị thường trọng lực bước sóng dài giữa các điểm đã biết dị thường trọng lực với nhau:

$$C(\Delta g^{dài}, \Delta g^{dài}) = \begin{bmatrix} C(\Delta g_1^{dài}, \Delta g_1^{dài}) & C(\Delta g_1^{dài}, \Delta g_2^{dài}) & \dots & C(\Delta g_1^{dài}, \Delta g_n^{dài}) \\ C(\Delta g_2^{dài}, \Delta g_1^{dài}) & C(\Delta g_2^{dài}, \Delta g_2^{dài}) & \dots & C(\Delta g_2^{dài}, \Delta g_n^{dài}) \\ \vdots & \vdots & \ddots & \vdots \\ C(\Delta g_n^{dài}, \Delta g_1^{dài}) & C(\Delta g_n^{dài}, \Delta g_2^{dài}) & \dots & C(\Delta g_n^{dài}, \Delta g_n^{dài}) \end{bmatrix} \quad (3)$$

$C_{\Delta g^{dài}}$ là ma trận hiệp phương sai độ chính xác dị thường trọng lực bước sóng dài:

$$C_{\Delta g^{dài}} = \begin{bmatrix} C_{\Delta g_1^{dài} \Delta g_1^{dài}} & C_{\Delta g_1^{dài} \Delta g_2^{dài}} & \dots & C_{\Delta g_1^{dài} \Delta g_n^{dài}} \\ C_{\Delta g_2^{dài} \Delta g_1^{dài}} & C_{\Delta g_2^{dài} \Delta g_2^{dài}} & \dots & C_{\Delta g_2^{dài} \Delta g_n^{dài}} \\ \vdots & \vdots & \ddots & \vdots \\ C_{\Delta g_n^{dài} \Delta g_1^{dài}} & C_{\Delta g_n^{dài} \Delta g_2^{dài}} & \dots & C_{\Delta g_n^{dài} \Delta g_n^{dài}} \end{bmatrix} \quad (4)$$

$\Delta g^{dài}$ là ma trận các giá trị dị thường trọng lực bước sóng dài đã biết.

2.2.2. Tính hiệp phương sai thực nghiệm

Hiệp phương sai thực nghiệm của dị thường trọng lực bước sóng dài được tính dựa vào số liệu thực nghiệm theo công thức (Nguyễn Văn Sáng và nnk., 2023):

$$\hat{C}_{\Delta g}(\psi_i) = \frac{1}{m_i} \sum_{n=1}^{m_i} [\Delta g^{dài}(P) \cdot \Delta g^{dài}(P')]_n, \quad (5)$$

trong đó: P và P' - là các điểm đã biết giá trị dị thường trọng lực bước sóng dài $\Delta g^{dài}$. Khoảng cách cầu giữa 2 điểm thỏa mãn điều kiện:

$$\psi_i - \frac{\Delta\psi}{2} \leq \psi \leq \psi_i + \frac{\Delta\psi}{2}, \quad (6)$$

ở đây: m_i - là số lượng cặp điểm thỏa mãn điều kiện, $\Delta\psi$ - là khoảng cách cầu giữa 2 điểm gần nhất trong bộ số liệu đầu vào và $\psi_i - \Delta\psi/2 = 0$, nếu $\psi_i < \Delta\psi/2$.

2.2.3. Xác định hàm hiệp phương sai lý thuyết

Hàm hiệp phương sai lý thuyết của dị thường trọng lực có dạng (Neiman, 2010):

$$C(\Delta g_i, \Delta g_j) = a \sum_{l=2}^N d_l \frac{(l-1)^2}{r_i r_j} \left(\frac{R^2}{r_i r_j} \right)^{l+1} P_l(\cos \psi) + \sum_{l=N+1}^{\infty} \frac{A}{(l-2)(l+b)} \frac{(l-1)}{r_i r_j} \left(\frac{R_B^2}{r_i r_j} \right)^{l+1} P_l(\cos \psi), \quad (7)$$

trong đó: a - tham số bổ sung; d_l - phương sai bậc N ; b - số tự nhiên; A - hằng số có đơn vị là $(\text{m/s})^4$; R - bán kính trung bình Trái đất; R_B - bán kính của mặt cầu nằm trong Trái đất; $P_l(\cos \psi)$ - hàm Legendre bậc l ; ψ - khoảng cách cầu giữa điểm i và j ; r_i và r_j - khoảng cách từ điểm đến gốc tọa độ.

Các tham số của hàm hiệp phương sai lý thuyết sẽ được xác định bằng cách làm khớp giá trị hiệp phương sai thực nghiệm với hàm lý thuyết (Nguyễn Thành Lê và nnk., 2021).

Các phương án tính toán nội suy dị thường trọng lực bước sóng dài

Để nghiên cứu ảnh hưởng của mật độ điểm đo trực tiếp đến kết quả nội suy dị thường trọng lực bước sóng dài, các tính toán nội suy được thực hiện ở 8 phương án (Bảng 1): PA1 - giữ nguyên các điểm như khi đo trực tiếp (số liệu phần 1, Hình 1); trong các phương án tiếp theo, các điểm đo trực tiếp được bỏ bớt (làm thưa) để khoảng cách tối thiểu tăng dần từ 183 m (tương ứng 0,1') đến 3667 m (tương ứng 2').

Bảng 1. Các phương án tính toán

PA	Khoảng cách tối thiểu (m)	PA	Khoảng cách tối thiểu (m)
PA1	Giữ nguyên như khi đo	PA5	1100
PA2	183	PA6	1467
PA3	367	PA7	1833
PA4	733	PA8	3667

Các kết quả tính toán nội suy được so sánh với các điểm số liệu phần 2 (Hình 1) để đánh giá. Các giá trị độ lệch Δ_i , độ lệch trung bình Δ_{TB} được tính theo công thức:

$$\Delta_i = \Delta g_i^{NS} - \Delta g_i^{ship}; \quad \Delta_{TB} = \frac{1}{n} \sum_{i=1}^n \Delta_i. \quad (8)$$

trong đó: Δg_i^{NS} là dị thường trọng lực bước sóng dài nội suy; Δg_i^{ship} là dị thường trọng lực bước sóng dài tính từ giá trị đo trực tiếp bằng tàu.

Độ lệch chuẩn STD được tính theo công thức (Đặng Nam Chính và nnk., 2015):

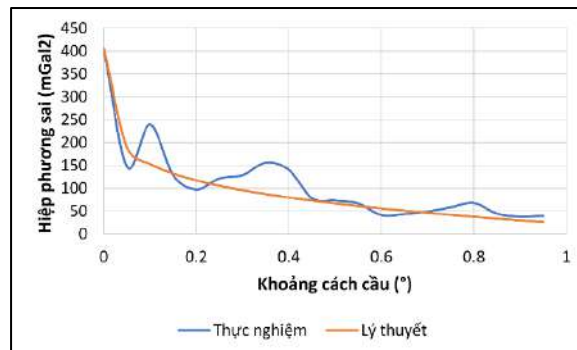
$$STD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\Delta_i - \Delta_{TB})^2}, \quad (9)$$

Độ lệch trung phương RMS được tính theo công thức (Đặng Nam Chính và nnk., 2015):

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (\Delta_i)^2}. \quad (10)$$

3. Kết quả và thảo luận

Trên Bảng 2 trình bày minh họa kết quả xác định hiệp phương sai thực nghiệm và lý thuyết của PA1 với khoảng cách cầu tính đến 0,95°. Trên Hình 2 là đồ thị của hàm hiệp phương sai thực nghiệm và lý thuyết. Hình này cho thấy đồ thị hiệp phương sai lý thuyết đã được làm khớp với các giá trị thực nghiệm.

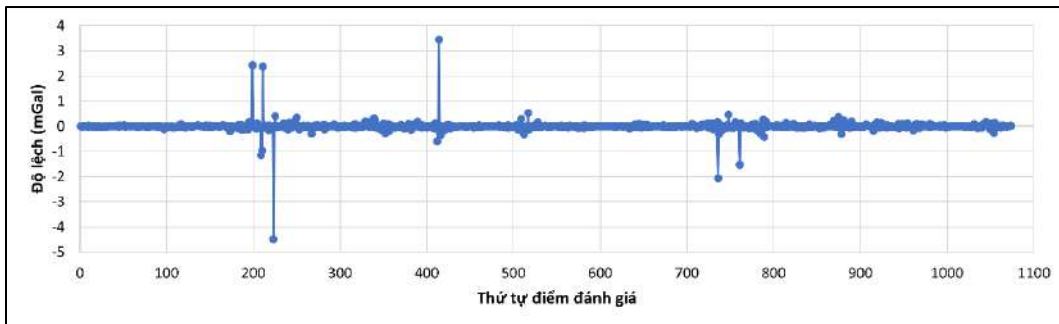


Hình 2. Đồ thị hàm hiệp phương sai

Bảng 2. Kết quả tính hiệp phương sai

S T T	Khoảng cách cầu (°)	Hiệp phương sai (mGal ²)		ST T	Khoảng cách cầu (°)	Hiệp phương sai (mGal ²)	
		Thực nghiệm	Lý thuyết			Thực nghiệm	Lý thuyết
1	0,00	405,3815	405,8097	11	0,50	72,9110	66,8104
2	0,05	146,7698	188,3606	12	0,55	66,6154	61,1589
3	0,10	239,1841	152,8680	13	0,60	41,2277	55,8837
4	0,15	128,2844	131,9489	14	0,65	43,2563	50,9248
5	0,20	97,2602	116,9620	15	0,70	48,7123	46,2375
6	0,25	121,1211	105,1911	16	0,75	58,2313	41,7886
7	0,30	128,5357	95,4265	17	0,80	67,5137	37,5533
8	0,35	155,6112	87,0247	18	0,85	44,3913	33,5129
9	0,40	140,2369	79,6040	19	0,90	37,5825	29,6539
10	0,45	78,0320	72,9206	20	0,95	39,5408	25,9664

Độ lệch giữa dị thường trọng lực bước sóng dài nội suy được và giá trị đã biết tại các điểm đánh giá của PA1 được trình bày trên Hình 3. Hình này cho thấy: hầu hết các độ lệch có giá trị rất nhỏ, chỉ có vài điểm có độ lệch lớn, các điểm này sẽ được loại bỏ. Độ lệch trung phương nhỏ, đạt 0,231 mGal (Bảng 3).



Hình 3. Độ lệch dị thường trọng lực bước sóng dài

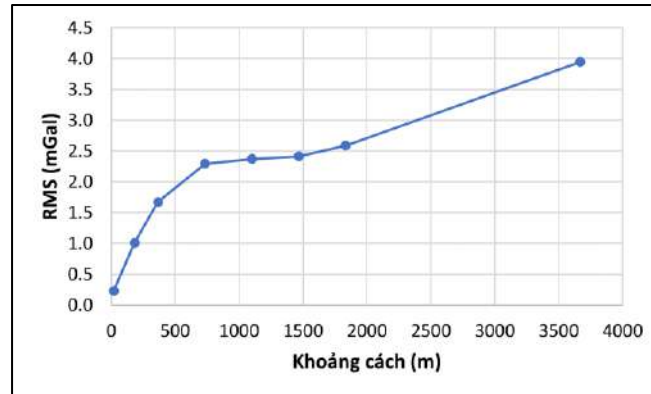
Kết quả đánh giá của 8 phương án được trình bày tổng hợp trên Bảng 3, trong đó: RMS là độ lệch trung phương, STD là độ lệch chuẩn, Mean là độ lệch trung bình. Kết quả trong bảng này cho thấy: Các phương án tính đều có độ lệch trung bình nhỏ, chứng tỏ không tồn tại độ lệch hệ thống giữa kết quả tính và số liệu đánh giá; Độ lệch trung phương tăng dần khi mật độ điểm đo giảm (khoảng cách tối thiểu tăng). Khi khoảng cách tối thiểu giữa các điểm đo là 0,1' (183 m) thì độ chính xác nội suy là $\pm 1,014$ mGal. Khi khoảng cách tối thiểu tăng lên là 2' (3667 m) thì độ chính xác nội suy là $\pm 3,945$ mGal.

Bảng 3. Kết quả đánh giá độ chính xác theo các phương án tính

PA	Khoảng cách tối thiểu (m)	Độ lệch (m)				
		Max	Min	Mean	STD	RMS
PA 1	Giữ nguyên như khi đo	3,446	-4,493	-0,002	$\pm 0,231$	$\pm 0,231$
PA 2	183	6,439	-13,131	-0,058	$\pm 1,012$	$\pm 1,014$
PA 3	367	13,144	-21,623	-0,108	$\pm 1,671$	$\pm 1,673$
PA 4	733	24,150	-23,533	-0,081	$\pm 2,295$	$\pm 2,295$
PA 5	1100	13,692	-15,670	-0,153	$\pm 2,367$	$\pm 2,371$
PA 6	1467	13,671	-14,955	-0,042	$\pm 2,414$	$\pm 2,414$
PA 7	1833	13,736	-14,070	-0,061	$\pm 2,590$	$\pm 2,589$
PA 8	3667	25,850	-19,060	0,095	$\pm 3,946$	$\pm 3,945$

Mối tương quan giữa khoảng cách tối thiểu và các điểm đo và độ chính xác nội suy được trình bày trên Hình 4. Từ hình này cho thấy: khi khoảng cách tối thiểu tăng từ 183 m lên 733 m thì độ lệch trung phương (RMS) tăng nhanh, đồ thị dốc; khi khoảng cách tối thiểu tăng từ 733 m lên 1833 m thì RMS tăng chậm, đồ thị

thoải; khi khoảng cách tối thiểu tăng từ 1833 m lên 3667 m thì RMS tăng ở mức trung bình. Như vậy, với khoảng cách tối thiểu giữa các điểm đo là 1833 m (1') sẽ đem lại hiệu quả tối ưu giữa mật độ đo và độ chính xác nội suy.



Hình 4. Mối quan hệ giữa mật độ điểm đo trực tiếp và độ chính xác nội suy

4. Kết luận

Bài báo đã nghiên cứu ảnh hưởng của mật độ điểm đo đến độ chính xác nội suy dị thường trọng lực bước sóng dài theo phương pháp Collocation. Tám phương án tính toán nội suy với mật độ điểm khác nhau đã được tính toán, kết quả nội suy được so sánh với giá trị đã biết để đánh giá. Khi khoảng cách tối thiểu tăng từ 183 m lên 733 m thì độ lệch trung phương (RMS) tăng nhanh từ $\pm 1,014$ mGal lên $\pm 2,295$ mGal. Khi khoảng cách tối thiểu tiếp tục tăng lên 1833 m thì RMS tăng chậm lên $\pm 2,589$ mGal. Khi khoảng cách tối thiểu tăng đến 3667 m thì RMS tăng trung bình lên $\pm 3,945$ mGal. Kết quả nghiên cứu này rất có ý nghĩa trong bài toán xác định độ sâu từ dị thường trọng lực.

Cần tiếp tục nghiên cứu ở những khu vực rộng lớn hơn để có thể đánh giá được ảnh hưởng ở mật độ điểm đo ở mức thưa hơn đến độ chính xác nội suy dị thường trọng lực bước sóng dài.

Lời cảm ơn

Nghiên cứu này được sự hỗ trợ của đề tài KHCN cấp bộ của Bộ Giáo dục và Đào tạo, mã số: B2021-MDA-06 và đề tài cấp Quốc gia Việt Nam ĐTDLCN.07/23. Tác giả cũng xin cảm ơn BGI đã hỗ trợ nhóm tác giả số liệu trong nghiên cứu này.

Tài liệu tham khảo

Đặng Nam Chinh, Nguyễn Xuân Bắc, Bùi Thị Hồng Thắm, Trần Thị Thu Trang, Ninh Thị Kim Anh, 2015. Giáo trình lý thuyết sai số. Trường Đại học Tài nguyên và môi trường, Hà Nội.

Forsberg R., Tscherning C.C., 2008. An overview manual for the GRAVSOF Geodetic Gravity Field Modelling Programs, DTU Space.

Kwang Bae Kim, Yu-Shen Hsiao, Jeong Woo Kim, Bang Yong Lee, Yi Kyun Kwon, Chang Hwan Kim, 2010. Bathymetry enhancement by altimetry-derived gravity anomalies in the East Sea (Sea of Japan). Mar Geophys Res (2010) 31:285-298. Doi 10.1007/s11001-010-9110-0.

Neiman. Y.M., 2010. Phương pháp hiệp phương sai trong trắc địa vật lý và Collocation (Sách chuyên khảo). Matxcova (tiếng Nga).

Nguyễn Thành Lê, Nguyễn Văn Sáng, Đinh Xuân Mạnh, 2021. Xây dựng giải pháp tối ưu việc xác định các tham số của hàm hiệp phương sai lý thuyết trong phương pháp LSC. Tạp chí Khoa học Đo đạc và Bản đồ, Số 49 - 9/2021, tr 15-20. Viện Khoa học Đo đạc và Bản đồ.

Nguyễn Văn Sáng và nnk., 2023. Nghiên cứu đề xuất phương pháp xác định độ sâu đáy biển từ số liệu dị thường trọng lực trên khu vực Biển Đông. Đề tài cấp bộ, Bộ Giáo dục và Đào tạo, Hà Nội.

Nguyen Van Sang, Khuong Van Long, Tran Tuan Dung, Lam Van Nguyen, Bui Cong Que, Do Van Mong, Bui Dang Quang, Ole Baltazar Andersen, Rene Forsberg, Dieu Tien Bui, 2023. Seafloor Depth Mapping of Central Vietnam's Sea area and its surrounding using gravity anomaly data and Gravity Geological Method. Advances in Space Research, Volume 72, Issue 5, 1 September 2023, p. 1721-1738. Doi: <https://doi.org/10.1016/j.asr.2023.04.033>.

Xueshuang Xiang, Xiaoyun Wan, Running Zhang, Yang Li, Xiaohong Sui and Wenbin Wang, 2017. Bathymetry inversion with Gravity-Geologic Method: A study of long-wavelength gravity modeling based on adaptive mesh. Marine Geodesy. ISSN: 0149-0419 (Print) 1521-060X (Online), Doi: 10.1080/01490419.2017.1335257.

Yeon Yeu, Jurng-Jae Yee, Hong Sik Yun and Kwang Bae Kim, 2018. Evaluation of the Accuracy of Bathymetry on the Nearshore Coastlines of Western Korea from Satellite Altimetry, Multi-Beam, and Airborne Bathymetric LiDAR. Sensors, MDPI. Doi:10.3390/s18092926.

Yu-Shen Hsiao, Jeong Woo Kim, Kwang Bae Kim, Bang Yong Lee, and Cheinway Hwang, 2011. Bathymetry Estimation Using the Gravity-Geologic Method: An Investigation of Density Contrast Predicted by the Downward Continuation Method. Terr. Atmos. Ocean. Sci., Vol. 22, No. 3, 347-358. Doi: 10.3319/TAO.2010.10.13.01.

ABSTRACT

Effect of the density of the shipborne points on interpolation accuracy of the long-wavelength gravity anomaly when calculating seabed depth from gravity anomaly

Nguyen Van Sang^{1,*}

¹*Hanoi University of Mining and Geology*

The purpose of this study is to determine how the density of shipborne points of gravity and depth affects the interpolation accuracy of long-wavelength gravity anomalies when determining seabed depth from gravity anomaly. The Collocation method was used to interpolate long-wavelength gravity anomalies. The experimental area has latitude from 14.5° to 16.0°; longitude from 115.0° to 116.5°, located in the East - South of Vietnam's Hoang Sa Archipelago. The calculations are performed in 8 options with different density of shipborne points. The experimental results show that: When the minimum distance between shipborne points increases from 183 m (0.1 x 0.1') to 3667 m (2' x 2'), the root mean square deviation of interpolated long-wavelength gravity anomaly increases from ± 1.014 mGal to ± 3.945 mGal. The results of this study are very meaningful when determining the needed density of shipborne points when calculating seabed depth from gravity anomaly data.

Keywords: Density of shipborne points, interpolation accuracy, long wavelength gravity anomaly, gravity derived-depth.

Kết hợp đa chỉ số phổ trong chiết tách đất xây dựng ở khu vực Hà Nội từ ảnh vệ tinh Sentinel-2

Nguyễn Văn Trung^{1,2*}, Lê Thị Thu Hà^{1,2}, Nguyễn Thành Len³, Lê Thị Minh Phương³

¹Trường Đại học Mỏ - Địa chất

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mỏ - Địa chất

³Trường Đại học Kiến trúc Hà Nội

TÓM TẮT

Việc lập bản đồ đất xây dựng ở các khu vực đô thị hầu hết đều gặp khó khăn lớn do các đối tượng sử dụng đất xen kẽ với nhau. Chính vì vậy, trong nghiên cứu này, nhóm tác giả cố gắng tăng cường độ chính xác của bản đồ đất xây dựng đô thị cho thành phố Hà Nội bằng việc tập trung xác định tổ hợp phù hợp của các chỉ số phổ tính từ dữ liệu ảnh vệ tinh. Để thực hiện nghiên cứu, bốn chỉ số phổ đã được lựa chọn bao gồm chỉ số đất đô thị chuẩn hóa (NDTI), chỉ số đất trống (BSI), chỉ số đất trống khô (DBSI) và chỉ số thực vật chuẩn hóa (NDVI). Tất cả các chỉ số phổ này đều được tính từ dữ liệu Sentinel-2 thu được trong mùa khô. Hai sự kết hợp được tạo ra từ sự xếp chồng các lớp chỉ số phổ NDTI/BSI/NDVI và NDTI/DBSI/NDVI. Việc sử dụng thuật toán “K-means” làm công cụ phân loại không giám sát mang lại khả năng phát hiện đất đô thị tự động và nhanh chóng. Kết quả cho thấy chỉ số BSI hoạt động tốt hơn so với việc sử dụng chỉ số DBSI. Do đó, chỉ số BSI mang lại những cải tiến: các loại đất trống và quá trình tích tụ được phân biệt rõ hơn, trong đó độ chính xác tổng thể tăng 5,82% và hệ số Kappa tăng bằng 11,1%. Kết quả cho thấy bộ dữ liệu đa chỉ số phổ NDTI/BSI/NDVI phù hợp cho việc lập bản đồ các khu vực đô thị với khả năng giúp quản lý đô thị tốt hơn trong mùa khô.

Từ khóa: Đất xây dựng, Sentinel-2, K-means, NDTI, BSI, DBSI.

1. Đặt vấn đề

Việc phân loại đất xây dựng trong đô thị từ ảnh vệ tinh gặp phải khó khăn do trong khu vực đô thị thường nhầm lẫn giữa đất trống và đất xây dựng. Gần đây, các nghiên cứu bao gồm (Pal và Antil, 2017; Vigneshwaran và Vasanth Kumar, 2018; Xi và cộng sự, 2019; Yilmaz và cộng sự, 2019) đã tận dụng được những ưu điểm của dữ liệu ảnh vệ tinh Sentinel-2 mới đề xuất nhiều ứng dụng chỉ số đất xây dựng trong phân loại đất đô thị. Tuy nhiên, sự hiệu quả đã được thể hiện khi sử dụng tích hợp các chỉ số phổ bao gồm chỉ số đất xây dựng chuẩn hóa NDBI (Zha và cộng sự, 2003), chỉ số đất xây dựng IBI (Xu, 2008), chỉ số đất xây dựng mới NBI (Jieli, 2010), chỉ số tỷ lệ kênh phổ cho đất xây dựng BRBA và chỉ số diện tích đất xây dựng Chuẩn hóa NBAI (Waqar và cộng sự, 2012), chỉ số thành phần sinh lý đất xây dựng BCI (Deng và Wu, 2012), chỉ số đất xây dựng sửa đổi MBI (Liu và cộng sự, 2014), chỉ số chiết xuất đất xây dựng BAEI (Bouzekri và cộng sự, 2015) và chỉ số kết hợp đất xây dựng CBI (Sun và cộng sự, 2016) chịu sự thay đổi theo mùa, đặc biệt là vào tháng khô ráo, nơi có sự tương đồng giữa quang phổ đất trống và đất xây dựng đô thị đang tăng. Các kết quả đã không tạo ra được bản đồ chính xác khu vực đô thị trong thời kỳ khô hạn trong khu vực đang nghiên cứu.

Trong vài năm gần đây, các dữ liệu mới đã được tạo ra bởi sự kết hợp của các chỉ số phổ khác nhau. Trong thực tế, việc sử dụng các chỉ số phổ như các kênh phổ hoạt động tốt hơn so với sử dụng các kênh phổ ban đầu (Xu, 2008). Lựa chọn này cho phép một số nhà nghiên cứu bao gồm (Pal và Antil, 2017; Leroux và cộng sự, 2018; Gašparović và cộng sự, 2019) đạt được một kết quả tốt hơn khi phân loại để lập bản đồ lớp phủ/sử dụng đất. Tương tự như vậy, các nghiên cứu khác tập trung vào phân loại khu vực đô thị như các tác giả (Xu, 2008; Bramhe và cộng sự, 2018; Nur Hidayati và cộng sự, 2018; Ettahadi Osgouei và cộng sự, 2019; Lynch và Blesius, 2019). Tuy nhiên, nghiên cứu của (Ettahadi Osgouei và cộng sự, 2019) đã áp dụng chỉ số đất chênh lệch chuẩn hóa (NDTI) được phát triển và cũng được sử dụng trong (Eskandari và cộng sự, 2016) đã sử dụng các kênh phổ SWIR của ảnh vệ tinh Sentinel-2 và đã thành công trong việc phân biệt các lớp đất trống và khu vực đất xây dựng tốt hơn so với các chỉ số phổ khác được sử dụng trong nghiên cứu.

Các kết quả được đề cập ở trên cho thấy cần thiết phải cải thiện độ chính xác thành lập bản đồ các khu vực đất xây dựng của các thành phố vào mùa khô dựa trên sự chồng xếp các chỉ số phổ, trong đó đa chỉ số bao gồm chỉ số NDTI kết hợp với chỉ số phổ khác có thể tăng cường phân biệt được đất xây dựng. Vì vậy, do vùng đô thị được đặc trưng bởi mật độ thực vật thấp và đất trống rộng, đặc biệt vào mùa khô, hai chỉ số đất trống đã được lựa chọn. Đầu tiên, chỉ số đất trống (BSI) được giới thiệu bởi (Rikimaru và Miyatake, 1997) nhằm tăng cường việc xác định

* Tác giả liên hệ

Email: nguyenvantrung@humg.edu.vn

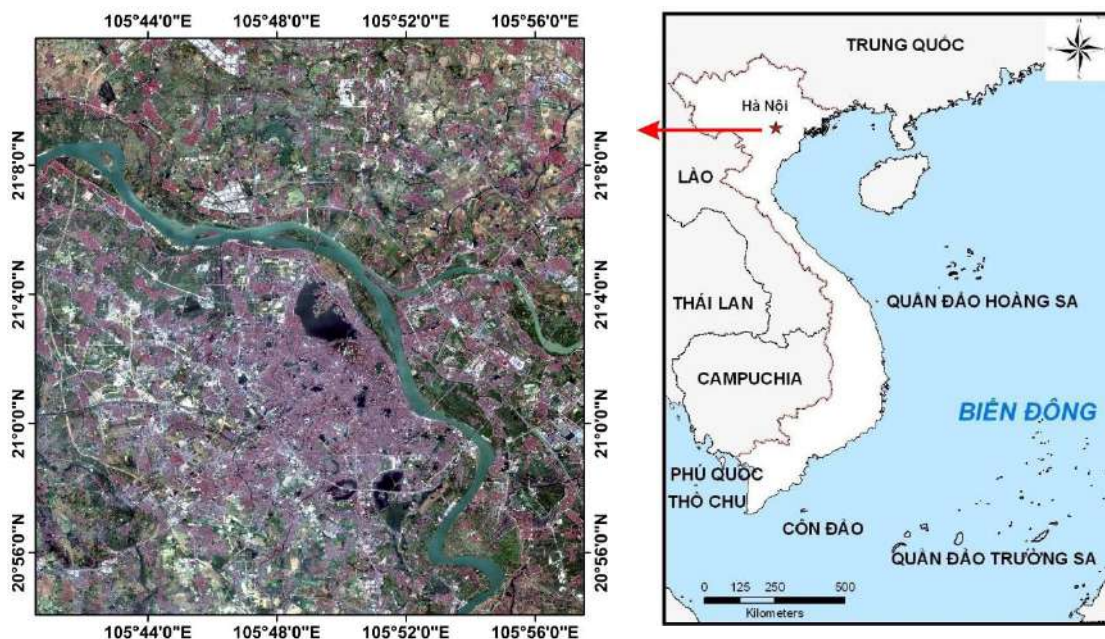
các khu vực đất trống và đất bỏ hoang; từ đó phân biệt nó với thảm thực vật và các loại khác các loại lớp phủ đất khác. Gần đây, chỉ số BSI đã được được lựa chọn rộng rãi để áp dụng trong một số nghiên cứu (Doumit và Sakr, 2015; Ettehad Osgouei và cộng sự, 2019; Tola và cộng sự, 2019; Useya và cộng sự, 2019). Tiếp theo là chỉ số đất trống khô (DBSI) được phát triển gần đây bởi (Rasul và cộng sự, 2018) và đã được đề cập trong phần tổng quan tài liệu của (Lynch và Blesius, 2019) là chỉ số diện tích trống phát hiện dành cho khí hậu khô.

Bài báo nghiên cứu việc lập bản đồ đất xây dựng nhanh chóng và chính xác trong mùa khô thông qua việc thử nghiệm một loại dữ liệu mới kết hợp một số chỉ số phổ có nguồn gốc từ ảnh vệ tinh Sentinel-2 với mục đích tìm ra sự kết hợp đa chỉ số phổ để chiết tách ra đất xây dựng đô thị. Trong thực nghiệm này, nhóm tác giả tận dụng ưu thế của dữ liệu ảnh vệ tinh Sentinel-2 mới với độ phân giải không gian 10 m miễn phí được sử dụng rộng rãi. Nhóm tác giả đã chọn thành phố Hà Nội là khu vực nghiên cứu để thực hiện quá trình tách đất trống và đất xây dựng đô thị trong mùa khô được bao quanh bởi diện tích rộng lớn và đất trống không đồng nhất.

2. Dữ liệu sử dụng và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Thành phố Hà Nội là thủ đô của nước ta với diện tích 3.359,82 km² (Bộ Tài nguyên và Môi trường, 2022) và dân số 8,4 triệu người (Tổng cục Thống kê, 2022). Khu vực có địa hình bao gồm vùng đồng bằng trung tâm và vùng đồi núi ở phía Bắc và phía Tây thành phố. Khí hậu của khu vực Hà Nội được chia làm hai mùa chính là mùa mưa (từ tháng 4 tới tháng 10) và mùa khô (từ tháng 11 tới tháng 3), nhưng thời tiết được chia thành bốn mùa nhờ các tháng giao mùa (Climate of Hanoi, 2021). Năm 2023, Hà Nội lập Đồ án Quy hoạch chung xây dựng Thủ đô Hà Nội đến năm 2030 và tầm nhìn đến năm 2050 cho một thành phố 9,1 triệu dân vào năm 2030 và trên 10 triệu người vào năm 2050 (An Trân, 2010). Vị trí khu vực nghiên cứu của thành phố được đưa ra trong Hình 1.



Hình 1. Tổ hợp màu tự nhiên (RGB=432) của ảnh Sentinel-2 chụp ngày 12/02/2024

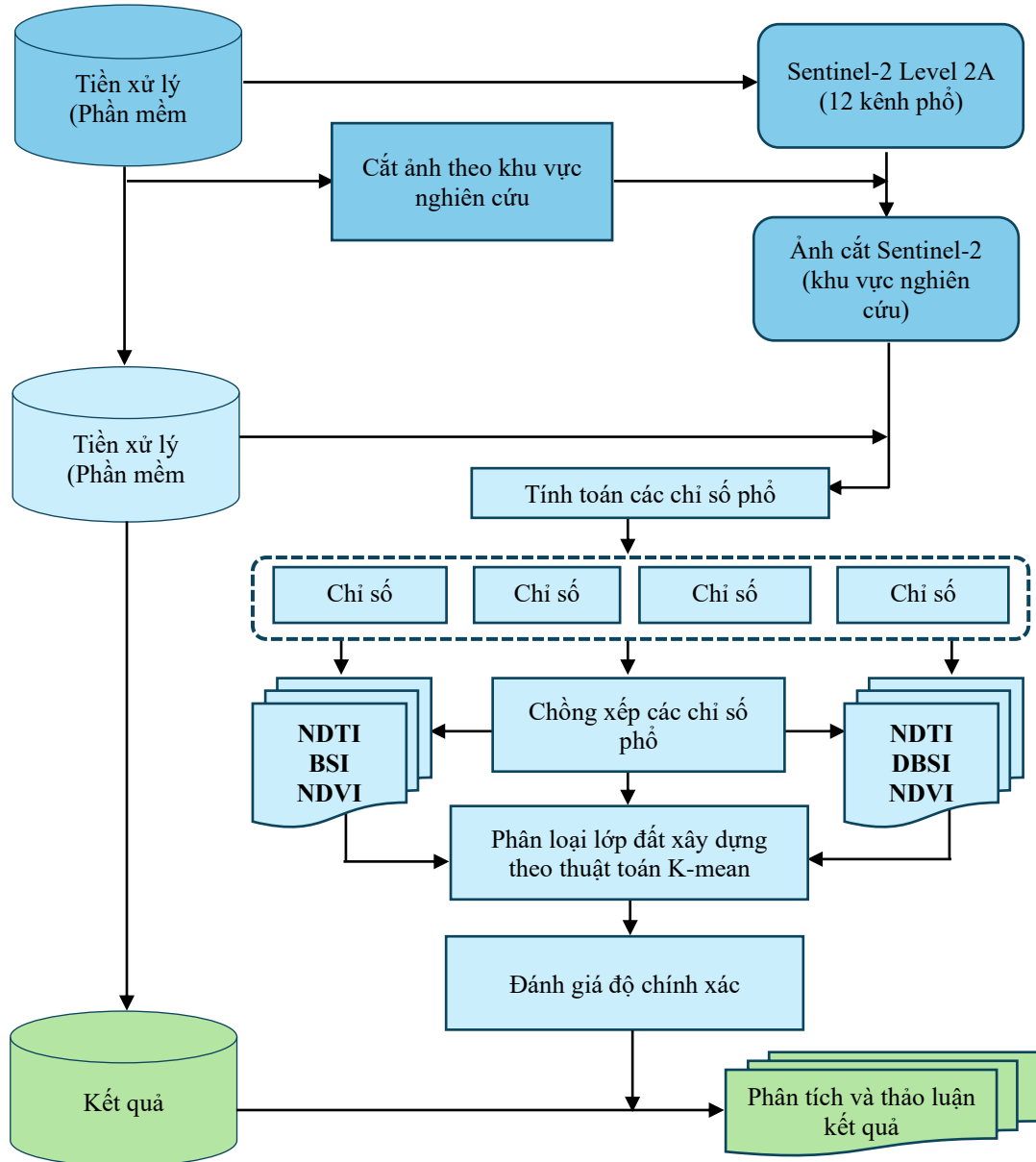
2.2. Dữ liệu sử dụng

Dữ liệu ảnh vệ tinh Sentinel-2 được chọn để làm khu vực nghiên cứu do dữ liệu ảnh vệ tinh này miễn phí có sẵn tại <https://scihub.copernicus.eu> with. Hơn nữa, một số nghiên cứu như công trình của (Xi và cộng sự, 2019) kết luận rằng Sentinel-2 mang lại hiệu quả cao về độ chính xác cho loại đất xây dựng đô thị hơn so với dữ liệu ảnh vệ tinh Landsat-8. Ngoài ra, một nghiên cứu của (Pal và Antil, 2017) cho rằng dữ liệu Sentinel-2 tốt hơn dữ liệu Landsat-8 khi lập bản đồ phân lớp đất xây dựng đô thị bằng cách sử dụng kết hợp các chỉ số phổ khác nhau.

Dữ liệu ảnh vệ tinh Sentinel-2 được thu nhận và xử lý ở mức độ 2A (tên sản phẩm: S2B_MSIL2A_20240212T032849_N0510_R018_T48QWJ_20240212T054533) đã được tham chiếu trước địa lý. Bộ dữ liệu được thu thập vào ngày 12 tháng 02 năm 2024 tương ứng cho tháng khô hạn nhất trong năm bao trùm khu vực nghiên cứu bao gồm thành phố Hà Nội. Dữ liệu ảnh vệ tinh trong ngày đã chọn cung cấp rằng đất trống và đất xây dựng đô thị có thể được quan sát tốt và có thể phân biệt bởi vì không có thảm thực vật hỗn hợp mọc bên trong đất trống do sự trùng hợp với thời điểm kết thúc thu hoạch trong vùng.

2.3. Phương pháp nghiên cứu

Hai sự kết hợp gồm ba chỉ số phổ được thực hiện. Hai kết quả này được sử dụng để phân loại đất xây dựng dựa vào thuật toán K-means. Sự lựa chọn thuật toán K-means phù hợp với việc phân loại để chiết tách một đối tượng đơn ra khỏi các đối tượng còn lại nhanh chóng và chính xác. Phương pháp tổng thể được áp dụng bao gồm ba bước chính: tiền xử lý, xử lý và đánh giá kết quả và được thể hiện trong Hình 2.



Hình 2. Sơ đồ phương pháp trích xuất đất xây dựng đô thị sử dụng tích hợp đa chỉ số phổ của dữ liệu ảnh vệ tinh Sentinel-2 bằng phân loại K-means

2.3.1. Tiền xử lý

Quá trình chồng lớp các kênh phổ và cắt khu vực nghiên cứu được thực hiện bằng cách sử dụng các công cụ trong phần mềm SNAP phiên bản 8.0. Các công cụ xử lý ảnh vệ tinh Sentinel-2 có thể được tìm thấy tại <http://step.esa.int/main/download/>.

2.3.2. Xử lý ảnh

2.3.2.1. Tính toán các chỉ số phổ

Tất cả các chỉ số phổ đã lựa chọn được mô tả cụ thể trong Bảng 1. Trong nghiên cứu này phần mềm ENVI 5.1

với công cụ “Band Math” được sử dụng để tính toán các chỉ số phổ đã được đưa ra để phục vụ cho các công việc tiếp theo. Kết quả tính các chỉ số phổ khu vực nghiên cứu được trình bày trong Hình 3.

Bảng 1. Các chỉ số phổ sử dụng trong nghiên cứu này đối với dữ liệu ảnh vệ tinh Sentinel-2

Chỉ số phổ	Ký hiệu	Tác giả	Công thức đối với dữ liệu ảnh Sentinel-2
Normalized Difference Tillage Index	NDTI	(Deventer, 1997)	$(B11-B12)/(B11+B12)$
Bare Soil Index	BSI	(Rikimaru & Miyatake, 1997)	$((B11+B4)-(B8+B2))/((B11+B4)+(B8+B2))$
Dry Bare-Soil Index	DBSI	(Rasul và cộng sự, 2018)	$((B11-B3)/(B11+B3)) - NDVI$
Normalized Difference Vegetation Index	NDVI	(Tucker, 1979)	$(B8-B4)/(B8+B4)$

2.3.2.2. Kết hợp đa chỉ số phổ

Hệ số tương quan giữa các chỉ số phổ đã chọn được thể hiện trong Bảng 2. Một mối tương quan cao được thấy rõ giữa cả hai chỉ số phổ đất trống BSI và DBSI trong khu vực nghiên cứu. Ngược lại, mối tương quan thấp đã được ghi nhận giữa chỉ số đất xây dựng NDTI và cả chỉ số đất trống BSI và DBSI. Trong khi đó, chỉ số BSI là chỉ số tương quan thấp nhất đối với chỉ số NDVI. Mối tương quan giữa NDVI và chỉ số NDTI là trung bình.

Dựa trên mức độ của mối quan hệ được thảo luận giữa bốn chỉ số phổ, hai sự kết hợp khác nhau được đưa ra thông qua quá trình xếp chồng lớp để gia tăng sự khác biệt phổ giữa ba loại lớp phủ đất chính (đất xây dựng, đất trống, thảm thực vật). Sự kết hợp đã được thực hiện bằng cách giữ chỉ số NDTI và NDVI, trong đó đa chỉ số phổ NDTI/BSI/NDVI là sự kết hợp đầu tiên và NDTI/DBSI/NDVI đa chỉ số là sự kết hợp thứ hai. Những sự kết hợp đa chỉ số phổ này được so sánh với chỉ số phổ NDTI để xác định một kết quả có độ chính xác cao nhất đối với việc trích xuất khu vực đất xây dựng. Ngoài ra, để đánh giá tác dụng của chỉ số đất trống với sự kết hợp ban đầu NDTI/NDVI hướng tới giảm sự phân loại sai giữa đất trống và đất xây dựng.

Bảng 2. Giá trị tương quan giữa các chỉ số phổ sử dụng tại khu vực nghiên cứu

Chỉ số phổ	BSI	DBSI	NDTI	NDVI
BSI	1,000000	0,952421	0,173467	-0,305834
DBSI	0,952421	1,000000	0,124518	-0,332879
NDTI	0,173467	0,124518	1,000000	0,534376
NDVI	-0,305834	-0,332879	0,534376	1,000000

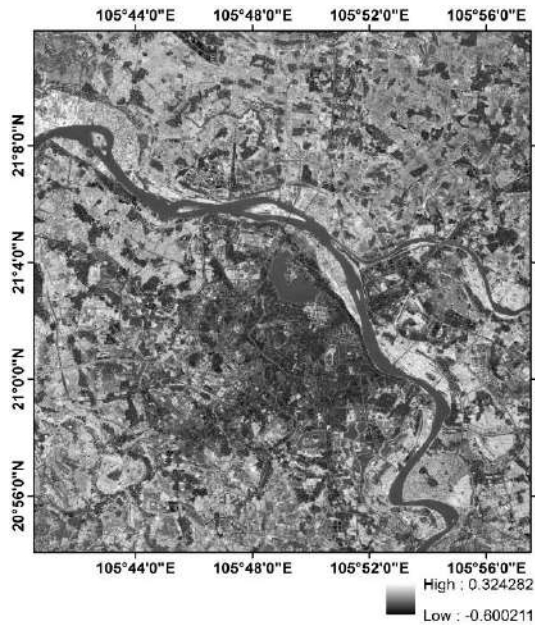
2.3.2.3. Chiết tách lớp đất xây dựng đô thị sử dụng phương pháp K-means

Đất xây dựng đô thị trong khu vực đang nghiên cứu chủ yếu bao gồm khu dân cư và công nghiệp, đường giao thông và các bề mặt không thấm nước khác. Tuy nhiên, lớp không phải đất xây dựng đô thị bao gồm rừng, đất đồng cỏ và đất trống. Để đạt được sự tách biệt chính xác cao giữa hai các lớp chính này (khu vực đất xây dựng và đất chưa xây dựng), sự nhị phân hóa đã được sử dụng bởi các tác giả (Zuur và cộng sự, 2007) để làm nổi bật một loại lớp phủ đất cụ thể thông qua việc tìm ra các ngưỡng tối ưu là một vấn đề lớn khó khăn.

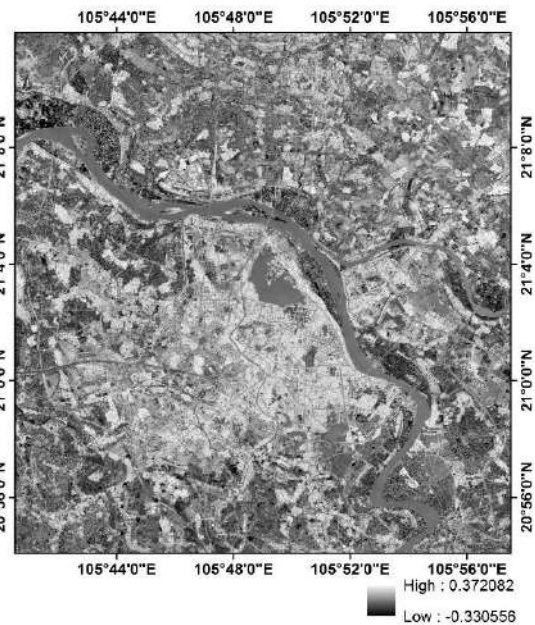
Bảng 3. Giá trị thống kê của bốn chỉ số phổ

Chỉ số phổ	Nhỏ nhất	Lớn nhất	Trung bình	Độ lệch chuẩn	Phạm vi giá trị
BSI	-0,445267	0,552146	0,184725	0,052132	Từ -0,445267 đến 0,552146
DBSI	-0,394781	0,573225	0,236719	0,073513	Từ -0,394781 đến 0,573225
NDTI	-0,326284	0,337416	0,071523	0,043426	Từ -0,326284 đến 0,337416
NDVI	-0,384516	0,901425	0,101642	0,072846	Từ -0,384516 đến 0,901425

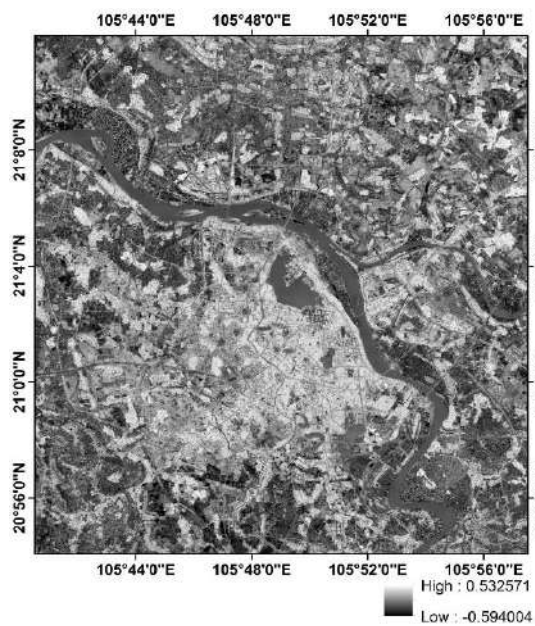
Mặt cắt phổ được trình bày trong Hình 4 cho thấy chỉ số NDVI làm nổi bật thảm thực vật ở vùng có giá trị dương lớn và được thể hiện bằng tông màu xám trắng có thể quan sát được trong Hình 3d. Ngoài ra, BSI và DBSI nêu bật những cánh đồng đất trống và bỏ hoang. Giá trị trung bình đề cập đến đất xây dựng và giá trị thấp đánh dấu thảm thực vật. Thảm thực vật và đất trống trong khu vực có giá trị NDTI dương và đất xây dựng đô thị được thể hiện bằng giá trị NDTI dương thấp. Mặt khác, các công trình xây dựng có kết quả ở giá trị âm được biểu thị bằng các pixel tối trong Hình 3. Ngoài ra, các giá trị thống kê của bốn chỉ số phổ có thể được tìm thấy trong Bảng 3, trong đó giá trị độ lệch chuẩn thấp đối với tất cả các chỉ số phổ được sử dụng.



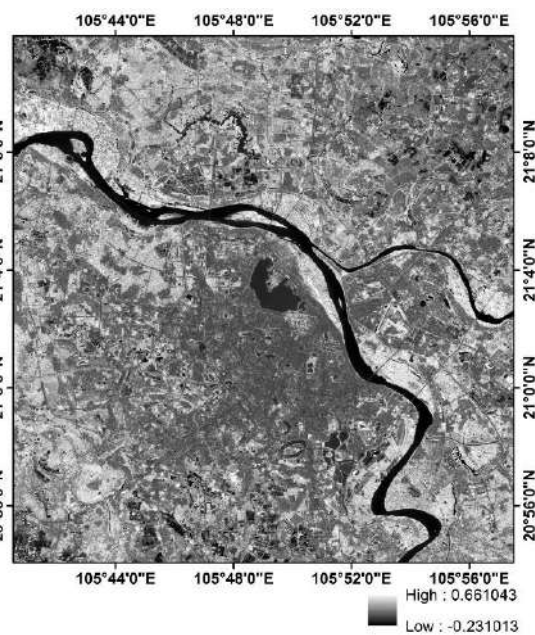
a) Kết quả chỉ số phổ NDTI



b) Kết quả chỉ số phổ BSI



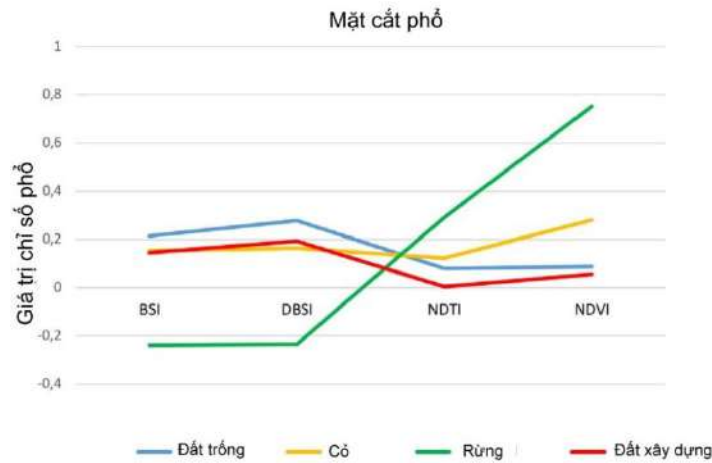
c) Kết quả chỉ số phổ DBSI



d) Kết quả chỉ số phổ NDVI

Hình 3. Kết quả các chỉ số phổ sử dụng trong khu vực nghiên cứu: a) Kết quả chỉ số phổ NDTI; b) Kết quả chỉ số phổ BSI; c) Kết quả chỉ số phổ DBSI và d) Kết quả chỉ số phổ NDVI

Dựa trên phân tích ở trên về mất cắt phổ được trình bày trong Hình 4, quá trình lập bản đồ đất xây dựng từ các loại lớp phủ đất chính khác là đất trống và thảm thực vật có thể được thực hiện dễ dàng bằng cách sử dụng thuật toán phân loại không giám sát K-means được phát triển bởi (MacQueen, 1967). Thuật toán được áp dụng cho một kết quả nhanh chóng và chính xác để phân cụm được chỉ định trước (Gillavata và cộng sự, 2004). Thuật toán K-means được sử dụng thành công để trích xuất một lớp phủ đất cụ thể ở khu vực từ một kết quả chỉ số phổ duy nhất (Li và cộng sự, 2017). Ngoài ra, công trình của (Gašparović và cộng sự, 2019) đã chứng minh hiệu suất của thuật toán cho phân loại lớp phủ đất dựa trên tập dữ liệu đa chỉ số phổ. Việc thực hiện chiết tách lớp đất xây dựng đô thị bằng thuật toán K-means được thực hiện trên phần mềm ENVI 5.1.



Hình 4. Mẫu phổ được biểu thị bằng giá trị trung bình của bốn loại lớp phủ mặt đối với đa chỉ số phổ

3. Kết quả và thảo luận

3.1. Kết quả

Kết quả lớp đất xây dựng đô thị được chiết xuất bằng cách sử dụng chỉ số NDTI và bộ dữ liệu đa chỉ số phổ đã lựa chọn được đưa ra trong Hình 5. Số liệu thống kê được tạo ra từ kết quả phân loại các chỉ số phổ sử dụng thuật toán K-means được thể hiện trong Bảng 4.

Dựa trên giá trị diện tích lớp phủ bề mặt của từng loại chiết tách từ đa chỉ số phổ trong Bảng 4, chúng ta thấy rằng diện tích đất xây dựng là kết quả phân loại từ đa chỉ số phổ NDTI/BSI/NDVI là 593,23 km², chiếm khoảng 57,29% tổng diện tích khu vực nghiên cứu; diện tích đất xây dựng này nhỏ hơn cùng một khu vực lớp này phân loại từ đa chỉ số phổ NDTI/ DBSI/NDVI là 2,32 km² và lớn hơn cùng loại diện tích lớp đất xây dựng khi sử dụng chỉ số NDTI là 61,89 km².

Bảng 4. Diện tích đất xây dựng được chiết tách từ đa chỉ số phổ ở thành phố Hà Nội vào mùa khô tháng 2

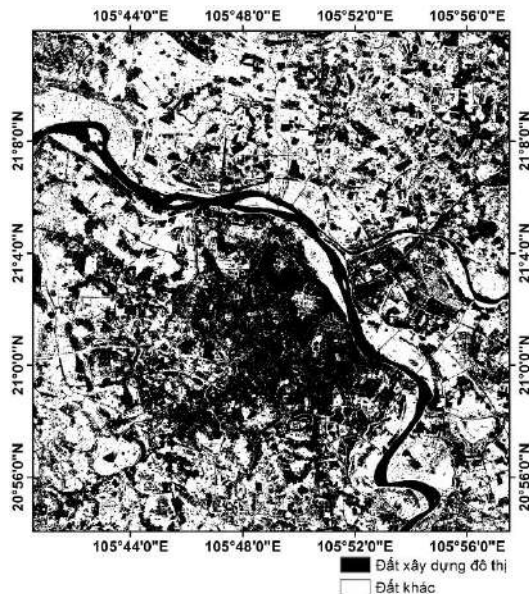
Đa chỉ số phổ	Diện tích đất xây dựng		Diện tích các loại đất khác		Tổng diện tích	
	Diện tích (km ²)	Diện tích (%)	Diện tích (km ²)	Diện tích (%)	Diện tích (km ²)	Diện tích (%)
NDTI/BSI/NDVI	593,23	57,29	442,24	42,71	1035,47	100
NDTI/DBSI/NDVI	595,56	57,52	439,91	42,48	1035,47	100
NDTI	531,34	51,31	504,13	48,69	1035,47	100

3.2. Đánh giá độ chính xác

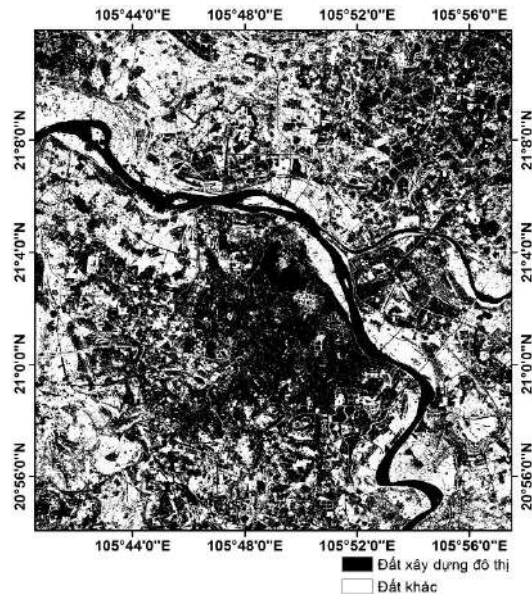
Việc đánh giá độ chính xác khi sử dụng đa chỉ số phổ để chiết tách lớp đất xây dựng cho vùng nghiên cứu cho thấy tính phù hợp của các sự kết hợp đa chỉ số phổ. Do đó, việc đánh giá độ chính xác được thực hiện bằng cách sử dụng phương pháp lấy mẫu ngẫu nhiên phân tầng (Congalton, 1991) đã khuyến nghị cần ít nhất 50 điểm dành cho mỗi lớp; 300 điểm ngẫu nhiên được phân phối giữa các các lớp xây dựng và không xây dựng. Ngoài ra, sử dụng công cụ miễn phí Google Earth lấy các ảnh chụp cùng ngày có độ phân giải cao làm dữ liệu tham khảo. Các ảnh độ phân giải cao được chọn cùng mùa và năm chụp để sự thay đổi các lớp phủ ở khu vực nghiên cứu là ít nhất có thể. Sau đó, ma trận sai số (Foody, 2002) được tính toán là kết quả của việc đánh giá độ chính xác quá trình phân loại bao gồm: độ chính xác tổng thể và hệ số Kappa, độ chính xác sản xuất (PA) và độ chính xác sử dụng (UA). Tất cả các loại độ chính xác được thể hiện trong Bảng 5.

Bảng 5. Đánh giá độ chính xác phân loại đất xây dựng từ đa chỉ số phổ của ảnh Sentinel-2

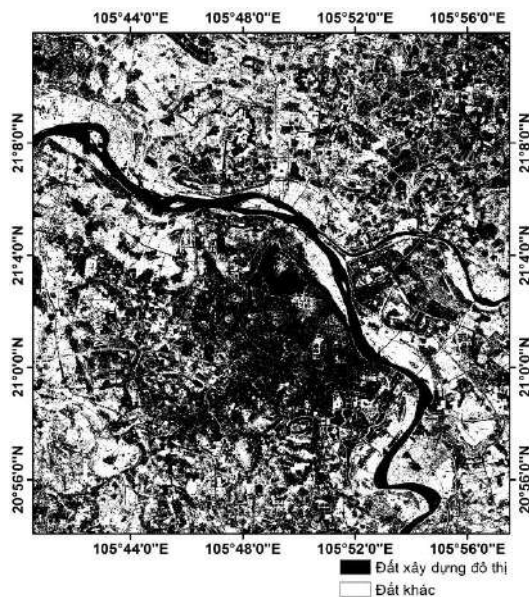
	Chỉ số phổ NDTI		Đa chỉ số phổ NDTI/BSI/NDVI		Đa chỉ số phổ NDTI/DBSI/NDVI	
	ĐCX sử dụng (UA) %	ĐCX sản xuất (PA) %	ĐCX sử dụng (UA) %	ĐCX sản xuất (PA) %	ĐCX sử dụng (UA) %	ĐCX sản xuất (PA) %
Đất xây dựng	79,32	89,32	85,84	92,76	81,38	83,33
Đất khác	91,22	83,84	95,26	91,22	89,28	88,46
Độ chính xác tổng thể (%)	86,43		91,96		86,14	
Hệ số Kappa (%)	72,24		82,34		71,24	



a) Kết quả phân loại K-means chỉ số phổ NDTI



b) Kết quả phân loại K-means đa chỉ số phổ NDTI/BSI/NDVI



c) Kết quả phân loại K-means đa chỉ số phổ NDTI/DBSI/NDVI

Hình 5. Ảnh nhị phân thu được từ phân loại K-means của:

- (a) Chỉ số phổ NDTI
- (b) Đa chỉ số phổ NDTI/BSI/NDVI
- (c) Đa chỉ số phổ NDTI/DBSI/NDVI

Ma trận sai số được trình bày trong Bảng 5 cho thấy khu vực đất xây dựng có độ chính xác tốt hơn khi so sánh sử dụng đa chỉ số phổ NDTI/BSI/NDVI với đa chỉ số phổ NDTI/DBSI/NDVI. Độ chính xác tổng thể đạt được lần lượt là 91,96% và 86,14%, hệ số Kappa đạt được là 82,34% và 71,24% tương ứng. Vì vậy, sự cải thiện độ chính xác nhận thấy ở hệ số Kappa tăng 11,1% và độ chính xác tổng thể tăng 5,82% có một ý nghĩa quan trọng. Trong khi đó, đa chỉ số phổ NDTI/DBSI/NDVI cung cấp kết quả phân loại lớp đất xây dựng có độ chính xác tương tự cả về độ chính xác tổng thể và hệ số Kappa so với kết quả phân loại sử dụng chỉ số NDTI trong nghiên cứu khu vực.

Ma trận sai số cho thấy rằng độ chính xác sử dụng của lớp đất xây dựng phân loại từ đa chỉ số phổ NDTI/BSI/NDVI đạt được 85,84% tăng 4,44% so độ chính xác sử dụng khi phân loại từ đa chỉ số phổ NDTI/DBSI/NDVI. Tương tự như vậy, độ chính xác sản xuất của lớp đất xây dựng phân loại từ chỉ số phổ NDTI/BSI/NDVI đạt 92,76% tăng 9,43% so với độ chính xác sản xuất phân loại từ đa chỉ số phổ NDTI/DBSI/NDVI. Qua quan sát trực quan các kết quả phân loại đất xây dựng được minh họa trong Hình 5, cho thấy việc sử dụng chỉ số NDTI và

phát triển đa chỉ số đã phân biệt được đất xây dựng với loại đất khác các loại. Các quan sát cho thấy rằng một vài pixel đất trống bị phân loại sai là đất được xây dựng, đặc biệt là đất sáng và ngược lại. Chỉ số DBSI tạo ra dải biến động lớn hơn so với chỉ số BSI. Do đó, chỉ số DBSI cung cấp bề mặt đất xây dựng nhiều hơn bề mặt đất trống do chỉ số BSI cung cấp. Điều này chứng tỏ rằng chỉ số BSI có ảnh hưởng đến việc tách đất trống và đất xây dựng bằng cách giảm các pixel bị phân loại sai bằng cách giảm phổ nhằm lẫn lộn tối đa của cả hai lớp. Như vậy, chỉ số BSI đã chứng minh được sự hiệu quả đối với việc lập bản đồ đất trống. Vì lớp rừng có giá trị âm đối với cả hai kết quả lớp đất trống được minh họa trên Hình 4. Nên việc bổ sung chỉ số NDVI làm nổi bật thảm thực vật có giá trị tích cực cao đối với đa chỉ số phổ, cho phép phân loại tự động tách thảm thực vật từ lớp đất trống nằm trong quá trình phân loại.

Các phân tích ở trên đã chứng minh rằng chỉ số BSI trong đa chỉ số phổ NDTI/BSI/NDVI giúp kết quả phân loại đất xây dựng có độ chính xác tốt nhất do tăng khả năng phân biệt đất trống với đất xây dựng. Vì vậy, sử dụng ảnh đa chỉ số phổ NDTI bằng cách thêm chỉ số BSI và NDVI có tác dụng tốt hơn khi chiết tách đất xây dựng với đất trống khi sử dụng phương pháp phân loại K-means.

Kết quả này hỗ trợ phương pháp do (Ettehadi Osgouei và cộng sự, 2019) đề xuất hướng tới cách tiếp cận đa chỉ số phổ dựa trên NDTI. Bên cạnh đó, kết quả nhận được cũng giúp phát triển những phát hiện của (Li và cộng sự, 2017) và (Gašparović và cộng sự, 2019) về phương pháp phân loại dựa trên việc sử dụng xếp chồng lớp chỉ số phổ làm tập dữ liệu đầu vào phân cụm K-means. Do đó, bài báo đã được kiểm tra lại cả hai đa chỉ số phổ trong phân loại đất xây dựng và đất trống và so sánh ảnh hưởng của nó kể từ khi được sử dụng cùng với chỉ số NDTI. Ngoài ra, kết quả cho thấy sự phối hợp tốt giữa ba phát hiện của ba công trình (Li và cộng sự, 2017; Ettehadi Osgouei và cộng sự, 2019; Gašparović và cộng sự, 2019) để tăng cường khả năng phân biệt giữa đất xây dựng và các loại đất khác đã được phát triển từ các dữ liệu được thử nghiệm trong khu vực đô thị vào mùa khô.

4. Kết luận

Thành phố Hà Nội (Thủ đô của Việt Nam) đặc trưng bởi khí hậu có mùa khô và mùa mưa được chọn làm khu vực nghiên cứu. Mục tiêu chính của công việc nghiên cứu này là tìm ra giải pháp tốt nhất sự kết hợp đa chỉ số phổ để chiết tách lớp đất xây dựng có độ chính xác cao ra khỏi các loại lớp phủ đất khác, chủ yếu là lớp phủ đất trống thông qua thực nghiệm phương pháp xếp chồng lớp các chỉ số phổ khác nhau đã được chọn trước đó. Các phần mềm viễn thám khác nhau được sử dụng bao gồm: SNAP, ENVI và ArcGIS để xử lý dữ liệu ảnh vệ tinh Sentinel-2.

Nhìn chung, cả hai sự kết hợp đa chỉ số phổ NDTI/BSI/NDVI và NDTI/DBSI/NDVI đối với dữ liệu ảnh vệ tinh Sentinel-2 đã được thử nghiệm để lập bản đồ đất xây dựng thông qua phân loại K-means. Kết quả nhận được cho thấy quá trình phân loại nhanh chóng cho lớp đất xây dựng. Tuy nhiên, dựa trên ma trận sai số bao gồm độ chính xác tổng thể và hệ số Kappa cho thấy kết quả tốt nhất khi sử dụng đa chỉ số phổ bao gồm chỉ số BSI kết hợp với chỉ số NDTI và NDVI so với sự kết hợp chỉ số DBSI.

Những kết quả như vậy cung cấp sự kết hợp dữ liệu đa chỉ số phổ rất hữu ích để xác định khu vực đất xây dựng đô thị với thông tin chính xác trong mùa khô. Vì vậy, việc lập kế hoạch và ra quyết định có thể nhanh chóng được thực hiện và ước tính khu vực đất xây dựng các thành phố trong mùa khô. Tuy nhiên, việc kiểm tra lại sự kết hợp đa chỉ số phổ thêm trong nghiên cứu này được yêu cầu ở các khu vực khác có sự không đồng nhất về lớp phủ đất xây dựng và đất trống trong cả trong mùa khô và mùa mưa. Ngoài ra, các chỉ số phổ khác cũng cần được nghiên cứu để đề xuất kết hợp đa chỉ số phổ nhằm cải thiện việc lập bản đồ đất xây dựng khu vực đô thị.

Tài liệu tham khảo

- An Trân, 2010. Quy hoạch chung xây dựng Thủ đô Hà Nội đến năm 2030 và tầm nhìn 2050. Hà Nội mới. <https://hanoimoi.vn/qh-chung-xd-thu-do-ha-noi-den-nam-2030-va-tam-nhin-2050-234935.html> (accessed on 24/7/2024).
- Climate of Hanoi, 2021. <https://web.archive.org/web/20210110061853/http://www.hanoi.climatemps.com/> (accessed on 24/7/2024).
- Bouzekri, S., Aziz Lasbet, A., Lachehab, A., 2015. A New Spectral Index for Extraction of Built Up Area Using Landsat-8 Data. *Journal of the Indian Society of Remote Sensing*, 43. <https://doi.org/10.1007/s12524-015-0460-6>.
- Bộ Tài nguyên và Môi trường, 2022. Quyết định số 387/QĐ-BTNMT ngày 02/3/2020 của về phê duyệt và công bố kết quả thống kê diện tích đất đai năm 2020.
- Bramhe, V., Ghosh, S., Garg, P. 2018. Extraction of builtup area by combining textural features and spectral indices from Landsat-8 multispectral image. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-5. <https://doi.org/10.5194/isprsarchives-XLII-5-727-2018>.
- Congalton, R., 1991. *A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data* (Vol. 37). [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B).

- Daughtry, C.S.T., Serbin, G., Reeves, J.B., Doraiswamy, P.C., Hunt, E.R., 2010. Spectral Reflectance of Wheat Residue during Decomposition and Remotely Sensed Estimates of Residue Cover. *Remote Sensing*, 2(2), 416-431. <https://doi.org/10.3390/rs2020416>.
- Deng, C., Wu, C., 2012. BCI: A biophysical composition index for remote sensing of urban environments. *Remote Sensing of Environment*, 127, 247-259. <https://doi.org/10.1016/j.rse.2012.09.009>.
- Deventer, A., Ward, A.D., Gowda, P.H., Lyon, J.G., 1997. Using Thematic Mapper Data to Identify Contrasting Soil Plains and Tillage Practices, *Photogrammetric Engineering and Remote Sensing*, 63(1), 87-93.
- Doumit, J., Sakr, S., 2015. La Cartographie du Sol nu dans la Vallée de la Bekaa à partir de la Teledetection. *InterCarto. InterGIS*, 1, 19-24. <https://doi.org/10.24057/2414-9179-2015-1-21-19-24>.
- Eskandari, I., Navid, H., Rangzan, K., 2016. Evaluating spectral indices for determining conservation and conventional tillage systems in a vetch-wheat rotation. *International Soil and Water Conservation Research*, 4(2), 93-98. <https://doi.org/10.1016/j.iswcr.2016.04.002>.
- Ettehad Osgouei, P., Kaya, S., Sertel, E., Alganci, U., 2019. Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote Sensing*, 11(3). <https://doi.org/10.3390/rs11030345>
- Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), 185-201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4).
- Gašparović, M., Zrinjski, M., Gudelj, M., 2019. Automatic cost-effective method for land cover classification (ALCC). *Computers, Environment and Urban Systems*, 76, 1-10. <https://doi.org/10.1016/j.compenvurbsys.2019.03.001>.
- Gilavata, J., Ewerth, R., Freisleben, B., 2004. Text detection in images based on unsupervised classification of high-frequency wavelet coefficients. In *Proceedings of the 17th International Conference on Pattern Recognition*, 2004. ICPR 2004. <https://doi.org/10.1109/ICPR.2004.1334146>.
- Jieli, C., Manchun, L., Yongxue, L., Chenglei, S., Wei, H., 2010. Extract residential areas automatically by New Built-up Index. In *Proceedings of the 18th International Conference on Geoinformatics*, Beijing, China. <https://doi.org/10.1109/GEOINFORMATICS.2010.5567823>.
- Leroux, L., Congedo, L., Bellón, B., Gaetano, R., Bégué, A., 2018. Land Cover Mapping Using Sentinel-2 Images and the Semi-Automatic Classification Plugin: A Northern Burkina Faso Case Study (pp. 131-165). <https://doi.org/10.1002/9781119457107.ch4>
- Li, H., Wang, C., Zhong, C., Su, A., Xiong, C., Wang, J., Liu, J., 2017. Mapping Urban Bare Land Automatically from Landsat Imagery with a Simple Index. *Remote Sensing*, 9(3), 249. <https://doi.org/10.3390/rs9030249>.
- Liu, Y., Chen, J., Cheng, W., Sun, C., Zhao, S., Yingxia, P., 2014. Spatiotemporal dynamics of the urban sprawl in a typical urban agglomeration: a case study on Southern Jiangsu, China (1983-2007). *Frontiers of Earth Science*, 8, 490-504. <https://doi.org/10.1007/s11707-014-0423-1>.
- Lynch, P., Blesius, L., 2019. Urban Remote Sensing: Feature Extraction, https://www.academia.edu/39831472/Lynch_P_2019_Urban_Remote_Sensing_Feature_Extraction (accessed on 24/7/2024)
- MacQueen, J., 1967. Some Methods for Classification and Analysis of Multivariate Observations. Paper presented at the In L.M. Le Cam & J. Neyman (eds.) *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, CA, USA.
- Nur Hidayati, I., Suharyadi, R., Danoedoro, P. 2018. Developing an Extraction Method of Urban BuiltUp Area Based on Remote Sensing Imagery Transformation Index. *Forum Geograf*, 32. <https://doi.org/10.23917/forgeo.v32i1.5907>.
- Pal, M., Antil, K. 2017. Comparison of Landsat-8 and Sentinel-2 data for Accurate Mapping of Built-Up Area and Bare Soil. *Paper presented at the 38th Asian Conference on Remote Sensing*, New Delhi, India.
- Rasul, A., Balzter, H., Ibrahim, G.R.F., Hameed, H.M., Wheeler, J., Adamu, B.,... Najmaddin, P.M., 2018. Applying Built-Up and Bare-Soil Indices from Landsat-8 to Cities in Dry Climates. *Land*, 7(3), 81. <https://doi.org/10.3390/land7030081>
- Rikimaru, A., Miyatake, S., 1997. Development of Forest Canopy Density Mapping and Monitoring Model using Indices of Vegetation, Bare soil and Shadow. *Paper presented at the Proceedings of the 18th Asian Conference on Remote Sensing (ACRS) 1997*, Kuala Lumpur, Malaysia.
- Sun, G., Chen, X., Jia, X., Yao, Y., Wang, Z., 2016. Combinational Build-Up Index (CBI) for Effective Impervious Surface Mapping in Urban Areas. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 9(5), 2081-2092. <https://doi.org/10.1109/JSTARS.2015.2478914>.
- Tola, E., Al-Gaadi, K.A., Madugundu, R., 2019. Employment of GIS techniques to assess the long-term impact of tillage on the soil organic carbon of agricultural fields under hyperarid conditions. *PLOS ONE*, 14, e0212521. <https://doi.org/10.1371/journal.pone.0212521>.
- Tổng cục Thống kê, 2022. Niên giám Thống kê Việt Nam năm 2021. Nhà xuất bản Thống kê, tr. 89.

- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127-150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0).
- Useya, J., Chen, S., Murefu, M., 2019. Cropland Mapping and Change Detection: Toward Zimbabwean Cropland Inventory. *IEEE Access*, 7, 53603-53620. <https://doi.org/10.1109/ACCESS.2019.2912807>.
- Vigneshwaran, S., Vasantha Kumar, S., 2018. Extraction of Built-Up Area Using High Resolution Sentinel-2a and Google Satellite Imagery. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W9, 165-169. <https://doi.org/10.5194/isprs-archives-XLII-4-W9-165-2018>.
- Waqar, M., Mirza, J., Mumtaz, R., Hussain, E., 2012. Development of New Indices for Extraction of Built-Up Area & Bare Soil from Landsat Data. *Open Access Scientific Reports*, 1(1), 01-04.
- Xi, Y., Thinh, N.X., Li, C., 2019. Preliminary comparative assessment of various spectral indices for built-up land derived from Landsat-8 OLI and Sentinel-2A MSI imagery. *European Journal of Remote Sensing*, 52, 240-252. <https://doi.org/10.1080/22797254.2019.1584737>.
- Xu, H., 2008. A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing*, 29, 4269-4276. <https://doi.org/10.1080/01431160802039957>.
- Yilmaz, E., Varol, B., Topaloğlu, R., Sertel, E., 2019. Object-Based Classification of Izmir Metropolitan City by Using Sentinel-2 Images. In *9th International Conference on Recent Advances in Space Technologies (RAST)*, 407-412. <https://doi.org/10.1109/RAST.2019.8767781>.
- Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24(3), 583-594. <https://doi.org/10.1080/01431160304987>.
- Zuur, A.F., Ieno, E.N., Smith, G.M., 2007. *Principal component analysis and redundancy analysis*. In *Analysing Ecological Data. Statistics for Biology and Health*. Springer, New York, NY, pp. 193-224. <https://doi.org/10.1007/978-0-387-45972-1>.

ABSTRACT

Application of multi-spectral index from Sentinel-2 data for extracting Build-Up land of Hanoi area in the dry season

Nguyen Van Trung^{1,2*}, Le Thi Thu Ha^{1,2}, Nguyen Thanh Len³, Le Thi Minh Phuong³

¹Hanoi University of Mining and Geology

²Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology

³Hanoi Architectural University

Mapping urban areas mostly faces great difficulties, especially in the dry season. Therefore, in this study, we try to increase the accuracy of the urban land map developed for Hanoi city by focusing on determining the appropriate combination of spectral indices calculated from satellite image data. To conduct the study, four spectral indices were selected including namely normalized difference tillage index (NDTI), bare soil index (BSI), dry bare soil index (DBSI) and the normalized difference vegetation index (NDVI). All these spectral indices are calculated from Sentinel-2 data acquired in the dry season. The two combinations are created from the superposition of NDTI/BSI/NDVI and NDTI/DBSI/NDVI spectral index layers. The use of the “K-means” algorithm as an unsupervised classifier provides rapid and automatic urban land detection. The results show that the BSI index performs better than using the DBSI index. As a result, the BSI index brings improvements: bare soil types and accumulation processes are better differentiated, with overall accuracy increasing by 5.82% and Kappa coefficient increasing by 11.1%. The results show that the NDTI/BSI/NDVI multi-spectral index dataset is suitable for mapping urban areas with the potential to help better urban management during the dry season.

Keywords: Sentinel-2, K-means, NDTI, BSI, DBSI.

Phân tích tự tương quan không gian trong nghiên cứu kiểu mẫu phân bố bệnh sốt xuất huyết Dengue: Thử nghiệm tại Thành phố Hồ Chí Minh

Vũ Danh Tuyên^{1*}, Vương Thị Hòe¹, Phạm Thị Thu Hương¹, Nguyễn Tiến Thành¹

¹Khoa Trắc địa, Bản đồ và Thông tin địa lý, Trường Đại học Tài nguyên và Môi trường Hà Nội

TÓM TẮT

Bài báo giới thiệu phương pháp phân tích tự tương quan không gian trong nghiên cứu kiểu mẫu phân bố bệnh sốt xuất huyết Dengue. Chỉ số Moran toàn cục và cục bộ được ứng dụng để xác định mức độ tự tương quan không gian về tỷ lệ nhiễm bệnh sốt xuất huyết Dengue. Kết quả thử nghiệm trên bộ số liệu sốt xuất huyết Dengue được thu thập trong tháng 8 năm 2023 tại Thành phố Hồ Chí Minh cho thấy 04 cụm dị thường không gian với ba low-high và một high-low đã được phát hiện. Các cụm low-high và high-low lần lượt xuất hiện tại các quận huyện phía Tây và phía Nam của thành phố (Quận 4, Bình Chánh, Cần Giờ và Tân Phú). Trong khi đó, không có cụm không gian (high-high và low-low) nào được phát hiện. Kết quả nghiên cứu có ý nghĩa quan trọng giúp xác định sự phân bố về không gian của bệnh sốt xuất huyết Dengue.

Từ khóa: Phân tích tự tương quan không gian, chỉ số Moran toàn cục, chỉ số Moran cục bộ, kiểu mẫu phân bố không gian, sốt xuất huyết Dengue, Thành phố Hồ Chí Minh.

1. Đặt vấn đề

Sốt xuất huyết Dengue là bệnh truyền nhiễm do virus Dengue gây ra. Bệnh có triệu chứng lâm sàng đặc trưng như sốt cao, xuất huyết, thoát huyết tương, nhanh chóng dẫn đến sốc và tử vong nếu không được điều trị kịp thời. Theo WHO, bệnh sốt xuất huyết Dengue được chia làm 3 mức độ: sốt xuất huyết Dengue, sốt xuất huyết Dengue có dấu hiệu cảnh báo và sốt xuất huyết Dengue nặng (WHO, 2009). Khảo sát dịch tễ học cho thấy rằng sự lây nhiễm sốt xuất huyết Dengue lên đến 2/5 dân số thế giới với khoảng 390 triệu ca mắc mới hàng năm, dẫn đến 500.000 ca nhập viện và khoảng 20.000 ca tử vong (Khan và cộng sự, 2023). Tại Việt Nam, trong những năm gần đây, sốt xuất huyết có biểu hiện lâm sàng đa dạng, diễn biến nhanh và phức tạp, tỷ lệ diễn biến nặng và tử vong cao. Do đó, nghiên cứu kiểu mẫu phân bố của bệnh sốt xuất huyết có vai trò quan trọng trong nghiên cứu đặc điểm dịch tễ sốt xuất huyết. Sốt xuất huyết Dengue là bệnh lưu hành chủ yếu tại khu vực phía Nam (Ngô Văn Đình và cộng sự, 2022). Trong đó, Thành phố Hồ Chí Minh là khu vực thường xuyên được ghi nhận các ca nhiễm sốt xuất huyết Dengue cao.

Trên thế giới, nhiều nghiên cứu đã thành công ứng dụng phương pháp phân tích tự tương quan không gian để nghiên cứu sự phân bố về không gian của dịch bệnh, đặc biệt đối với các bệnh truyền nhiễm như đại dịch Covid-19 do SARS-CoV-2 gây ra (Alphonso và cộng sự, 2024; Santos và cộng sự, 2024), bệnh lao (Gemechu và cộng sự, 2023) hay bệnh AIDS/HIV (Cao Thị Hiền và cộng sự, 2024; Johnson và cộng sự, 2024). Trong bài báo này, phương pháp phân tích tự tương quan không gian được thử nghiệm để xác định kiểu mẫu phân bố bệnh sốt xuất huyết Dengue ở Thành phố Hồ Chí Minh.

2. Dữ liệu sử dụng và phương pháp nghiên cứu

2.1. Dữ liệu sử dụng

Theo thống kê của Trung tâm kiểm soát bệnh tật Thành phố Hồ Chí Minh, mặc dù trong 6 tháng đầu năm 2023, số lượng ổ dịch sốt xuất huyết được ghi nhận giảm 62% so với cùng kỳ năm 2022. Tuy nhiên, theo quy luật diễn tiến dịch bệnh hàng năm, mùa cao điểm của bệnh sốt xuất huyết đã bắt đầu tăng vào tuần thứ 24, sẽ tăng cao trong tháng 7 và dự kiến sẽ kéo dài đến hết tháng 10 hàng năm. Trong khi đó, qua giám sát điểm nguy cơ, Trung tâm Kiểm soát bệnh tật Thành phố Hồ Chí Minh phát hiện có lăng quăng tại các điểm nguy cơ được giám sát là 47,6% (49/103 điểm). Đây là con số đáng báo động, tỷ lệ này sẽ cao hơn nữa khi Thành phố mưa nhiều hơn và không có những biện pháp quyết liệt để kiểm soát các điểm nguy cơ. Kết quả trong nghiên cứu của Viên Chinh Chiến và cộng sự (2022) chỉ ra rằng những tháng mùa mưa thường có số ca mắc cao hơn những tháng mùa khô, trong đó số mắc cao thường được ghi nhận từ tháng 8 đến tháng 10 hàng năm. Do đó, trong nghiên cứu này, số liệu về tỷ lệ nhiễm bệnh sốt xuất huyết Dengue trong tháng 8 năm 2023 đã được lựa chọn và thu thập tại Thành phố Hồ Chí Minh để thử nghiệm đánh giá kiểu mẫu phân bố bệnh sốt xuất huyết Dengue.

* Tác giả liên hệ

Email: vdtuyen@hunre.edu.vn

2.2. Chỉ số Moran toàn cục

Nghiên cứu sử dụng chỉ số tự tương quan không gian Moran toàn cục để xác định kiểu mẫu không gian của bệnh sốt xuất huyết Dengue tổng thể. Chỉ số Moran toàn cục do Patrick Alfred Pierce Moran (1950) đề xuất và được Anselin Luc (1995) phát triển theo công thức (1) dưới đây:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

trong đó: x_i và x_j là tỷ lệ nhiễm bệnh sốt xuất huyết Dengue tại quận/huyện i và j ; \bar{x} là trị trung bình số học của tỷ lệ nhiễm bệnh sốt xuất huyết Dengue và được xác định bằng công thức: $\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$; n là tổng số quận/huyện trên toàn bộ khu vực nghiên cứu và W_{ij} là ma trận trọng số không gian ($n \times n$).

Giống với hệ số tương quan Pearson, giá trị chỉ số Moran toàn cục nằm trong phạm vi $[-1; 1]$. Nếu $0 < I < 1$ phản ánh về tổng thể có sự tương quan thuận (giá trị quan sát được tập trung cao (hoặc thấp)). Nếu $-1 < I < 0$ phản ánh tổng thể có sự tương quan nghịch (giá trị cao tập trung xung quanh giá trị thấp hay ngược lại). Nếu $I = 0$ thể hiện kiểu mẫu không gian được phân bố một cách ngẫu nhiên (random spatial patterns).

Chỉ số tự tương quan Moran toàn cục chỉ xem xét mức độ tương quan không gian tổng thể trên toàn khu vực nghiên cứu. Tuy nhiên, trong nhiều trường hợp, Moran toàn cục phản ánh sự phân bố của các giá trị quan sát được là ngẫu nhiên, nhưng trong một phạm vi cục bộ nào trên khu vực nghiên cứu lại tồn tại kiểu mẫu tập trung (hay phân tán) về mặt không gian của các trị quan sát được. Đây là nhược điểm của Moran toàn cục. Nhược điểm này có thể được khắc phục bằng cách sử dụng chỉ số Moran cục bộ.

2.3. Chỉ số Moran cục bộ

Chỉ số tự tương quan Moran cục bộ (I_i) đo lường mức độ tự tương quan không gian cục bộ của tỷ lệ nhiễm sốt xuất huyết Dengue tại vùng i theo công thức (2):

$$I_i = \frac{(x_i - \bar{x})}{\sigma^2} \sum_{j \# i, j \in J_i}^N W_{ij} (x_j - \bar{x}) \quad (2)$$

trong đó: x_i , x_j , \bar{x} , và W_{ij} đã được định nghĩa trong công thức (1); N là tổng số vùng lân cận (26); J_i là tập hợp tất cả các vùng lân cận của vùng i ; $j \# i$ ám chỉ tổng của $(x_j - \bar{x})$ tại các vùng lân cận của vùng i nhưng không bao gồm x_i ; and σ^2 là phương sai và được xác định bằng công thức (3). W_{ij} là ma trận kết nối vùng lân cận và có thể được xây dựng trên cơ sở nguyên lý sự tiếp giáp các vùng lân cận.

$$\sigma^2 = \frac{1}{N} \sum_{j=1}^N (x_j - \bar{x})^2 \quad (3)$$

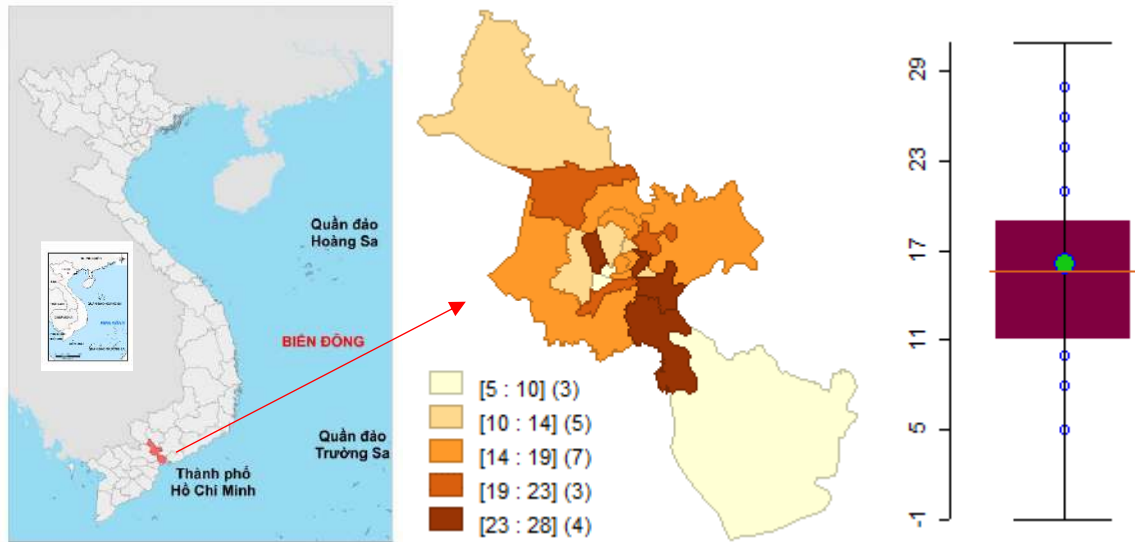
Nếu $p(I_i) < \alpha$ ($\alpha = 0,05$ là giá trị được chọn trong nghiên cứu này), có nghĩa là khó có khả năng kiểu mẫu không gian quan sát được do quá trình ngẫu nhiên sinh ra, bác bỏ giả thiết gốc H_0 , chấp nhận giả thiết đối H_1 : kiểu mẫu không gian quan sát được là phân tán hoặc tập trung và có ý nghĩa về mặt thống kê. Nếu $p(I_i) < \alpha$ và $I_i > 0$ và $x_i - \bar{x} > 0$, thì giá trị quan sát được x_i và $x_{j \in J_i}$ đều là các giá trị cao và có sự tập trung cao về không gian (high-high cluster) của tỷ lệ nhiễm sốt xuất huyết Dengue ở mức cao, gọi là cụm cao - cao. Nếu $p(I_i) < \alpha$ và $I_i > 0$ và $x_i - \bar{x} < 0$, thì giá trị quan sát được x_i và $x_{j \in J_i}$ đều là các giá trị thấp và có sự tập trung cao về không gian (low-low cluster) của tỷ lệ nhiễm sốt xuất huyết Dengue ở mức thấp, gọi là cụm thấp - thấp. Nếu $p(I_i) < \alpha$ và $I_i < 0$ và $x_i - \bar{x} > 0$, thì các giá trị quan sát $x_{j \in J_i}$ thấp hơn hẳn x_i và có sự phân tán về không gian (high-low cluster), gọi là cụm cao - thấp. Trong khi đó, nếu $p(I_i) < \alpha$ và $I_i < 0$ và $x_i - \bar{x} < 0$, thì các giá trị quan sát $x_{j \in J_i}$ cao hơn hẳn x_i và có sự phân tán về không gian (low-high cluster), gọi là cụm thấp - cao. Cụm cao - cao và cụm thấp - thấp được gọi là cụm không gian (spatial clusters), cụm cao thấp và cụm thấp - cao gọi là dị thường không gian (spatial outliers) (Nguyen và cộng sự, 2016).

3. Kết quả và thảo luận

3.1. Sự phân bố của các ca nhiễm sốt xuất huyết

Sự phân bố về không gian của tỷ lệ nhiễm bệnh sốt xuất huyết/100.000 dân trong tháng 8 năm 2023 tại Thành phố Hồ Chí Minh được thể hiện trong Hình 1. Dữ liệu trong Hình 1 (phải) cho thấy tỷ lệ nhiễm bệnh sốt xuất huyết ở mức cao có xu hướng tập trung về tại khu vực phía trung tâm của thành phố, đặc biệt là tại các quận phía Nam của Thành phố. Trong tháng 8, số lượng ca nhiễm sốt xuất huyết thấp nhất và cao nhất lần lượt là 3 và 28 ca. Giá trị trung bình và trung vị lần lượt đạt 16,1 và 15,5. Biểu đồ boxplot trong Hình 1 (trái) cho thấy phần lớn là các quận huyện với tỷ lệ nhiễm sốt xuất huyết ở mức khá cân bằng do giá trị trung vị xuất hiện gần giá trị trung

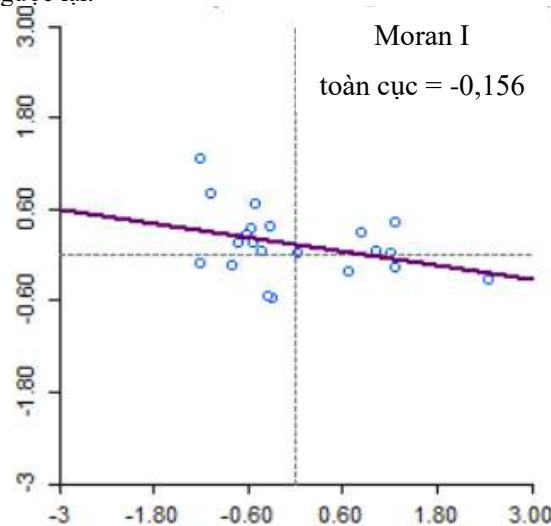
binh. Trong khi đó, với biểu đồ boxplot, khoảng tứ vị phân thứ nhất và thứ ba có giá trị lần lượt là 11 và 19. Giá trị tham số đo độ phân tán IQR và SDEV lần lượt là 8 và 6,1.



Hình 1. Số lượng ca nhiễm sốt xuất huyết trong tháng 8 năm 2023 tại Thành phố Hồ Chí Minh: sự phân bố không gian (trái) và biểu đồ boxplot (phải)

3.2. Sự phân bố cụm không gian của tỷ lệ nhiễm sốt xuất huyết

Giá trị tương quan Moran I toàn cục phản ánh mức độ tự tương quan không gian của tỷ lệ số lượng ca nhiễm sốt xuất huyết trong tháng 8 năm 2023 của Thành phố Hồ Chí Minh. Biểu đồ phân tán Moran I trong Hình 2 cho thấy chỉ số tự tương quan Moran I toàn cục có giá trị -0,156 (nhỏ hơn 0). Về tổng thể, chỉ số tự tương quan Moran I toàn cục thể hiện tồn tại mối tương quan nghịch giữa các quận/huyện có số lượng ca nhiễm sốt xuất huyết thể hiện bằng các quận/huyện có số lượng ca nhiễm lớn có xu hướng tập trung xung quanh các quận huyện có số lượng ca nhiễm thấp và ngược lại.

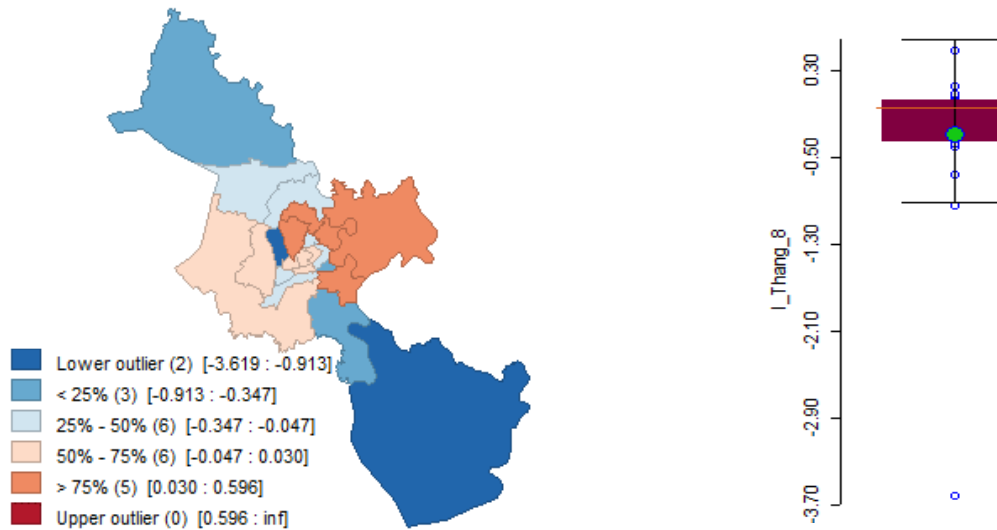


Hình 2. Biểu đồ phân tán Moran I và chỉ số Moran toàn cục

Sự phân bố về không gian của giá trị Moran I cục bộ xác định từ tỷ lệ nhiễm sốt xuất huyết trong tháng 8 năm 2023 thể hiện trong Hình 3. Các số liệu thống kê mô tả của Moran cục bộ như sau: giá trị thấp nhất và cao nhất lần lượt là -3,6 và 0,49, trị trung bình và trung vị lần lượt đạt -0,28 và -0,05. Dữ liệu trong Hình 3 cho thấy các quận/huyện có chỉ số Moran lớn hơn 0 tập trung ở một số quận huyện phía Đông và phía Tây của trung tâm thành phố. So sánh với kết quả từ việc chỉ số Moran I toàn cục có giá trị -0,156 cho thấy chỉ số tự tương quan Moran toàn cục chỉ xem xét mức độ tương quan không gian tổng thể trên toàn khu vực nghiên cứu. Tuy nhiên, từ số liệu về tỷ lệ ca nhiễm sốt xuất huyết Dengue của tháng 8 năm 2023 tại Thành phố Hồ Chí Minh có thể xác nhận rằng chỉ số Moran toàn cục phản ánh sự phân bố của các giá trị quan sát được có thể coi là ngẫu nhiên,

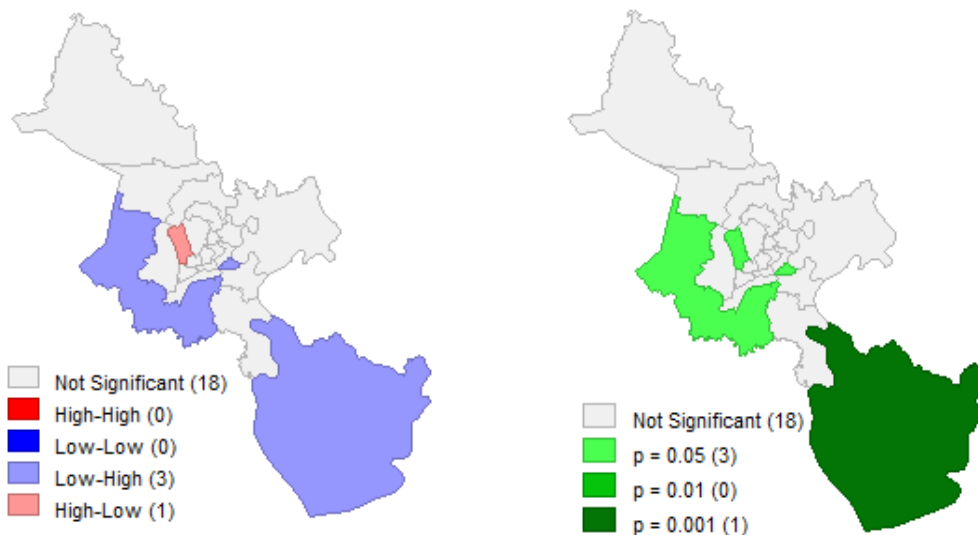
nhưng trong một phạm vi cục bộ (như tại phía Đông và phía Tây của thành phố) lại tồn tại kiểu mẫu tập trung (hay phân tán) về mặt không gian của tỷ lệ ca nhiễm sốt xuất huyết Dengue.

Biểu đồ boxplot từ Hình 3 (phải) cho thấy khoảng tứ vị phân thứ nhất và thứ ba có giá trị lần lượt là -0,35 và 0,03. Giá trị đo độ phân tán IQR và SDEV lần lượt là 0,38 và 0,8. Dữ liệu từ biểu đồ Boxplot trong Hình 3 (phải) cho thấy chỉ số Moran I trong tháng 8 bị lệch về bên trên (phía có giá trị lớn hơn không) chứng tỏ tỷ lệ nhiễm của sốt xuất huyết trong tháng 8 ở mức cao. Dữ liệu về sự phân bố của chỉ số Moran cục bộ trong Hình 3 (phải) cho thấy, trong tháng 8 năm 2023, Moran cục bộ có giá trị thấp cũng chủ yếu được phát hiện tại các quận/huyện phía Bắc và Nam của thành phố, trong khi đó chỉ số Moran có giá trị trung bình và cao lần lượt được phát hiện tại các quận huyện phía Tây và Đông của thành phố. Đặc biệt, giá trị Moran cục bộ có giá trị cao phân bố tập trung ở khu vực phía Đông của Thành phố Hồ Chí Minh.



Hình 3. Phân bố không gian của chỉ số Moran cục bộ (trái) và biểu đồ boxplot (phải) về mức độ tự tương quan không gian của tỷ lệ nhiễm sốt xuất huyết Dengue

Kết quả xác định cụm không gian về tỷ lệ nhiễm sốt xuất huyết được thể hiện trong Hình 4. Về tổng thể, dữ liệu từ Hình 4 cho thấy, dị thường không gian chủ yếu được phát hiện ở trung tâm Thành phố. Trong tháng 8, chỉ số Moran cục bộ phát hiện 4 cụm dị thường với 03 low-high và 01 high-low. Các cụm low-high và high-low lần lượt xuất hiện tại các quận huyện phía Tây và phía Nam của Thành phố gồm Quận 4 (11 ca/100.000 dân), Bình Chánh (16 ca/100.000 dân), Cần Giờ (5 ca/100.000 dân) và Tân Phú (24 ca/100.000 dân). Không có cụm không gian (high-high và low-low) được phát hiện trong tháng 8 mặc dù vẫn có một số quận huyện có ca nhiễm cao như Nhà Bè (28 ca/100.000 dân), Quận 1 (26 ca/100.000 dân), và Quận 7 (24 ca/100.000 dân). Kết quả so sánh với với số liệu về tỷ lệ nhiễm sốt xuất huyết Dengue của tháng 7 cho thấy cụm không gian trong tháng 8 có xu hướng dịch chuyển về phía Tây và phía Nam của Thành phố Hồ Chí Minh.



Hình 4. Phân bố cụm không gian (trái) và mức đạt ý nghĩa thống kê về tỷ lệ nhiễm sốt xuất huyết Dengue

Dữ liệu từ Hình 4 (phải) thể hiện sự phân bố không gian về mức đạt về ý nghĩa thống kê cho từng quận huyện của Thành phố Hồ Chí Minh. Các mức thống kê cũng được thể hiện theo 4 thang đo từ không đạt ý nghĩa thống kê ($p\text{-value} > 0,05$) và đạt ý nghĩa thống kê ở các mức 0,05; 0,01 và 0,001. Hình 3 (phải) cho thấy chỉ có một quận đạt mức ý nghĩa rất cao tại mức 0,001 tại huyện Cần Giờ. Ba quận/huyện đạt mức ý nghĩa thống kê tại mức 0,05 gồm Bình Chánh, Tân Phú và Quận 4.

4. Kết luận

Bài báo đã tiến hành thử nghiệm áp dụng phương pháp phân tích tự tương quan không gian trong nghiên cứu kiểu mẫu phân bố bệnh sốt xuất huyết Dengue. Trong nghiên cứu, chỉ số Moran toàn cục và cục bộ đã được ứng dụng để xác định mức độ tự tương quan không gian về tỷ lệ nhiễm bệnh sốt xuất huyết Dengue tại Thành phố Hồ Chí Minh. Kết quả nghiên cứu từ số liệu sốt xuất huyết Dengue thu thập trong tháng 8 năm 2023 cho thấy 04 cụm dị thường không gian với ba low-high và một high-low đã được phát hiện. Các cụm low-high và high-low lần lượt xuất hiện tại các quận huyện phía Tây và phía Nam của Thành phố (Quận 4, Bình Chánh, Cần Giờ và Tân Phú). Ngoài ra, không phát hiện cụm không gian (high-high và low-low) nào được phát hiện. Kết quả nghiên cứu có ý nghĩa quan trọng giúp xác định sự phân bố về không gian của bệnh sốt xuất huyết Dengue, đặc biệt là dịch tễ sốt xuất huyết Dengue.

Tài liệu tham khảo

- World Health Organization, 2009. Dengue guidelines for diagnosis, treatment, prevention and control: new edition.
- Khan MB, Yang Z-S, Lin C-Y et al., 2023. Dengue overview: An updated systemic review. *J Infect Public Health*, 16: 1625-42.
- Viên Chinh Chiến, Phạm Ngọc Thanh, Nguyễn Lê Mạnh Hùng, Lý Thị Thùy Trang và Vũ Sinh Nam, 2022. Thực trạng bệnh sốt xuất huyết Dengue tại khu vực Tây Nguyên, giai đoạn 2000-2020. *Tạp chí Y học Dự phòng*, 32, số 2, 46-52.
- Alphonso, Sophie R., Marcus R. Andrews, Seann D. Regan, Alyssa Shishkov, Jonathan H. Cantor, Tiffany M. Powell-Wiley, and Kosuke Tamura, 2024. Geospatially clustered low Covid-19 vaccine rates among adolescents in socially vulnerable US counties. *Preventive Medicine Reports* 37: 102545.
- Santos, V. S., T. S. Siqueira, J. R. S. Silva, and R. Q. Gurgel, 2023. Spatial clustering of low rates of Covid-19 vaccination among children and adolescents and their relationship with social determinants of health in Brazil: A nationwide population-based ecological study. *Public Health*, 214: 38-41.
- Gemechu, Leta Lencha, and Legesse Kassa Debusho, 2023. Spatial Clustering of Tuberculosis-HIV Coinfection in Ethiopia at Districts Level. *AIDS Research and Treatment*.
- Cao Thi Hien, Dao Thi Hong, 2024. Identification of spatial clusters and spatial outliers of HIV/AIDS using local Moran's I statistic. *World Journal of Biology Pharmacy and Health Sciences*, 17, No. 1: 101-108.
- Johnson, K.A., McDaniel, J.T., Graham, H.K., Robertson, E.T., McIntosh, S., Wallace, J.P. and Albright, D.L., 2024. A Geospatial Analysis of Social and Structural Determinants of Health and High HIV Prevalence in Alabama, USA. *Journal of Community Health*, 49(3), pp. 385-393.
- Ngô Văn Đình, Lương Chấn Quang, Diệp Thanh Hải, Đỗ Kiến Quốc, Nguyễn Thị Thanh Thảo, Võ Tố Uyên, Nguyễn Vũ Thượng, Vũ Hải Hà và Phan Trọng Lân, 2022. Đặc điểm dịch tễ học bệnh sốt xuất huyết Dengue tại khu vực phía Nam giai đoạn 2001-2020. *Tạp chí Y học Dự phòng* 32(2): 25-35.
- Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*. 37(1): 17-23. Doi:10.2307/2332142.
- Anselin, Luc, 1995. Local indicators of spatial association-LISA. *Geographical Analysis* 27(2): 93-115.
- Nguyen, Tien Thanh, Vu Danh Tuyen, and Nguyen Thi Le Hang, 2016. Spatial cluster and outlier identification of geochemical association of elements: A case study in Juirui copper mining area. *Bulletin of the Mineral Research and Exploration* 153, 153: 159-167.

ABSTRACT

Spatial autocorrelation analysis in the study of spatial patterns of dengue hemorrhagic fever: A case study in Ho Chi Minh City

Vu Danh Tuyen^{1*}, Vuong Thi Hoe¹, Pham Thi Thu Huong¹, Nguyen Tien Thanh¹

¹*Faculty of Surveying, Mapping and Geographic Information,
Hanoi University of Natural Resources and Environment*

This paper introduces the method of spatial autocorrelation analysis in the study of spatial patterns of dengue hemorrhagic fever. The global and local Moran's I statistic were applied to measure the degree of spatial autocorrelation in dengue hemorrhagic fever infection rates. Results from the dengue hemorrhagic fever dataset collected in August 2023 in Ho Chi Minh City showed that four spatial outliers with three low-high and one high-low outliers were successfully detected. These low-high and high-low spatial outliers were mainly detected in the western and southern districts of the city (District 4, Binh Chanh, Can Gio and Tan Phu), respectively. No high-high and low-low spatial clusters were detected. The research results have important implications for identifying the spatial distribution of dengue hemorrhagic fever.

Keywords: Spatial autocorrelation analysis, global Moran I statistic, local Moran I statistic, spatial patterns, Dengue hemorrhagic fever, Ho Chi Minh City.

Sử dụng dữ liệu vệ tinh Sentinel-2 thành lập bản đồ sinh vật đáy khu vực quần đảo Côn Đảo, tỉnh Bà Rịa - Vũng Tàu, Việt Nam

Nguyễn Văn Trung^{1,2*}, Lê Thị Thu Hà^{1,2}, Đoàn Thị Nam Phương^{1,2}, Phạm Thị Hương¹

Nguyễn Quang Minh³, Phạm Vọng Thành⁴

¹Trường Đại học Mở - Địa chất

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

³Cục Bản đồ - Bộ Tổng tham mưu

⁴Hội Trắc địa - Bản đồ - Viễn thám Việt Nam

TÓM TẮT

Dữ liệu vệ tinh đa phổ Sentinel-2A và 2B (MSI) cung cấp thông số kỹ thuật về giá trị tiềm năng hướng tới mục tiêu trong thành lập bản đồ các sinh vật đáy. Quan trắc các rạn san hô và cỏ biển là một thách thức đặc biệt đối với công nghệ viễn thám do sự không đồng nhất phản xạ phổ xảy ra ở các độ sâu khác nhau, các thành phần vật chất trong nước. Tuy nhiên, các giải pháp mới đã giúp cho dữ liệu viễn thám có thể làm tăng thêm khả năng cung cấp các thông tin giám sát môi trường quy mô lớn. Để thực hiện công việc này, một số dữ liệu vệ tinh đa phổ đã được sử dụng bao gồm: Landsat, Sentinel-2 và dữ liệu có độ phân giải không gian cao như WorldView-2 để quan trắc các rạn san hô phục vụ việc phát hiện các sự kiện tẩy trắng san hô.

Dữ liệu vệ tinh đa phổ Sentinel-2 được chứng minh có tính hiệu quả trong việc thành lập bản đồ các sinh vật đáy với khả năng loại trừ các phản xạ phổ của các tia sáng trên bề mặt nước. Độ phân giải không gian 10 m là ngưỡng quan trọng để phân định các đặc điểm sinh vật đáy được quan tâm như các rạn san hô với các dữ liệu kiểm chứng từ hình ảnh và thực địa ở quần đảo Côn Đảo. Ngoài ra, dữ liệu được cung cấp bởi vùng rộng 290 km mang lại lợi thế so với hình ảnh có độ phân giải không gian siêu cao.

Từ khóa: Sentinel-2, MSI, rạn san hô, quần đảo Côn Đảo.

1. Đặt vấn đề

Các rạn san hô hiện đang bị đe dọa trên toàn thế giới từ nhiều nguồn ảnh hưởng về môi trường và con người (Hoegh-Guldberg và cộng sự, 2007; Pandolfi và cộng sự, 2003). Do đó, chính sách và quản lý hiệu quả sự hình thành đòi hỏi phải có sự giám sát trên tất cả các quy mô từ địa phương đến toàn cầu. Mặc dù không chi tiết như khảo sát tại hiện trường, nhưng công nghệ viễn thám cung cấp một phương pháp quan trắc có giá trị, đặc biệt là ở quy mô lớn khi việc khảo sát trực tiếp sẽ rất tốn kém và khó thực hiện (Hedley và cộng sự, 2016). Mục tiêu lập bản đồ sinh vật đáy và bao gồm các loại sinh vật đáy như san hô và tảo, phát hiện sự thay đổi của sinh vật đáy và hiện tượng tẩy trắng san hô yêu cầu dữ liệu viễn thám trong điều kiện phạm vi không gian và thời gian cụ thể (Hock và cộng sự, 2014).

Để lập bản đồ rạn san hô, các lựa chọn hiện nay dựa trên vệ tinh đang chiếm ưu thế bằng các dữ liệu đa phổ có độ phân giải không gian cao (pixel ≤ 5 m) và độ phân giải trung bình (pixel 10-50 m). Các nỗ lực với các dữ liệu thương mại chủ yếu hướng tới dữ liệu đa phổ có độ phân giải không gian rất cao (ví dụ: Worldview-4, Pleiades, pixel kích thước 2 m, chòm sao Planet Labs, 3-5 m) (Mumby và cộng sự, 1997; Hochberg và Atkinson, 2003). Tuy nhiên, các rạn san hô rất cao không đồng nhất ở quy mô lớn, các loài sinh vật đáy khác nhau các loại như san hô, cỏ biển hoặc tảo vĩ mô thường xuất hiện ở các quy mô không gian khác nhau và do đó độ phân giải không gian tối thiểu tối ưu có thể khác nhau tùy thuộc vào mục tiêu và đặc điểm khu vực nghiên cứu (Phinn và cộng sự, 2010; Yamano, 2013).

Về cơ sở khoa học, các rạn san hô nằm ở giữa lớp phủ đất và đại dương. Chính vì vậy, các dữ liệu vệ tinh có sẵn thường đáp ứng những yêu cầu cụ thể. Dữ liệu Landsat có sẵn miễn phí đã được sử dụng ở rạn san hô đáp ứng cho tiềm năng lập dự án bản đồ môi trường sống rạn san hô thiên niên kỷ (Smith và cộng sự, 1975; Bour, 1988; Andréfouët và cộng sự, 2004; Palandro và cộng sự, 2008). Landsat-8 nói riêng, với độ phân giải bức xạ 12 bit và dải màu xanh bổ sung ở bước sóng 443 nm mang lại sự thâm nhập sâu trong vùng nước (El-Askary và cộng sự, 2014; Giardino và cộng sự, 2016). Dữ liệu được thu nhận trong một vùng rộng 185 km có ý nghĩa lớn khi xử lý mà không gặp khó khăn khi ghép hình ảnh thu được trong các điều kiện khác nhau. Các ứng dụng siêu phổ như Hyperion hoặc HICO cũng đã được thử nghiệm (Kutser và cộng sự, 2006; Garcia và cộng sự, 2014a) nhưng chưa có cam kết sứ mệnh lâu dài. Các dữ liệu khác như HypIRI hoặc EnMAP (Devred và cộng sự, 2013; Guanter và cộng sự, 2015) có thể ứng dụng siêu phổ cho giám sát rạn san hô nhưng các lựa chọn sử dụng dữ liệu vệ tinh này vẫn còn là các dịch vụ thương mại.

* Tác giả liên hệ

Email: nguyenvantrung@humg.edu.vn

Dữ liệu Sentinel-2 khác biệt đáng kể so với các dữ liệu đã thảo luận trước đó ở cả độ phân giải không gian và thời gian. Độ phân giải không gian 10 m ở bốn dải nhìn thấy và cận hồng ngoại (NIR) nằm giữa dữ liệu đa phổ có độ phân giải không gian cao 2 m và dữ liệu đa phổ Landsat-8 ở 30 m. Việc chụp lặp lại sau 5 ngày của cặp vệ tinh Sentinel-2A và 2B cung cấp khả năng lựa chọn dữ liệu ít mây và có nắng. Phạm vi dữ liệu của Sentinel-2 là 290 km rộng hơn phạm vi ảnh của Landsat là 185 km mang lại lợi thế về tính đồng nhất của khu vực rộng hơn (Roy và cộng sự, 2017). Trong mô phỏng phản xạ phổ ở độ phân giải 10 m trên sinh vật đáy tối các đặc điểm như san hô có giá trị phổ trên kênh 1 tương ứng quá cao, giá trị ở mức trung bình trên một diện tích lớn hơn (độ phân giải 60 m). Sau đó trong một phân tích độ sâu, những pixel này được dùng để ước tính sai số ở vùng nước sâu (Hedley và cộng sự, 2012).

Trong phạm vi nghiên cứu này, tại Việt Nam hiện nay còn ít nghiên cứu quan tâm đến việc ứng dụng sự ưu việt của các loại dữ liệu viễn thám sẵn có trong quan trắc, giám sát và thành lập bản đồ sinh vật đáy cho các khu vực trên lãnh thổ Việt Nam. Nghiên cứu (Hoàng và cộng sự, 2020) đã sử dụng ảnh vệ tinh Landsat 7 TM và Landsat-8 OLI để tiến hành theo dõi hiện trạng và biến động của các hệ sinh thái các thảm cỏ biển tại đầm Lăng Cô, tỉnh Thừa Thiên Huế. Năm 2012, nghiên cứu (Cao và cộng sự, 2013) đã nghiên cứu các thảm cỏ tại vùng ven biển Việt Nam bằng cách thành lập bản đồ phân bố của thảm cỏ biển. Năm 2016, nghiên cứu (Chen và cộng sự, 2016) đã đánh giá sự biến động của thảm cỏ biển theo thời gian tại vịnh Cam Ranh, Việt Nam. Do đặc tính kỹ thuật của dữ liệu viễn thám về độ phân giải không gian, phổ và bức xạ nên các nghiên cứu có phạm vi giới hạn ở một vài loại sinh vật đáy biển. Trong nghiên cứu này đối với dữ liệu đa phổ Sentinel-2 tập trung vào hai loại sinh vật bao gồm rạn san hô và cỏ biển.

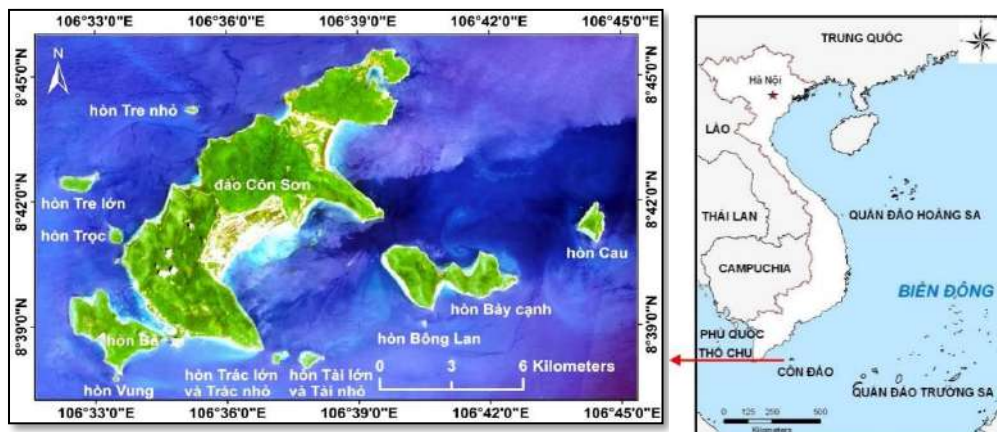
Bài viết này tập trung sử dụng các kỹ thuật xử lý dữ liệu đa phổ Sentinel-2 để thành lập bản đồ sinh vật đáy, đặc biệt là các rạn san hô và cỏ biển để phục vụ quản lý, giám sát các hiện tượng biến đổi của chúng trong bối cảnh nước biển ấm lên do biến đổi khí hậu. Thuận lợi của dữ liệu Sentinel-2 với độ phân giải không gian phù hợp và chế độ chụp lặp 5 ngày giúp đáp ứng được các điều kiện sinh vật đáy cụ thể ở khu vực Côn Đảo. Các kỹ thuật loại trừ phản xạ phổ của tia sáng trên bề mặt nước, tính toán chỉ số độ sâu không đổi và phân loại các sinh vật đáy được áp dụng để thành lập bản đồ sinh vật đáy với sự kiểm chứng của các dữ liệu ảnh vệ tinh độ phân giải rất cao.

2. Dữ liệu sử dụng và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Quần đảo Côn Đảo nằm cách thành phố Vũng Tàu 97 hải lý có kinh độ địa lý 106°36' Đông và vĩ độ địa lý 8°36' Bắc (Hình 1). Quần đảo gồm có 16 hòn đảo với tổng diện tích đất nổi là 76 km², trong đó đảo lớn nhất là đảo Côn Sơn 51,52 km². Đảo này có địa hình đồi núi, ưu thế bởi các dãy đá granit chạy từ phía Tây Nam đến Đông Bắc, che chở cho các vùng vịnh của quần đảo cả về hai phía khỏi những luồng gió mạnh. Quần đảo Côn Đảo nằm ở vùng giao nhau giữa luồng hải lưu ấm từ phía Nam và luồng hải lưu lạnh từ phía Bắc. Nhiệt độ nước biển từ 25,7 °C đến 29,2 °C (Trần Đức Thanh và cộng sự, 2012).

Vùng biển của quần đảo có 1.383 loài sinh vật biển, trong đó có 127 loài rong biển, 11 loài cỏ biển, 157 loài thực vật phù du, 115 loài động vật phù du, 202 loài cá, 8 loài thú và bò sát biển. Các rạn san hô nơi đây do 219 loài hợp thành; độ phủ trung bình là 42,6 %. Côn Đảo không chỉ là vùng có nhiều rùa biển nhất nước Việt Nam mà còn là nơi duy nhất ở Việt Nam còn tồn tại một quần thể bò biển có cuộc sống không tách rời các thảm cỏ biển (<https://www.condaopark.com.vn/vn/?vnTRUST=mod:about>).



Hình 1. Tổ hợp màu tự nhiên (RGB=432) của ảnh Sentinel-2 chụp ngày 29/9/2023 khu vực Côn Đảo

2.2. Dữ liệu sử dụng

Dữ liệu ảnh vệ tinh Sentinel-2 được chọn để làm khu vực nghiên cứu do dữ liệu ảnh vệ tinh này miễn phí có sẵn tại <https://scihub.copernicus.eu>. Các kênh phổ của dữ liệu Sentinel-2 được đưa ra trong Bảng 1. Để định lượng cho phạm vi toàn cầu, một phân tích dựa trên việc chia ô lưới dữ liệu Sentinel-2 đã được tiến hành (SUHET, 2015), trong đó danh sách tất cả các ô lưới 110 km × 110 km dữ liệu Sentinel-2 chứa các rạn san hô đã được tạo ra. Các dữ liệu còn thiếu sẽ được thu nhận bổ sung để cung cấp đầy đủ dữ liệu để giám sát rạn san hô trên toàn cầu.

Dữ liệu ảnh vệ tinh Sentinel-2 được thu nhận và xử lý ở mức độ 2A (tên sản phẩm: S2B_MSIL2A_20230929T030529_N0509_R075_T48PXQ_20230929T063909) đã được tham chiếu địa lý. Bộ dữ liệu được thu thập vào ngày 29 tháng 9 năm 2023 tương ứng cho tháng mùa mưa trong năm bao trùm khu vực nghiên cứu quần đảo Côn Đảo. Dữ liệu ảnh vệ tinh trong ngày đã chọn cung cấp rằng lượng mây che phủ rất ít và nằm ở khu vực đất trên quần đảo giúp cho việc quan trắc các sinh vật đáy ở các khu vực nước nông ven bờ được tốt nhất.

Bảng 1. Các kênh phổ của dữ liệu đa phổ Sentinel-2

(Thông tin từ: <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/radiometric>)

Sentinel-2		
Kênh phổ	Bước sóng trung tâm (nm)	Độ phân giải (m)
B1 – Coastal aerosol	443	60
B2 – Blue (B)	490	10
B3 – Green (G)	560	10
B4 – Red (R)	665	10
B5 – Red edge (RE1)	705	20
B6 – Red edge (RE2)	740	20
B7 – Near infrared narrow1 (NIRn1)	783	20
B8 – Near infrared (NIR)	842	10
B8A – Near infrared narrow2 (NIRn2)	865	20
B9 – Water vapor	945	60
B10 – SWIR Cirrus	1375	60
B11 – Shortwave infrared (SWIR1)	1610	20
B12 – Shortwave infrared (SWIR2)	2190	20

2.3. Phương pháp nghiên cứu

Phương pháp tổng thể được áp dụng bao gồm ba bước chính: Tiền xử lý, xử lý dữ liệu và đánh giá kết quả và được thể hiện trong Hình 2.

2.3.1. Tiền xử lý

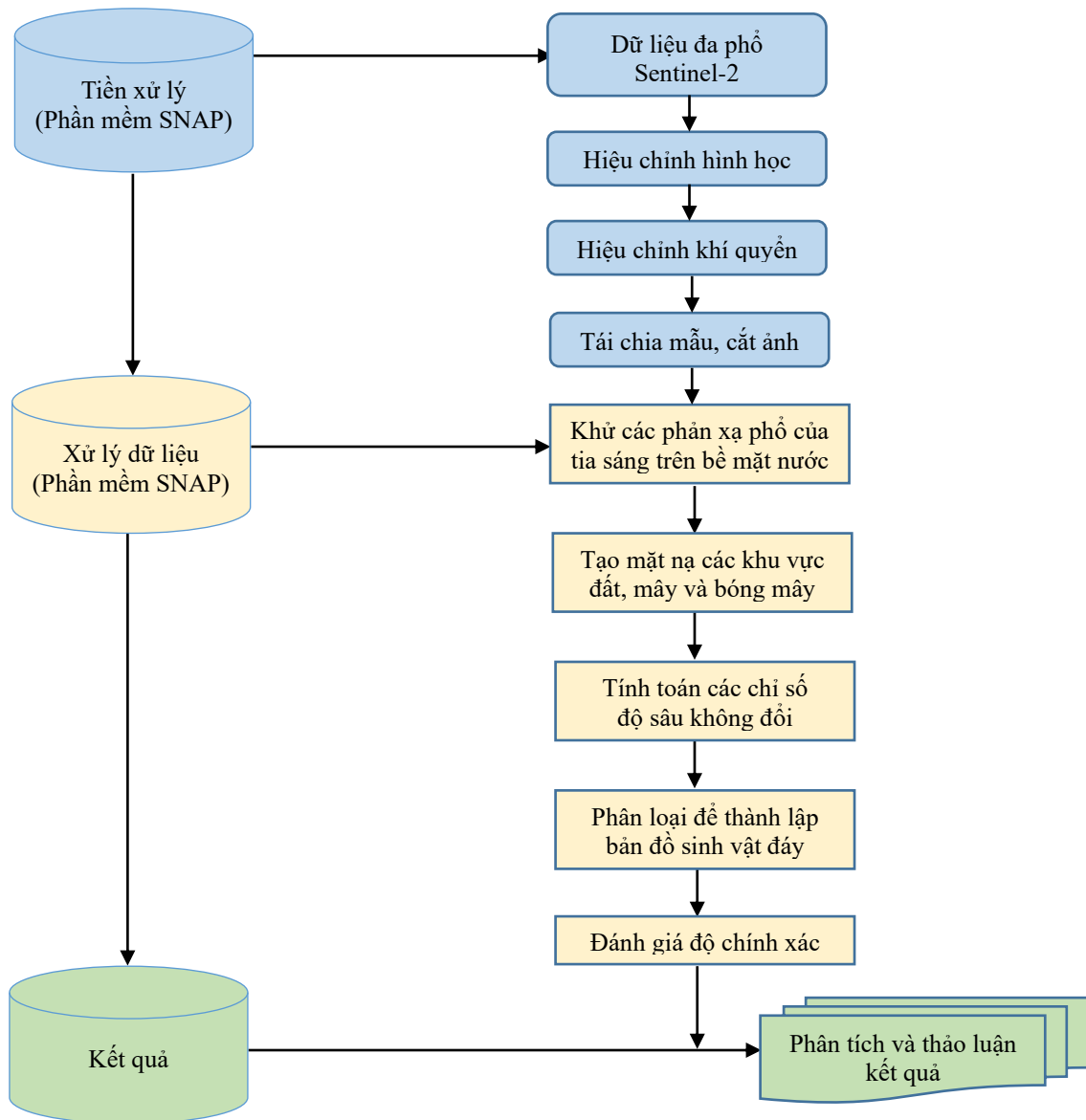
Quá trình chồng lớp các kênh phổ và cắt khu vực nghiên cứu được thực hiện bằng cách sử dụng các công cụ trong phần mềm SNAP phiên bản 8.0. Các công cụ xử lý ảnh vệ tinh Sentinel-2 có thể được tìm thấy tại <http://step.esa.int/main/download/>.

2.3.2. Xử lý ảnh

2.3.2.1. Khử các phản xạ phổ của tia sáng trên bề mặt nước (Deglint)

Các phản xạ phổ của tia sáng mặt trời là một hiện tượng phổ biến trong ảnh vệ tinh đề cập đến sự phản chiếu sự phản chiếu của mặt trời trên bề mặt nước. Độ phản xạ từ bề mặt nước có thể khó quan sát do sự phản xạ của ánh sáng mặt trời trực tiếp trên bề mặt phân cách giữa không khí và nước theo hướng vệ tinh. Dữ liệu đa phổ của vệ tinh Sentinel-2 dễ bị xảy ra các phản xạ phổ của tia sáng mặt trời. Khi có phản xạ phổ của tia sáng mặt trời, chúng ta cần áp dụng thuật toán loại bỏ phổ phản xạ này để có thể quan sát được đáy biển nhằm mục đích lập bản đồ sinh vật đáy.

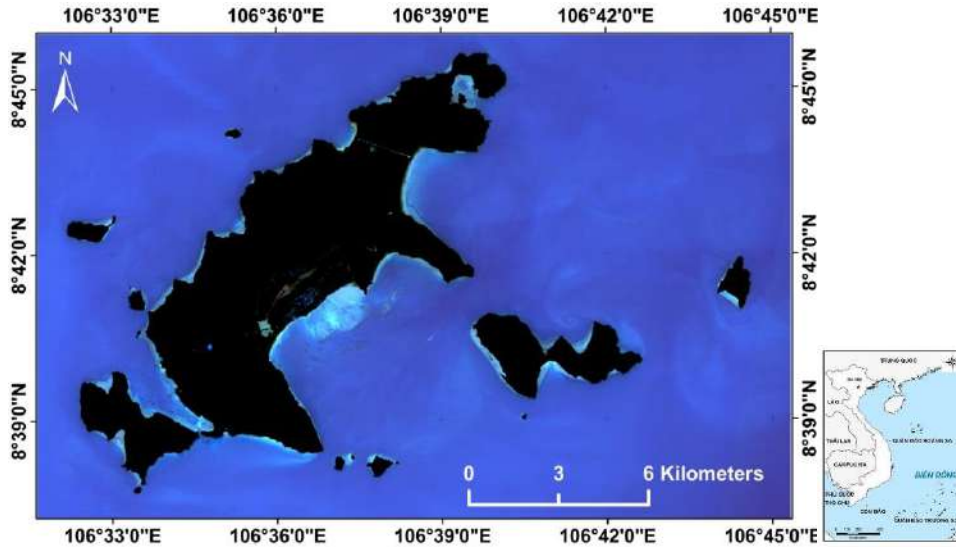
Phương pháp loại bỏ các phản xạ phổ của tia sáng mặt trời cho dữ liệu vệ tinh có độ phân giải cao áp dụng cho các vùng ven biển. Thuật toán được phát triển trong công cụ Sen2Coral bởi (Hedley và cộng sự, 2005). Mặc dù thuật toán cho kết quả khá tốt nhưng nên sử dụng dữ liệu có các phản xạ phổ của tia sáng tối thiểu.



Hình 2. Sơ đồ phương pháp thành lập bản đồ rạn san hô từ dữ liệu đa phổ Sentinel-2

2.3.2.2. Tạo mặt nạ các khu vực đất, mây và bóng mây

Một bước tiền xử lý cần thiết khác là tạo mặt nạ che phủ đất, đỉnh bọt biển trên sóng, mây và bóng mây. Ba cái đầu tiên đề cập đến các đặc trưng độ sáng có thể được xác định dựa vào các bước sóng hồng ngoại gần (NIR). Kênh phổ này không xâm nhập vào nước nên sau khi khử phản xạ phổ của tia sáng các vùng nước trong rất tối. Mây, đỉnh bọt biển trên sóng và đất thường có nhiệt độ cao phản xạ trong NIR, do đó chúng ta có thể đặt ngưỡng để tạo mặt nạ để che chúng. Tuy nhiên, mặt nạ này không thể che các đặc điểm tối như bóng mây hoặc bóng địa hình. Các đặc trưng tối có thể được phân định khi sử dụng ngưỡng cho kênh phổ 2 để tạo mặt nạ bổ sung. Kết quả của quá trình xử lý loại bỏ các phản xạ phổ của tia sáng mặt trời trên bề mặt nước và tạo mặt nạ các khu vực đất, đỉnh bọt biển trên sóng, mây và bóng mây được đưa ra trong Hình 3.



Hình 3. Tổ hợp màu tự nhiên (RGB=432) của ảnh Sentinel-2 sau khi Deglint và tạo mặt nạ (màu đen)

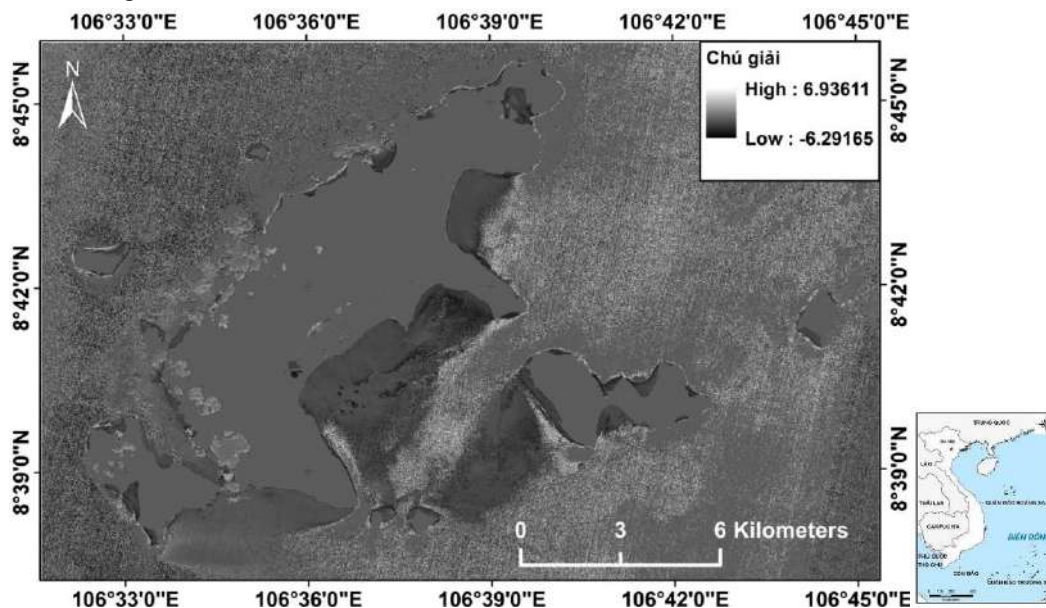
2.3.2.3. Tính toán các chỉ số độ sâu không đổi

Việc tính toán các chỉ số độ sâu không đổi là bước hiệu chỉnh cột nước rất hữu ích của quá trình xử lý dữ liệu để phân loại sinh vật đáy. Việc phát triển công cụ Sen2Coral dựa theo (Lyzenga, 1978; Green và cộng sự, 2000). Dựa vào cơ sở của sự suy giảm của độ phản xạ xấp xỉ theo hàm mũ nghịch đảo với độ sâu của nước, do đó sự biến đổi:

$$X_i = -\ln(R_i - R_i^{deep}) \quad (1)$$

Xấp xỉ tuyến tính hóa ảnh hưởng của độ sâu lên hệ số phản xạ, R_i là độ phản xạ pixel trong kênh phổ i và R_i^{deep} là độ phản xạ nước sâu trong kênh phổ i đó. Biểu đồ các giá trị pixel của hai biến đổi các kênh phổ trên một loại sinh vật đáy có hệ số phản xạ không đổi nhưng ở các độ sâu khác nhau sẽ nằm trên một đường thẳng.

Để tính các chỉ số độ sâu không đổi, chúng ta cần chọn hai kênh phổ có khả năng xuyên qua nước để quan sát các đặc trưng dưới bề mặt đáy cần quan tâm. Đối với Sentinel-2, các kênh phổ 1-3 là hữu ích nhất, mặc dù các kênh phổ 1-5 có thể đáp ứng điều kiện. Chúng ta không cần biết độ suy giảm nước riêng lẻ các hệ số trong hai kênh phổ đầu vào đã chọn, nhưng chúng ta cần biết tỷ lệ của chúng. Điều này có thể được tính toán từ các pixel mẫu trên cùng loại đáy ở các độ sâu khác nhau. Kết quả của quá trình xử lý tính toán các chỉ số độ sâu không đổi được đưa ra trong Hình 4.

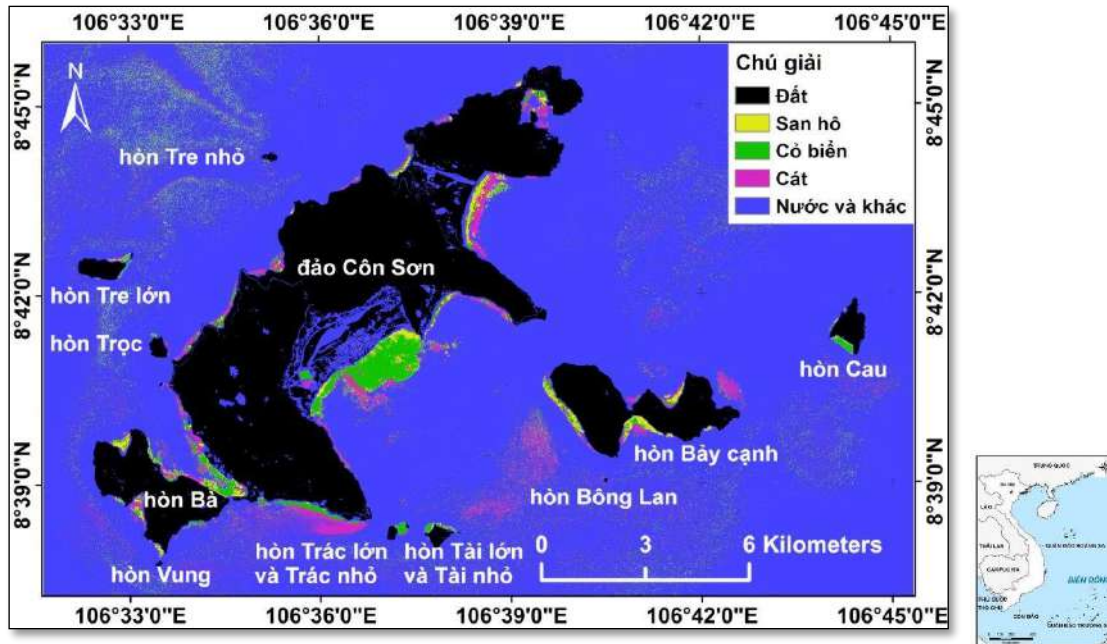


Hình 4. Chỉ số độ sâu không đổi cho vùng biển tính từ dữ liệu đa phổ Sentinel-2

3. Kết quả và thảo luận

3.1. Kết quả phân loại để thành lập bản đồ sinh vật đáy

Sử dụng phương pháp phân loại xác suất cực đại để phân loại thành các lớp gồm san hô, cỏ, cát ở khu vực nước nông ven bờ từ kết quả bản đồ chỉ số độ sâu không đổi với các vùng mẫu lấy từ các dữ liệu vệ tinh độ phân giải siêu cao. Các lớp đất trên đảo hoặc vùng nước sâu cũng được phân thành lớp riêng biệt để tạo thành bản đồ sinh vật đáy trong Hình 5.



Hình 5. Kết quả phân loại các sinh vật đáy cho vùng biển chiết tách từ dữ liệu đa phổ Sentinel-2

Dựa trên giá trị diện tích các sinh vật đáy chiết tách từ đa chỉ số phổ trong Bảng 2, chúng ta thấy rằng diện tích rạn san hô là 5,72 km², chiếm khoảng 24,51% tổng diện tích. Diện tích lớp cỏ biển lớn nhất 9,06 km², chiếm khoảng 38,82% tổng diện tích. Phần còn lại là diện tích lớp cát 8,56 km², chiếm khoảng 36,67% tổng diện tích các sinh vật đáy ở khu vực nghiên cứu.

Bảng 2. Diện tích các sinh vật đáy cho vùng biển chiết tách từ dữ liệu đa phổ Sentinel-2

Diện tích san hô		Diện tích cỏ biển		Diện tích cát		Tổng diện tích	
(km ²)	(%)	(km ²)	(%)	(km ²)	(%)	(km ²)	(%)
5,72	24,51	9,06	38,82	8,56	36,67	23,34	100

3.2. Đánh giá độ chính xác

Việc đánh giá độ chính xác khi phân loại các sinh vật đáy từ kết quả tính toán các chỉ số độ sâu không đổi cần có các dữ liệu kiểm chứng. Do đó, việc đánh giá độ chính xác được thực hiện bằng cách sử dụng phương pháp lấy mẫu ngẫu nhiên cần ít nhất 50 điểm dành cho mỗi lớp; 231 điểm ngẫu nhiên được phân phối giữa các lớp sinh vật đáy. Ngoài ra, sử dụng công cụ miễn phí Google Earth lấy các ảnh chụp cùng ngày có độ phân giải cao làm dữ liệu tham khảo. Sau đó, ma trận sai số (Foody, 2002) được tính toán là kết quả của việc đánh giá độ chính xác quá trình phân loại bao gồm: độ chính xác tổng thể và hệ số Kappa, độ chính xác sản xuất (PA) và độ chính xác sử dụng (UA). Các công thức tính toán cho các loại độ chính xác được đưa ra cụ thể sau đây:

Độ chính xác tổng thể được xác định như sau:

$$OA = \frac{\sum_{i=1}^k p_{ii}}{N} \quad (2)$$

Hệ số Kappa được xác định như sau:

$$Kappa = \frac{p_o - p_c}{1 - p_c} \quad (3)$$

Độ chính xác sản xuất (PA) và độ chính xác sử dụng (UA) được tính toán như sau:

$$PA = \frac{p_{ii}}{\sum_{i=1}^k p_{i+}} \quad (4)$$

$$UA = \frac{p_{ii}}{\sum_{j=1}^k p_{+j}} \quad (5)$$

trong đó:

$$p_0 = \frac{\sum_{i=1}^k p_{ii}}{N}; \quad p_c = \frac{\sum_{i,j=1}^k p_{i+} p_{+j}}{N}$$

p_{ii} là các phần tử trong đường chéo chính của ma trận sai số,

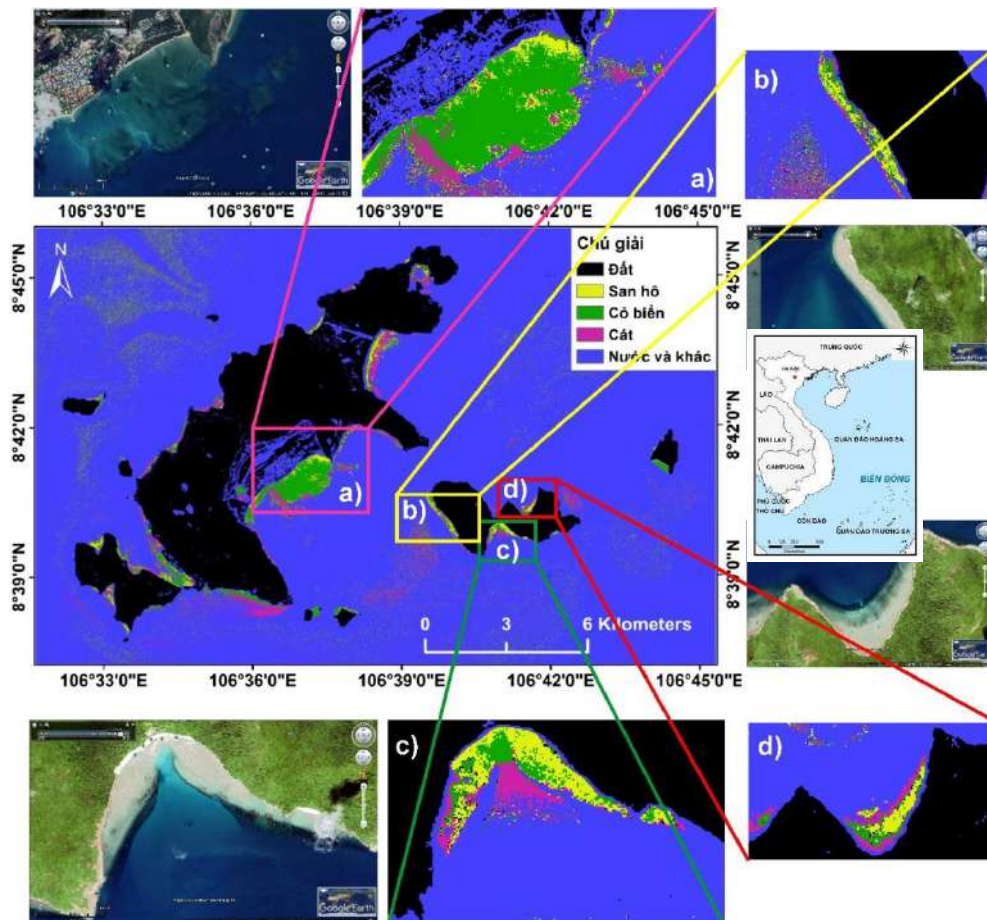
p_{i+} và p_{+j} là các phần tử đối xứng qua đường chéo chính của ma trận sai số,

N là tổng số hàng và cột của ma trận sai số.

Ma trận sai số được trình bày trong Bảng 3 cho thấy lớp rạn san hô có độ chính xác sản xuất 79,80% và độ chính xác sử dụng 81,44%. Độ chính xác tổng thể đạt 79,66%, hệ số Kappa đạt được là 69,42% tương ứng. Độ chính xác của lớp cỏ biển thấp hơn độ chính xác lớp san hô do diện tích bao phủ của lớp này lớn hơn lớp san hô. Kết quả độ chính xác đạt được thể hiện trong Bảng 3 do có sự phân loại nhầm lẫn 12 mẫu của lớp san hô sang lớp cỏ biển và phân loại bỏ sót 13 mẫu của lớp cát thành lớp san hô. Mặc dù đã lựa chọn các mẫu cho cả ba lớp có tính khác biệt về đặc trưng, nhưng sự nhầm lẫn và bỏ sót trong quá trình phân loại là không thể tránh khỏi do sự phân biệt phổ giữa các sinh vật đáy không rõ rệt.

Bảng 3. Ma trận sai số kết quả phân loại các sinh vật đáy từ đa chỉ số phổ của ảnh Sentinel-2

Mẫu	San hô	Cỏ biển	Cát	Tổng hàng	Độ chính xác sử dụng
San hô	79	12	6	97	81,44%
Cỏ biển	7	68	11	86	79,07%
Cát	13	10	84	107	78,50%
Tổng cột	99	90	101	290	
Độ chính xác sản xuất	79,80%	75,56%	83,17%		
Độ chính xác toàn bộ	79,66%				
Kappa	69,42%				



Hình 6. Kết quả phân loại các sinh vật đáy và dữ liệu có độ phân giải siêu cao

Các kết quả phân loại các sinh vật đáy và dữ liệu vệ tinh có độ phân giải siêu cao được minh họa trong Hình 6. Đối với khu vực bờ biển có dạng thẳng kéo dài và có địa hình đáy thoải cho thấy sự phát triển của cỏ biển tập trung phổ biến ở khu vực nước nông nhất và có bề mặt đáy tương đối bằng phẳng và lớp cát biển thường bao quanh khu vực cỏ biển như trong Hình 6a. Tuy nhiên, đối với địa hình dốc hơn, sự xuất hiện nhiều rạn san hô nhiều hơn xen lẫn với cát biển thay vì cỏ biển kéo dài dọc theo vùng nước nông trong Hình 6b. Đặc biệt ở phía Tây Nam của Hình 6b có xuất hiện khu vực cát xen lẫn san hô ở khu vực xa bờ hơn cho thấy khả năng xuyên thấu qua nước biển của một số kênh phổ vệ tinh Sentinel-2.

Đối với khu vực bờ biển có dạng lõm vào sâu, khu vực có rạn san hô dày đặc xen lẫn với cỏ biển xuất hiện ở vùng nước nông có độ dốc lớn hơn và khu vực cát biển ở vùng đáy bằng phẳng và độ sâu lớn hơn được nhận biết trong Hình 6c. Rạn san hô có xu hướng ở gần bờ và xen lẫn với cỏ biển tạo thành hệ sinh thái chủ yếu của sinh vật đáy ở vùng biển có đáy dốc, cát biển thường phân bố ở các khu vực đáy biển có địa hình tương đối bằng phẳng ở gần bờ nhất hoặc ở khu vực nước sâu nhất như trong Hình 6d.

4. Kết luận

Quần đảo Côn Đảo đặc trưng bởi khí hậu á xích đạo, nơi giao nhau của cả hải lưu ấm và hải lưu lạnh tỏa nên một hệ sinh thái sinh vật biển đa dạng, đặc biệt là các sinh vật đáy biển như san hô và cỏ biển. Mục tiêu chính của công việc nghiên cứu này là sử dụng các phương pháp xử lý dữ liệu đa phổ Sentinel-2 bao gồm khử phản xạ phổ của các tia sáng mặt trời trên mặt biển, tính toán các chỉ số độ sâu không đổi để hiệu chỉnh cột nước và phân loại các sinh vật biển. Độ chính xác tổng thể của kết quả phân loại các sinh vật đáy đạt 79,66%, hệ số Kappa đạt được là 69,42% tương ứng. Trong đó, lớp rạn san hô có độ chính xác sản xuất 79,80% và độ chính xác sử dụng 81,44%.

Những kết quả của bài báo cung cấp sự áp dụng dữ liệu đa phổ Sentinel-2 rất hữu ích để xác định các sinh vật đáy thông qua khả năng xuyên thấu nước biển của một số kênh phổ. Vì vậy, việc lập kế hoạch có thể nhanh chóng để giám sát rạn san hô và ra quyết định khi xảy ra các hiện tượng thay đổi tiêu cực đối với rạn san hô do sự ấm lên của nước biển trong bối cảnh biến đổi khí hậu. Các nghiên cứu với dữ liệu ảnh vệ tinh có độ phân giải không gian, độ phân giải phổ và độ phân giải bức xạ cao hơn sẽ góp phần tăng cường sự hiểu biết về sinh vật đáy với quy mô lớn hơn.

Tài liệu tham khảo

- Andréfouët, S., Kramer, P., Torres-Pulliza, D., Joyce, K.E., Hochberg, E.J., Garza-Pérez, R., et al., 2003. Multi-site evaluation of IKONOS for classification of tropical coral reef environments. *Remote Sens. Environ.* 88, 128-143.
- Bour, W., 1988. SPOT images for a coral reef mapping in New Caledonia. A fruitful approach for classic and new topics. In: *Proceedings of the 6th International Coral Reef Symposium, Australia*, vol. 2, pp. 445-448.
- Cao Văn Lương, Đàm Đức Tiến, Nguyễn Đức Thế, Nguyễn Văn Quân. 2013. Thành phần loài và phân bố của cỏ biển tại đầm Nai - Ninh Thuận, *Tuyển tập Hội nghị Khoa học toàn quốc về sinh học biển và phát triển bền vững lần thứ II*, 131-137.
- Chen, C.F.; Lau, V.K.; Chang, N.B.; Nguyen, T.S.; Tong, P. H. T.; Chiang, S.H. 2016. Multi-temporal change detection of seagrass beds using integrated Landsat TM/ETM+/OLI imageries in Cam Ranh bay, Vietnam, *Ecological Informatics*, 35, 43-54, Doi: 10.1016/j.ecoinf.2016.07.005.
- Devred, E., Turpie, K.R., Moses, W., Klemas, V.V., Moisan, T., Babin, M., Toro-Farmer, G., Forget, M.-H., Young-Heon, J., 2013. Future retrievals of water column bio-optical properties using the hyperspectral infrared imager (HyspIRI). *Remote Sens.* 5, 6812-6837. <https://doi.org/10.3390/rs5126812>.
- El-Askary, H., Abd El-Mawla, S.H., Li, J., El-Hattab, M.M., El-Raey, M., 2014. Change detection of coral reef habitat using Landsat-5 TM, Landsat 7 ETM+ and Landsat-8 OLI data in the Red Sea (Hurghada, Egypt). *Int. J. Remote Sens.* 35, 2327-2346.
- Foody, G.M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), 185-20.
- Garcia, R.A., McKinna, L.I.W., Hedley, J.D., Fearn, P.R.C.S., 2014. Improving the optimization solution for a semi-analytical shallow water inversion model in the presence of spectrally correlated noise. *Limnol. Oceanogr. Methods* 12, 651-669. <https://doi.org/10.1016/j.rse.2014.03.010>.
- Giardino, C., Bresciani, M., Fava, F., Matta, E., Brando, V.E., Colombo, R., 2016. Mapping submerged habitats and mangroves of Lampi Island Marine National Park (Myanmar) from in situ and satellite observations. *Remote Sens.* 8, 2. <https://doi.org/10.3390/rs8010002>.
- Green, E., Mumby, P., Edwards, A., and Clark, C. (2000). *Remote Sensing: Handbook for Tropical Coastal Management* (United Nations Educational, Scientific and Cultural Organization (UNESCO)).
- Guanter, L., Kauffmann, H., Segl, K., Foerster, S., Rogass, C., et al., 2015. The EnMAP spaceborne imaging spectroscopy mission for Earth observation. *Remote Sens.* 7, 8830-8857. <https://doi.org/10.3390/rs70708830>

- Hedley, J.D., Harborne, A.R., and Mumby, P.J., 2005. Technical note: Simple and robust removal of sun glint for mapping shallow - water benthos. *Int. J. Remote Sens.* 26, 2107-2112.
- Hedley, J.D., Roelfsema, C., Koetz, B., Phinn, S., 2012. Capability of the Sentinel-2 mission for tropical coral reef mapping and coral bleaching detection. *Remote Sens. Environ.* 120, 145-155. <https://doi.org/10.1016/j.rse.2011.06.028>.
- Hoàng Công Tín, Nguyễn Hữu Chí Tư, Nguyễn Tú Uyên, 2020. Hiện trạng và biến động thảm cỏ biển ở đầm Lăng Cô tỉnh Thừa Thiên Huế giai đoạn 2001-2020. *Tạp chí Khoa học Đại học Huế: Nông nghiệp và Phát triển Nông thôn*, 129(3D), 81-97.
- Hochberg, E.J., Atkinson, M.J., 2003. Capabilities of remote sensors to classify coral, algae, and sand as pure and mixed spectra. *Remote Sens. Environ.* 85, 174-189.
- Hock, K., Wolff, N.H., Condie, S.A., Anthony, K.R.N., Mumby, P.J., Paynter, Q., 2014. Connectivity networks reveal the risks of crown-of-thorns starfish outbreaks on the Great Barrier Reef. *J. Appl. Ecol.* 51, 1188-1196.
- Hoegh-Guldberg, O., Mumby, P.J., Hooten, A.J., Steneck, R.S., Greenfield, P., Gomez, E., et al., 2007. Coral reefs under rapid climate change and ocean acidification. *Science* 318, 1737-1742.
- Kutser, T., Miller, I.R., Jupp, D.L.B., 2006. Mapping coral reef benthic substrates using hyperspectral space-borne images and spectral libraries. *Estuar. Coast. Shelf Sci.* 70, 449-460.
- Lyzenga, D.R. (1978). Passive remote sensing techniques for mapping water depth and bottom features. *Appl. Opt.* 17, 379-383.
- Mumby, P.J., Green, E.P., Edwards, A.J., Clark, C.D., 1997. Coral reef habitat-mapping: how much detail can remote sensing provide? *Mar. Biol.* 130, 193-202. <https://doi.org/10.1007/s002270050238>.
- Palandro, D.A., Andréfouët, S., Hu, C., Hallock, P., Muller-Karger, F., Dustan, P., Callahan, M.K., Kranenburg, C., Beaver, C.R., 2008. Quantification of two decades of shallow-water coral reef habitat decline in the Florida Keys National Marine Sanctuary using Landsat data (1984-2002). *Remote Sens. Environ.* 112, 3388-3399. <https://doi.org/10.1016/j.rse.2008.02.015>.
- Pandolfi, J.M., Bradbury, R.H., Sala, E., Hughes, T.P., Bjorndal, K.A., Cooke, R.G., McArdle, D., McClenachan, L., Newman, M.J.H., Paredes, G., et al., 2003. Global trajectories of the long-term decline of coral reef ecosystems. *Science* 301, 955-958. <https://doi.org/10.1126/science.1085706>.
- Phinn, S.R., Roelfsema, C.M., Stumpf, R., 2010. Remote sensing: discerning the promise from the reality. In: Longstaff, B.J., Carruthers, T.J.B., Dennison, W.C., Lookingbill, T.R., Hawkey, J.M., Thomas, J.E., Wicks, E.C., Woerner, J. (Eds.), *Integrating And Applying Science: A Handbook For Effective Coastal Ecosystem Assessment*. IAN Press, Cambridge, Maryland.
- Roy, D.P., Li, J., Zhang, H.K., Yan, L., Huang, H., Li, Z., 2017. Examination of sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance. *Remote Sens. Environ.* 199, 25-38.
- Smith, V.E., Rogers, R.H., Reed, L.E., 1975. Automated mapping and inventory of Great Barrier Reef zonation with Landsat data. In: *Ocean 75 Conference Record*. Institute of Electrical and Electronics Engineers, Inc, New York.
- SUHET, 2015. Sentinel-2 User Handbook, Issue 1 Revision 2. *European Space Agency*. <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi>.
- Trần Đức Thanh (Chủ biên), Lê Đức An, Nguyễn Hữu Cừ, Trần Đình Lân, Nguyễn Văn Quân, Tạ Hoà Phương (2012). *Biển đảo Việt Nam - Tài nguyên vị thế và những kì quan địa chất, sinh thái tiêu biểu*, Hà Nội. Nhà xuất bản Khoa học tự nhiên và Công nghệ, ISBN 978-604-913-063-2.
- Yamano, H., Tamura, M., 2004. Detection limits of coral reef bleaching by satellite remote sensing: simulation and data analysis. *Remote Sens. Environ.* 90, 86-103.

ABSTRACT

Using Sentinel-2 data to create a benthic map Con Dao archipelago area, Ba Ria - Vung Tau province, Vietnam

Nguyen Van Trung^{1,2}, Le Thi Thu Ha^{1,2}, Doan Thi Nam Phuong^{1,2}, Pham Thi Huong¹
Nguyen Quang Minh³, Pham Vong Thanh⁴

¹*Hanoi University of Mining and Geology*

²*Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology*

³*Department of Cartography, The General Staff*

⁴*Vietnam Association of Georaphy, Cartography and Remote Sensing*

Sentinel-2A and 2B multispectral satellite (MSI) data provide specifications of potential value towards targeting benthic mapping. Monitoring coral reefs and seagrasses is a special challenge for remote sensing technology due to spectral reflectance inhomogeneities occurring at different depths and material compositions in the water. However, new solutions have made it possible for remote sensing data to increase the ability to provide large-scale environmental monitoring information. To carry out this work, a number of multispectral satellite data were used including: Landsat, Sentinel-2 and high spatial resolution data such as WorldView-2 to monitor coral reefs for detecting coral bleaching events. Sentinel-2 multispectral satellite data has been shown to be effective in mapping benthic organisms with the ability to eliminate spectral reflections of light rays on the water surface. A spatial resolution of 10 m is an important threshold to delineate benthic features of interest such as coral reefs with validation data from imagery and fieldwork in the Con Dao archipelago. Additionally, the data provided by the 290 km wide region offers advantages over ultra-high spatial resolution imagery.

Keywords: Sentinel-2, MSI, Coral reef, Con Dao archipelago.

Nghiên cứu giải pháp tự động chiết xuất đối tượng đường giao thông trên ảnh vệ tinh Sentinel-2 MSI

Lê Văn Phú¹, Trịnh Lê Hùng^{1,2*}, Tô Thị Phụng³

¹Học viện Kỹ thuật Quân sự

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

³Trường Đại học Thành Đông, Hải Dương

TÓM TẮT

Phân loại đường giao thông từ dữ liệu viễn thám thường gặp nhiều khó khăn do sự phức tạp của đối tượng cũng như hạn chế về độ phân giải không gian của dữ liệu ảnh. Bài báo này trình bày giải pháp tự động chiết xuất đối tượng đường giao thông từ ảnh vệ tinh quang học Sentinel-2 MSI khu vực thành phố Hà Nội trên cơ sở chỉ số đường (RI - Road Index). Toán tử Bottom-Hat được sử dụng để làm nổi bật đặc điểm của đường giao thông và loại bỏ ảnh hưởng của các đối tượng khác trên ảnh. Sau đó, phương pháp phân loại không giám sát và thuật toán phân đoạn Contrast-Split được áp dụng để trích xuất mạng lưới đường giao thông. Cuối cùng, phương pháp lọc nhiễu sử dụng hàm Gaussian được tích hợp để loại bỏ nhiễu trong ảnh phân loại. Kết quả nhận được trong nghiên cứu có thể sử dụng phục vụ công tác quy hoạch giao thông, giám sát sự thay đổi hệ thống đường giao thông đô thị.

Từ khóa: Viễn thám, chỉ số đường (road index), Sentinel-2 MSI, Hà Nội.

1. Đặt vấn đề

Việc phát triển hệ thống đường sá có ý nghĩa lớn trong hoạt động giao thông, kết nối kinh tế, phát triển đô thị và bảo vệ môi trường, góp phần quan trọng trong phát triển kinh tế - xã hội (Han và cộng sự, 2009). Thông tin về mạng lưới đường giao thông có ý nghĩa quan trọng phục vụ xây dựng các bản đồ địa hình, bản đồ chuyên đề (Shi và cộng sự, 2014; Jiao và cộng sự, 2021). Dữ liệu viễn thám quang học với độ phân giải không gian cao cho phép chiết xuất nhanh thông tin về mạng lưới đường giao thông, giúp tiết kiệm thời gian và công sức so với các phương pháp đo đạc truyền thống (Kennedy và cộng sự, 2009).

Hiện nay, việc phân loại đối tượng đường giao thông từ ảnh viễn thám gặp rất nhiều khó khăn. Trước hết, về đặc điểm hình học, đường giao thông có dạng hình tuyến, dễ nhầm lẫn với các đối tượng thủy văn. Bên cạnh đó, phản xạ phổ của các đối tượng trên đường giao thông có thể không đồng nhất do nhiều yếu tố khác nhau như vật liệu bề mặt, trạng thái bề mặt và ánh sáng. Các phương pháp phân loại truyền thống để trích xuất đường giao thông từ ảnh viễn thám quang học thường có độ chính xác không cao do những hạn chế trong quá trình lấy mẫu phân loại (Liu và cộng sự, 2015). Một số nghiên cứu đã sử dụng phương pháp phân loại dựa trên đối tượng, bao gồm cả phân loại có giám sát (Miao và cộng sự, 2015) và không giám sát (Grinias và cộng sự, 2016) để trích xuất thông tin về hệ thống đường giao thông từ ảnh vệ tinh độ phân giải cao. Bên cạnh đó, một số nghiên cứu khác đã sử dụng phương pháp dựa trên thông tin về kết cấu (Sghaier và Lepage, 2016), hình học (He và cộng sự, 2012) trong phân loại đối tượng đường giao thông (Cheng và cộng sự, 2012).

Nhìn chung, các phương pháp và kỹ thuật trích xuất đối tượng đường giao thông từ ảnh viễn thám có thể được phân loại thành hai loại: (1) phương pháp phân loại tự động và bán tự động và (2) phương pháp trích xuất khu vực đường và tìm đường (Alshehhi và Marpu, 2017). Gần đây, các thuật toán trí tuệ nhân tạo đã được phát triển và ứng dụng trong việc trích xuất đặc trưng và phân loại đối tượng đường giao thông từ hình ảnh viễn thám. Các thuật toán như SVM (Simler, 2011), ANN (George và cộng sự, 2013; Li và Chen, 2014) đã được chứng minh có hiệu quả cao trong phân loại đối tượng đường giao thông từ ảnh viễn thám độ phân giải cao.

Nghiên cứu này trình bày kết quả trích xuất thông tin đối tượng đường giao thông từ dữ liệu viễn thám Sentinel-2 MSI bằng phương pháp phân loại không giám sát trên cơ sở chỉ số đường RI - Road Index. Các kênh hồng ngoại sóng ngắn (kênh 11), cận hồng ngoại (kênh 8) và xanh lam (kênh 2) ảnh Sentinel-2 MSI được sử dụng để tính chỉ số RI, sau đó sử dụng toán tử Bottom-Hat để làm nổi bật đặc điểm của đường giao thông và loại bỏ ảnh hưởng của các đối tượng khác trên ảnh. Kết quả nhận được tiếp được phân mảnh (segmentation) và phân loại không giám sát, sau đó vector hóa để tạo thông tin về đối tượng đường giao thông.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Dữ liệu và khu vực nghiên cứu

* Tác giả liên hệ

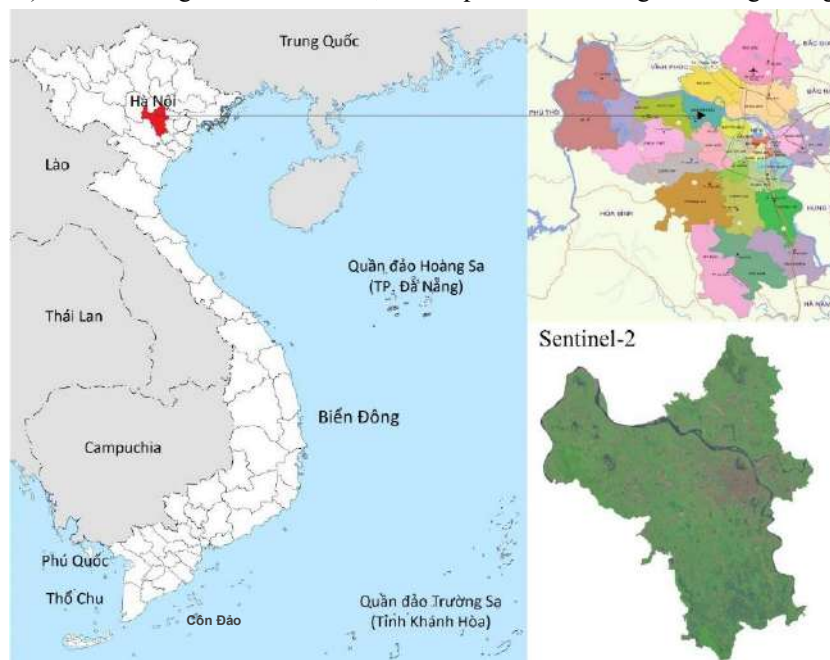
Email: tringlehung125@gmail.com

a) Khu vực nghiên cứu

Khu vực nghiên cứu được lựa chọn là thành phố Hà Nội, một trong hai thành phố lớn nhất của Việt Nam - Hình 1. Hà Nội là một trung tâm kinh tế, chính trị, văn hóa, xã hội và là thủ đô của đất nước với dân số đông, mật độ xây dựng cao. Các tuyến đường chính trong thành phố thường tập trung ở trung tâm, trong khi các khu vực ngoại ô và vùng lân cận có mật độ đường thấp hơn. Bên cạnh đó, Hà Nội có một hệ thống đường sá phức tạp, bao gồm các đường phố chính, đường phố phụ và các con phố nhỏ.

b) Dữ liệu sử dụng

Dữ liệu viễn thám sử dụng trong nghiên cứu là ảnh vệ tinh Sentinel-2 MSI chụp ngày 18/5/2023. Đặc điểm chung của ảnh vệ tinh Sentinel-2 MSI, bao gồm: 02 vệ tinh (Sentinel-2A và Sentinel 2B) được thể hiện trong Bảng 1. Ảnh Sentinel-2 MSI được thu thập ở mức xử lý L2A, trong đó đã được hiệu chỉnh bức xạ và đưa về giá trị phản xạ phổ. Ảnh thu thập trên nền tảng GEE được hiệu chỉnh hình học nhằm loại bỏ những sai số về hình dạng và vị trí của ảnh một cách tự động. Sau khi tiền xử lý, dữ liệu ảnh Sentinel-2 MSI được tổ hợp màu và cắt theo khu vực nghiên cứu. Các kênh phổ ở dải sóng xanh lam (kênh 2), cận hồng ngoại (kênh 8) và hồng ngoại sóng ngắn (kênh 11) được sử dụng để tính chỉ số RI, sau đó phân loại tự động đối tượng đường.



Hình 1. Khu vực nghiên cứu

Bảng 1. Đặc điểm dữ liệu ảnh vệ tinh Sentinel-2A MSI sử dụng trong nghiên cứu

Kênh ảnh	Tên kênh	Bước sóng trung tâm Sentinel-2A (nm)	Độ phân giải (m)
Kênh 1	Coastal Aerosol	442,7	60
Kênh 2	Blue	492,4	10
Kênh 3	Green	559,8	10
Kênh 4	Red	664,6	10
Kênh 5	Red Edge 1	704,1	20
Kênh 6	Red Edge 2	740,5	20
Kênh 7	Red Edge 3	782,8	20
Kênh 8	Near Infrared (NIR)	832,8	10
Kênh 8A	Narrow NIR	864,7	20
Kênh 9	Water Vapour	945,1	60
Kênh 10	SWIR Cirrus	1373,5	60
Kênh 11	SWIR ₁	1613,7	20
Kênh 12	SWIR ₂	2202,4	20

2.2. Phương pháp nghiên cứu

Quy trình xử lý ảnh viễn thám Sentinel-2 MSI nhằm tự động trích xuất thông tin đối tượng đường được trình bày cụ thể trong Hình 2. Trước hết, trong nghiên cứu tiến hành thu thập dữ liệu ảnh Sentinel-2 MSI trên nền tảng

điện toán đám mây Google Earth Engine (GEE). Hình ảnh tiền xử lý, sau đó được cắt theo khu vực nghiên cứu.

Tiếp theo, chỉ số đường Road Index (RI) được tính dựa trên các kênh ảnh hồng ngoại sóng ngắn (kênh 11), cận hồng ngoại (kênh 8) và xanh lam (kênh 2). RI được giới thiệu và phát triển cho dữ liệu ảnh vệ tinh Landsat-8 (Reddy và Garg, 2020). Chỉ số này mang lại kết quả tốt hơn so với các phương pháp trích xuất đường thông thường đối với hình ảnh đa phổ. Dữ liệu đầu ra là một hình ảnh được chuẩn hóa có quang phổ với các đối tượng đường được nổi bật so với với các đối tượng khác. Chỉ số đường (RI) được tính toán từ phương trình (1):

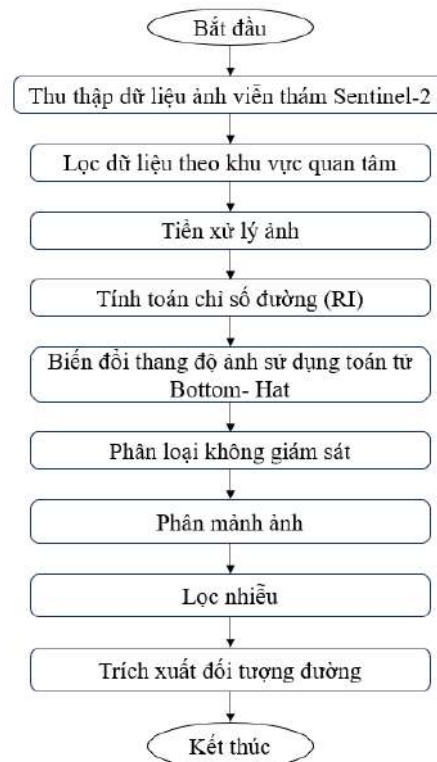
$$RI = 1 - \frac{3 * \min (B11_{10m}, B8, B2)}{B11_{10m} + B8 + B2} \quad (1)$$

trong đó: RI là chỉ số đường, $B11_{10m}$, B8, B2 lần lượt là kênh hồng ngoại sóng ngắn được chuẩn hoá về độ phân giải 10 m, kênh cận hồng ngoại và kênh xanh lam.

Sau đó, hình ảnh RI được xử lý thông qua bộ lọc Bottom-Hat. Toán tử Bottom-Hat thực hiện chuyển đổi hình thái trên ảnh đầu vào thang độ xám hoặc nhị phân. Nó tính toán độ đóng hình thái của hình ảnh đầu vào và sau đó trừ kết quả cho hình ảnh gốc bằng cách sử dụng phần tử cấu trúc (Thapar và cộng sự, 2012). Phép biến đổi Bottom-Hat có thể được xác định bằng phương trình sau:

$$T_b(f) = (f * b) - f \quad (2)$$

trong đó: * là thao tác đóng, b là phần tử cấu trúc ma trận vuông, $T_b(f)$ biểu thị toán tử Bottom-Hat, f là hình ảnh được sử dụng. Với việc áp dụng phép biến đổi Bottom-Hat cho hình ảnh RI, các đối tượng đường được làm nổi bật trong khi các đối tượng khác bị hạn chế.



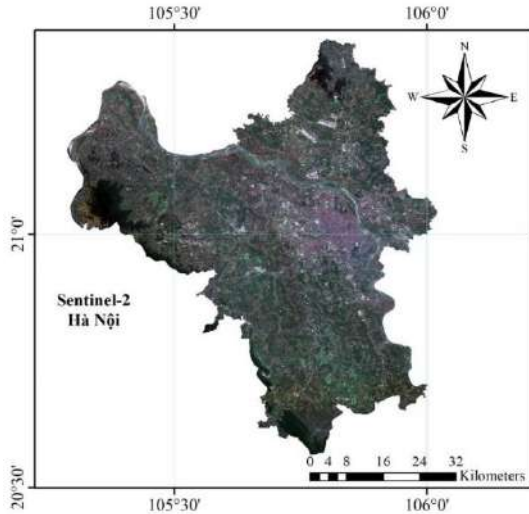
Hình 2. Quy trình xử lý dữ liệu Sentinel-2

Sau khi áp dụng bộ lọc Bottom-Hat, hình ảnh tiếp tục được phân loại không giám sát sử dụng thuật toán ee.Clusterer.wekaKMeans trong GEE. WekaKMeans là một thuật toán gom cụm dựa trên thuật toán K-means được cung cấp bởi Weka. Phân loại không giám sát hình ảnh chỉ số đường cho phép đánh giá ban đầu về khu vực nghiên cứu và xác định các mẫu quang phổ phân lớp tồn tại trong mỗi lớp. Hình ảnh phân cụm thu được từ wekaKMeans sau đó được xử lý thông qua thuật toán phân đoạn Contrast-Split (CSS). Thuật toán CSS tạo các đối tượng dựa trên những thay đổi về mức độ tương phản. Sự thay đổi độ tương phản (tỉ lệ chiều dài và chiều rộng) được xử lý để thu được các đặc điểm của mạng lưới đường.

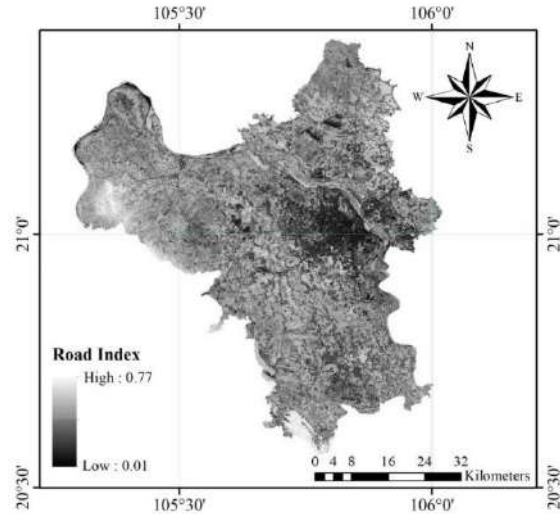
Cuối cùng, phương pháp lọc nhiễu sử dụng hàm Gaussian được sử dụng để loại bỏ nhiễu trong ảnh phân loại. Khi áp dụng bộ lọc Gauss lên ảnh, mỗi điểm ảnh mới sẽ được tính dựa trên trung bình có trọng số của các điểm ảnh lân cận. Quá trình này làm loại bỏ các đối tượng nhỏ, nhiễu khi phân loại. Hình ảnh cuối cùng chứa các đối tượng đường đã được trích xuất dựa trên hình ảnh vệ tinh.

3. Kết quả và thảo luận

Khu vực nghiên cứu được lựa chọn là toàn bộ thành phố Hà Nội. Dữ liệu ảnh Sentinel-2 được thu thập và tiền xử lý trên nền tảng GEE, sau đó cắt theo ranh giới khu vực nghiên cứu. Dữ liệu hình ảnh vệ tinh Sentinel-2 khu vực thành phố Hà Nội được trình bày trên Hình 3.

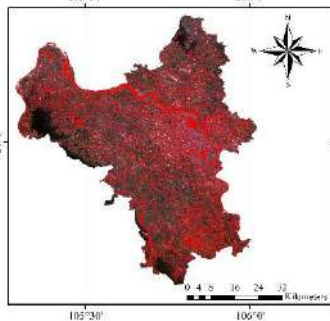


Hình 3. Ảnh Sentinel-2 khu vực thành phố Hà Nội ngày 18/5/2023

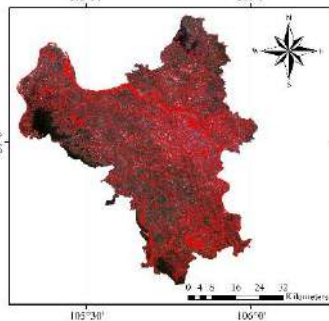


Hình 4. Chỉ số RI xác định từ ảnh Sentinel ngày 18/5/2023

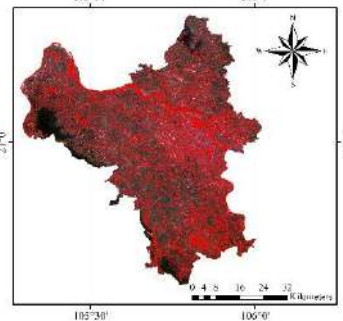
Hình 4 trình bày hình ảnh chỉ số đường được tính từ ảnh Sentinel-2 đã thu thập từ bước trước. Có thể thấy rằng, ảnh RI có giá trị nhỏ nhất là 0,01 và giá trị lớn nhất là 0,77. Bên cạnh đó, các đối tượng đường có giá trị thấp, tương ứng với sắc màu tối hơn so với các đối tượng khác.



(a) 3x3



(b) 4x4

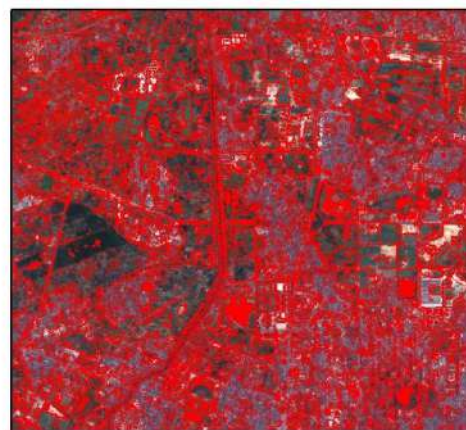


(c) 5x5

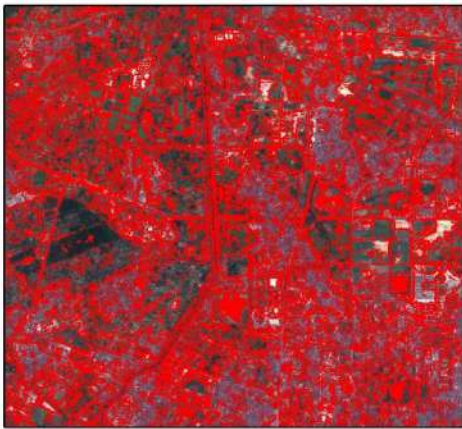
Hình 5. Các đối tượng đường được trích xuất với các phân tử cấu trúc khác nhau



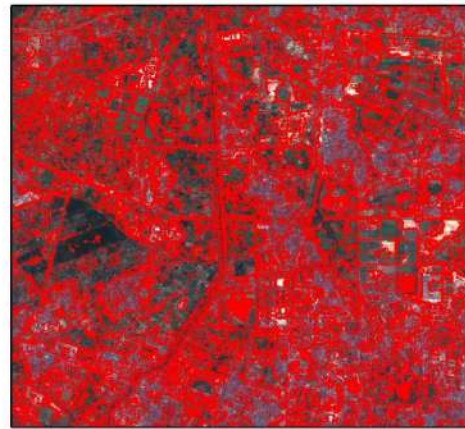
(a) Ảnh gốc



(b) 3x3



(c) 4x4



(d) 5x5

Hình 6. Các đối tượng đường được trích xuất khu vực quận Bắc Từ Liêm với các phân tử cấu trúc khác nhau

Hình 5 và Hình 6 lần lượt trình bày kết quả trích xuất các đối tượng đường cho toàn bộ thành phố Hà Nội và quận Bắc Từ Liêm. Có thể thấy phân tử cấu trúc 5x5 trong bước hiệu chỉnh hình thái cho kết quả tốt hơn so với các phân tử 3x3 và 4x4. Hai phân tử này đưa ra kết quả chứa các đối tượng nhiễu khác như mái nhà xây dựng, bồn cây.



Hình 7. Các đối tượng đường được trích xuất khu vực đường Quốc lộ 21B, Hà Đông trước và sau khi lọc Gauss



Hình 7 thể hiện các đối tượng đường dọc Quốc lộ 21B (điểm đầu bắc tại Hà Đông) sau khi lọc qua thuật toán CSS và lọc nhiễu Gauss. Có thể thấy rằng, kết quả trích xuất được toàn bộ đường Quốc lộ 21B và một số con đường dân sinh lớn, nối trực tiếp với các nút giao với đường quốc lộ. Các đối tượng đường nhỏ, đường đi qua

khu dân cư không được trích xuất do ảnh hưởng bởi các đối tượng khác xung quanh như bóng cây, nhà và bóng các tòa nhà lớn.

Nhìn chung, dữ liệu sau khi tính toán với thuật toán CSS và lọc nhiễu Gauss, các đối tượng nhỏ và nhiễu đã được loại bỏ mà không ảnh hưởng tới quá trình trích xuất các đối tượng đường trong ảnh vệ tinh.

4. Kết luận

Nghiên cứu đề xuất phương pháp để trích xuất các đối tượng đường trên ảnh chỉ số đường (RI) chuẩn hóa. Phương pháp được áp dụng trên dữ liệu Sentinel-2 MSI. Sau khi tính toán chỉ số đường, hình ảnh được hiệu chỉnh hình thái thông qua bộ lọc Bottom-Hat, phân loại không giám sát bằng thuật toán WekaKMeans, lọc đối tượng thông qua thuật toán phân đoạn Contrast-Split và lọc nhiễu qua bộ lọc Gauss. Kết quả thu được các đối tượng đường khu vực thành phố Hà Nội với sự áp dụng của các phần tử cấu trúc của bộ lọc Bottom-Hat khác nhau. Việc trích xuất các đối tượng đường từ dữ liệu hình ảnh vệ tinh có thể cung cấp thông tin quan trọng và hữu ích cho nhà hoạch định chính sách trong nhiều lĩnh vực như quản lý giao thông, kế hoạch đô thị, quản lý môi trường và quy hoạch hạ tầng.

Tài liệu tham khảo

- Alshehhi, R., & Marpu, P. R. (2017). Hierarchical graph-based segmentation for extracting road networks from high-resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 126, 245-260.
- Cheng, J., Ding, W., Ku, X., & Sun, J. (2012). Road extraction from high-resolution SAR images via automatic local detecting and human-guided global tracking. *International Journal of Antennas and Propagation*.
- George, J., Mary, L., & Riyas, K. S. (2013, December). Vehicle detection and classification from acoustic signal using ANN and KNN. In: *2013 International Conference on Control Communication and Computing (ICCC)*, pp. 436-439. IEEE.
- Grinias, I., Panagiotakis, C., & Tziritas, G. (2016). MRF-based segmentation and unsupervised classification for building and road detection in peri-urban areas of high-resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 145-166.
- Jiao C, Heitzler M, Hurni L. (2021). A survey of road feature extraction methods from raster maps. *Trans GIS*; 25: 2734-63.
- Han J, Hayashi Y, Cao X, Imura H. (2009). Application of an integrated system dynamics and cellular automata model for urban growth assessment: A case study of Shanghai, China. *Landsc Urban Plann*; 91:133-41.
- He, C., Liao, Z. X., Yang, F., Deng, X. P., & Liao, M. S. (2012). Road extraction from SAR imagery based on multiscale geometric analysis of detector responses. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(5), 1373-1382.
- Kennedy RE, Townsend PA, Gross JE, Cohen WB, Bolstad P, Wang YQ, Adams P. (2009). Remote sensing change detection tools for natural resource managers: understanding concepts and tradeoffs in the design of landscape monitoring projects. *Remote Sens Environ*; 113:1382-96.
- Li, J., & Chen, M. (2014, August). On-road multiple obstacles detection in dynamical background. In *2014 Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics*, Vol. 1, pp. 102-105. IEEE.
- Liu, B., Wu, H., Wang, Y., & Liu, W. (2015). Main road extraction from zy-3 grayscale imagery based on directional mathematical morphology and vgi prior knowledge in urban areas. *PloS one*, 10(9), e0138071.
- Miao, Z., Shi, W., Gamba, P., & Li, Z. (2015). An object-based method for road network extraction in VHR satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(10), 4853-4862.
- Reddy S, Rao C, Kumar P, Anjaneyulu RVG, Gopala Krishna B. (2020). An index-based road feature extraction from LANDSAT-8 OLI images. *Int J Electr Comput Eng*; 11: 1319-36.
- Sghaier, M. O., & Lepage, R. (2015). Road extraction from very high-resolution remote sensing optical images based on texture analysis and beamlet transform. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5), 1946-1958.
- Shi, W., Miao, Z., & Debayle, J. (2013). An integrated method for urban main-road centerline extraction from optical remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(6), 3359-3372.
- Simler, C. (2011, July). An improved road and building detector on VHR images. In: *2011 IEEE International Geoscience and Remote Sensing Symposium*, pp. 507-510. IEEE.
- Thapar S, Garg S. (2012). Study and implementation of various morphology-based image contrast enhancement techniques. *Int. J. Comput. Bus. Res*; 128: 2229-6166. <https://sentinels.copernicus.eu/web/>.

ABSTRACT

Development method for automatic extraction of road using Sentinel-2 MSI data

Le Van Phu¹, Trinh Le Hung^{1,2*}, To Thi Phuong³

¹*Le Quy Don Technical University, Hanoi, Vietnam*

²*Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology*

³*Thanh Dong University, Hai Duong*

Road extraction from remote sensing data often faces many difficulties due to the complexity of the object as well as the limited spatial resolution of the image data. This article presents a development method for automatic extraction of road from Sentinel-2 MSI data, a case study in Hanoi city area based on road index (RI). The Bottom-Hat operator is used to highlight road features and remove the influence of other objects on the image. Then, the unsupervised classification method and Contrast-Split segmentation algorithm are applied to extract the road. Finally, the noise filtering method using the Gaussian function is used to remove noise in the classified image. The obtained results can be used for traffic planning and monitor changes to the urban road system.

Keywords: Remote Sensing, road index, Sentinel-2 MSI, Hanoi.

Phát hiện cháy rừng đang xảy ra sử dụng dữ liệu vệ tinh Sentinel-3 SLSTR khu vực xã Khánh Bình Tây, huyện Trần Văn Thời, tỉnh Cà Mau

Đoàn Thị Nam Phương^{1,2*}, Nguyễn Văn Trung^{1,2}, Lê Thị Thu Hà^{1,2}, Nguyễn Quang Minh³

¹Trường Đại học Mở - Địa chất

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

³Cục Bản đồ - Bộ Tổng tham mưu

TÓM TẮT

Dữ liệu vệ tinh đa phổ Sentinel-3A và 3B SLSTR cung cấp khả năng giám sát nhiệt độ bề mặt Trái đất để phát hiện các khu vực xảy ra cháy rừng. Phát hiện cháy rừng dựa vào sự chênh lệch nhiệt độ bề mặt khu vực cháy và nhiệt độ môi trường trung bình khu vực xung quanh (300 K). Công việc này được thực hiện khi sử dụng hai kênh phổ Sentinel-3 SLSTR MIR (F1 có bước sóng 3,74 μm) và kênh phổ Sentinel-3 SLSTR TIR (F2 có bước sóng 10,85 μm). Hai kênh phổ này được sử dụng do sự khác biệt tăng đối với khu vực không xảy ra cháy rừng và khác biệt giảm đối với khu vực xảy ra cháy rừng. Ngoài ra, các kênh phổ khác của dữ liệu Sentinel-3 SLSTR giúp loại trừ các khu vực bị mây che và các dữ liệu bổ trợ về các lớp phủ đất trong đó có các lớp thực phủ giúp khoanh vùng các khu vực rừng cũng được kết hợp để đưa ra kết quả khu vực xảy ra cháy rừng cuối cùng. Dữ liệu vệ tinh đa phổ Sentinel-3 có tính hiệu quả trong việc phát hiện cháy rừng cả ban ngày và ban đêm do sử dụng các kênh phổ hồng ngoại và khả năng cung cấp dữ liệu hàng ngày đêm với chiều rộng dải quét (từ 740 km đến 1400 km) bao phủ một phạm vi rộng lớn. Trong nghiên cứu này các ngưỡng giá trị được đưa ra đối với kênh F1 (325 và 315 đối với ban ngày và ban đêm lần lượt) và chênh lệch giữa kênh F2 và F1 (18 và 15 đối với ban ngày và ban đêm lần lượt) cho các khu vực rừng đã được chiết tách từ các lớp phủ khác để xác định khu vực cháy.

Từ khóa: Sentinel-3, SLSTR, cháy rừng, Cà Mau.

1. Đặt vấn đề

Cháy rừng làm xáo trộn một khu vực lớn với nhiều loại quần xã sinh vật hơn bất kỳ tác nhân nào khác (Bowman và cộng sự, 2009; Johnston và cộng sự, 2011; Zhang và cộng sự, 2018). Cháy rừng cũng thải ra khói có tác động đáng kể đến Trái đất gồm thành phần khí quyển và hóa học, chất lượng không khí và cân bằng bức xạ (Natarajan và cộng sự, 2012). Quan sát Trái đất qua vệ tinh (EO) trở nên quan trọng đối với việc định lượng một cách linh hoạt các sự kiện cháy rừng ở quy mô từ nhỏ đến lớn, rời rạc (xem đánh giá gần đây của Wooster và cộng sự, 2021). Các yêu cầu về dữ liệu cháy rừng không chỉ mở rộng cho các nghiên cứu khoa học mà còn cho việc tạo ra và cung cấp thông tin hữu ích hỗ trợ giám sát và ra quyết định theo thời gian thực. Ngoài ra, các kết quả quan trắc được mô tả các bức tranh tự nhiên về cảnh quan và tác động của khói ảnh hưởng tới chất lượng không khí mà con người cần quan tâm.

Các sản phẩm hoạt động cháy (Active fire - AF) là một trong những bộ dữ liệu vệ tinh EO được sử dụng rộng rãi nhất, mô tả các vị trí cháy rừng vào thời điểm đó. Thông thường các sản phẩm như vậy cung cấp ước tính của đầu ra công suất bức xạ tần số cao (FRP) liên quan đến tần số, cường độ, mức tiêu thụ nhiên liệu và lượng khói thải (Freeborn và cộng sự, 2008; Kaiser và cộng sự, 2012; Nguyen và Wooster, 2020; Nguyen và cộng sự, 2023). Dữ liệu từ bộ cảm đo bức xạ nhiệt độ bề mặt đất và biển (SLSTR) trên các vệ tinh Sentinel-3 của châu Âu (Coppo và cộng sự, 2010) được sử dụng để tạo ra sản phẩm AF chứa các vị trí, thời gian và ước tính của FRP. SLSTR sử dụng tia hồng ngoại sóng dài (LWIR) trong bộ cảm đo bức xạ có một số điểm tương đồng với bộ cảm đo bức xạ hồng ngoại trung bình.

Bộ cảm MODIS hoạt động trên vệ tinh Terra và Aqua của NASA. Các sản phẩm AF được tạo từ dữ liệu SLSTR của vệ tinh Sentinel-3 hiện được sử dụng cùng với bộ cảm MODIS và cuối cùng sẽ thay thế bộ cảm MODIS trên vệ tinh Terra khi nhiệm vụ đó kết thúc vì chức năng giống nhau. Phát hiện cháy AF vào ban đêm của dữ liệu Sentinel-3 và sản phẩm FRP đã được tạo ra từ tháng 3 năm 2020, dựa trên dữ liệu với việc thu thập khoảng 22:00 giờ địa phương (xem Bảng thời gian ban đêm, thuật toán và đặc tính hiệu suất sản phẩm được trình bày chi tiết trong (Xu và cộng sự, 2020). Thuật toán được sử dụng để tạo ra Active Fire và FRP vào ban ngày bổ sung sản phẩm Sentinel-3 được thực hiện vào khoảng 10:00 giờ địa phương. Trong đó, những khác biệt về thuật toán giữa ban ngày so với ban đêm được trình bày chi tiết, điều này cho thấy hiệu năng tăng lên rất đáng kể khi sử dụng kênh phổ MIR "F1" cho sản phẩm ban ngày thu được. Điều này thể hiện khả năng phát hiện cháy

* Tác giả liên hệ:

Email: doanthinamphuong@humg.edu.vn

của dữ liệu AF từ bộ cảm Sentinel-3 SLSTR cao hơn khi so sánh với dữ liệu AF từ bộ cảm MODIS Terra, dữ liệu AF từ bộ cảm VIIRS và cao hơn nhiều dữ liệu AF từ bộ cảm OLI của vệ tinh Landsat-8. Phát hiện cháy AF Sentinel-3 và FRP là các sản phẩm được tạo ra bởi cả thuật toán ban đêm và ban ngày đều có sẵn thông qua Trung tâm dữ liệu Copernicus (<https://scihub.copernicus.eu/>).

Vệ tinh Sentinel-3A và Sentinel-3B được phóng vào năm 2016 và năm 2018 tương ứng. Mỗi chiếc đều mang cảm biến SLSTR và các nhiệm vụ này sẽ được thực hiện bởi các vệ tinh Sentinel-3C và Sentinel-3D trong thời gian tới năm. SLSTR được xây dựng dựa trên nguyên lý đo bức xạ quét dọc theo đường (Delderfeld và cộng sự 1986; Coppo và cộng sự, 2010) cung cấp chi tiết toàn diện về nhiều khía cạnh công cụ. Hai vệ tinh Sentinel-3 hoạt động đồng thời cùng nhau cung cấp vùng phủ sóng SLSTR toàn cầu một lần vào ban ngày và một lần vào ban đêm ở xích đạo và tần suất lặp lại lớn hơn ở vĩ độ cao hơn. Kênh phổ MIR được sử dụng phổ biến nhất đối với việc phát hiện AF (Wooster và cộng sự, 2021), nhưng MIR quan sát ban ngày do SLSTR thực hiện bị cản trở bởi sự bắt đầu bão hòa ở nhiệt độ nóng hơn khoảng 311 K (Xu và cộng sự, 2020; 2021). Bên cạnh đó, kênh SLSTR MIR có mức tăng thấp hơn (được gọi là “F1” hoạt động với bước sóng trung tâm 3,742 μm) bão hòa trên ~ 450 K và việc sử dụng nó có thể giảm thiểu đáng kể những vấn đề như vậy. Tuy nhiên, so với kênh phổ S7, F1 có một số đặc điểm khiến dữ liệu của nó trở nên phong phú hơn khi sử dụng cùng với các kênh hồng ngoại nhiệt SLSTR khác và các kênh hồng ngoại nhìn thấy được sử dụng trong quá trình phát hiện AF (Xu và cộng sự, 2020; 2021), dẫn đến việc phát hiện AF từ Sentinel-3 ban ngày thuật toán phức tạp hơn đáng kể so với thuật toán ban đêm. Dưới đây nhóm tác giả cung cấp chi tiết về các điều chỉnh thuật toán cần thiết để giải quyết sự phức tạp này.

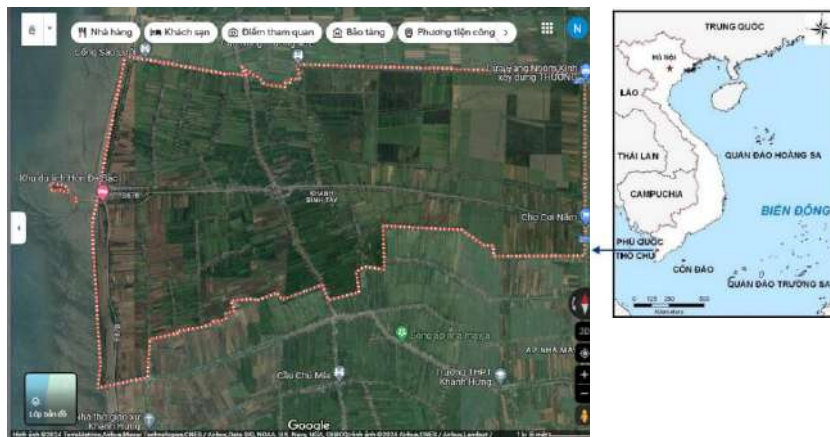
Các kênh phổ SLSTR S7 và F1 của Sentinel-3 xuất hiện khá giống nhau nhưng không thực sự giống nhau (Xu và cộng sự, 2021). Dữ liệu kênh phổ F1 có phạm vi động rộng hơn có ý nghĩa đáng kể về nhiễu, hình dạng và diện tích pixel mặt đất khác với dữ liệu từ kênh phổ S7 (Xu và cộng sự, 2021). Trong quá trình quan sát ban ngày ở một số vùng ẩm hơn, nhiệt lượng mặt trời và bức xạ phản xạ mặt trời cùng nhau có thể tạo ra độ bão hòa khá rộng trong kênh phổ S7. Kết quả là thuật toán phát hiện SLSTR AF yêu cầu mở rộng hơn bằng việc sử dụng dữ liệu kênh phổ F1 vào ban ngày so với ban đêm, tập trung vào đặc biệt trên các vùng bề mặt đất nơi bão hòa trong kênh phổ S7 xảy ra ở cả những pixel. Phương pháp được sử dụng để phát triển thuật toán sản phẩm Sentinel-3 AF ban ngày là giữ tất cả các giai đoạn của thuật toán ban đêm được trình bày chi tiết trong (Xu và cộng sự, 2020), nhưng để thay đổi ngưỡng và đầu vào ban ngày so với vào ban đêm.

2. Dữ liệu sử dụng và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Xã Khánh Bình Tây, huyện Trần Văn Thời, tỉnh Cà Mau có diện tích tự nhiên 5.456,59 ha (Hình 1), dân số 3.265 hộ, với 13.396 khẩu, có 03 dân tộc anh em sinh sống gồm Kinh, Hoa và Khmer, trong đó dân tộc Khmer 441 hộ, với 1.863 khẩu. Với vị trí địa lý có 4,8 km đường bờ biển. Khai thác, nuôi trồng thủy sản, trồng rừng tràm và du lịch phát triển khá mạnh tại địa phương. Xã có diện tích rừng tràm 150,04 ha. Thời gian qua, lực lượng kiểm lâm đóng trên địa bàn, thường xuyên tuyên truyền, vận động người dân sống trong khu vực rừng tràm luôn nâng cao ý thức trong việc giữ gìn và bảo vệ rừng (Anh Thư, 2015).

Thời gian phát hiện cháy khoảng 12 giờ, ngày 10/4/2024. Địa điểm cháy tại Đội Quản lý đất Quốc phòng, Cục Hậu cần, Quân khu 9 (Nông trường 402 cũ), thuộc địa giới hành chính ấp Cơi 6B, xã Khánh Bình Tây, huyện Trần Văn Thời. Diện tích bị thiệt hại khoảng 40 ha (chưa tiến hành đo đạc thực địa), không thuộc diện tích quy hoạch đất lâm nghiệp trên địa bàn tỉnh. Hiện trạng bị cháy: Cây tràm tái sinh, cây tràm trồng năm 2019 và thực bì gồm cây sậy, dón, choai... (Trang thông tin điện tử Đảng bộ huyện Trần Văn Thời).



Hình 1. Ảnh Google Map của xã Khánh Bình Tây, huyện Trần Văn Thời, tỉnh Cà Mau

2.2. Dữ liệu sử dụng

Dữ liệu ảnh vệ tinh Sentinel-3 được chọn để làm khu vực nghiên cứu do dữ liệu ảnh vệ tinh này miễn phí có sẵn tại <https://scihub.copernicus.eu>. Các kênh phổ của dữ liệu Sentinel-3 SLSTR được đưa ra trong Bảng 1. Dữ liệu được cung cấp mỗi ngày một cảnh ảnh ban ngày và một cảnh ảnh ban đêm ở xích đạo và tần suất nhiều hơn ở vĩ độ lớn hơn.

Dữ liệu ảnh vệ tinh Sentinel-3 khu vực nghiên cứu được thu nhận và xử lý ở mức độ 1 vào ban đêm và ban ngày tương ứng với tên các sản phẩm: S3A_SL_1_RBT_20240410T150621_20240410T150921_20240412T010755_0179_111_111_0000_PS1_O_NT_004.SEN3 và S3B_SL_1_RBT_20240411T025916_20240411T030216_20240411T235027_0179_091_360_2700_PS2_O_NT_004.SEN3.

Bảng 1. Các kênh phổ của dữ liệu đa phổ Sentinel-3 SLSTR
(Thông tin từ: <https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/slstr/>)

Kênh	Bước sóng trung bình (nm)	Độ rộng (nm)	Chức năng	Chú thích		Độ phân giải (m)
S1	554,27	19,26		Nhìn thấy và cận hồng ngoại	Các kênh phổ phản xạ năng lượng mặt trời	500
S2	659,47	19,25	Chỉ số NDVI, giám sát thực vật, sol khí			
S3	868,00	20,60	Chỉ số NDVI, xác định mây, hiệu chỉnh pixel			
S4	1374,80	20,80	Xác định mây che phủ mặt đất	Hồng ngoại ngắn		
S5	1613,40	60,68	Giám sát băng, tuyết và thực vật			
S6	2255,70	50,15	Trạng thái thực vật và loại bỏ mây			
S7	3742,00	398,00	Nhiệt độ bề mặt đất và biển, hoạt động cháy	Các kênh phổ hồng ngoại nhiệt môi trường xung quanh		1000
S8	10854,00	776,00	Nhiệt độ bề mặt đất và biển, hoạt động cháy			
S9	12022,50	905,00	Nhiệt độ bề mặt đất và biển			
F1	3742,00	398,00	Hoạt động cháy	Các kênh hồng ngoại nhiệt phát xạ phổ do cháy		
F2	10854,00	776,00	Hoạt động cháy			

2.3. Phương pháp nghiên cứu

Phương pháp tổng thể được áp dụng bao gồm ba bước chính: Tiền xử lý, xử lý dữ liệu và đánh giá kết quả và được thể hiện trong Hình 2.

2.3.1. Tiền xử lý

Quá trình chuyển giá trị bức xạ sang hệ số phản xạ, tái chia mẫu, hiệu chỉnh hình học và cắt khu vực nghiên cứu được thực hiện bằng cách sử dụng các công cụ trong phần mềm SNAP phiên bản 8.0. Các công cụ xử lý ảnh vệ tinh Sentinel-3 có thể được tìm thấy tại <http://step.esa.int/main/download/>.

Công thức tính chuyển giá trị bức xạ L_{TOA} sang hệ số phản xạ R_{TOA} như sau:

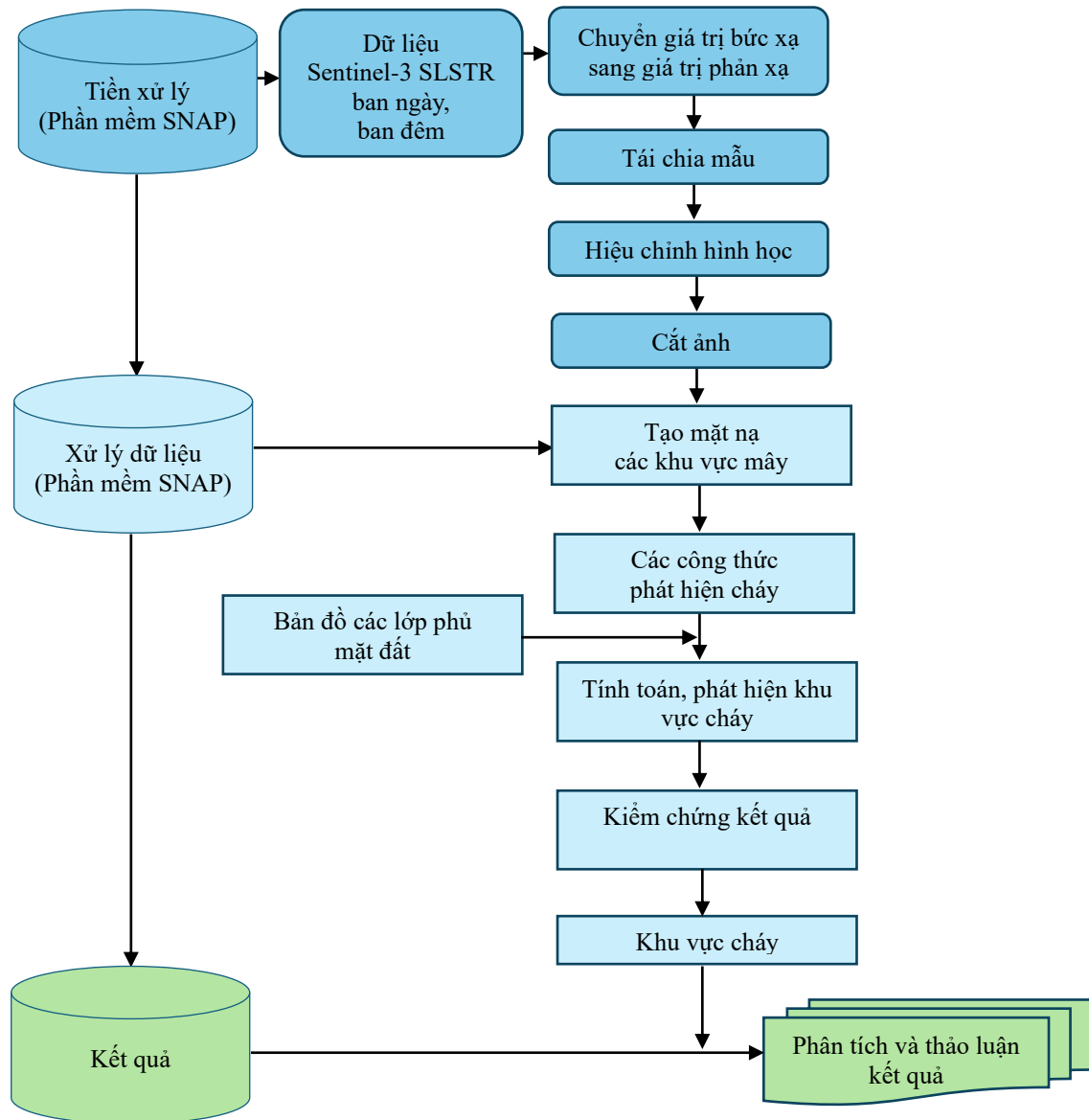
$$R_{TOA}(\lambda) = \frac{\pi L_{TOA}}{E_0(\lambda) \cos(\theta)} \quad (1)$$

trong đó:

E_0 và θ là bức xạ quang phổ mặt trời và góc thiên đỉnh của mặt trời tại thời điểm thu nhận, tương ứng. Các giá trị bức xạ quang phổ mặt trời được lấy từ siêu dữ liệu sản phẩm mức độ 1 của sản phẩm SLSTR.

Giá trị bức xạ L_{TOA} là biến được đo trực tiếp bằng các thiết bị viễn thám. Đó là lượng ánh sáng được đo bằng dụng cụ từ bề mặt của vật thể. Trong sản phẩm SLSTR được biểu thị dưới dạng độ sáng của bề mặt trên mỗi bước sóng đơn vị [$\text{mW} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \cdot \text{nm}^{-1}$ = watt trên mét vuông trên nanomet].

Hệ số phản xạ R_{TOA} là tỷ lệ (phần trăm) của lượng ánh sáng rời khỏi mục tiêu và lượng ánh sáng tiến tới mục tiêu. Nó không có đơn vị. Nó là thuộc tính của đối tượng/vật liệu được quan sát.



Hình 2. Sơ đồ phương pháp phát hiện cháy rừng từ dữ liệu đa phổ Sentinel-3 SLSTR

2.3.2. Xử lý ảnh

2.3.2.1. Tạo mặt nạ các khu vực mây

Dựa vào các đặc trưng phản xạ phổ của của mây trên các kênh phổ của sản phẩm SLSTR sau khi đã chuyển đổi sang hệ số phản xạ phổ. Các ngưỡng được sử dụng để khoanh vùng chính xác các đám mây trong khu vực nghiên cứu được áp dụng theo (Serco Italia SPA, 2018) như sau:

Đối với các dữ liệu SLSTR ban ngày sử dụng các ngưỡng theo một trong các công thức sau:

$$(S2_reflectance_an + S3_reflectance_an) > 0.9 \quad (2)$$

$$S9_BT_in < 265 \quad (3)$$

$$((S2_reflectance_an + S3_reflectance_an) > 0.7 \text{ and } S9_BT_in < 285) \quad (4)$$

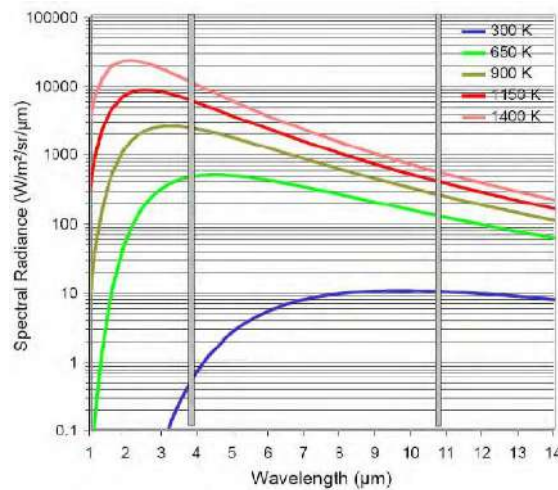
Đối với các dữ liệu SLSTR ban đêm sử dụng các ngưỡng theo công thức (3).

2.3.2.2. Phương pháp phát hiện hoạt động cháy

Trong bước này, một thuật toán đơn giản được áp dụng để xác định các pixel chứa một hoặc nhiều đám cháy đang hoạt động tại thời điểm thu nhận dữ liệu của vệ tinh Sentinel-3. Cách tiếp cận này khai thác sự khác biệt về

phản xạ phổ giữa kênh phổ hồng ngoại trung (MIR) và kênh phổ hồng ngoại nhiệt (TIR) đến các mục tiêu pixel có nhiệt độ cao bất thường. Đặc biệt, thuật toán tìm kiếm một độ bức xạ (nhiệt độ sáng - BT) tăng đáng kể ở bước sóng 3,74 μm (kênh phổ F1) so với giá trị đo được độ bức xạ ở bước sóng 10,85 μm (kênh phổ F2) được mô tả theo công thức (5) và (6) dưới đây.

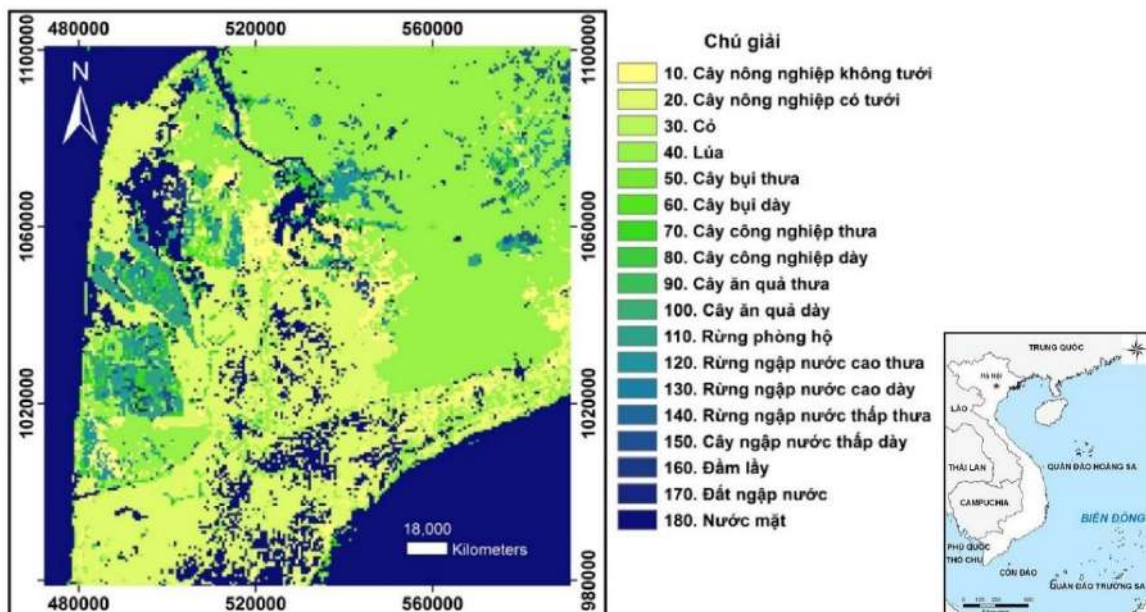
Dấu hiệu hoạt động cháy đặc trưng này là kết quả của sự khác biệt lớn về vật đen bức xạ ở bước sóng 4 μm và 11 μm phát ra ở nhiệt độ đốt cháy thực vật được mô tả bởi hàm số Planck. Hình 3 dưới đây cho thấy bức xạ quang phổ phát ra từ các vật thể đen ở môi trường xung quanh Trái đất nhiệt độ (300 K) và phạm vi nhiệt độ có thể xảy ra cháy do thảm thực vật (650 - 1400 K). Các bước sóng trung tâm của hai kênh phổ Sentinel-3 SLSTR MIR (F1 có bước sóng 3,74 μm) và TIR (F2 có bước sóng 10,85 μm) là cũng được chỉ ra đúng trong trường hợp này. Theo đó, bức xạ của lớp phủ bề mặt bao gồm cả thực vật ở kênh MIR nhỏ hơn kênh TIR ở nhiệt độ bề mặt 300 K. Tuy nhiên, khi có hoạt động cháy với nhiệt độ lớn hơn 600 K thì xảy ra ngược lại, bức xạ của lớp phủ bề mặt ở kênh MIR lớn hơn kênh TIR.



Hình 3. Sentinel-3 Active Fire: Phát hiện cháy và đánh giá năng lượng bức xạ cháy (ESA)

2.3.2.3. Bổ sung bản đồ các lớp phủ bề mặt

Ngoài ra, trong quá trình xử lý cần sử dụng mặt nạ các đám mây để loại trừ các khu vực có hoạt động cháy. Bên cạnh đó, cần bổ sung bản đồ các lớp phủ bề mặt để khoanh vùng các khu vực có thực vật nơi có thể xảy ra hoạt động cháy rừng. Trong nghiên cứu này, dữ liệu các lớp phủ bề mặt “land_cover_CCILandCover-2015” được cung cấp để lựa chọn các vùng có thực phủ (Hình 4).



Hình 4. Bản đồ các lớp phủ bề mặt khu vực nghiên cứu

3. Kết quả và thảo luận

3.1. Tính toán, phát hiện khu vực cháy

Việc tính toán để phát hiện khu vực cháy sử dụng các ngưỡng đối với phương pháp đã đề xuất ở tiêu Mục 2.3.2.2 kết hợp với số liệu mặt nạ các đám mây và sự khoanh vùng các loại thực phủ lấy từ bản đồ các lớp phủ bề mặt để chiết tách các khu vực cháy trên dữ liệu thu nhận ở thời điểm xảy ra cháy trong khu vực nghiên cứu.

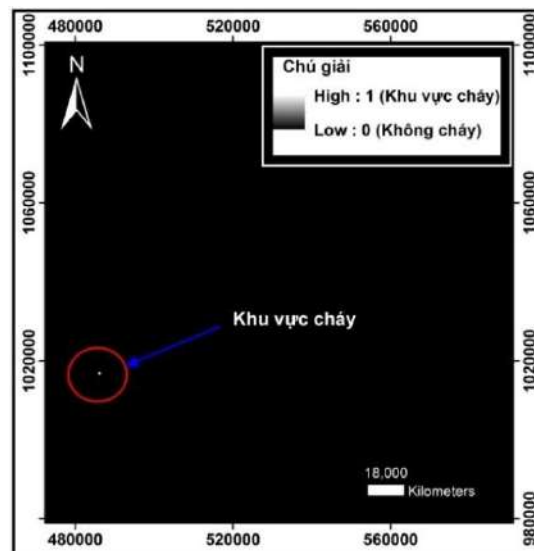
Đối với các dữ liệu SLSTR ban ngày sử dụng các ngưỡng theo một trong các công thức sau:

$$F1_BT_in > 325 \text{ and } (F1_BT_in - F2_BT_in) > 18 \text{ and } cloud_mask == 0 \text{ and } 'land_cover_CCI LandCover-2015' >= 50 \text{ and } 'land_cover_CCI LandCover-2015' <= 110 \quad (5)$$

Đối với các dữ liệu SLSTR ban đêm sử dụng các ngưỡng theo một trong các công thức sau:

$$F1_BT_in > 315 \text{ and } (F1_BT_in - F2_BT_in) > 15 \text{ and } cloud_mask == 0 \text{ and } 'land_cover_CCI LandCover-2015' >= 50 \text{ and } 'land_cover_CCI LandCover-2015' <= 110 \quad (6)$$

Kết quả của quá trình xử lý tính toán theo công thức (5) và công thức (6) đã đưa ra khu vực xảy ra cháy rừng ở Hình 5.



Hình 5. Kết quả khu vực cháy tính toán từ dữ liệu Sentinel-3 SLSTR đêm 10/4/2024 và ngày 11/4/2024

3.2. Đánh giá độ chính xác

Việc đánh giá độ chính xác được so sánh vị trí phát hiện cháy chiết tách từ dữ liệu Sentinel-3 SLSTR ban đêm và ban ngày thu nhận trong quá trình xảy ra cháy và vị trí diện tích rừng trầm bị cháy trên thực địa (Hình 6). Kết quả đánh giá cho thấy hoàn toàn trùng khớp về vị trí tọa độ khu vực đã xảy ra cháy trên dữ liệu tính toán và thực địa.



Hình 6. Vụ cháy đã thiêu rụi hoàn toàn 40 ha rừng trầm tại Đội quản lý đất quốc phòng (Nông trường 402 cũ, xã Khánh Bình Tây, huyện Trần Văn Thời) (Trần An, 2024)

Mặc dù đã có các khó khăn phát hiện cháy rừng đối với dữ liệu Sentinel-3 SLSTR ban ngày do sự phản xạ phổ mạnh của ánh sáng mặt trời lên các kênh phổ. Nhưng các kết quả tính toán phát hiện cháy rừng đối với dữ liệu Sentinel-3 SLSTR ban ngày đã góp phần tăng thêm thông tin hữu ích bổ sung về phát hiện cháy cho dữ liệu Sentinel-3 SLSTR ban đêm. Kết quả này khẳng định dữ liệu Sentinel-3 SLSTR ban ngày và ban đêm có thể phát hiện được sự tăng nhiệt độ của đối tượng thực vật để cảnh báo cho kiểm lâm có thể kiểm tra có khả năng xuất hiện đám cháy sớm nhất có thể nhằm ngăn chặn cháy rừng và ảnh hưởng môi trường đất, không khí.

4. Kết luận

Phát hiện cháy rừng sử dụng dữ liệu Sentinel-3 SLSTR ban ngày và ban đêm liên tục hàng ngày đêm là thông tin ban đầu quan trọng. Bên cạnh đó, sự kết hợp với các nguồn thông tin khác với các phát hiện cháy với sự kiểm chứng tại thực địa sẽ giúp phát hiện sớm khả năng cháy rừng nhằm giảm thiểu thiệt hại để bảo vệ rừng và môi trường sống. Các kết quả phát hiện cháy và kiểm chứng ở thực địa đã khẳng định tính hiệu quả của dữ liệu sử dụng trong khu vực nghiên cứu.

Các sự kiện cháy rừng có hậu quả lớn do sự phát hiện muộn và phương pháp phòng cháy chữa cháy chưa thực sự hiệu quả. Kết quả nghiên cứu của bài báo giúp ích cho việc phát hiện cháy ở các khu vực rừng ở vùng sâu, xa ít người biết trong điều kiện có dữ liệu cung cấp sớm nhất có thể. Các vệ tinh Sentinel-3C và 3D sẽ sớm được đưa lên quỹ đạo để tăng cường và thay thế các vệ tinh Sentinel-3A và 3B đảm bảo sự giám sát liên tục.

Tài liệu tham khảo

- Anh Thu, 2015. Xã Khánh Bình Tây: tiềm năng phát triển kinh tế và du lịch biển đảo. https://tranvanthoi.camau.gov.vn/wps/portal/?1dmy&page=htvt.trangchitiet&urile=wcm%3Apath%3A/tvtlibrary/tvtsite/tintucsukien/vanhoaxahoi/admin_00324 (truy cập ngày 01/6/2024).
- Bowman, D.M.J.S., Balch, J.K., Artaxo, P., Bond, W.J., Carlson, J.M., Cochrane, M.A., D'Antonio, C.M., DeFries, R.S., Doyle, J.C., Harrison, S.P., Johnston, F.H., Keeley, J. E., Krawchuk, M.A., Kull, C.A., Marston, J.B., Moritz, M.A., Prentice, I.C., Roos, C.I., Scott, A.C., Swetnam, T.W., Van der Werf, G.R., Pyne, S.J., 2009. Fire in the Earth system. *Science* 324, 481-484.
- Coppo, P., Ricciarelli, B., Brandani, F., Delderfeld, J., Ferlet, M., Mutlow, C., et al., 2010. SLSTR: a high accuracy dual scan temperature radiometer for sea and land surface monitoring from space. *J. Mod. Opt.* 57(18), 1815-1830, October 2010.
- Delderfeld, J., Llewellyn-Jones, D.T., Bernard, R., De Javel, Y., Williamson, E.J., Mason, I., Pick, D.R., Barton, I.J., 1986. The along track scanning radiometer (ATSR) for ERS1. In: *Instrumentation for Optical Remote Sensing from Space*, vol. 589. International Society for Optics and Photonics, pp. 114-120.
- Freeborn, P.H., Wooster, M.J., Hao, W.M., Ryan, C.A., Nordgren, B.L., Baker, S.P., Ichoku, C., 2008. Relationships between energy release, fuel mass loss, and trace gas and aerosol emissions during laboratory biomass fires. *J. Geophys. Res. Atmos.* 113(D1).
- Johnston, F., Hanigan, I., Henderson, S., Morgan, G., Bowman, D., 2011. Extreme air pollution events from bushfires and dust storms and their association with mortality in Sydney, Australia 1994-2007. *Environ. Res.*, 111, 811-816.
- Kaiser, J.W., Heil, A., Andreae, M.O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.J., Razinger, M., Schultz, M.G., Suttie, M., van der Werf, G.R., 2012. Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences* 9, 527-554.
- Natarajan, M., Pierce, R.B., Schaack, T.K., Lenzen, A.J., Al-Saadi, J.A., Soja, A.J., Charlock, T.P., Rose, F.G., Winker, D.M., Worden, J.R., 2012. Radiative forcing due to enhancements in tropospheric ozone and carbonaceous aerosols caused by Asian fires during spring 2008. *J. Geophys. Res. Atmos.*, 117.
- Nguyen, H.M., Wooster, M.J., 2020. Advances in the estimation of high Spatio-temporal resolution pan-African top-down biomass burning emissions made using geostationary fire radiative power (FRP) and MAIAC aerosol optical depth (AOD) data. *Rem. Sens. Environ.* 248, 111971.
- Nguyen, H.M., He, J., & Wooster, M.J. (2023). Biomass burning CO, PM and fuel consumption per unit burned area estimates derived across Africa using geostationary SEVIRI fire radiative power and Sentinel-5P CO data. *Atmospheric Chemistry and Physics*, 23(3), 2089-2118.
- Polehampton, E., C. Cox, D. Smith (RAL Space), D. Ghent (ULEIC), M. Wooster, W. Xu (KCL), J. Bruniquel, C. Henocq (ACRI-ST), S. Dransfeld (ESA), 2023. Sentinel-3 SLSTR Land User Guide, ESA
- Serco Italia SPA, 2018. Active Fire Detection with Sentinel-3 SLSTR (SNAP) - Portugal (version 1.2). Retrieved from RUS Lectures at <https://rus-copernicus.eu/portal/the-rus-library/learn-by-yourself/>.
- Trang thông tin điện tử đảng bộ huyện Trần Văn Thời, 2024. Thông tin về vụ cháy rừng tại Nông trường 402 cũ. <https://tranvanthoi.camau.dcs.vn/chinh-quyen/thong-tin-ve-vu-chay-rung-tai-nong-truong-402-cu-123492> (available online 01/6/2024).

Trần An, 2024. Vụ cháy rừng ở Cà Mau: Đơn vị quản lý rừng chưa thực hiện tốt công tác PCCC. <https://congty.vn/vu-chay-rung-o-ca-mau-don-vi-quan-ly-rung-chua-thuc-hien-tot-cong-tac-pccc-425893.html> (truy cập ngày 01/6/2024).

Wooster, M.J., Roberts, G.J., Giglio, L., Roy, D.P., Freeborn, P.H., Boschetti, L., et al., 2021. Satellite remote sensing of active fires: history and current status, applications and future requirements. *Rem. Sens. Environ.*, 267, 112694.

Xu, W., Wooster, M., He, J., Zhang, T., 2020. First study of sentinel-3 SLSTR active fire detection and FRP retrieval: night-time algorithm enhancements and global intercomparison to MODIS and VIIRS AF products. *Rem. Sens. Environ.*, 248, 111947.

Xu, W., Wooster, M.J., Polehampton, E., Yemelyanova, R., Zhang, T., 2021. Sentinel-3 active fire detection and FRP product performance-Impact of scan angle and SLSTR middle infrared channel selection. *Rem. Sens. Environ.* 261, 112460.

Zhang, T., Wooster, M., de Jong, M., Xu, W., 2018. How well does the “small fire boost” methodology used within the GFED4.1s fire emissions database represent the timing, location and magnitude of agricultural burning? *Rem. Sens.* 10(6), 823.

ABSTRACT

Detection of active forest fire using Sentinel-3 SLSTR data in Khanh Binh Tay commune, Tran Van Thoi district, Ca Mau province

Doan Thi Nam Phuong^{1,2*}, Nguyen Van Trung^{1,2}, Le Thi Thu Ha^{1,2}, Nguyen Quang Minh³

¹*Hanoi University of Mining and Geology*

²*Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology*

³*Department of Cartography, The General Staff*

Sentinel-3A and 3B SLSTR multispectral satellite data provide the ability to monitor the earth's surface temperature to detect areas of wildfires. Forest fire detection is based on the difference in surface temperature of the fire area and the average ambient temperature of the surrounding area (300 K). This work was performed using two spectral channels Sentinel-3 SLSTR MIR (F1 band with wavelength 3.74 μm) and spectral channels Sentinel-3 SLSTR TIR (F2 band with wavelength 10.85 μm). These two spectral channels are used because the difference increases for areas without wildfires and the difference decreases for areas with wildfires. In addition, other spectral channels of Sentinel-3 SLSTR data exclude areas covered by clouds and supporting data on land cover classes including vegetation classes help localize forest areas are combined to give the final wildfire area result. Sentinel-3 multispectral satellite data is effective in detecting forest fires during both day and night due to the use of infrared spectral channels and the ability to provide daily and nightly data with a wide scanning range (from 740 km to 1400 km) covering a large range. In this study, the value thresholds are given for F1 bands (325 and 315 for day and night, respectively) and the difference between bands F2 and F1 (18 and 15 for day and night, respectively). For forest areas that have been extracted from other land use classes to determine the active fire area.

Keywords: Sentinel-3, SLSTR, forest fire, Ca Mau.

Công nghệ SBAS xử lý dữ liệu lớn cho xác định biến dạng bề mặt địa hình vùng Thành phố Hồ Chí Minh

Nguyễn Minh Hải^{1,*}

¹Trường Đại học Mỏ - Địa chất, Hà Nội, Việt Nam

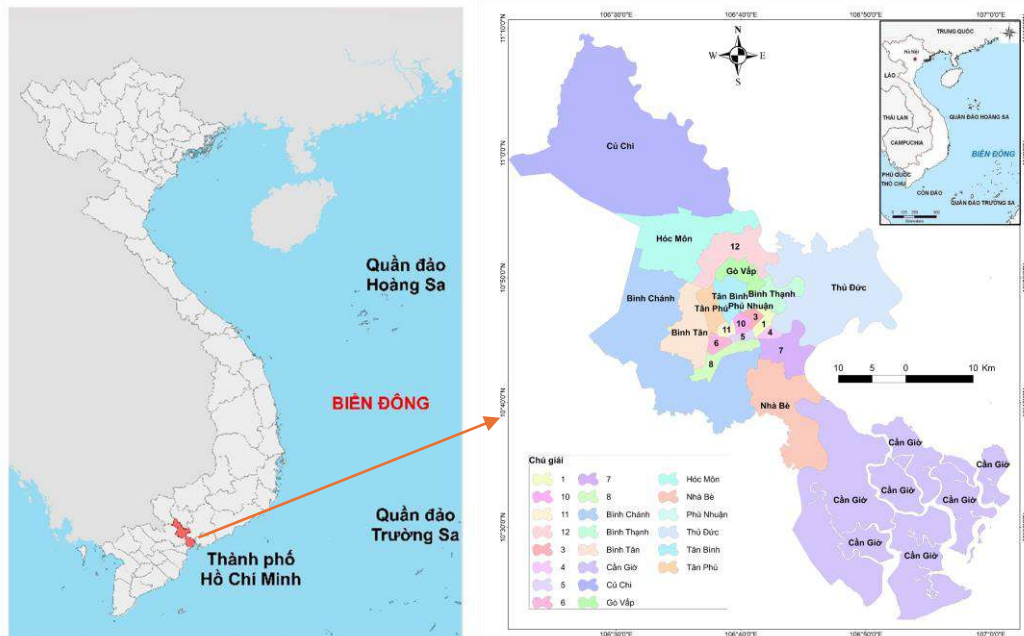
TÓM TẮT

Thành phố Hồ Chí Minh là trung tâm kinh tế trọng điểm phía Nam và hàng đầu cả nước, với diện tích chiếm 0,6% và dân số chiếm 6,6% so với cả nước, có tốc độ tăng trưởng kinh tế cao nhất nước. Với sự phát triển nhanh về cơ sở hạ tầng, nền địa chất yếu đã khiến cho bề mặt địa hình của khu vực Thành phố bị biến dạng. Sụt lún đất ở khu vực Thành phố Hồ Chí Minh dẫn đến nhiều nguy cơ tiềm ẩn và đe dọa sự ổn định của môi trường và cuộc sống của con người. Trong bài báo này, tác giả đã áp dụng kỹ thuật Small Baseline Subset (SBAS) để theo dõi sụt lún đất của Thành phố Hồ Chí Minh. Bộ dữ liệu ảnh radar từ năm 2006 đến năm 2022 được sử dụng để tạo ra tốc độ biến dạng trung bình và chuỗi thời gian biến dạng. Kết quả xác định cho thấy, (1) tốc độ biến dạng trung bình hàng năm xác định từ ảnh vệ tinh ALOS 1 (2006-2011) thì diện tích biến dạng tập trung chủ yếu ở giá trị lún (-20)-0 mm/năm (598,493898 km²) và giá trị trôi 0-20 mm/năm (1.263,241792 km²); (2) tốc độ biến dạng trung bình hàng năm xác định từ ảnh vệ tinh Sentinel (sử dụng ảnh từ năm 2014-2022 để xác định biến dạng) thì diện tích biến dạng có sự thay đổi nhiều giá trị lún < -40 (25,017193 km²), giá trị lún (-40)-(-20) mm/năm (158,605032 km²), giá trị lún (-20)-0 (1.052,709457 km²) và giá trị trôi 0-20 mm/năm (186,858293 km²).

Từ khoá: Hồ Chí Minh, SBAS, InSAR.

1. Giới thiệu

Thành phố Hồ Chí Minh là trung tâm kinh tế trọng điểm phía Nam và hàng đầu cả nước. Với tổng diện tích tự nhiên 2.095,2 km² chiếm 0,6% so với cả nước nhưng mức GDP lớn nhất trong cả nước.



Hình 1. Khu vực nghiên cứu

Nền đất thành phố Hồ Chí Minh được hình thành trên hai tầng trầm tích: Pleistocen và Holocen. Trong đó, trầm tích Pleistocen (phù sa cổ) chiếm phần hết phía Bắc - Tây Bắc và Đông Bắc Thành phố; Trầm tích Holocen (trầm tích phù sa trẻ) có nguồn gốc ven biển, vũng vịnh, sông biển, lòng sông và bãi bồi..., đã hình thành nên

* Tác giả liên hệ

Email: nguyminhhai@humg.edu.vn

nhiều loại đất khác nhau (Tìm hiểu Việt Nam, 2019). Địa hình thành phố thấp dần từ Bắc xuống Nam và từ Tây sang Đông.

Sụt lún là một hiện tượng chung của đồng bằng, nguyên nhân gây ra có thể do tự nhiên (nén và hợp nhất trầm tích, dịch chuyển trầm tích,...) và do nhân tạo (phát triển đô thị, khai thác nước ngầm,...). Các hoạt động của tự nhiên và nhân tạo đã gây ra sụt lún đất, được xác định bằng các phép đo trực tiếp, mô hình và InSAR (Urban và cộng sự, 2014; Minderhoud và cộng sự, 2017). Các phép đo này thực hiện cho các khu vực nhỏ, do đó việc quan trắc và giám sát cho toàn bộ khu vực trên diện rộng cho khu vực nhằm làm rõ sự thay đổi không gian-thời gian, phân tích nguyên nhân để khám phá cơ chế của sụt lún đất. Từ đó có thể góp phần phòng ngừa và giảm nhẹ thiệt hại do sụt lún đất gây ra.

Quan trắc và giám sát sụt lún thực hiện dựa trên các phép đo tại chỗ (hệ thống định vị toàn cầu-GPS, thủy chuẩn, toàn đạc). Tuy nhiên, để thực hiện công việc này thì chi phí tốn kém (nhân công, thiết bị, thời gian,...) và số liệu thu được chỉ là ở dạng điểm không đặc trưng thể hiện cho vùng cần quan trắc. Ngày nay với sự phát triển của khoa học, kỹ thuật InSAR cung cấp các phép đo biến dạng với độ phân giải không gian cao, độ chính xác cao được ứng dụng nhiều để theo dõi biến dạng (Zhang và cộng sự, 2018; Ferretti và cộng sự, 2000; Berardino và cộng sự, 2002).

Một trong những phương pháp quan trắc biến dạng bề mặt đất được tiến hành sớm ở miền Bắc đó là phương pháp thủy chuẩn. Phương pháp đo GPS hai tần số độ chính xác cao cũng có thể được sử dụng trong nghiên cứu biến dạng bề mặt đất.

Đối với phương pháp radar giao thoa vệ tinh ứng dụng cho giám sát biến dạng bề mặt đất tại Việt Nam mới chỉ được thực hiện ở hai thành phố lớn là Thủ đô Hà Nội và Thành phố Hồ Chí Minh.

Tại Thành phố Hồ Chí Minh, việc ứng dụng phương pháp radar giao thoa trong nghiên cứu biến dạng bề mặt đất không thể không đề cập tới kết quả của nhóm nghiên cứu do Lê Văn Trung và Hồ Tổng Minh Định (2008). Trong nghiên cứu năm 2008 (Lê Văn Trung và Hồ Tổng Minh Định, 2008), nhóm đã sử dụng 02 ảnh ENVISAT ASAR ngày 16/12/2003 và 04/5/2004 với dữ liệu DEM-SRTM được xây dựng theo kỹ thuật InSAR. Nhóm cũng đã thử nghiệm tiếp kỹ thuật xử lý PSInSAR với các ảnh vệ tinh ERS-1 và ERS-2 (được cung cấp bởi cơ quan không gian châu Âu) cho khu vực Thành phố Hồ Chí Minh. Các kết quả thử nghiệm ban đầu đã cho thấy bề mặt địa hình ở tất cả các quận huyện tại Thành phố Hồ Chí Minh đều bị biến dạng với các mức độ khác nhau.

2. Dữ liệu và phương pháp

2.1. Dữ liệu

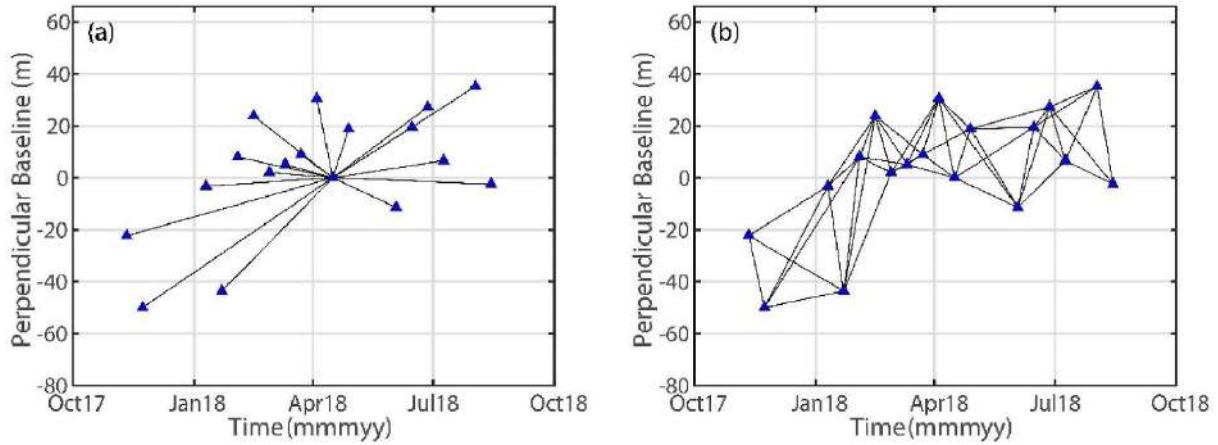
Để xác định biến dạng bề mặt của khu vực Đồng bằng sông Cửu Long, nhóm tác giả đã thu thập 1037 bộ dữ liệu hình ảnh SAR từ năm 2006 đến năm 2022 bởi các vệ tinh ALOS1/PALSAR băng tần L gồm và vệ tinh Sentinel-1A băng tần C gồm 2 quỹ đạo cho khu vực nghiên cứu. Bảng 1 cung cấp các thông số hình ảnh của từng quỹ đạo, bao gồm số, khoảng thời gian và các thông số có hệ thống (dữ liệu được tải tại trang web <https://search.asf.alaska.edu/>).

Bảng 1. Bảng dữ liệu ảnh radar của khu vực nghiên cứu

STT	Dữ liệu	Dải	Khung	Số lượng ảnh	Hướng	Thời gian
1	Alos Palsar	476	190, 200, 210	36	Tăng dần	5/16/2006 - 4/22/2011
		477		65		
		110	3390, 3400, 3410	9	Giảm dần	
		111		6		
2	Sentinel-1A	26	26, 28, 32	285	Tăng dần	6/15/2014 - 12/31/2022
		128	29	176		
		18	554, 556, 558	250	Giảm dần	
		120	552, 554, 556	210		

2.2. Phương pháp

Việc sử dụng và phát triển kỹ thuật InSAR để phân tách độ cao mặt đất và tín hiệu biến dạng (Li và cộng sự, 2022). Các kỹ thuật D-InSAR, PSI,... đã được phát triển phục vụ cho việc xác định biến dạng bề mặt đất, tuy nhiên các phương pháp này vẫn còn một số hạn chế nhất định (ví dụ: D-InSAR thu thông tin biến dạng tại tất cả các điểm trong toàn bộ hình ảnh, PSI dễ bị ảnh hưởng của mối quan hệ không gian-thời gian). Từ đó Berardino và cộng sự.(2002) đã nghiên cứu và giới thiệu thuật toán phân tách giá trị số ít (Singular value decomposition-SVD) và phát triển nó thành phương pháp SBAS cổ điển.



Hình 2. Sơ đồ đường chảy cơ sở cho (a) phương pháp PSI và (b) SBAS (Li và cộng sự, 2022)

Đã có một số tác giả nghiên cứu và chỉ ra rằng phương pháp SBAS có độ chính xác cao và đáng tin cậy để theo dõi sự dịch chuyển chậm của bề mặt đất (Casu và cộng sự, 2006). Trong những năm gần đây, kỹ thuật InSAR-SBAS (Small Baseline Subset) được sử dụng rộng rãi trong việc xác định các dịch chuyển của bề mặt đất (Li và cộng sự, 2022).

Trên khu vực nghiên cứu có $(N+1)$ ảnh SAR tương ứng với thời gian (t_0, \dots, t_N) , cứ 2 ảnh SAR có thể tạo thành một cặp ảnh giao thoa M thỏa mãn điều kiện (Berardino và cộng sự, 2002):

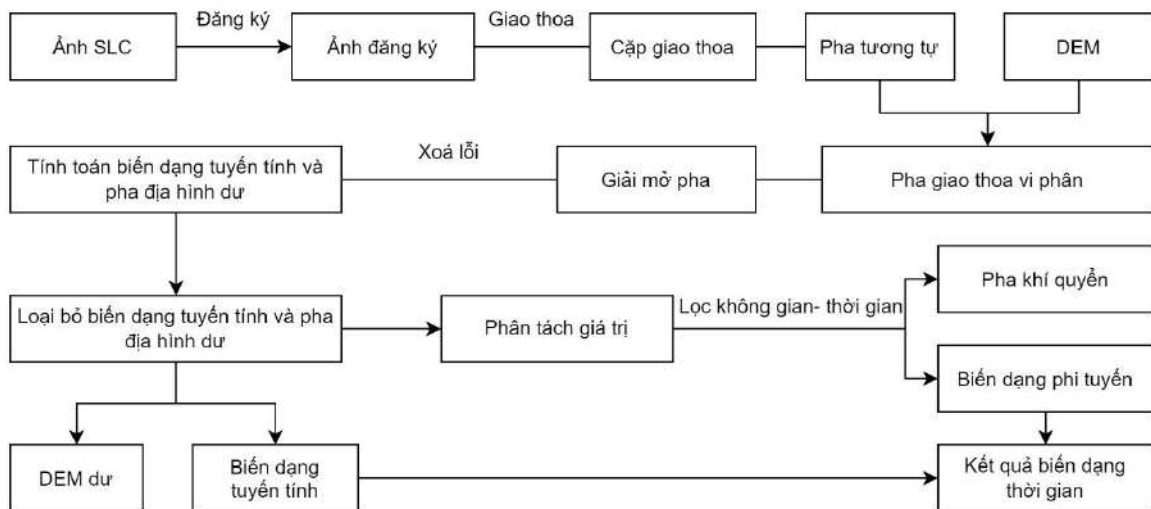
$$\frac{N+1}{2} \leq M \leq N \left(\frac{N+1}{1} \right) \quad (1)$$

Các điểm mà sự gắn kết pha trung bình cao hơn một ngưỡng nhất định được gọi là điểm gắn kết cao. Giai đoạn giải mở pha của một điểm gắn kết cao (x, r) của một điểm giải mở pha i (tương ứng với hai khoảng thời gian t_1 và t_2) sau khi loại bỏ các nhiễu, pha địa hình được biểu diễn theo công thức (Berardino và cộng sự, 2002):

$$\Delta\varphi_i(x, r) = \varphi(t_1, x, r) - \varphi(t_2, x, r) \approx \frac{4\pi}{\lambda} [d(t_1, x, r) - d(t_2, x, r)] \quad (2)$$

trong đó: λ - là bước sóng truyền đi; $d(t_1, x, r)$ và $d(t_2, x, r)$ là các biến dạng tích lũy của line of sight (LOS) tại thời gian t_1 và t_2 so với thời gian tham chiếu t_0 .

Từ những phân tích trên, tác giả lựa chọn phương pháp SBAS để xác định biến dạng bề mặt địa hình cho khu vực thực nghiệm. Trong quá trình lý số liệu ảnh vệ tinh radar, nhóm tác giả sử dụng phần mềm GAMA theo sơ đồ hình vẽ dưới đây:

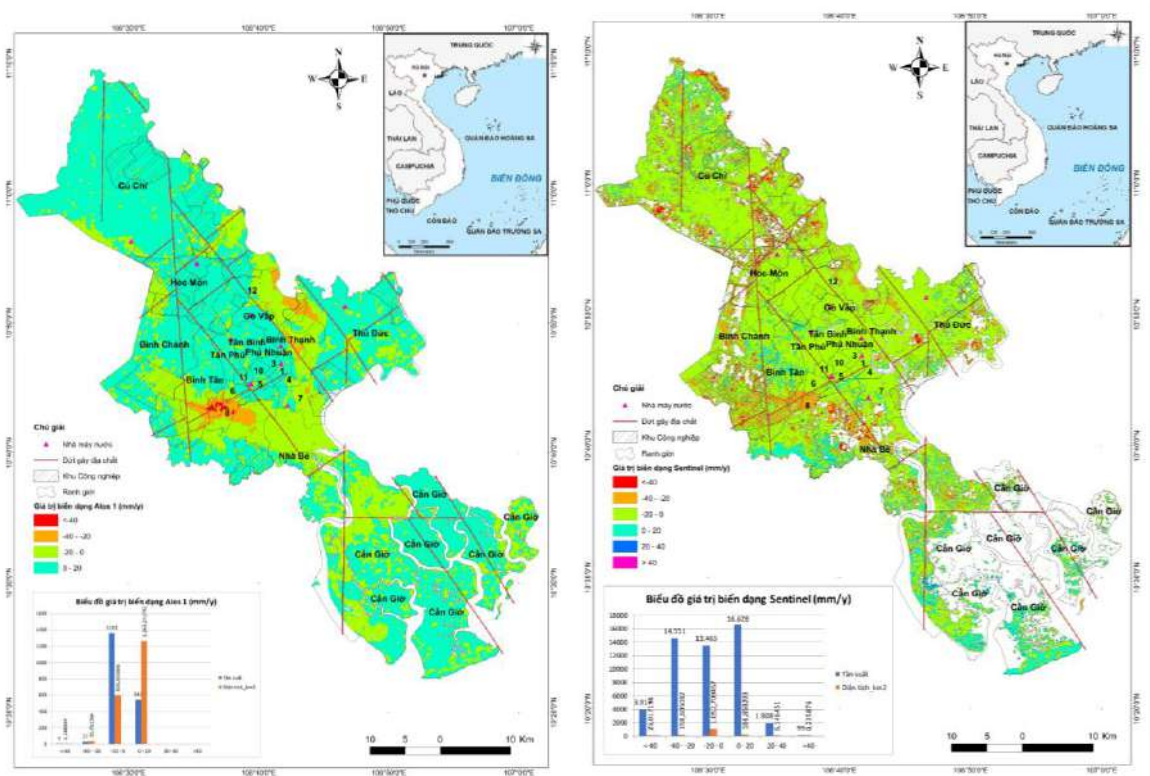


Hình 3. Sơ đồ công nghệ SBAS xử lý ảnh radar sử dụng phần mềm GAMA (Guangcai Feng, 2019)

Khi sử dụng phần mềm GAMA, nhóm tác giả đã cài đặt một số thông số cho quá trình xử lý ảnh radar: tỷ lệ góc nhìn 3:8; ngưỡng kết hợp 0,4; biên độ giá trị ngưỡng 0,3 và sử dụng phương pháp lọc cải tiến Goldstein.

3. Kết quả

Kết quả của quá trình xử lý ảnh sử dụng phần mềm GAMA được biên tập trên phần mềm ArcGIS(ArcGIS Desktop, n.d.), GMT(Wessel & Smith, 2006; GMT, 2019) để thành lập các bản đồ chuyên đề.



Hình 4. Kết quả xác định biến dạng bề mặt địa hình; Trái- Alos 1, Phải- Sentinel

Bảng 2. Thống kê biến dạng bề mặt địa hình

Mã	Giá trị	Alos 1		Sentinel	
		Tần suất	Diện tích (km ²)	Tần suất	Diện tích (km ²)
1	<-40	4	2,166964	3917	25,017193
2	-40 - -20	27	35,763364	14551	158,605032
3	-20 - 0	1363	598,493898	13465	1.052,709457
4	0 - 20	542	1.263,241792	16628	186,858293
5	20- 40			1908	6,146451
6	>40			99	0,235876

Dựa vào Bảng 2 thống kê diện tích biến dạng bề mặt địa hình ta thấy:

- Tốc độ biến dạng trung bình hàng năm xác định từ ảnh vệ tinh ALOS 1 (sử dụng ảnh từ năm 2006-2011 để xác định biến dạng) thì diện tích biến dạng tập trung chủ yếu ở giá trị lún (-20)-0 mm/năm (598,493898 km²) và giá trị trồi 0-20 mm/năm (1.263,241792 km²).
- Tốc độ biến dạng trung bình hàng năm xác định từ ảnh vệ tinh Sentinel (sử dụng ảnh từ năm 2014-2022 để xác định biến dạng) thì diện tích biến dạng có sự thay đổi nhiều giá trị lún < -40 (25,017193 km²), giá trị lún (-40)-(-20) mm/năm (158,605032 km²), giá trị lún (-20)-0 (1.052,709457 km²) và giá trị trồi 0-20 mm/năm (186,858293 km²).

4. Thảo luận

Kết quả của khu vực nghiên cứu đã khẳng định được tính thực tiễn của công nghệ SBAS xử lý dữ liệu lớn ảnh vệ tinh radar cho xác định biến dạng bề mặt địa hình. Việc kết hợp hai loại ảnh radar là ALOS-1 và Sentinel-1 ta thấy rằng dữ liệu miễn phí hoàn toàn phù hợp cho mục đích nghiên cứu và xác định biến dạng bề mặt địa hình.

Kết quả được sử dụng tiếp cho các nghiên cứu sau này, cũng như giúp các nhà quản lý có được cái nhìn tổng quan khi quy hoạch phát triển cho Thành phố Hồ Chí Minh.

Quá trình xử lý ảnh radar sử dụng phần mềm GAMA đòi hỏi trình độ của người xử lý và cần lựa chọn các thông số phù hợp cho quá trình xử lý. Để có thể đánh giá được độ chính xác thì cần thêm nhiều nguồn dữ liệu khác (ví dụ: đo thủy chuẩn, số liệu từ các trạm quan trắc, đo GPS,...) để so sánh với giá trị biến dạng bề mặt địa hình được xác định trong nghiên cứu. Nên trong bài báo này, tác giả chỉ dừng lại ở việc xác định biến dạng bề mặt địa hình bằng việc sử dụng phương pháp SBAS để làm kết quả tham khảo cho các nghiên cứu sau này.

Tài liệu tham khảo

- ArcGIS Desktop (n.d.). *ArcGIS tutorial*. <https://desktop.arcgis.com/en/arcmap/10.6/get-started/introduction/arcgis-tutorials.htm>
- Berardino, P., Fornaro, G., Lanari, R., & Sansosti, E. (2002). A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11), 2375-2383. <https://doi.org/10.1109/TGRS.2002.803792>
- Casu, F., Manzo, M., & Lanari, R. (2006). A quantitative assessment of the SBAS algorithm performance for surface deformation retrieval from DInSAR data. *Remote Sensing of Environment*, 102(3-4), 195-210. <https://doi.org/10.1016/j.rse.2006.01.023>
- Erban, L. E., Gorelick, S. M., & Zebker, H. A. (2014). Groundwater extraction, land subsidence, and sea-level rise in the Mekong Delta, Vietnam. *Environmental Research Letters*, 9(8), 084010. <https://doi.org/10.1088/1748-9326/9/8/084010>
- Ferretti, A., Prati, C., & Rocca, F. (2000). Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5), 2202-2212. <https://doi.org/10.1109/36.868878>
- GMT (2019). *GMT Chinese manual v5.4*.
- Guangcai Feng, Z. L. (2019). *Topography and deformation monitoring in the Pearl River Delta based on radar interferometry*. Central South University.
- Lê Văn Trung và Hồ Tổng Minh Định. (2008). Ứng dụng kỹ thuật InSAR vi phân trong quan trắc biến dạng mặt đất khu vực thành phố Hồ Chí Minh. *Tạp Chí Phát Triển KH&CN*, 11(12), 121-130.
- Li, S., Xu, W., & Li, Z. (2022). Review of the SBAS InSAR Time-series algorithms, applications, and challenges. *Geodesy and Geodynamics*, 13(2), 114-126. <https://doi.org/10.1016/j.geog.2021.09.007>
- Minderhoud, P. S. J., Erkens, G., Pham, V. H., Bui, V. T., Erban, L., Kooi, H., & Stouthamer, E. (2017). Impacts of 25 years of groundwater extraction on subsidence in the Mekong delta, Vietnam. *Environmental Research Letters*, 12(6), 064006. <https://doi.org/10.1088/1748-9326/aa7146>
- Tim hiểu Việt Nam (2019). *Giới thiệu tổng quan về Thành phố Hồ Chí Minh*. <https://www.timhieuvietnam.vn/gioi-thieu-tong-quan-ve-thanh-pho-ho-chi-minh>
- Wessel, P., & Smith, W. H. F. (2006). *Generic Mapping Tools Graphics - 4. June*, 180.
- Zhang, Z., Wang, C., Zhang, H., Tang, Y., & Liu, X. (2018). Analysis of Permafrost Region Coherence Variation in the Qinghai-Tibet Plateau with a High-Resolution TerraSAR-X Image. *Remote Sensing*, 10(2), 298. <https://doi.org/10.3390/rs10020298>

ABSTRACT

SBAS technology processes big data to determine terrain surface deformation in Ho Chi Minh City area

Nguyen Minh Hai^{1*}

¹*Hanoi University of Mining and Geology, Hanoi, Vietnam*

Ho Chi Minh City is the key economic center in the South and the country's leading, with an area of 0.6% and a population of 6.6% of the country, with the highest economic growth rate water. With the rapid development of infrastructure, weak geology has caused the terrain surface of the city area to be deformed. Land subsidence in the Ho Chi Minh City area leads to many potential risks and threatens the stability of the environment and human life. In this article, the author applied the Small BASeline Subset (SBAS) technique to monitor land subsidence in Ho Chi Minh City. Radar image datasets from 2006 to 2022 were used to generate average strain rates and strain time series. The determination results show that (1) the average annual deformation rate determined from ALOS 1 satellite images (2006-2011), the deformation area is mainly concentrated at the settlement value (-20)-0 mm/ year (598,493898 km²) and upwelling value 0-20 mm/year (1.263,241792 km²); (2) The average annual deformation rate determined from Sentinel satellite images (using images from 2014-2022 to determine deformation), the deformation area has changed a lot with settlement values <-40 (25,017193 km²), subsidence value (-40)-(-20) mm/year (158,605032 km²), subsidence value (-20)-0 (1.052,709457 km²) and heave value 0-20 mm/year (186, 858293 km²).

Keywords: Ho Chi Minh, SBAS, InSAR.

Nghiên cứu phân bố không gian gió vùng biển vịnh Bắc Bộ giai đoạn 1981-2020 từ dữ liệu viễn thám

Nguyễn Ngọc Tuấn¹, Nguyễn Hoàng Minh¹, Đỗ Thị Phương Thảo^{2*}

¹Viện Nghiên cứu Hải sản

²Trường Đại học Mở - Địa chất

TÓM TẮT

Vịnh Bắc Bộ (VBB) nằm trong đới khí hậu có mùa đông lạnh, mùa hè nóng trên nền nhiệt chung của khí hậu nóng ẩm, các đặc trưng của thời tiết bị chi phối bởi hoàn lưu khí quyển dạng gió mùa luân phiên nhau. Trong nghiên cứu này sử dụng dữ liệu đã thu thập về chế độ gió biển VBB giai đoạn 1981-2020 để phân tích, đánh giá chi tiết về hướng gió, tốc độ gió và phân bố gió biển. Kết quả nghiên cứu cho thấy từ tháng 11 đến tháng 3 năm sau có hướng gió chính là Đông Bắc (NE) và Đông (E), từ tháng 5 đến tháng 9 có hướng gió chính là Nam (S) và Tây Nam (SW), tháng 4 và tháng 10 là các tháng thường giao mùa chuyển tiếp giữa hai mùa gió nên có hướng gió không ổn định. Tốc độ gió trung bình năm dao động từ 4,3-6,4 m/s, trung bình tháng dao động từ 3,8-7,5 m/s, tuy tốc độ gió trung bình không cao nhưng tốc độ gió bão từ 25-50 m/s với 150/407 cơn bão, trung bình mỗi năm có 4 cơn bão đổ bộ. Tháng 1 và tháng 7 là hai tháng chính mùa đại diện cho hai mùa gió Đông Bắc và Tây Nam, mùa gió Đông Bắc có hướng gió chủ yếu là hướng Đông Bắc (NE) với tốc độ gió cao nhất là 7,5 m/s (khu vực giữa vịnh). Vào mùa gió Tây Nam có hướng gió Nam (S) với tốc độ gió cao nhất là 5,5 m/s. Cả hai mùa gió tốc độ gió đều có xu hướng giảm dần từ giữa vịnh vào ven bờ.

Từ khóa: Hướng gió, tốc độ gió, phân bố gió biển, vịnh Bắc Bộ.

1. Mở đầu

Chế độ gió vùng biển Việt Nam là tổng hợp các điều kiện gió mùa nhiệt đới Đông Nam Á, hệ thống các xoáy nghịch Đông Nam Á và là một bộ phận của gió mùa Ấn Độ Dương. Toàn vùng biển Việt Nam nói chung và vùng biển vịnh Bắc Bộ (VBB) nói riêng chịu ảnh hưởng của hai chế độ gió mùa là gió mùa Đông Bắc (từ tháng 11 đến tháng 3 năm sau) và gió mùa Tây Nam (từ tháng 5 đến tháng 9). Cho đến nay các nghiên cứu về chế độ gió khu vực vùng biển Việt Nam đã được tiến hành từ lâu qua các chương trình nhà nước, hợp tác quốc tế. Một số kết quả nghiên cứu tiêu biểu nêu ra nguyên nhân hình thành gió, quy luật gió mùa, đặc điểm biến đổi trường gió Biển Đông (Phạm Sĩ Hoàn, 2012; Từ Thị Năm và Trần Văn Sơn, 2018),... Để cập nhật và bổ sung thêm các thông tin khoa học về chế độ gió biển vùng VBB, trong nghiên cứu này tổng hợp dữ liệu gió biển từ dữ liệu viễn thám giai đoạn 1981-2020 để phân tích, đánh giá đặc điểm chế độ gió biển nhằm hỗ trợ cho ngư dân khai thác hải sản trên vùng biển, các nhà khoa học và các cơ quan quản lý có những thông tin khoa học về chế độ gió biển tại VBB để có những biện pháp phòng tránh và giảm nhẹ thiệt hại do thiên tai gió biển gây ra.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Dữ liệu

Bộ dữ liệu gió biển được thu thập, giải đoán, trích lọc từ dữ liệu viễn thám NASA với độ phân giải $0,5^\circ \times 0,5^\circ$ vùng VBB trong giai đoạn 1981-2020. Dữ liệu gió bão được thu thập từ trang web <http://weather24h.freevnn.com/> theo các khung giờ (0h, 6h, 12h, 18h). Dữ liệu bản đồ nền hành chính tỷ lệ 1:3.000.000 có ranh giới đường bờ biển cho vùng VBB với giới hạn khung trong bản đồ từ $16,5^\circ$ - $22,5^\circ$ N và 105° - 110° E, dữ liệu này phục vụ cho việc phân tích gió biển theo không gian.

Dữ liệu quan trắc gió ở 3/8 trạm tại các đảo Cô Tô, Bạch Long Vĩ và Cồn Cỏ của VBB với chuẩn độ cao tốc độ gió dùng cho nghiên cứu là 10 m.

Thời gian nghiên cứu: Từ năm 1981 đến năm 2020.

2.2. Phương pháp nghiên cứu

Phương pháp thống kê: Sử dụng phương pháp thống kê trong toán học để tính toán, xác định các giá trị đặc trưng cơ bản của chuỗi số liệu. Các giá trị đặc trưng này được xác định cho các chuỗi số liệu tháng, mùa, năm và nhiều năm.

Phương pháp đồ họa: Sử dụng phần mềm WRPLOT View được phát triển bởi Công ty Lakes Environmental

* Tác giả liên hệ:

Email: dothiphuongthao@humg.edu.vn

(Canada) để tính toán, thống kê và hiển thị tần suất xuất hiện của hướng gió (8 hướng hoặc 16 hướng) và tốc độ gió biến theo knot hoặc m/s. Ngoài ra phần mềm WRPLOT View còn phân tích dữ liệu gió biến dưới dạng biểu đồ tần suất hình cột từ các file số liệu đầu vào *.xlsx.

Phương pháp bản đồ và GIS: Các số liệu được tiến hành xây dựng biến trình theo thời gian và không gian các đặc trưng thống kê của chuỗi số liệu. Các kết quả được trình bày dưới dạng biểu đồ, đồ thị, bản đồ cho việc phân tích đánh giá kết quả. Để đảm bảo tính thống nhất của tất cả các bản đồ được thể hiện và tính khách quan, chính xác của các ranh giới phạm vi, địa danh, thang màu,... cần phải kết hợp với các công cụ, phần mềm của hệ thống tin địa lý.

3. Kết quả nghiên cứu

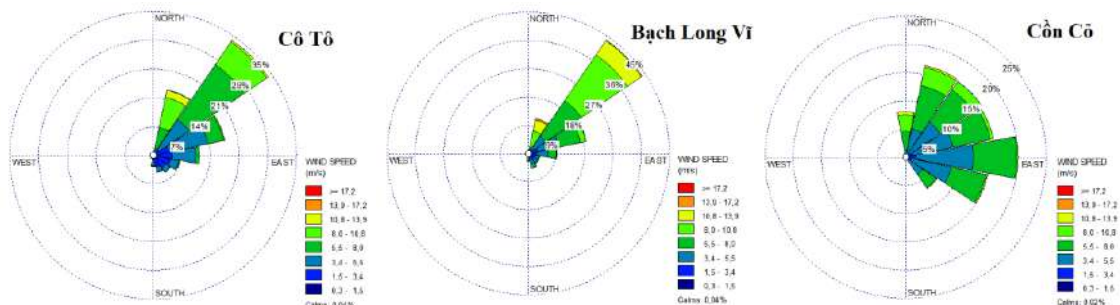
3.1. Hướng gió

Ở nước ta hướng gió thường được biểu thị bằng phương vị và góc, nhưng chủ yếu biểu thị bằng phương vị với 8 hướng chính là Đông, Tây, Nam, Bắc, Đông Bắc, Đông Nam, Tây Nam, Tây Bắc (Bảng 1). Ngoài các hướng chính còn có các hướng phụ được chia nhỏ như sau: Bắc Đông Bắc (NNE), Đông Đông Nam (ESE), Tây Tây Nam (WSW), Bắc Tây Bắc (NNW), Đông Đông Bắc (ENE), Nam Đông Nam (SSE), Nam Tây Nam (SSW) và Tây Tây Bắc (WNW).

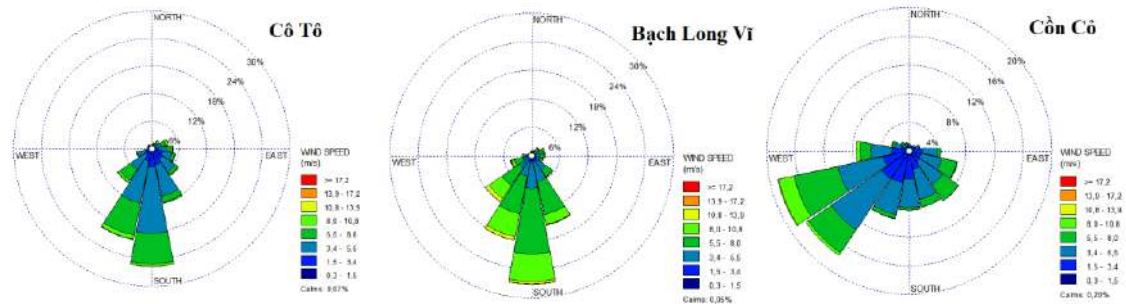
Bảng 3. Các hướng gió chính

Hướng gió	Tên Tiếng Anh	Kí hiệu	Giới hạn
Bắc	North	N	337,5°-22,5°
Đông Bắc	Northeast	NE	22,5°-67,5°
Đông	East	E	67,5°-112,5°
Đông Nam	Southeast	SE	112,5°-157,5°
Nam	South	S	157,5°-202,5
Tây Nam	Southwest	SW	202,5°-247,5°
Tây	West	W	247,5°-292,5°
Tây Bắc	Northwest	NW	292,5°-337,5°

Vào thời kỳ khoảng từ tháng 11 đến tháng 3 năm sau, khu vực VBB chịu một khối không khí lạnh hình thành từ trung tâm áp cao Xibia di chuyển xuống khu vực có khối không khí ẩm tại Việt Nam. Lúc này thời tiết miền Bắc nói chung và VBB nói riêng có tính chất lạnh và khô. Trong giai đoạn 1981-2020 vào các tháng này, tại một số vị trí VBB như Cô Tô và Bạch Long Vĩ có hướng gió thịnh hành là hướng Đông Bắc (NE) với tần suất xuất hiện từ 33,9-73,4%, còn tại Cồn Cỏ có các hướng gió thịnh hành là hướng Đông (E), Đông Nam (SE) và Đông Bắc (NE) với tần suất xuất hiện từ 35,5-46,1%. Thời gian này chủ yếu gió mùa Đông Bắc hoạt động, khi gió mùa Đông Bắc suy yếu cũng là lúc gió mùa Tây Nam bắt đầu hoạt động, vào tháng 4 và tháng 10 là các tháng giao mùa chuyển tiếp giữa hai mùa gió: Đông Bắc và Tây Nam. Khi đó tại Cô Tô và Bạch Long Vĩ có hướng gió thịnh hành là hướng Đông Bắc (NE) hoặc hướng Nam (S), tại Cồn Cỏ có hướng gió thịnh hành là Đông Nam (SE) hoặc Đông (S). Từ tháng 5 đến tháng 9, một khối khí được hình thành từ cao áp cận chí tuyến bán cầu Nam hoạt động mạnh, khi vượt qua vùng biển xích đạo khối khí này trở nên nóng ẩm, gây mưa lớn cho toàn Nam Bộ, Tây Nguyên và khiến cho Bắc Bộ khô, ít mưa. Lúc này tại Cô Tô và Bạch Long Vĩ có hướng gió thịnh hành chủ yếu là hướng Nam (S) với tần suất xuất hiện từ 27,9-60,7%, còn tại Cồn Cỏ có hướng gió thịnh hành chủ yếu là hướng Tây Nam (SW) với tần suất xuất hiện khoảng 27,8-33,7% (Hình 1, 2; Bảng 2).



Hình 1. Hoa gió tháng 1 tại một số vị trí VBB



Hình 2. Hoa gió tháng 7 tại một số vị trí VBB

Bảng 4. Hướng gió thịnh hành và tần suất xuất hiện tại một số vị trí VBB

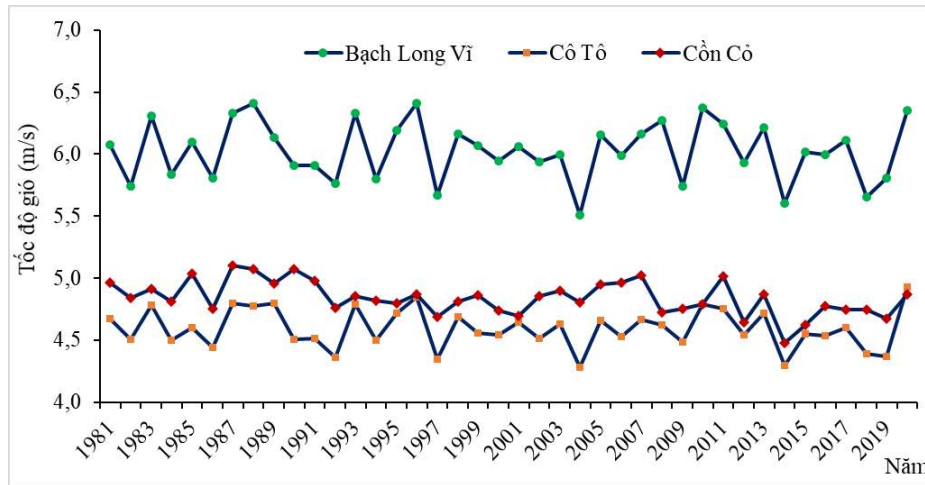
Khu vực Tháng	Cô Tô		Bạch Long Vĩ		Cồn Cỏ	
	Hướng chính	Tần suất (%)	Hướng chính	Tần suất (%)	Hướng chính	Tần suất (%)
1	NE	58,7	NE	64,9	E	36,4
2	NE	41,2	NE	48,2	E	35,5
3	NE	33,9	NE	39,9	SE	39,2
4	S	33,5	S	33,4	SE	44,2
5	S	44,2	S	60,7	SE	39,1
6	S	52,2	S	56,5	SW	33,7
7	S	45,4	S	46,7	SW	33,3
8	S	27,9	S	30,5	SW	27,8
9	NE	27,7	NE	32,0	E	21,8
10	NE	47,3	NE	52,8	E	41,9
11	NE	55,0	NE	63,5	NE	42,6
12	NE	67,1	NE	73,4	NE	46,1

Như vậy có thể nói, chế độ gió ở vùng VBB thể hiện hai mùa rõ rệt, mùa đông thịnh hành là hướng Đông Bắc (NE), mùa hè thịnh hành là hướng Tây Nam (SW). Nhưng tùy thuộc từng vị trí địa lý, địa hình, gió thịnh hành ngay trong cùng một mùa cũng có thể khác nhau về hướng. Ngoài sự thay đổi hướng gió qua các mùa theo chu kỳ năm còn có sự thay đổi theo chu kỳ ngày và đêm. Ban ngày ở lục địa hấp thụ nhiệt nhanh, nóng hơn mặt nước biển nên trên lục địa hình thành áp thấp; ở trên biển mát hơn hình thành áp cao, gió thổi từ áp cao tới áp thấp gọi là gió biển (gió biển ẩm và mát). Nhưng ban đêm, lục địa toả nhiệt nhanh hơn hình thành áp cao, trên biển toả nhiệt chậm nên hình thành áp thấp nên gió thổi từ đất liền thổi ra biển nên gọi là gió đất (gió đất khô và nóng). Gió đất, gió biển tuần hoàn quanh năm nhưng chỉ thể hiện rõ nét nhất khi hoàn lưu chung suy yếu.

3.2. Tốc độ gió

Trong 40 năm qua ở vùng biển VBB tốc độ gió trung bình năm dao động từ 4,3-6,4 m/s, tại Bạch Long Vĩ có tốc độ gió trung bình năm cao nhất dao động trong khoảng 5,5-6,4 m/s, cao hơn so với Cồn Cỏ và Cô Tô khoảng từ 1,0-1,5 m/s. Tốc độ gió trung bình tháng tại VBB dao động trong khoảng 3,8-7,5 m/s, khu vực Bạch Long Vĩ có vị trí địa lý gần như nằm giữa VBB nên ít ảnh hưởng bởi các yếu tố địa hình và đường bờ nên có tốc độ gió ổn định và cao nhất so với các vị trí khác. Tốc độ gió trung bình vào các tháng thời kỳ gió mùa đông lớn hơn so với thời kỳ gió mùa hè và lớn nhất vào tháng 11, tháng 12. Trong các tháng chính mùa gió Tây Nam tốc độ gió trung bình tại Bạch Long Vĩ dao động trong khoảng 6,5-7 m/s trong khi tại một số vị trí khác gần bờ như Cô Tô và Cồn Cỏ tốc độ gió trung bình chỉ khoảng 4,0-4,5 m/s, thấp hơn khoảng 2,5-3 m/s (Hình 3, Bảng 5). Thực tế gió mùa không phải hoạt động ổn định mà phát triển từng đợt, xen kẽ các đợt gió mùa hoặc thời kỳ giao mùa thì tốc độ gió thường rất nhỏ, tỷ lệ lặng gió cũng đạt tần suất cao.

Tốc độ gió trung bình tháng, trung bình năm chỉ phản ánh tính chất trung bình một cách tương đối nhất về sức gió tại VBB. Cho nên, nếu chỉ dừng lại ở tốc độ gió trung bình tháng, trung bình năm thì việc sử dụng để tính toán, đánh giá tiềm năng cũng như tác hại của gió biển gây ra sẽ gặp nhiều khó khăn. Theo số liệu 40 năm từ năm 1981-2020 tại VBB, tốc độ gió trung bình ngày mạnh nhất ở Bạch Long Vĩ vào khoảng 26,7 m/s vào ngày 10/11/2013 (tương đương gió bão cấp 10 theo thang sức gió Beaufort), tại Cồn Cỏ tốc độ gió mạnh nhất là 20,2 m/s vào ngày 15/09/2017 và 29/09/2009, tại Cô Tô tốc độ gió mạnh nhất tháng là 22,1 m/s vào các ngày 27/09/2005 (Bảng 4).



Hình 3. Tốc độ gió trung bình năm

Bảng 5. Tốc độ gió trung bình tháng

Vị trí Tháng	Cô Tô		Bạch Long Vĩ		Cồn Cỏ	
	V_{TB} (m/s)	Hướng chính	V_{TB} (m/s)	Hướng chính	V_{TB} (m/s)	Hướng chính
1	5,2	NE	6,7	NE, E	5,4	NE, E
2	5,0	NE	6,3	NE	5,1	NE, E
3	5,1	NE, S	6,3	S	5,2	NE, SE
4	4,4	NE, E	5,3	NE	4,6	NE, E, SE
5	4,3	S	5,6	S	4,3	E
6	4,5	S	6,5	SW, S	4,0	SE
7	4,7	S, SW	6,9	SE, SW	4,6	SW
8	3,8	S, SW	5,2	S	4,6	SW
9	4,5	NE, N	5,7	NE, N	4,7	N
10	5,3	NE, E	7,2	NE, E	5,7	NE
11	5,4	NE, E	7,4	NE, E	6,0	NE
12	5,5	NE, E	7,5	NE, E	5,9	NE

(Chú thích: V_{TB} : Vận tốc trung bình) Bảng 6. Tốc độ gió trung bình ngày mạnh nhất tháng và năm (Đơn vị: m/s)

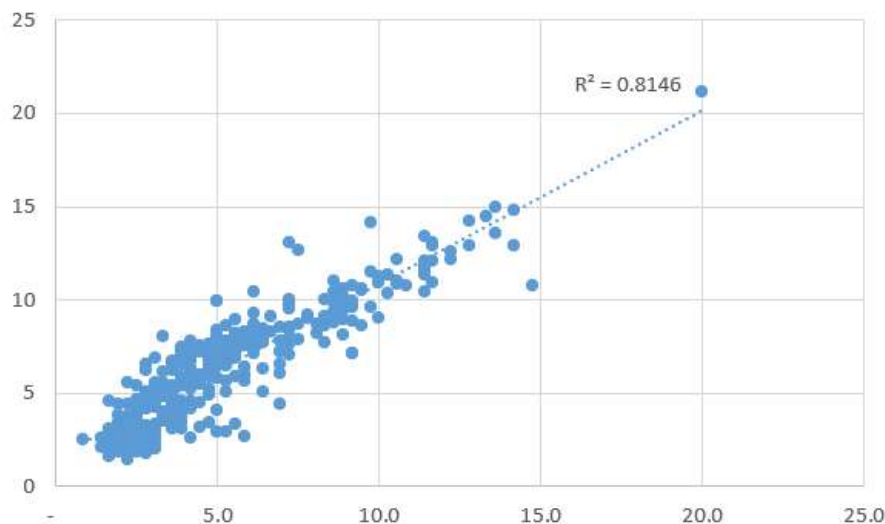
Vị trí Tháng	Bạch Long Vĩ	Cô Tô	Cồn Cỏ
1	16,7	14,8	12,9
2	17,0	15,3	13,7
3	17,2	15,3	15,5
4	17,7	15,8	12,9
5	15,2	13,1	19,1
6	24,6	18,3	16,2
7	23,8	18,4	14,9
8	23,2	18,4	18,7
9	23,4	22,1	20,2
10	22,5	17,3	18,0
11	26,7	21,1	15,9
12	18,1	15,9	15,3
Năm	26,7	22,1	20,2

3.3. Gió bão

Gió mạnh thường xảy ra trong cơn dông hoặc do ảnh hưởng của bão, áp thấp nhiệt đới, gió mùa Đông Bắc hoặc Tây Nam hoạt động mạnh, song nhìn chung tốc độ gió mạnh nhất chủ yếu xảy ra khi có các cơn bão mạnh từ Biển Đông hoặc biển Thái Bình Dương đổ bộ ảnh hưởng trực tiếp. Các nghiên cứu cho thấy vùng bờ biển Bắc Bộ bão hoạt động nhiều nhất về số lượng và mạnh nhất về cường độ. Cường độ gió mạnh do bão gây ra có thể lên đến cấp 10, giạt cấp 11, cấp 12 (Vũ Thanh Hằng và cộng sự, 2010; Nguyễn Văn Thắng và Mai Văn Khiên, 2016). Kết quả nghiên cứu của Phạm Thước trong thời gian 50 năm (1958-2009) (Phạm Thước, 2010) đã chỉ ra tốc độ gió bão vào khoảng 25 - 30 m/s, trong bão có thể quan trắc được gió giạt lên tới 50 m/s. Theo số liệu đã thống kê trong thời kỳ 1981-2020 đã có 150/407 cơn bão đổ bộ vào VBB với số ngày hoạt động là 239 ngày, trung bình mỗi năm tại VBB xuất hiện 4 cơn bão. Thời gian bão hoạt động ở vùng biển VBB kéo dài từ tháng 5 đến tháng 11, đặc biệt trong 3 tháng 8, 9, 10 có tần suất lớn. Diễn hình có một số cơn bão gió rất mạnh và mưa lớn trên diện rộng gây ra những thiệt hại nặng nề như làm đắm tàu thuyền đang hoạt động trên biển hoặc đang neo đậu gần bờ, đổ vỡ các công trình xây dựng, sụt lún vỡ đê ảnh hưởng đến các hoạt động đời sống cũng như kinh tế dân sinh như: bão Damrey (2005), bão Sơn Tinh (10/2012), bão Hải Yến (11/2013), bão Mirinae (7/2016), bão Doksuri (9/2017),... Những năm gần đây, các cơn bão mạnh có xu thế tăng nhẹ, vì vậy cần có kiến thức phòng tránh đảm bảo an toàn tối đa có thể khi có gió mạnh xảy ra.

3.4. Phân bố gió biển

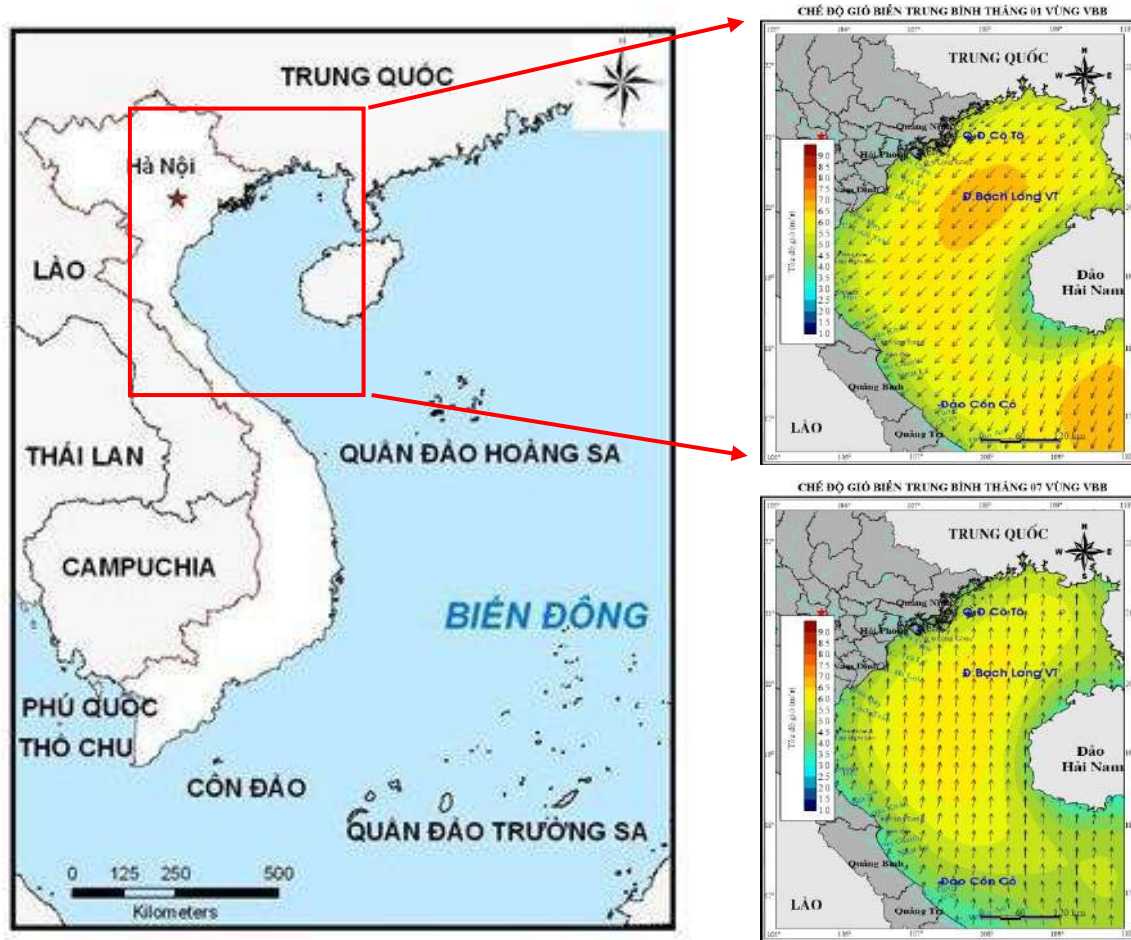
Để đánh giá tiềm năng năng lượng gió cần biết giá trị tốc độ gió tại độ cao nào đó mà số lượng trạm quan trắc gió trên cao. Tại Việt Nam số lượng trạm quan trắc này không nhiều, phần lớn được xác định gián tiếp bằng một hàm số từ tốc độ gió mặt đất ở các trạm khí tượng. Để đảm bảo kết quả nghiên cứu cần đánh giá chất lượng của số liệu xem kết quả tính toán có độ tin cậy không. Hình 4 sử dụng tương quan giữa số liệu thu được tại trạm quan trắc Bạch Long Vĩ và dữ liệu viễn thám theo toàn bộ ngày của năm 2020. Giá trị $R^2 = 0,8146$ cho thấy mối tương quan rất tốt giữa 2 loại dữ liệu thu được. Đa số dữ liệu vệ tinh của năm 2020 cao hơn dữ liệu được cung cấp tại trạm đo, điều này là do dữ liệu trên ảnh đã được tính toán thông qua các hàm với các tham số xấp xỉ nào đó, mặt khác giá trị trên ảnh vệ tinh thường là giá trị trung bình của các pixel ảnh. Để khắc phục tình trạng này cần nâng cao độ chính xác của thiết bị đo và nâng cao độ phân giải của dữ liệu vệ tinh.



Hình 4: So sánh tốc độ gió (m/s) từ dữ liệu trên ảnh vệ tinh và thu thập tại trạm quan trắc Bạch Long Vĩ năm 2020

Tháng 1 là tháng đại diện cho chính mùa gió Đông Bắc, tháng 7 là tháng đại diện cho chính mùa gió Tây Nam. Qua bản đồ phân bố gió biển giai đoạn 1981-2020 trong Hình 5 cho thấy, vào mùa gió Đông Bắc toàn VBB có hướng gió chủ yếu là hướng Đông Bắc (NE) và hướng Bắc Đông Bắc (NNE), tốc độ trung bình là 6-6,6 m/s. Khu vực Bạch Long Vĩ và phía cửa vịnh có tốc độ gió cao nhất (khoảng 7,5 m/s), xu hướng tốc độ gió giảm dần khi vào ven bờ. Tại đây đã có trạm điện gió cung cấp một phần năng lượng điện cho đảo.

Vào mùa gió Tây Nam toàn VBB có hướng gió chủ yếu là hướng Nam (S), tốc độ gió cao nhất ở giữa vịnh (khoảng 5,5 m/s) và cũng có xu hướng giảm dần vào ven bờ. Chênh lệch tốc độ gió giữa hai mùa khoảng 2 m/s. Tại đây, tiềm năng gió lớn nhất với tốc độ lần lượt là 6-11 m/s (tháng 1) và 6-10 m/s (tháng 10); tiềm năng gió thấp nhất trong các tháng 4, 7 với tốc độ gió lần lượt là 4-7 m/s và 5-7 m/s.



Hình 5. Phân bố gió biển theo mùa ở vùng biển VBB

4. Kết luận

Sử dụng dữ liệu viễn thám NASA có thể phân tích được phân bố và tốc độ gió theo không gian một cách nhanh chóng trên diện rộng, dữ liệu có độ tin cậy cao. Bản đồ phân bố gió có thể tham khảo trong khai thác tiềm năng năng lượng gió ngoài khơi.

Chế độ gió biển vùng VBB chủ yếu là gió mùa với hai hướng gió chính là Đông Bắc và Tây Nam. Từ tháng 11 đến tháng 3 năm sau có hướng thịnh hành là hướng Đông Bắc (NE) và hướng Đông (E) với tần suất xuất hiện từ 33,9-73,4%. Khu vực Bạch Long Vĩ có hướng gió ổn định và tốc độ gió mạnh hơn so với Côn Cỏ và Cô Tô tại cùng một thời điểm. Từ tháng 5 đến tháng 9 có hướng gió thịnh hành là hướng Tây Nam (SW) và hướng Nam (S) với tần suất xuất hiện từ 27,9-60,7%. Tháng 4 và tháng 10 là thời kỳ giao mùa chuyển tiếp giữa hai mùa gió nên hướng gió tại VBB lúc này không ổn định.

Tốc độ gió trung bình tháng dao động từ 3,8-7,5 m/s, tốc độ gió trung bình năm dao động từ 4,3-6,4 m/s. Tốc độ gió trung bình năm tại Bạch Long Vĩ cao hơn tốc độ gió trung bình năm tại Cô Tô và Côn Cỏ khoảng từ 2,5-3,0 m/s. Tốc độ gió mạnh nhất thường xảy ra vào khoảng tháng 9 hoặc tháng 11 với tốc độ gió khoảng từ 22-27 m/s. Gió mạnh thường xảy ra trong cơn dông hoặc do ảnh hưởng của bão, áp thấp nhiệt đới từ Biển Đông hoặc Tây Thái Bình Dương đổ bộ vào.

Phân bố gió vào mùa gió Đông Bắc toàn VBB có hướng gió thịnh hành là hướng Đông Bắc (NE) với tốc độ gió cao nhất là 7,5 m/s (khu vực giữa vịnh). Vào mùa gió Tây Nam VBB có hướng gió Nam (S) với tốc độ gió cao nhất là 5,5 m/s. Cả hai mùa gió tốc độ gió đều có xu hướng giảm dần từ ngoài khơi vào ven bờ.

Kết quả nghiên cứu cho thấy khu vực biển vịnh Bắc Bộ có tiềm năng năng lượng gió nhưng gần như ít được khai thác (mới có trạm gió tại Bạch Long Vĩ). Cần phát triển thêm các mô hình khai thác tiềm năng năng lượng gió tại các huyện đảo để đáp ứng phần nào nhu cầu sử dụng điện gió tại chỗ cho dân cư sinh sống ngoài đảo.

Tài liệu tham khảo

- Nguyễn Văn Thắng, Mai Văn Khiên (2016). Ảnh hưởng của bão ở Việt Nam thời kỳ 1961-2014. Tạp chí Khoa học ĐHQGHN, Các Khoa học Trái đất và Môi trường, tập 32, số 3S(2016), tr. 210-216.
- Phạm Sĩ Hoàn (2012). Sự biến động của trường gió tại vùng biển ven bờ Việt Nam. Kỷ yếu Hội nghị Quốc tế “Biển Đông 2012,” tháng 9/2012, tr. 45-55.
- Phạm Thược (2010). Nghề cá vịnh Bắc Bộ qua những chặng điều tra nghiên cứu (1958-2009). Nhà xuất bản Nông nghiệp, Hà Nội.
- Từ Thị Năm và Trần Văn Sơn (2018). Nghiên cứu chế độ gió khu vực Nam Trung Bộ. Tạp chí Khí tượng thủy văn, số tháng 7/2018, tr. 50-59.
- Vũ Thanh Hằng, Ngô Thị Thanh Hương và Phan Văn Tân (2010). Đặc điểm hoạt động của bão ở vùng biển gần bờ Việt Nam giai đoạn 1945 - 2007. Tạp chí Khoa học Đại học Quốc gia Hà Nội, Khoa học tự nhiên và Công nghệ số 26, số 3S(2010), tr. 344-353.

ABSTRACT

Using satellite data to study distribution of wind in the Gulf of Tonkin for the period 1981 - 2020

Nguyen Ngoc Tuan¹, Nguyen Hoang Minh¹, Do Thi Phuong Thao^{2*}

¹ *Research Institute for Marine Fisheries*

² *Hanoi University of Mining and Geology*

The Gulf of Tonkin (VBB) is located in a climate zone with cold winters and hot summers on the general heat of a hot and humid climate, the weather characteristics are dominated by alternating monsoonal atmospheric circulation. In this study, data collected on the VBB sea wind regime in the period 1981-2020 is used to analyze and evaluate in detail the wind direction, wind speed and distribution of sea breeze. Research results show that from November to March, the main wind directions are Northeast (NE) and East (E), from May to September, the main wind directions are South (S) and Southwest (SW). April and October are the transitional months between the two wind seasons, so the wind direction is not stable. The annual average wind speed ranges from 4.3 to 6.4 m/s, the monthly average ranges from 3.8 to 7.5 m/s, although the average wind speed is not high, the storm wind speed is from 25 to 50 m/s with 150/407 storms, an average of 4 storms per year. January and July are the two main months representing the two northeast and southwest monsoon seasons, the northeast monsoon season has the predominant wind direction from the Northeast (NE) with the highest wind speed of 7.5 m/s (the area in the middle of the bay). In the southwest monsoon season, there is a Southerly wind (S) with the highest wind speed of 5.5 m/s. In both monsoon seasons, wind speed tends to decrease gradually from the middle of the bay to the coast.

Keywords: Wind direction, wind speed, sea breeze space, gulf of Tonkin.

Nghiên cứu ứng dụng tư liệu ảnh Sentinel-5P trong giám sát chất lượng không khí tại các đô thị lớn

Lê Minh Hằng^{1,2}, Đỗ Thị Hoài^{3*}

¹Học viện Kỹ thuật quân sự

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

³Viện Khoa học đo đạc và Bản đồ

TÓM TẮT

Ô nhiễm không khí hiện nay ngày càng trở nên nghiêm trọng, đặc biệt tại các thành phố lớn của Việt Nam. Mật độ dân cư tăng nhanh, sự gia tăng của các phương tiện cá nhân và phát triển các ngành công nghiệp đã làm cho môi trường bị ô nhiễm. Tháng 10 năm 2017, Cơ quan Hàng không vũ trụ châu Âu (European Space Agency-ESA) đã phóng thành công vệ tinh Sentinel-5P (Sentinel-5 Precursor) với bộ cảm biến TROPOspheric (TROPOMI). Vệ tinh Sentinel-5P có mục tiêu chính là thực hiện theo dõi chất lượng không khí, bức xạ ozon và tia cực tím thông qua các phép đo khí quyển với độ phân giải thời gian cao. Trong nội dung bài báo, các tác giả trình bày các kết quả nghiên cứu về khả năng sử dụng tư liệu Sentinel-5P và nền tảng Google Earth Engine (GEE) trong giám sát chất lượng không khí tại các đô thị lớn tại Việt Nam. Khu vực nghiên cứu bao gồm toàn bộ vùng đất liền Việt Nam và một số thành phố lớn gồm Hà Nội, Hải Phòng, Vinh, Đà Nẵng, Nha Trang, Thành phố Hồ Chí Minh và Cần Thơ. Thời gian theo dõi chất lượng không khí là từ tháng 01/2022 đến tháng 12/2023. Kết quả nghiên cứu thể hiện phân bố không gian của nồng độ trung bình trong một năm của các chất khí SO₂, NO₂, CO trên vùng đất liền Việt Nam năm 2022 và năm 2023. Kết quả nghiên cứu cho thấy nồng độ của chất khí NO₂ và CO năm 2023 tăng hơn so với năm 2022 tại một số khu vực của Việt Nam. Bên cạnh đó, chất lượng không khí của các thành phố lớn sẽ được đánh giá dựa trên sự thay đổi của nồng độ trung bình các chất khí gây ô nhiễm SO₂, NO₂, CO trong từng tháng. Kết quả nghiên cứu của bài báo góp phần nâng cao khả năng ứng dụng tư liệu viễn thám phục vụ giám sát chất lượng không khí trên diện rộng tại Việt Nam.

Từ khóa: Sentinel-5P, ô nhiễm không khí, đô thị lớn, Google Earth Engine.

1. Đặt vấn đề

Chất lượng không khí đang là vấn đề được quan tâm tại nhiều quốc gia trên thế giới, đặc biệt là tại Việt Nam. Với sự phát triển mạnh mẽ của các ngành công nghiệp, sự gia tăng dân số và tốc độ đô thị hóa đã làm thay đổi chất lượng môi trường không khí tại các thành phố lớn và khu công nghiệp (Bộ TN&MT, 2022). Theo báo cáo của tổ chức Y tế Thế giới (WHO) năm 2022 thì một trong những mối nguy hiểm lớn nhất đối với sức khỏe toàn cầu là ô nhiễm không khí (WHO, 2022). Chất dạng hạt (PM), lưu huỳnh dioxide (SO₂), nitrogen dioxide (NO₂), ozone (O₃), carbon monoxide (CO) và carbon dioxide (CO₂) là những chất gây ô nhiễm không khí phổ biến nhất mà con người gặp phải hàng ngày (Chen và nnk., 2007). Tại các khu vực đô thị lớn, nồng độ các chất khí gây ô nhiễm không khí đã thay đổi bởi nhiều yếu tố khác nhau phụ thuộc vào đặc điểm của bề mặt đất, kiểu thời tiết, các loại chất ô nhiễm không khí và nguồn ô nhiễm. Các yếu tố tự nhiên như áp suất không khí, nhiệt độ, hướng gió và tốc độ gió ảnh hưởng đến những đặc điểm này (Hu và Yoshie, 2020). Tuy nhiên, những tác nhân do con người gây ra như đốt than sinh hoạt, khí thải xe cộ, khí thải công nghiệp là những nguyên nhân chủ yếu ảnh hưởng đến chất lượng không khí tại các khu đô thị (Liu và nnk., 2018).

Việc giám sát liên tục và chính xác chất lượng không khí là cần thiết để xây dựng mô hình dự báo và ra quyết định nhằm giảm thiểu tác động của ô nhiễm không khí. Hiện nay, việc tiến hành giám sát ô nhiễm mặt đất trên phạm vi rộng thường sử dụng các trạm cố định hoặc cảm biến di động. Tuy nhiên, các trạm giám sát cố định và cảm biến trên mặt đất có chi phí bảo trì cao và thu thập điểm dữ liệu bị giới hạn ở các khu vực nhỏ hoặc chỉ vài mét xung quanh cảm biến, dẫn đến dữ liệu không liên tục để cung cấp kết quả nội suy trên diện rộng (Amanollahi và nnk., 2013). Với sự phát triển của công nghệ viễn thám, các phương pháp viễn thám đang được nghiên cứu và ứng dụng trong quan sát Trái đất như giám sát nhiệt độ bề mặt, thảm thực vật, lớp phủ và chất lượng không khí. Kết quả quan sát bằng tư liệu viễn thám giúp cung cấp, bổ sung thông tin cho các khu vực thiếu các trạm giám sát cố định và thiếu các dữ liệu lịch sử (Nagy và nnk., 2022).

Nhiều nghiên cứu đã sử dụng thiết bị đo bức xạ quang phổ trên ảnh MODIS để thu được độ sâu quang học của khí dung (AOD) và xác định được nồng độ PM_{2.5} trên mặt đất (Zhang và nnk., 2021). Vào tháng 10 năm

* Tác giả liên hệ

Email: dthoai_dc@monre.gov.vn

2017, Cơ quan Hàng không vũ trụ châu Âu (ESA) đã phóng thành công vệ tinh Sentinel-5P mang theo bộ cảm TROPOspheric Monitoring Instrument (TROPOMI) phục vụ giám sát khí quyển, chất lượng không khí và trạng thái của tầng ozone. Nhiều nghiên cứu trên thế giới đang ứng dụng tư liệu viễn thám Sentinel-5P trong giám sát chất lượng không khí trên diện rộng, đặc biệt tại các thành phố lớn (Hassaan và nnk., 2024; Jeong và nnk., 2021; Sivaramakrishnan và nnk., 2021; Morozova và nnk., 2021) và các khu công nghiệp (Dimitrova và nnk., 2023). Các nghiên cứu đã xác định sự biến thiên của nồng độ các chất khí gây ô nhiễm như SO₂, NO₂, CO trên dữ liệu Sentinel-5P đa thời gian (Dimitrova và nnk., 2023; Davybida và nnk., 2023). Chất lượng không khí tại các thành phố lớn phụ thuộc vào nhiều yếu tố như tốc độ đô thị hóa và sự gia tăng của phương tiện giao thông. Nhiều nghiên cứu đã chứng minh rằng trong thời gian giãn cách xã hội do đại dịch Covid-19 thì các khí gây ô nhiễm như NO₂, O₃, CH₄ giảm đáng kể (Mejia và nnk., 2024; Phan và Fukui, 2023; Virghileamu và nnk., 2020). Trong thời gian giãn cách, việc hạn chế các phương tiện giao thông đã giảm đáng kể nồng độ các chất khí thải như NO₂ ra môi trường (Phan và Fukui, 2023, Oxoli và nnk., 2020). Ngoài ra, ô nhiễm không khí còn có mối liên hệ với đặc điểm lớp phủ bề mặt (Kaplan và Avdan, 2020; Hassaan và nnk., 2024; Nagy và nnk., 2024), có ảnh hưởng bởi các hoạt động quân sự (Rehmat và nnk., 2023; Wiecezorek, 2023). Tác giả Wang và cộng sự (2022) (Wang và nnk., 2022) đã chứng minh vai trò của chất lượng không khí trong đánh giá chất lượng môi trường sinh thái đô thị tại thành phố Bắc Kinh, Trung Quốc. Kết quả nghiên cứu đã sử dụng tư liệu Landsat-8, Sentinel-2 và Sentine-5P để xác định các chỉ số tác động đến sinh thái đô thị như NDVI, NDSBI, ISI, LST và chỉ số chất lượng không khí (AQ) (Wang và nnk., 2022). Tác giả Wang và cộng sự (2022) đã so sánh chỉ số RSEI (chỉ số đánh giá chất lượng sinh thái đô thị) và chỉ số AQRSEI (chỉ số đánh giá chất lượng sinh thái đô thị có thêm chỉ số đánh giá chất lượng không khí). Kết quả nghiên cứu đã chỉ ra rằng chỉ số AQRSEI đánh giá chất lượng sinh thái đô thị gần với điều kiện thực tế hơn chỉ số RSEI (Wang và nnk., 2022).

Tại Việt Nam, các nghiên cứu về chất lượng không khí chủ yếu dựa trên các kết quả đo trực tiếp tại các trạm quan trắc cố định (Đỗ Thị Phương Thảo và nnk., 2022; Dương Hồng Sơn và nnk., 2006). Một số nghiên cứu tại Việt Nam đã thử nghiệm sử dụng tư liệu Sentinel-5P trong đánh giá chất lượng không khí tại một số khu vực (Luu Thị Diệu Chinh và nnk., 2023). Theo báo cáo hiện trạng ô nhiễm quốc gia 2021 thì chất lượng không khí tại các thành phố lớn và các khu công nghiệp được ghi nhận là ô nhiễm tại một số thời điểm trong năm, đặc biệt là ô nhiễm bụi (Bộ TN&MT, 2022). Các tác động của điều kiện khí hậu làm thay đổi mức độ ô nhiễm không khí trong một năm. Đối với khu vực miền Bắc, ô nhiễm bụi tập trung vào các tháng mùa đông, mưa ít (Bộ TN&MT, 2022). Đối với khu vực miền Nam thì mức độ ô nhiễm giảm rõ rệt vào các tháng mùa mưa và cao hơn vào mùa khô (Bộ TN&MT, 2022). Các thông số về chất lượng không khí được ghi nhận tại các trạm quan trắc cố định liên tục trên cả nước, thông qua nồng độ một số chất khí gây ô nhiễm như SO₂, CO, NO_x, O₃, bụi, TPS (tổng lượng bụi lơ lửng), bụi PM₁₀, bụi PM_{2.5} và Pb. Trong giai đoạn 2016-2021, không khí tại các thành phố lớn như Hà Nội, Thành phố Hồ Chí Minh và một số khu công nghiệp lớn như tỉnh Bắc Ninh có mức độ ô nhiễm lớn tại một số thời điểm trong năm (Bộ TN&MT, 2022). Nguyên nhân gây ra ô nhiễm không khí tại Việt Nam do nhiều nguyên nhân khác nhau như điều kiện khí hậu, thời tiết trong năm, thay đổi diện tích rừng và cây xanh đô thị, phát triển các ngành công nghiệp, xây dựng và dịch vụ, tốc độ đô thị hóa, hoạt động giao thông vận tải, hoạt động sản xuất nông nghiệp (Bộ TN&MT, 2022).

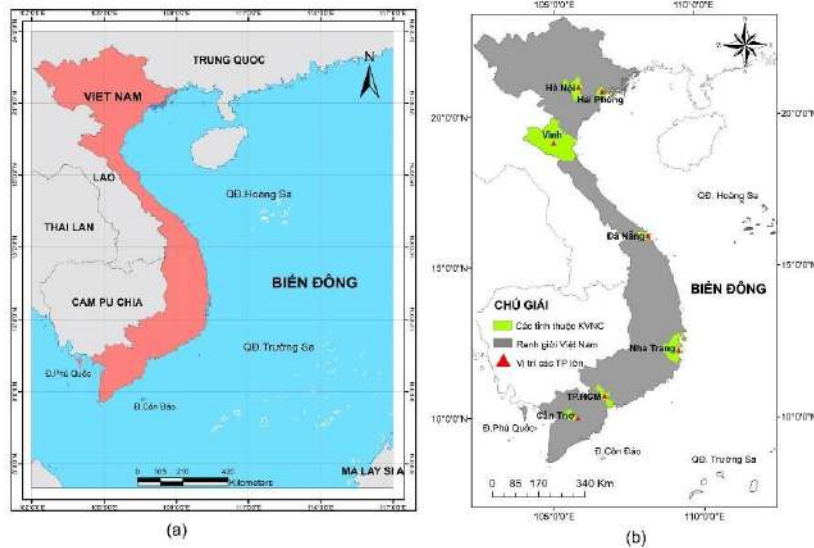
Như vậy, nhiều nghiên cứu tại Việt Nam và trên thế giới đã và đang nghiên cứu sử dụng tư liệu Sentinel-5P đa thời gian để đánh giá chất lượng không khí tại các khu vực khác nhau. Kết quả nghiên cứu cho thấy khả năng ứng dụng và ưu điểm của tư liệu Sentinel-5P trong giám sát chất lượng không khí. Trong nội dung bài báo, các tác giả trình bày kết quả nghiên cứu tư liệu viễn thám Sentinel-5P đa thời gian và nền tảng Google Earth Engine trong giám sát chất lượng không khí trên khu vực đất liền Việt Nam và tại các thành phố lớn trong giai đoạn từ tháng 01/2022 đến tháng 12/2023.

2. Vị trí nghiên cứu và dữ liệu thử nghiệm

2.1. Vị trí nghiên cứu

Khu vực nghiên cứu được thể hiện trên Hình 1a. Khu vực nghiên cứu là khu vực đất liền của Việt Nam. Diện tích của Việt Nam là 331.212 km², trong đó diện tích đất liền chiếm 93,62%. Đường bờ biển 3.444 km. Biên giới Việt Nam trên đất liền dài 4.639 km, giáp với vịnh Thái Lan ở phía Tây Nam, vịnh Bắc Bộ và Biển Đông ở phía Đông, Trung Quốc ở phía Bắc, Lào và Campuchia ở phía Tây.

Việt Nam thuộc khu vực có đặc điểm khí hậu nhiệt đới. Địa hình của Việt Nam phần lớn là đồi núi (chiếm 3/4 diện tích lãnh thổ), khu vực đồi núi thấp, khu vực đồng bằng chiếm 1/4 diện tích. Dựa trên các đặc điểm địa hình, khí hậu, Việt Nam được chia thành các vùng khác nhau như Đồng bằng Sông Hồng, Trung du và miền núi phía Bắc, Bắc Trung Bộ và Duyên hải, Tây Nguyên, Đông Nam Bộ và Đồng bằng sông Cửu Long (Tổng cục Thống kê, 2022). Trong bài báo, các tác giả đánh giá chất lượng không khí tại một số thành phố lớn như thành phố Hải Phòng, Hà Nội, Vinh, Đà Nẵng, Nha Trang, Thành phố Hồ Chí Minh và Cần Thơ (Hình 1b). Các thành phố thuộc khu vực nghiên cứu là khu vực phát triển kinh tế của vùng (Tổng cục Thống kê, 2022).



Hình 1. Vị trí khu vực nghiên cứu. (a) Vị trí phần đất liền Việt Nam; (b) Vị trí một số thành phố lớn thử nghiệm đánh giá chất lượng không khí

2.2. Dữ liệu thử nghiệm

2.2.1. Vệ tinh Sentinel-5P

Vệ tinh Sentinel-5 Precursor hay còn được gọi là vệ tinh Sentinel-5P, là một trong những vệ tinh thuộc nhiệm vụ Copernicus của Cơ quan Hàng không vũ trụ châu Âu (ESA - European Space Agency). Vệ tinh Sentinel-5P được phóng thành công vào ngày 13 tháng 10 năm 2017 và có nhiệm vụ chuyên theo dõi bầu khí quyển (Babic và nnk., 2022). Vệ tinh mang theo thiết bị cảm biến TROPOMI để lập bản đồ của nhiều loại khí vi lượng như nitơ dioxide (NO_2), ozone (O_3), formaldehyde (HCHO), sulfur dioxide (SO_2), metan (CH_4), carbon monoxide (CO) và sol khí. Các loại khí này đều có ảnh hưởng đến sức khỏe và khí hậu của chúng ta. Với chiều rộng 2600 km, vệ tinh Sentinel-5P lập bản đồ toàn bộ Trái đất mỗi ngày. Độ phân giải không gian của khí SO_2 là $5,5 \times 3,5 \text{ km}^2$ (Theys và nnk., 2022), khí CO là $7,0 \times 5,5 \text{ km}^2$ (Landgraf và nnk., 2022) và khí NO_2 là $7 \times 3,5 \text{ km}^2$ (Geffen và nnk., 2022). Ở độ phân giải này, ô nhiễm không khí trên các thành phố có thể được phát hiện.

Vệ tinh Sentinel-5P còn được ứng dụng trong giám sát tro núi lửa vì an toàn hàng không, các dịch vụ cảnh báo mức độ bức xạ tia cực tím cao có thể gây tổn thương da. Ngoài ra, các nhà khoa học cũng sử dụng dữ liệu Sentinel-5P để nghiên cứu các quá trình quan trọng trong khí quyển liên quan đến khí hậu và sự hình thành các lỗ hổng trên tầng ozone (Langen và nnk., 2017). Vệ tinh Sentinel-5P hiện đang đóng vai trò quan trọng trong việc giám sát và theo dõi tình trạng ô nhiễm không khí.

Trong nội dung bài báo, các tác giả đã sử dụng ảnh vệ tinh Sentinel-5P bao gồm ảnh nồng độ khí NO_2 , SO_2 và CO trên vùng đất liền của Việt Nam trong khoảng thời gian từ tháng 01 đến tháng 12 năm 2022 và năm 2023.

2.2.2. Nền tảng Google Earth Engine

Các dữ liệu Sentinel-5P sẽ được tiến hành khai thác và tiền xử lý ảnh từ nền tảng Google Earth Engine. Google Earth Engine (GEE) là một nền tảng để phân tích khoa học và trực quan hóa các bộ dữ liệu không gian địa lý, dành cho người dùng học thuật, phi lợi nhuận, doanh nghiệp và Chính phủ. GEE lưu trữ hình ảnh vệ tinh và lưu trữ nó trong kho lưu trữ dữ liệu công cộng bao gồm các hình ảnh lịch sử về Trái đất cách đây hơn bốn mươi năm. Các hình ảnh được ghi lại hàng ngày sau đó sẽ được cung cấp để khai thác dữ liệu trên quy mô toàn cầu. GEE cũng cung cấp API và các công cụ khác để cho phép phân tích các tập dữ liệu lớn (<https://earthengine.google.com/platform/>).

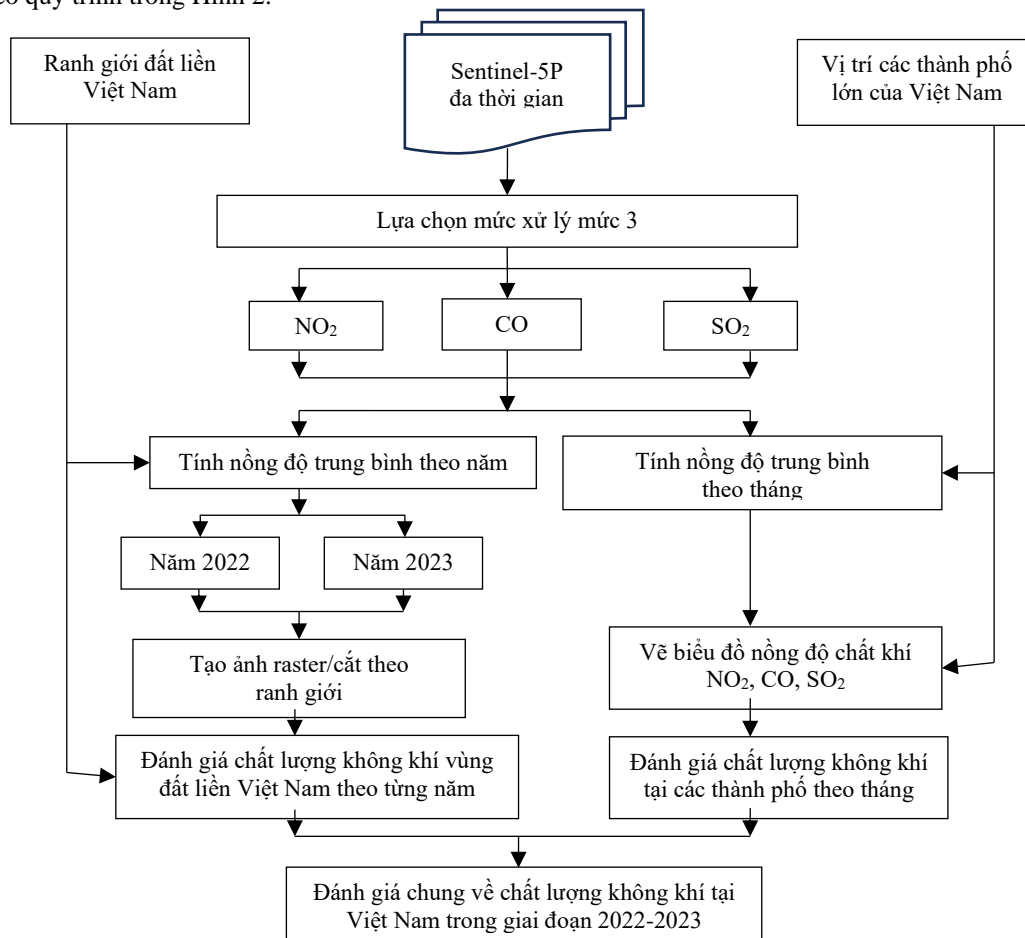
2.2.3. Quy chuẩn kỹ thuật quốc gia về chất lượng không khí

Vào ngày 13/3/2023, Bộ Tài nguyên và Môi trường ban hành Thông tư số 01/2023/TT-BTNMT. Trong đó, ban hành kèm Thông tư là quy chuẩn kỹ thuật quốc gia về chất lượng môi trường xung quanh. Chất lượng không khí được đánh giá theo quy chuẩn QCVN 05:2023/BTNMT “Quy chuẩn kỹ thuật quốc gia về chất lượng không khí” (Bộ KH&CN, 2023). Quy chuẩn QCVN 05:2023/BTNMT quy định những thông số đánh giá chất lượng không khí gồm các khí SO_2 (lưu huỳnh (sulfur) dioxide), CO (carbon monoxide), NO_2 (nitơ (nitrogen) dioxide),

O₃ (ozone), TSP (tổng bụi lơ lửng), bụi PM₁₀, bụi PM_{2.5}. Do đó, trong nội dung bài báo, các tác giả sử dụng thông số nồng độ của các chất khí gồm SO₂, CO và NO₂ để đánh giá chất lượng không khí của khu vực nghiên cứu.

3. Phương pháp nghiên cứu

Quy trình xác định chất lượng không khí trên khu vực nghiên cứu sử dụng ảnh vệ tinh Sentinel-5P được thực hiện theo quy trình trong Hình 2.



Hình 2. Quy trình đánh giá chất lượng không khí sử dụng ảnh Sentinel-5P trên nền tảng GEE

Quy trình đánh giá chất lượng không khí sử dụng ảnh Sentinel-5P được thực hiện trên nền tảng GEE và bao gồm các giai đoạn chính như sau:

- Bước 1: Thu thập dữ liệu ảnh Sentinel-5P đa thời gian trong phạm vi khu vực nghiên cứu. Dữ liệu đánh giá chất lượng không khí trong nội dung bài báo gồm nồng độ các chất khí SO₂, NO₂, CO từ tư liệu ảnh Sentinel-5P đa thời gian với mức xử lý L3.

- Bước 2: Đánh giá chất lượng không khí vùng đất liền Việt Nam trong năm 2022 và năm 2023. Các dữ liệu nồng độ các chất khí SO₂, NO₂, CO sẽ được tính trung bình trong thời gian một năm của năm 2022 và năm 2023. Kết quả được lưu dưới dạng ảnh raster và trong hệ tọa độ WGS84. Để tiến hành so sánh sự thay đổi về phân bố các chất khí trên vùng đất liền Việt Nam, các tác giả tính chuyển giá trị thang nồng độ năm 2022 về cùng thang nồng độ năm 2023 dựa trên giá trị lớn nhất và giá trị nhỏ nhất.

- Bước 3: Đánh giá chất lượng không khí tại vị trí các thành phố lớn (Hình 1b) theo từng tháng (01/2022 đến 12/2023). Các dữ liệu nồng độ chất khí SO₂, NO₂, CO sẽ được tính trung bình cho từng tháng. Kết quả nồng độ được biểu thị trên biểu đồ diễn biến của nồng độ chất khí tại các khu vực thành phố lớn.

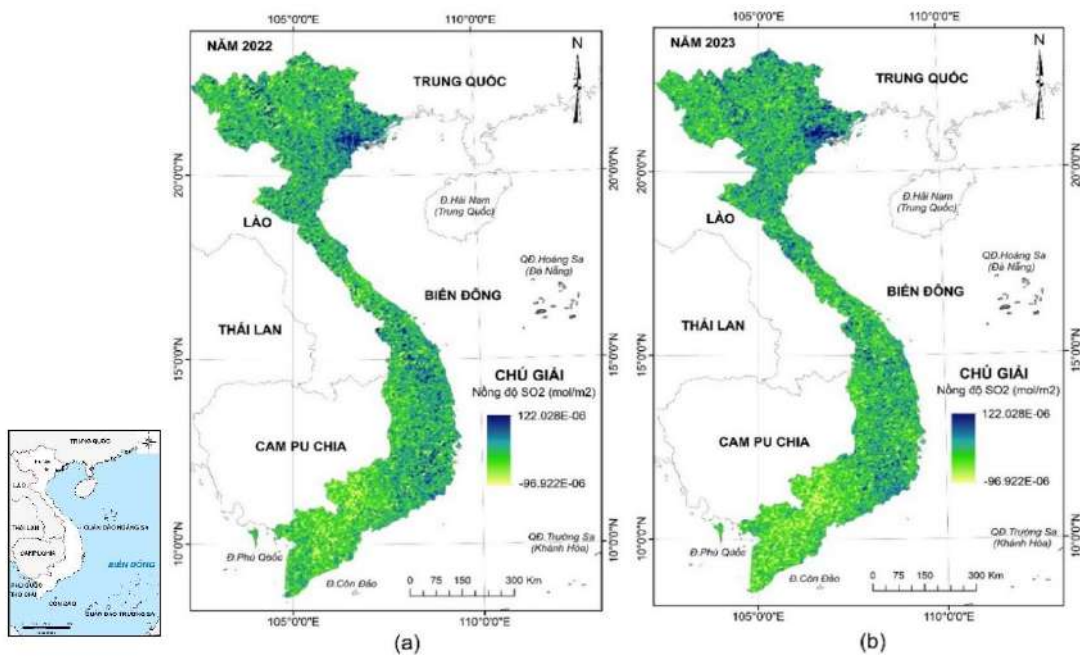
- Bước 3. Đánh giá chung về chất lượng không khí tại Việt Nam và tại các thành phố lớn trong giai đoạn 2022-2023.

4. Kết quả và thảo luận

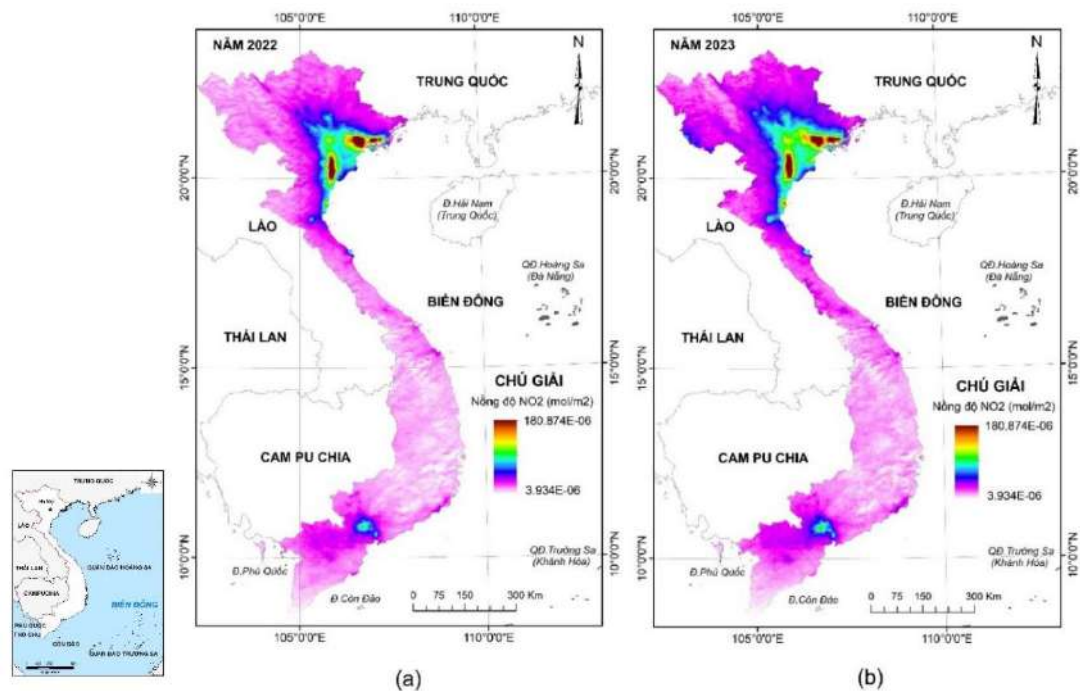
4.1. Đánh giá chất lượng không khí tại vùng đất liền Việt Nam

Dựa trên quy trình được thể hiện trong Hình 2, các tác giả xác định nồng độ trung bình trong năm 2022 và năm 2023 của các chất khí SO_2 , NO_2 , CO từ tư liệu ảnh Sentinel-5P trên phần đất liền của Việt Nam và được thể hiện trên Hình 3, 4 và 5.

Hình 3 thể hiện nồng độ trung bình năm 2022 và năm 2023 của chất khí SO_2 với đơn vị là (mol/m^2). Kết quả Hình 3a và Hình 3b cho thấy nồng độ trung bình năm 2022 và 2023 trên phần đất liền của Việt Nam đã có sự thay đổi về giá trị và vị trí phân bố. Nồng độ khí SO_2 cao nhất tập trung tại các tỉnh phía Bắc gồm Bắc Giang, Bắc Ninh, Hải Dương và một phần của tỉnh Quảng Ninh. Khu vực miền Nam là các tỉnh Ninh Thuận, Bình Thuận, Lâm Đồng. Vị trí các điểm có nồng độ SO_2 cao năm 2022 tập trung tại khu vực Nam Trung Bộ. Xét trong năm 2023 thì vị trí có mức nồng độ cao giảm trên toàn vùng đất liền của Việt Nam. Khu vực có nồng độ SO_2 tăng lên tại một số điểm phía Tây tỉnh Hà Tĩnh trong năm 2023.



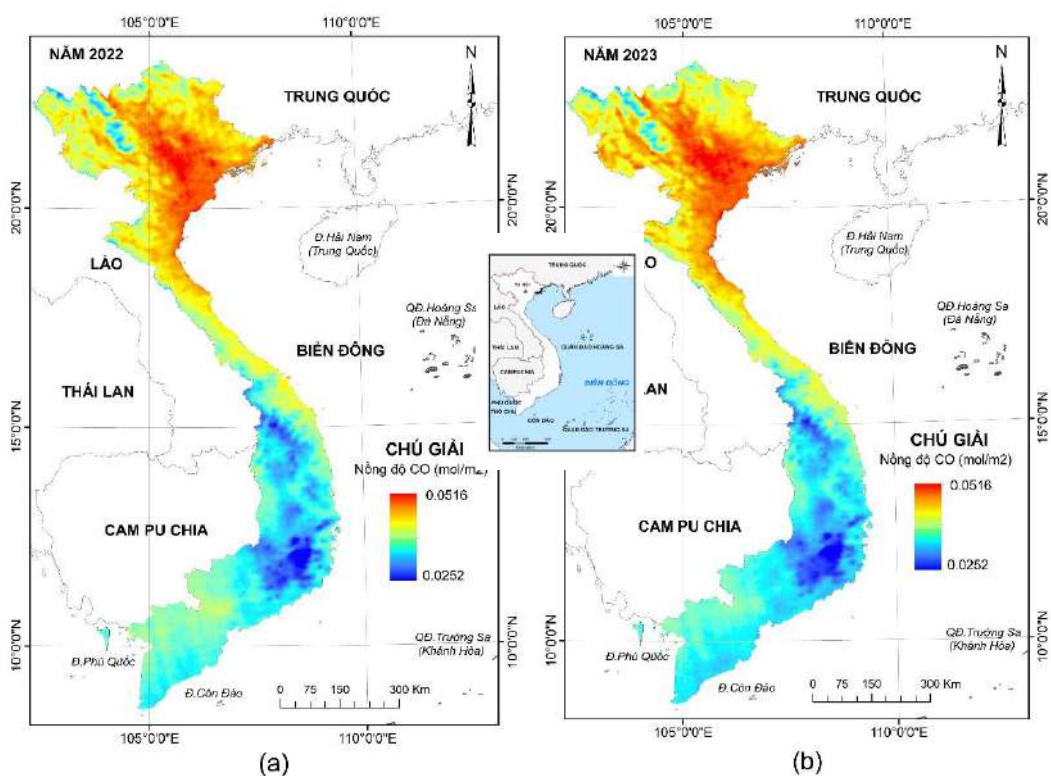
Hình 3. Nồng độ khí SO_2 năm 2022 và năm 2023



Hình 4. Nồng độ khí NO_2 năm 2022 và năm 2023

Nồng độ trung bình năm 2022 và năm 2023 của chất khí NO_2 được thể hiện trong Hình 4. Xét trên toàn bộ vùng đất liền của Việt Nam thì nồng độ NO_2 là tương đối thấp. Khu vực có nồng độ NO_2 cao nhất là vùng Đồng bằng Sông Hồng ở phía Bắc và khu vực Thành phố Hồ Chí Minh ở phía Nam. Xét vị trí khu vực phía Bắc của Việt Nam, tại khu vực nằm giữa ba tỉnh Hải Phòng, Quảng Ninh, Hải Dương có nồng độ khí NO_2 cao nhất. Đây là vị trí tập trung các khu công nghiệp lớn của miền Bắc. Ngoài ra, khu vực có nồng độ khí NO_2 cao tại miền Bắc là vị trí khu công nghiệp và đường Quốc lộ 1A đi qua các tỉnh Hà Nam, Ninh Bình và Thanh Hóa. Khu vực thành phố Hà Nội và Thành phố Hồ Chí Minh cũng được ghi nhận là nơi có nồng độ khí NO_2 cao. Năm 2023, nồng độ chất khí NO_2 tăng lên tại vùng phía Bắc và vùng Bắc Trung Bộ so với năm 2022. Tại khu vực miền Bắc, diện tích vùng có nồng độ NO_2 cao đã mở rộng về phía Tây Bắc. Khu vực Bắc Trung Bộ, vị trí có nồng độ NO_2 cao đã mở rộng hơn về phía Đông tỉnh Thanh Hóa và Nghệ An. Tại vị trí phía Tây tỉnh Sơn La và Điện Biên, gần biên giới với Lào đã xuất hiện khu vực có NO_2 cao trong năm 2023. NO_2 là chất khí gây ô nhiễm không khí và được thải ra từ các phương tiện giao thông và các hoạt động các khu công nghiệp. Dưới tác động của ánh sáng mặt trời khí NO_2 có thể chuyển thành khí NO và tạo ra khí ozone (O_3) cũng là chất khí gây ô nhiễm môi trường. Do đó, việc kiểm soát nồng độ khí NO_2 sẽ giúp giảm ảnh hưởng của khí NO_2 đến chất lượng không khí và sức khỏe con người.

Hình 5 thể hiện nồng độ trung bình năm 2022 và năm 2023 của chất khí CO. Hình 5 cho thấy nồng độ CO cao tập trung tại khu vực miền Bắc và Bắc Trung Bộ. Vị trí có nồng độ cao nhất tại vùng Đồng bằng Sông Hồng và một số tỉnh Thái Nguyên, Tuyên Quang, Yên Bái, Lào Cai. Khu vực phía đông các tỉnh Thanh Hóa, Nghệ An, Hà Tĩnh cũng ghi nhận nồng độ CO ở mức cao. Các tỉnh miền Nam Trung Bộ có nồng độ CO ở mức trung bình. Khu vực có nồng độ CO ở mức thấp là các tỉnh thuộc khu vực Tây Nguyên. So sánh hình ảnh trên Hình 5a và Hình 5b cho thấy khu vực có nồng độ CO ở mức cao năm 2023 đã tăng lên so với năm 2022 và mở rộng phía khu vực Tây Bắc, gần biên giới với Lào.



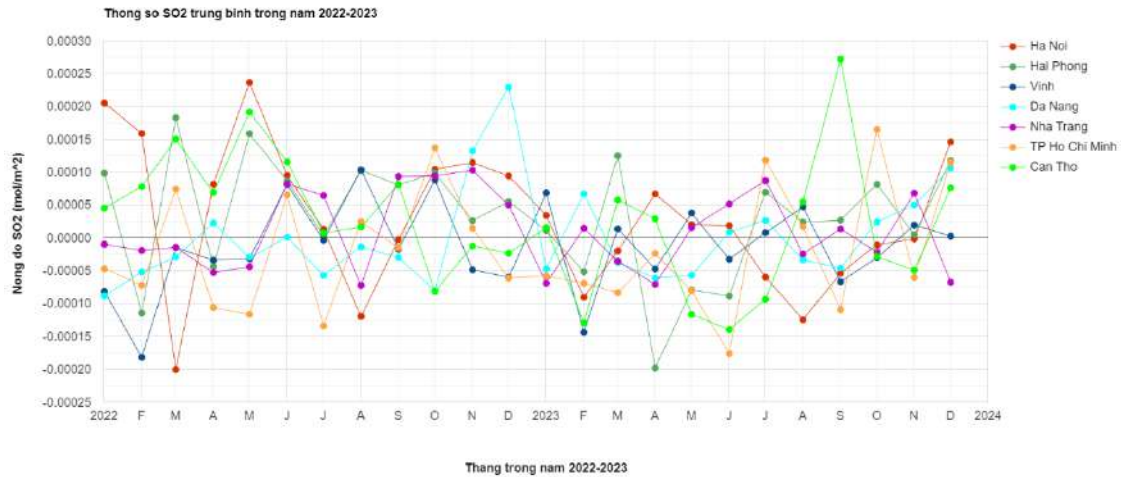
Hình 5. Nồng độ khí CO năm 2022 và năm 2023

4.2. Đánh giá chất lượng không khí tại các khu vực thành phố lớn

Các thành phố lớn được thử nghiệm đánh giá chất lượng không khí bao gồm thành phố Hà Nội, Hải Phòng, Vinh, Đà Nẵng, Nha Trang, Thành phố Hồ Chí Minh và Cần Thơ. Hình 6, 7 và 8 thể hiện biểu đồ biểu diễn diễn biến nồng độ trung bình theo từng tháng từ tháng 01/2022 đến tháng 12/2023 tại vị trí các thành phố thuộc khu vực nghiên cứu.

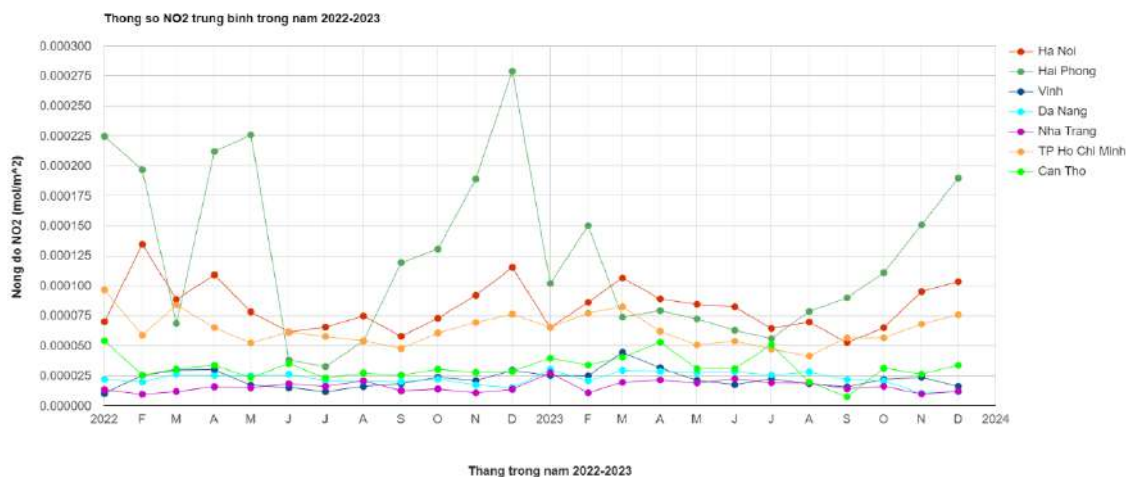
Xét nồng độ trung bình của khí SO_2 tại Hình 6. Kết quả trên biểu đồ cho thấy nồng độ chất khí SO_2 thay đổi theo từng vị trí địa lý và mức độ hoạt động của con người. Các vị trí có nồng độ SO_2 nhỏ hơn 0 là thời điểm

không khí rất sạch và không ghi nhận nồng độ của khí SO_2 . Các tháng có nồng độ SO_2 thấp thường là tháng 2, tháng 3 trong năm. Khu vực Hà Nội và Hải Phòng có mức nồng độ cao vào các tháng 5/2022 và tháng 12/2023. Khu vực thành phố Vinh, Đà Nẵng, Nha Trang có xu hướng diễn biến nồng độ SO_2 trong năm tương tự như nhau. Thành phố Hồ Chí Minh có mức nồng độ SO_2 thấp hơn so với thành phố Hà Nội. Thành phố Cần Thơ là khu vực có nhiều thời điểm có nồng độ SO_2 cao như tháng 3/2022, 5/2022 và tháng 9/2023. Qua kết quả phân tích trên biểu đồ cho thấy, tại khu vực miền Bắc, thời điểm nồng độ SO_2 cao là thời điểm mùa hè và mùa đông. Đối với các thành phố khu vực miền Đông Nam Bộ và vùng Đồng bằng sông Cửu Long thì nồng độ SO_2 cao thường vào các tháng mùa khô, ít mưa.



Hình 6. Nồng độ SO_2 trung bình theo tháng từ 01/2022 đến 12/2023 của các thành phố lớn

Hình 7 biểu diễn nồng độ trung bình NO_2 theo tháng từ tháng 01/2022 đến tháng 12/2023 tại các thành phố lớn. Diễn biến nồng độ NO_2 tại các thành phố Cần Thơ, Nha Trang, Đà Nẵng và Vinh ở mức thấp và không có diễn biến bất thường trong năm 2022 và 2023. Nồng độ NO_2 tại thành phố Hà Nội, Thành phố Hồ Chí Minh và Hải Phòng luôn ở mức cao. Thành phố Hải Phòng có mức nồng độ NO_2 diễn biến bất thường vào các tháng 1, 2, 4, 5 và tháng 12 của năm 2022. Nồng độ NO_2 của thành phố Hà Nội luôn cao hơn Thành phố Hồ Chí Minh.

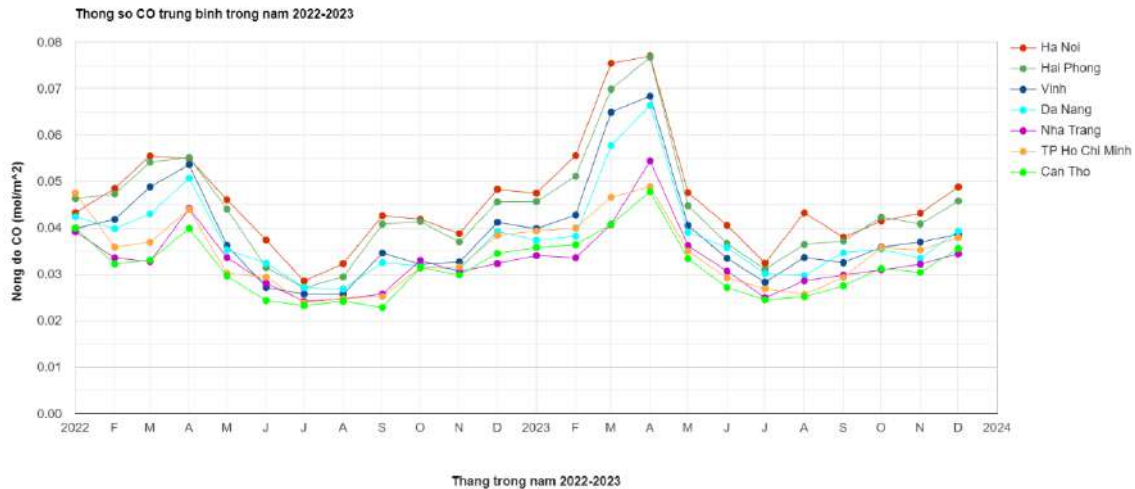


Hình 7. Nồng độ NO_2 trung bình theo tháng từ 01/2022 đến 12/2023 của các thành phố lớn

Hình 8 biểu diễn nồng độ trung bình của CO tại các thành phố lớn trong giai đoạn 2022-2023. Diễn biến nồng độ CO tại các thành phố tương tự như nhau. Trong đó, thành phố có nồng độ CO cao là Hà Nội, Hải Phòng, Vinh và Đà Nẵng. Tháng có nồng độ CO cao nhất là tháng 3, 4 hàng năm. Tuy nhiên, tháng 3, 4 năm 2023, nồng độ khí CO tại Hà Nội, Hải Phòng và Đà Nẵng tăng cao hơn so với các tháng trong năm 2022 và năm 2023.

Như vậy, qua kết quả phân tích trên ảnh và trên biểu đồ cho thấy các thành phố nằm ở miền Bắc có nồng độ trung bình các chất khí SO_2 , NO_2 , CO cao hơn các thành phố ở miền Trung và miền Nam. Đặc biệt là thành phố Hải Phòng luôn có mức nồng độ các chất khí cao nhất trong khu vực nghiên cứu. Thành phố Hải Phòng là khu vực có mạng lưới giao thông phức tạp, các khu công nghiệp, nhà máy, cảng biển, nuôi trồng thủy sản và đánh

bất hải sản đã làm gia tăng lượng khí thải các chất ô nhiễm. Đây cũng là những nguyên nhân làm tăng nồng độ các chất khí SO_2 , NO_2 , CO tại Hải Phòng. Các thành phố tập trung đông dân cư như Hà Nội, Thành phố Hồ Chí Minh thì nồng độ các chất khí SO_2 , NO_2 , CO luôn cao hơn so với khu vực xung quanh.



Hình 8. Nồng độ CO trung bình theo tháng từ 01/2022 đến 12/2023 của các thành phố lớn

5. Kết luận

Qua kết quả nghiên cứu cho thấy tư liệu ảnh Sentinel-5P cho phép giám sát chất lượng không khí trên diện rộng thông qua giá trị nồng độ của các chất khí gây ô nhiễm như SO_2 , NO_2 , CO. Đồng thời, tư liệu ảnh Sentinel-5P đa thời gian cho phép theo dõi sự thay đổi nồng độ các chất khí gây ô nhiễm trong nhiều năm và xác định các vị trí thay đổi theo không gian. Kết quả so sánh nồng độ trung bình năm của chất khí SO_2 , NO_2 , CO trên vùng đất liền Việt Nam trong giai đoạn 2022-2023 cho thấy nồng độ các chất NO_2 và CO đã tăng lên tại một số khu vực miền Bắc trong năm 2023. Bên cạnh đó, ảnh Sentinel-5P cho phép biểu diễn nồng độ các chất khí theo không gian địa lý và phân bố thang giá trị trên từng vùng. Trong nội dung bài báo, các tác giả đã thể hiện diễn biến nồng độ trung bình theo tháng trong khoảng thời gian từ tháng 01/2022 đến tháng 12/2023 của các chất khí SO_2 , NO_2 , CO tại một số thành phố lớn của Việt Nam gồm Hà Nội, Hải Phòng, Vinh, Đà Nẵng, Nha Trang, Thành phố Hồ Chí Minh và Cần Thơ. Kết quả cũng cho thấy diễn biến nồng độ các chất khí trên từng tháng tại các thành phố khác nhau và có tháng diễn biến bất thường. Thành phố Hải Phòng ghi nhận có nồng độ trung bình các chất khí SO_2 , NO_2 , CO cao tại một số thời điểm trong năm. Các khu vực có nồng độ chất khí SO_2 , NO_2 , CO cao thường là những khu vực có mật độ giao thông cao, tập trung nhiều khu công nghiệp. Kết quả nghiên cứu của bài báo cung cấp thêm thông tin về chất lượng không khí trong năm trên vùng đất liền Việt Nam và thời điểm từng tháng tại một số thành phố lớn của Việt Nam trong giai đoạn 2022-2023. Đồng thời, bài báo cũng bổ sung kết quả ứng dụng tư liệu Sentinel-5P trong đánh giá chất lượng không khí tại Việt Nam.

Tài liệu tham khảo

- Amanollahi, J., Tzanis, C., Abdullah, A. M., Ramli, M. F., Pirasteh, S., 2013. Development of the models to estimate particulate matter from thermal infrared band of landsat enhanced thematic mapper. *Int J Environ Sci Technol*, 10: 1245-1254.
- Babic, L., Braak, R., Dierssen, W., Kissi-Ameyaw, J., Kleipool, Q., Leloux, J., Loots, E., Ludewig, A., Rozemeijer, N., Smeets, J., Vacanti, G., 2022. Algorithm theoretical basis document for the TROPOMI L01b data processor. Issue 10.0.0, Creative Commons, California, USA: 14-56.
- Bộ Khoa học và Công nghệ, 2023. QCVN 05:2023/BTNMT: Quy chuẩn kỹ thuật quốc gia về chất lượng không khí, Hà Nội, 1-12.
- Bộ Tài nguyên và Môi trường, 2022. Báo cáo hiện trạng môi trường quốc gia 2021: Môi trường không khí thực trạng và giải pháp, Hà Nội, 1-95.
- Chen, T-M., Kuschner, W.G., Gokhale, J., Shofer, S., 2007. Outdoor air pollution: nitrogen dioxide, sulfur dioxide, and carbon monoxide health effects. *Am J Med Sci*, 333(4): 249-256.
- Lưu Thị Diệu Chinh, Hà Thị Hằng, Bùi Duy Quỳnh, 2023. Ứng dụng dữ liệu Sentinel-5P Tropomi trên nền tảng Google Earth Engine trong theo dõi ô nhiễm không khí tại thành phố Thái Nguyên. *Tạp chí Khoa học Công nghệ Xây dựng*, ĐHXDHN, 17(2V), 78-94.

- Davybida, L.I., 2023. Air quality impacts of war detected from the Sentinel-5P satellite over Ukraine. *IOP Conference Series: Earth and Environmental Science*, 22-26, May, Ukraine, Vol. 1254, No. 1, 012112: 1-7.
- Dimitrova, M., Trenchev, P., Avetisyan, D., Spasova, T., 2023. Spatio-temporal monitoring of air pollution over Bulgaria's largest industrial area using Sentinel-5p TROPOMI data. *Ninth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2023)*, 3-5 April, Cyprus, Vol. 12786: 615-622.
- Jeong, U., Hong, H., 2021. Assessment of Tropospheric Concentrations of NO₂ from the TROPOMI/Sentinel-5 Precursor for the Estimation of Long-Term Exposure to Surface NO₂ over South Korea. *Remote Sens.* 13, 1877: 1-16.
- Hassaan, M.A., Abdallah, S.M., Shalaby, E-S. A., Ibrahim, A.A., 2024. Assessing vulnerability of densely populated areas to air pollution using Sentinel-5P imageries: a case study of the Nile Delta, Egypt. *Scientific reports, Nature portfolio*, (2023) 13:17406: 1-11.
- Hu, T., Yoshie, R., 2020. Effect of atmospheric stability on air pollutant concentration and its generalization for real and idealized urban block models based on field observation data and wind tunnel experiments. *Journal of Wind Engineering and Industrial Aerodynamics*, 207, 104380.
- van Geffen, J.H.G.M., Eskes, H.J., Boersma, K.F., Veeffkind, J.P., 2022. TROPOMI ATBD tropospheric and total NO₂. Issue 2.4.0: 1-88.
- Kaplan, G., Avdan, Z.Y., 2020. Space-borne air pollution observation from sentinel-5p tropomi: Relationship between pollutants, geographical and demographic data. *International Journal of Engineering and Geosciences*, 5(3): 130-137.
- Landgraf, J., aan de Brugh, J., Scheepmaker, R. A., Borsdorff, T., Houweling, S., Hasekamp, O.P., 2022. Algorithm Theoretical Baseline Document for Sentinel-5 Precursor: Carbon Monoxide Total Column Retrieval. Issue 2.4.0: 1-83.
- Langen, J., Meijer, Y., Veihelmann, B., Ingmann, P., 2017. Copernicus Sentinels 4 and 5 mission requirements traceability document. Issue 2, European Space Research and Technology Centre, Keplerlaan, The Netherlands: 1-166.
- Liu, B., Wu, S-D., Shen, L-J., Zhao, T-X., Wei, Y., Tang, X-L., Long, C., Zhou, Y., He, D-W., Lin, T., 2019. Spermatogenesis dysfunction induced by PM2.5 from automobile exhaust via the ROS-mediated MAPK signaling pathway. *Ecotoxicol Environ Saf*, 167: 161-168.
- Mejia, D., Faican, G., Zalakeviciute, R., Matovelle, C., Bonilla, S., Sobrino, J.A., 2024. Spatio-temporal evaluation of air pollution using ground-based and satellite data during Covid-19 in Ecuador. *Heliyon*, 10(7).
- Morozova, A.E., Sizov, O.S., Elagin, P.O., Agzamov, N.A., Fedash, A.V., Lobzhanidze, N.E., 2022. Integral Assessment of Atmospheric Air Quality in the Largest Cities of Russia Based on TROPOMI (Sentinel-5P) Data for 2019-2020. *Cosmic Research*, 60 (Suppl 1): 57-68.
- Nagy, A., El-Zeiny, A., Sowilem, M., Atwa, W., Elshaier, M., 2022. Mapping mosquito larval densities and assessing area vulnerable to diseases transmission in Nile valley of Giza. *Egypt Egyptian J Remote Sens Space Sci*, 25(1): 63-71.
- Nagy, A., El-Zeiny, A., Sowilem, M., Moselhi, W., Elshaier, M., 2024. Monitoring air quality of Sharkia Governorate using remote sensing. *Acta Geophysica*, Research Article-Atmospheric & Space sciences: 1-14.
- Oxoli, D., Cedeno Jimenez, J. R., Brovelli, M. A., 2020. Assessment of SENTINEL-5P performance for ground-level air quality monitoring: preparatory experiments over the Covid-19 lockdown period. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 44: 111-116.
- Phan, A., Fukui, H., 2024. Unusual response of O₃ and CH₄ to NO₂ emissions reduction in Japan during the COVID-19 pandemic. *International Journal of Digital Earth*, 17(1), 2297844: 1-20.
- Rehmat, R., Rafique, L., Irfan, M., Ahmed, S. R., Lahori, A. H., Muhammadi, A., ... Shulga, M., 2023. Space-borne Air Quality Monitoring of Nitrogen dioxide (NO₂) over Karachi and Lahore using Remote Sensing Tools. *Proceedings of the Pakistan Academy of Sciences: B. Life and Environmental Sciences*, 60(3): 505-513.
- Sivaramakrishnan, K. N., Deka, L., Gupta, M., 2021. Use of Remote Sensing Data to Identify Air Pollution Signatures in India. *Geo-intelligence for Sustainable Development*: 109-125.
- Dương Hồng Sơn, Trương Anh Sơn, Phạm Văn Sỹ, Phan Ban Mai, Nguyễn Văn Tiến, Trần Thị Diệu Hằng, Lê Văn Thiện, 2006. Nghiên cứu thử nghiệm dự báo thời hạn ngắn chất lượng không khí vùng Đồng bằng Bắc Bộ. *Kỷ yếu "Các đề tài nghiên cứu khoa học-công nghệ ngành Tài nguyên và Môi trường giai đoạn 2002-2006"*: 752-763.
- Đỗ Thị Phương Thảo, Nghiêm Văn Ngọc, Vũ Kim Sơn, 2022. Thành lập bản đồ phân bố nồng độ khí NO₂, SO₂ khu vực Hà Nội từ dữ liệu viễn thám Sentinel-5P. *Tạp chí Khoa học Đo đạc và Bản đồ*, 54-12/2022: 46-53.
- Theys, N., De Smedt, I., Lerot, C., Yu, H., Van Roozendaal, M., 2022. S5P/Tropomi SO₂ ATBD. Issue 2.5.0: 1-75.
- Tổng cục Thống kê, 2022. Niên giám Thống kê Việt Nam năm 2022. Nhà xuất bản Thống kê, Hà Nội, Việt Nam, 1-1268.

Vîrghileanu, M., Săvulescu, I., Mihai, B.A., Nistor, C., Dobre, R., 2020. Nitrogen dioxide (NO₂) pollution monitoring with Sentinel-5P satellite imagery over Europe during the coronavirus pandemic outbreak. *Remote Sensing*, 12(21), 3575: 1-28.

Wang, Y., Cai, G., Yang, L., Zhang, N., Du, M., 2022. Monitoring of urban ecological environment including air quality using satellite imagery. *Plos one*, 17(8), e0266759: 1-15.

Wieczorek, B., 2023. Air Pollution Patterns Mapping of SO₂, NO₂, and CO Derived from TROPOMI over Central-East Europe. *Remote Sensing*, 15(6), 1565: 1-28.

World Health Organization (WHO), 2022. WHO ambient air pollution database, 2022 update. *Status Report*, World Health Organization, Geneva, Switzerland, 1-121.

Zhang, Y., Li, Z., Bai, K., Wei, Y., Xie, Y., Zhang, Y., Ou, Y., Cohen, J., Zhang, Y., Peng, Z., 2021. Satellite remote sensing of atmospheric particulate matter mass concentration: advances, challenges, and perspectives. *Fund Res*, 1(3): 240-258.

ABSTRACT

Application of Sentinel-5P imagery for monitoring the air quality of large cities

Le Minh Hang^{1,2}, Do Thi Hoai^{3*}

¹Le Quy Don Technical University, Hanoi, Vietnam

²Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology

³Vietnam Institute of Surveying and Mapping

Air pollution is becoming increasingly significant, particularly in the major cities of Vietnam. The rapid increase in population density, the emergence of personal automobiles, and the development of industries have all contributed significantly to environmental pollution. The Sentinel-5P (Sentinel-5 Precursor) satellite, equipped with the TROPospheric sensor (TROPOMI), was successfully launched by the European Space Agency (ESA) in October 2017. The primary objective of the Sentinel-5P satellite is to detect ultraviolet radiation, ozone, and air quality by conducting atmospheric measurements with high temporal resolution. In the article, the authors demonstrate the utilization of Sentinel-5P data and the Google Earth Engine (GEE) platform to monitor air quality in major cities in Vietnam. The study area includes the entire mainland of Vietnam and several major communities, such as Hanoi, Hai Phong, Vinh, Da Nang, Nha Trang, Ho Chi Minh City, and Can Tho. The period of air quality monitoring is from January 2022 to December 2023. In the period of 2022-2023, the spatial distribution of the annual average concentration of SO₂, NO₂, and CO on the mainland of Vietnam was demonstrated by the research results. In some areas, the amounts of CO and NO₂ in 2023 are higher than in 2022. In addition, we monitor the air quality of large cities through determining the average concentration of polluting gases such as SO₂, NO₂, and CO on a monthly basis. The research results of the article enhance the capacity to utilize remote sensing data for the purpose of large-scale air quality monitoring in Vietnam.

Keywords: Sentinel-5P, air pollution, large cities, Google Earth Engine.

Thành lập mô hình lún để chẩn đoán Nhà máy lọc dầu Dung Quất

Nguyễn Thị Kim Thanh^{1*}, Nguyễn Hà¹

¹Trường Đại học Mở - Địa chất

TÓM TẮT

Đề chẩn đoán là một hạng mục quan trọng đảm bảo an toàn cho nhà máy lọc dầu nhưng lại dễ bị chuyển dịch theo thời gian do chịu tác động của các yếu tố như sóng biển, nhiệt độ,... do đó chúng được giám sát thường xuyên thông qua công tác quan trắc. Hiện tại phần lớn các công trình nói chung và đề chẩn đoán nói riêng ở Việt Nam đều được quan trắc lún bằng phương pháp trắc địa, nên kết quả thu được đơn giản chỉ là giá trị lún của công trình. Chính vì vậy, để vận hành tuyến đề an toàn và hiệu quả, cần đánh giá tổng quan xu hướng lún của công trình dựa trên kết quả phân tích lún trong không gian. Phân tích lún trong không gian từ dữ liệu trắc địa bản chất là thành lập mô hình lún của công trình thông qua các hàm toán học. Bài báo đã ứng dụng phương pháp này cho trường hợp thành lập mô hình lún tuyến đề chẩn đoán tại Nhà máy lọc dầu Dung Quất. Kết quả thu được là một đường thẳng xác suất lún - thành lập trên cơ sở kết quả đo lún thực tế của tuyến đề. Theo dõi mô hình này, các nhà quản lý có thể giám sát công trình một cách trực quan và toàn diện hơn. Đồng thời, cũng giúp nhận biết và khắc phục các xu hướng chuyển biến xấu của công trình, từ đó duy trì hoạt động sản xuất an toàn và hiệu quả.

Từ khóa: Mô hình lún, quan trắc biến dạng, đề chẩn đoán, Nhà máy lọc dầu Dung Quất.

1. Đặt vấn đề

Phân tích dữ liệu quan trắc biến dạng công trình nói chung và quan trắc lún nói riêng là giải pháp nâng cao hiệu quả xử lý số liệu, giúp công tác quản lý vận hành công trình an toàn hơn (Trần Khánh, Trần Ngọc Đông, 2018; Nguyễn Thị Kim Thanh, 2021). Hiện tại, trên thế giới, xu hướng xác định và phân tích giá trị sụt lún của mặt đất do các thảm họa thiên nhiên như động đất, núi lửa phun trào, trượt lở,... hoặc độ lún của các công trình như đập thủy điện,... là dựa vào dữ liệu ảnh vệ tinh quang học hoặc ảnh vệ tinh radar. Những kỹ thuật hiện đại này không chỉ giúp tiết kiệm sức lao động của con người mà còn cho kết quả phân tích rất đa chiều và trực quan, được cập nhật liên tục hàng ngày, hàng tháng (Guiseppe Modoni et al., 2013; Zhiguo Pang et al., 2023). Bên cạnh đó, cũng có nghiên cứu sử dụng hệ thông tin địa lý (GIS) tích hợp các thông số gây sụt lún để phân tích và thành lập mô hình đánh giá tình trạng lún của công trình (Haya M. Alogayell et al., 2021). Tuy nhiên, các công nghệ, kỹ thuật hiện đại này cũng có những hạn chế nhất định đó là chưa được hệ thống hoá một cách rộng rãi và toàn diện (Federico Raspini et al., 2022). Ở những khu vực, những công trình có tốc độ lún chậm, hoặc không xác định được các thông số gây lún thì khó có thể dùng các phương pháp nói trên để phân tích lún (đây chính là trường hợp ở Việt Nam). Tại Việt Nam, các cơ sở sản xuất chủ yếu vẫn quan trắc lún bằng phương pháp trắc địa truyền thống bởi chúng ta chưa phóng được vệ tinh, cũng không dễ dàng lấy được dữ liệu ảnh vệ tinh một cách tự do, miễn phí. Hơn nữa, nếu có thu thập được dữ liệu vệ tinh thì cũng cần kỹ thuật chuyên môn để xử lý và ứng dụng thuận lợi. Việc này phải mất nhiều thời gian và công sức, rất khó thực hiện rộng rãi. Chính vì vậy, giải pháp phù hợp nhất để ứng dụng phân tích số liệu quan trắc lún được đo bằng phương pháp trắc địa là các phương pháp truyền thống như phân tích hình học (phân tích không gian), phân tích thống kê và phân tích theo thời gian (dự báo) (US.Army Corps of Engineers, 2018). Trong phạm vi nghiên cứu của bài báo này, tác giả đề xuất giải pháp phân tích dữ liệu quan trắc lún trong không gian từ số liệu trắc địa, dưới các dạng hình học như đường thẳng, đường cong, mặt phẳng,... (Trần Khánh, Nguyễn Quang Phúc, 2010) áp dụng để thành lập mô hình lún tuyến đề chẩn đoán tại Nhà máy lọc dầu Dung Quất. Dự án quan trắc lún tuyến đề chẩn đoán vẫn luôn được các nhà quản lý công trình quan tâm thực hiện thường xuyên theo các chu kỳ mỗi năm. Tuy nhiên, số liệu thu được đơn giản chỉ là các giá trị độ lún rời rạc của các điểm quan trắc trên tuyến đề. Do đó, việc đề xuất thành lập mô hình lún trong không gian cho công trình này là cần thiết và có tính thực tiễn cao. Phương pháp thực hiện đơn giản, kết quả thu được mang lại hiệu quả cao trong việc giám sát công trình ngay tại thời điểm quan trắc (Trần Ngọc Đông, 2009; Loizos Pelecanos et al., 2018). Mô hình thu được giúp các nhà quản lý đánh giá được tổng quan xu hướng chuyển dịch của công trình đồng thời có thể nội suy được giá trị chuyển dịch của các mốc quan trắc bị mất hoặc bị hư hỏng (Weizhou et al., 2011; Lê Thanh Lâm, 2023).

* Tác giả liên hệ

Email: nguyenthikimthanh@hmg.edu.vn

2. Cơ sở lý thuyết

Xây dựng mô hình chuyển dịch (chuyển dịch ngang và chuyển dịch thẳng đứng) là một trong những nhiệm vụ quan trọng của công tác quan trắc biến dạng công trình, đặc biệt là đối với các tổ hợp kỹ thuật nhạy cảm như nhà máy thủy điện, các tuyến đập tràn, đập chắn nước,...

Các nhiệm vụ được đặt ra trong khi phân tích, đánh giá chuyển dịch biến dạng công trình bao gồm:

- Xác định xu hướng chuyển dịch của công trình trong không gian, theo thời gian.
- Xác định và đánh giá mức độ phụ thuộc độ chuyển dịch công trình vào một số yếu tố ngoại cảnh.

Mô hình chuyển dịch có thể được phân chia thành 3 loại chủ yếu là: Mô hình chuyển dịch của đối tượng quan trắc trong không gian, mô hình chuyển dịch theo thời gian và mô hình thống kê.

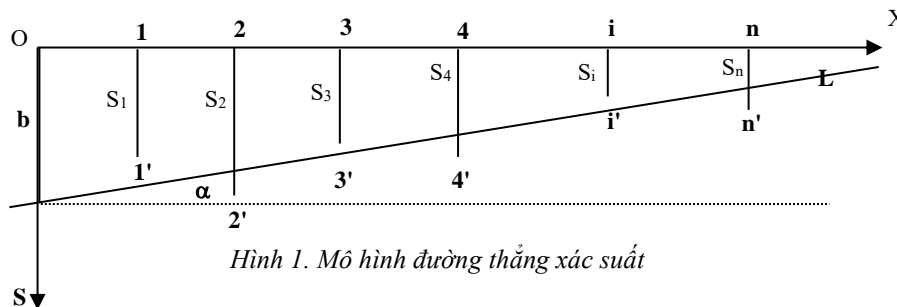
Trong phạm vi nghiên cứu của bài báo, đồng thời căn cứ vào số liệu đo đạc thực tế (quan trắc bằng phương pháp trắc địa) tác giả xem xét một số mô hình lún trong không gian như: Phương trình đường thẳng xác suất, phương trình đường cong, mặt phẳng xác suất. Những mô hình này được xây dựng giúp các nhà quản lý vận hành công trình đánh giá một cách tổng thể xu hướng lún của công trình. Ngoài ra, các mô hình lún trong không gian còn có thể nội suy độ lún của những mốc quan trắc của công trình bị hỏng, bị mất hoặc gặp khó khăn trong quá trình đo đạc.

2.1. Phương pháp đường thẳng xác suất

Mô hình này được ứng dụng phù hợp cho tuyến đập có kết cấu bê tông cốt thép như tuyến đập tràn, đập bê tông trọng lực.

Nếu dọc theo tuyến công trình bố trí và thực hiện quan trắc n mốc lún như Hình 1 với vector tọa độ theo hướng ngang của các điểm quan trắc là $X = (x_1, x_2, \dots, x_n)$, vector độ lún tương ứng là $S = (S_1, S_2, \dots, S_n)$. Tại chu kỳ đang xét các mốc đã bị lún xuống một đại lượng nào đó và có vị trí tương ứng là $1', 2', 3', \dots, n'$. Vậy đường gấp khúc $G = \{1'-2'-3'-\dots-n'\}$ biểu thị cho đặc tính lún của công trình dọc theo trục OX .

Nếu xấp xỉ (G) bằng một đường thẳng (L) sao cho tổng bình phương độ lệch các đỉnh của G so với đường thẳng L là nhỏ nhất: $[V_S^2] \rightarrow \text{Min}$ thì L được gọi là đường thẳng lún xác suất hoặc đơn giản là đường thẳng lún.



Hình 1. Mô hình đường thẳng xác suất

Phương trình đường thẳng L có thể được viết dưới dạng:

$$S = aX + b \text{ với } a = \tan \alpha \quad (1)$$

trong đó:

α - Góc nghiêng của đường thẳng so với phương nằm ngang (thể hiện độ nghiêng tổng thể của công trình).

b - Giá trị lún của công trình tại điểm gốc tọa độ, nếu gốc tọa độ lấy trùng với trọng tâm công trình thì b sẽ là độ lún tại trọng tâm công trình.

Cần có ít nhất 2 điểm quan trắc lún để xác định một hệ gồm 2 phương trình tuyến tính (1) và tìm được 2 tham số a, b . Nếu số điểm quan trắc nhiều hơn 2, thì đường thẳng L xác suất nhất được xác định theo nguyên lý số bình phương nhỏ nhất với điều kiện $[V_S^2] \rightarrow \text{Min}$, trong trường hợp này bài toán được giải theo trình tự sau:

Trên cơ sở các vector X và S lập được hệ phương trình số hiệu chỉnh

$$V_S = Az + S \quad (2)$$

với:

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \dots & \dots \\ x_n & 1 \end{bmatrix}; z = \begin{bmatrix} a \\ b \end{bmatrix}; S = \begin{bmatrix} -S_1 \\ -S_2 \\ \dots \\ -S_n \end{bmatrix}; V_S = \begin{bmatrix} V_{S_1} \\ V_{S_2} \\ \dots \\ V_{S_n} \end{bmatrix}$$

Từ đó thành lập được hệ phương trình chuẩn:

$$A^T A z + A^T S = 0 \quad (3)$$

và xác định được vector tham số z theo công thức:

$$z = -(A^T A)^{-1} A^T S \quad (4)$$

Nếu chuyển điểm gốc O về vị trí trọng tâm công trình (sao cho $[x]=0$), khi đó sẽ tính được:

$$(A^T A) = \begin{pmatrix} [x^2] & [x] \\ [x] & n \end{pmatrix} = \begin{pmatrix} [x^2] & 0 \\ 0 & n \end{pmatrix} \quad (5)$$

$$(A^T S) = \begin{pmatrix} [xS] \\ [S] \end{pmatrix}$$

Trong trường hợp này vector tham số z được xác định theo công thức:

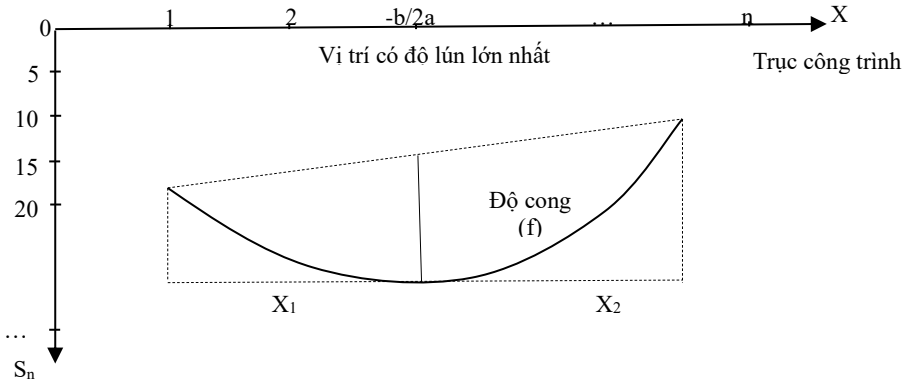
$$a = \frac{[xS]}{[x^2]}, \quad b = \frac{[S]}{n} \quad (6)$$

2.2. Phương pháp đường cong

Phương trình đường cong là mô hình đặc trưng của công trình tuyến đập đất đá đổ. Vị trí ở giữa của những tuyến đập dạng này sẽ chịu tải trọng và áp lực lớn nhất vì vậy độ lún tại đó sẽ lớn hơn nhiều so với hai đầu tuyến đập nên xảy ra hiện tượng bị cong võng của công trình.

Giả thiết S_1, S_2, \dots, S_n là độ lún của các mốc quan trắc được gắn trên trục công trình thì phương trình đường cong thể hiện độ lún của các điểm đó được viết dưới dạng tổng quát sau:

$$S_n = ax_n^2 + bx_n + c \quad (7)$$



Hình 2. Độ cong của tuyến đập được biểu diễn bằng đường parabol

Các hệ số a, b, c của phương trình (7) được xác định dựa theo nguyên lý số bình phương nhỏ nhất bởi mỗi trị đo lún lập được một phương trình số hiệu chỉnh dạng:

$$v_i = ax_i^2 + bx_i + c - S_i \quad (8)$$

- Viết dưới dạng ma trận với n điểm quan trắc

$$V = AX + L \quad (9)$$

trong đó

$$V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad A = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix} \quad X = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad L = \begin{bmatrix} -S_1 \\ -S_2 \\ \vdots \\ -S_n \end{bmatrix} \quad (10)$$

- Lập hệ phương trình chuẩn

$$A^T AX + A^T L = 0 \quad (11)$$

- Tham số của phương trình đường cong

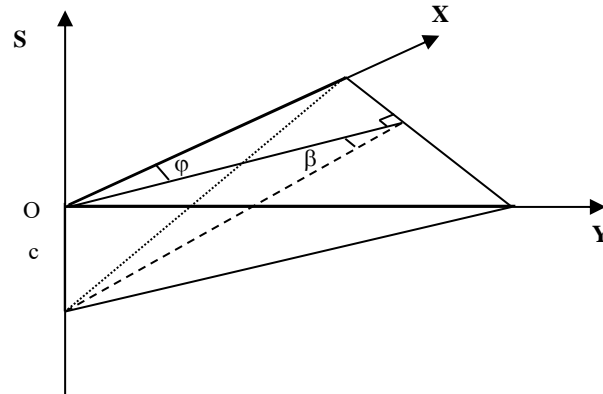
$$X = -(A^T A)^{-1} A^T L \quad (12)$$

- Sai số thành lập mô hình

$$\mu = \pm \sqrt{\frac{[vv]}{n-3}} \quad (13)$$

2.3. Mặt phẳng xác suất

Mặt phẳng xác suất là mô hình lún trong không gian biểu diễn độ lún của công trình có dạng vùng, các điểm quan trắc phân bố không cùng trên một đường thẳng mà trên toàn bộ diện tích công trình. Vì vậy ở mỗi chu kỳ quan trắc sẽ tồn tại một bề mặt lún thể hiện đặc tính lún của các điểm đó.



Hình 3. Tham số lún của công trình dạng vùng

Phương trình của mặt phẳng lún có thể được viết dưới dạng:

$$S_i = ax_i + by_i + c \quad (14)$$

Với n điểm quan trắc ($n \geq 3$), thành lập được hệ phương trình số hiệu chỉnh

$$V = AZ + L \quad (15)$$

trong đó:

$$Z = (a \quad b \quad c)^T$$

$$A = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \dots & \dots & \dots \\ x_n & y_n & 1 \end{bmatrix} \quad (16)$$

$$L = (S_1 \quad S_2 \quad \dots \quad S_n)^T$$

Áp dụng nguyên lý số bình phương nhỏ nhất, xác định được vector Z

$$Z = (A^T A)^{-1} A^T L \quad (17)$$

Từ đó tính được hệ tham số lún tổng quát của mặt phẳng lún

Độ lún tại gốc tọa độ (thường là điểm trọng tâm công trình)

$$S_{TT} = c \quad (18)$$

Góc nghiêng lớn nhất của mặt phẳng lún

$$\beta = \arctan(\sqrt{a^2 + b^2}) \quad (19)$$

Hướng nghiêng lớn nhất (tính từ trục OX theo chiều kim đồng hồ)

$$\varphi = \arctan \frac{b}{a} \quad (20)$$

3. Thành lập mô hình lún để chấn sóng Nhà máy lọc dầu Dung Quất

3.1. Giới thiệu công trình



Hình 4. Đê chắn sóng Nhà máy lọc dầu Dung Quất

Công trình Nhà máy lọc dầu Dung Quất được xây dựng tại địa bàn hai xã Bình Trị và Bình Thuận, huyện Bình Sơn, tỉnh Quảng Ngãi có công suất 6 triệu tấn/năm là công trình quan trọng liên quan đến an ninh quốc gia, có ý nghĩa đặc biệt quan trọng trong việc thúc đẩy phát triển kinh tế - xã hội của khu vực miền Trung, Tây Nguyên.

Đê chắn sóng là một trong những hạng mục quan trọng của công trình, chu kỳ 2 được thực hiện quan trắc theo gói thầu chung của toàn bộ công trình từ 17/02/2023 đến ngày 01/3/2023.

Đê chắn sóng được xây dựng là loại đê bê tông (gồm nhiều khối bê tông xếp chồng nhau) dài gần 1,6 km; cao trình đỉnh đê từ 8-10 m trên mặt nước biển, bề rộng lớn nhất tại chân đê khoảng 110 m và tại đỉnh là từ 10-11 m.

Bảng 1. Dữ liệu quan trắc đê chắn sóng chu kỳ 2

Tên điểm	Hệ tọa độ VN2000			Độ lún	Tên điểm	Hệ tọa độ VN2000			Độ lún
	Kinh tuyến trực 108°, múi 3°					Kinh tuyến trực 108°, múi 3°			
	Chu kỳ 2 (02-2023)					Chu kỳ 2 (02-2023)			
	N (m)	E (m)	H (m)			S (m)	N (m)	E (m)	
1R	1705336,894	585145,228	10,907	-0,013	8L	1705329,445	584445,071	10,651	T
1C	1705334,000	585145,431	10,899	-0,008	9R	1705337,846	584345,448	10,997	-0,014
1L	1705330,738	585145,590	10,873	-0,015	9C	1705333,318	584344,383	11,089	-0,010
2R	1705337,708	585045,024	11,426	-0,010	9L	1705329,224	584346,035	11,082	-0,009
2C	1705334,028	585045,710	11,409	-0,011	10R	1705337,585	584245,301	11,334	-0,007
2L	1705329,963	585044,530	11,390	-0,015	10C	1705334,024	584245,628	11,532	-0,012
3R	1705337,311	584945,858	11,420	-0,009	10L	1705331,631	584245,854	11,556	-0,009
3C	1705334,343	584945,414	11,453	-0,01	11R	1705337,878	584146,515	11,252	-0,007
3L	1705329,933	584946,302	11,470	-0,012	11C	1705334,275	584144,737	9,987	-0,015
4R	1705337,054	584845,337	11,214	-0,010	11L	1705330,422	584145,058	11,297	-0,008
4C	1705333,918	584845,334	11,237	-0,014	12R	1705339,094	584045,645	10,881	-0,013
4L	1705330,913	584846,572	11,239	-0,013	12C	1705334,487	584045,505	10,968	-0,010
5R	1705337,212	584745,278	11,140	-0,011	12L	1705331,260	584045,952	11,088	-0,011
5C	1705334,067	584745,393	11,191	-0,015	13R	1705336,488	583941,848	10,954	-0,008
5L	1705331,238	584745,519	11,170	-0,013	13C	1705334,128	583944,678	11,062	-0,015
6R	1705336,681	584645,461	10,550	-0,012	13L	1705330,728	583944,807	11,147	-0,011
6C	1705334,434	584645,569	10,576	-0,011	14R	1705337,349	583846,507	10,910	-0,013
6L	1705330,828	584645,596	10,601	-0,013	14C	1705334,956	583845,942	11,050	-0,012
7R	1705338,160	584546,409	10,355	-0,009	14L	1705330,960	583845,998	11,107	-0,008
7C	1705334,151	584545,375	10,666	-0,013	15R	1705337,795	583744,245	10,931	-0,009
7L	1705330,370	584544,937	10,577	-0,014	15C	1705333,809	583745,662	11,021	-0,010
8R	1705338,198	584446,277	10,392	-0,015	15L	1705330,996	583745,166	10,972	-0,014
8C	1705334,043	584445,717	10,643	-0,010					

3.2. Kết quả thực nghiệm

Do đê chắn sóng có dạng thẳng và có cấu tạo bê tông nên mô hình lún có dạng đường thẳng (Nguyễn Thị Kim Thanh, 2021) như sau:

$$S = ax + b$$

Như vậy, 45 mốc quan trắc trên tuyến đê đều có độ lún thỏa mãn quy luật của mô hình trên. Để xác định tham số a, b của đường thẳng xác suất, áp dụng nguyên lý số bình phương nhỏ nhất, được:

$$\begin{cases} a = 0.000091 \\ b = -155,605 \end{cases} \quad (S \text{ tính theo đơn vị m})$$

Để thuận lợi cho việc biểu diễn đồ thị mô hình lún, dịch chuyển gốc về trọng tâm của công trình sao cho $[x] = 0$, khi đó tọa độ X của các mốc quan trắc được tính lại như sau:

Bảng 2. Toạ độ X và độ lún các mốc quan trắc (thay đổi gốc toạ độ)

Tên điểm	Toạ độ X (m)	Độ lún S (m)	Tên điểm	Toạ độ X (m)	Độ lún S (m)
1R	2,808	-0,013	8L	-4,641	-0,008
1C	-0,086	-0,008	9R	3,760	-0,014
1L	-3,348	-0,015	9C	-0,768	-0,010
2R	3,622	-0,010	9L	-4,862	-0,009
2C	-0,058	-0,011	10R	3,499	-0,007
2L	-4,123	-0,015	10C	-0,062	-0,012
3R	3,225	-0,009	10L	-2,455	-0,009
3C	0,257	-0,010	11R	3,792	-0,007
3L	-4,153	-0,012	11C	0,189	-0,015
4R	2,968	-0,010	11L	-3,664	-0,008
4C	-0,168	-0,014	12R	5,008	-0,013
4L	-3,173	-0,013	12C	0,401	-0,010
5R	3,126	-0,011	12L	-2,826	-0,011
5C	-0,019	-0,015	13R	2,402	-0,008
5L	-2,848	-0,013	13C	0,042	-0,015
6R	2,595	-0,012	13L	-3,358	-0,011
6C	0,348	-0,011	14R	3,263	-0,013
6L	-3,258	-0,013	14C	0,870	-0,012
7R	4,074	-0,009	14L	-3,126	-0,008
7C	0,065	-0,013	15R	3,709	-0,009
7L	-3,716	-0,014	15C	-0,277	-0,010
8R	4,112	-0,015	15L	-3,090	-0,014
8C	-0,043	-0,010			

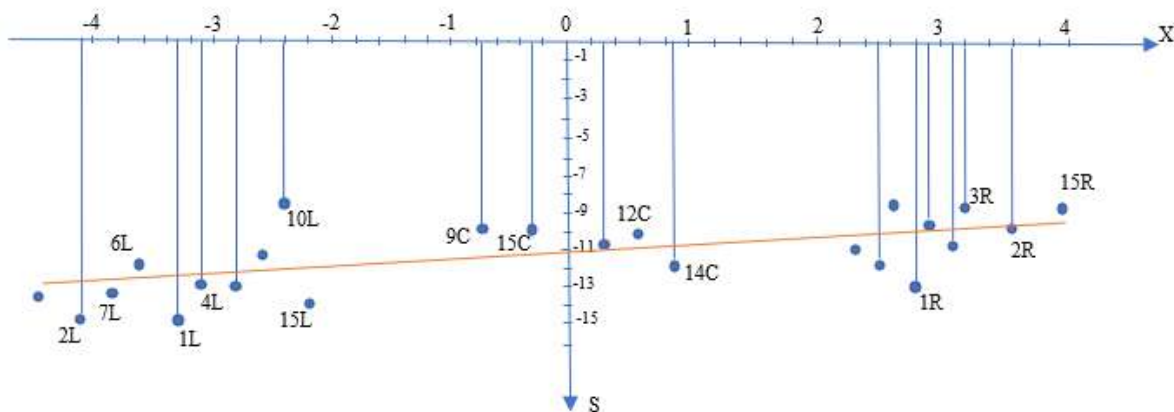
Khi đó hệ số a, b của phương trình đường thẳng được tính lại là:

$$\begin{cases} a = 0,000091 \\ b = -0,0113 \end{cases}$$

Đường thẳng xác suất thể hiện độ lún của đèn chấn sóng Nhà máy lọc dầu Dung Quất là:

$$S = 0,000091 \cdot X - 0,0113 \quad (21)$$

Đồ thị lún của các điểm quan trắc trên tuyến đèn chấn sóng và đường thẳng xác suất nhất (mô hình lún) được thể hiện trong Hình 5.



Hình 5. Mô hình lún tuyến đèn chấn sóng Nhà máy lọc dầu dung Quất

Nhận xét: Từ kết quả phân tích số liệu quan trắc trong không gian, xác định được mô hình lún của tuyến đèn chấn sóng Nhà máy lọc dầu Dung Quất có dạng đường thẳng như công thức (21) và được thể hiện là đường màu cam trong Hình 5, có góc nghiêng $\alpha = 19''$, độ lún trọng tâm công trình là 11,3 mm.

4. Kết luận

Sau khi nghiên cứu lý thuyết và áp dụng xây dựng mô hình lún cho tuyến đề chắn sóng tại Nhà máy lọc dầu Dung Quất, tác giả rút ra một số kết luận như sau:

Phân tích độ lún trong không gian (phân tích hình học) cho cái nhìn trực quan và tổng thể nhất về xu hướng chuyển dịch của công trình vì vậy thành lập các mô hình lún là giải pháp để công tác phân tích hình học này trở nên hiệu quả và toàn diện.

Việc lựa chọn mô hình lún trong không gian phần nào phụ thuộc vào đặc điểm hình dạng và kết cấu của công trình. Với công trình kết cấu bê tông và dầm thép thì thường lựa chọn mô hình lún dạng đường thẳng, còn các tuyến đề đập có kết cấu đất đá, áp lực sẽ tập trung nhiều ở giữa tạo độ võng vì vậy mô hình lún có xu hướng là đường cong, trong khi đó các công trình dạng khối, dạng vùng (nhà) thì nên lựa chọn mô hình mặt phẳng.

Mô hình lún của đề chắn sóng Nhà máy lọc dầu Dung Quất có dạng đường thẳng. Kết quả này có thể được ứng dụng trong thực tế sản xuất, làm cơ sở để các nhà quản lý giám sát tình trạng của tuyến đề. Trong trường hợp các mốc quan trắc trên tuyến đề vì lý do nào đó (sóng đánh, va chạm tàu thuyền,...) mà bị hư hỏng, không thuận tiện cho công tác đo đạc, quan trắc thì kiến nghị có thể dùng mô hình lún vừa thành lập nội suy độ lún và khôi phục lại của các mốc đó (dựa vào phương trình đường thẳng tính độ lún của mốc).

Tài liệu tham khảo

Trần Ngọc Đông, 2009. Phân tích đánh giá kết quả quan trắc độ lún công trình. *Tạp chí Khoa học Công nghệ Xây dựng*, số 1.

Federico Raspini et al., 2022. Review of satellite radar interferometry for subsidence analysis. *Earth Science Reviews* volume 235, <https://doi.org/10.1016/j.earscirev.2022.104239>.

Guiseppe Modoni et al., 2013. Spatial analysis of land subsidence induced by groundwater withdrawal. *Engineering Geology* 167, p. 59-71. Doi: 10.1016/j.enggeo.2013.10.014.

Haya M. Alogayell et al., 2021. Spatial modeling of ground subsidence susceptibility along Al-Shamal train pathway in Saudi Arabia. *Open Geosciences*, Vol. 13, issue 1. <https://doi.org/10.1515/geo-2020-0296>.

John Olusegun Ogundare, 2019. *Understanding least squares estimation and geomatics data analysis*. John Wiley & Sons, Inc, Published 2019 by John Wiley & Sons, Inc.

Trần Khánh, Trần Ngọc Đông, 2018. Giải pháp nâng cao chất lượng công tác quan trắc độ lún công trình. *Tạp chí Khoa học Công nghệ Xây dựng*, số (1), 2.

Trần Khánh, Nguyễn Quang Phúc, 2010. *Quan trắc chuyển dịch và biến dạng công trình*. Nhà xuất bản Giao thông Vận tải, Hà Nội.

Lê Thanh Lâm, 2023. Nghiên cứu phương pháp phân tích độ lún công trình thủy điện từ số liệu đo trắc địa. Luận văn thạc sĩ, Trường Đại học Mở - Địa chất.

Loizos Pelecanos, Dimitrios Skarlatos, George Pantazis, 2018. Finite element analysis of the monitored long-term settlement behaviour of Kouris earth dam in Cyprus. *9th European Conference on Numerical Methods in Geotechnical Engineering (IX NUMGE)*, At Porto, Portugal.

Nguyễn Thị Kim Thanh, 2021. Nghiên cứu giải pháp nâng cao hiệu quả xử lý số liệu quan trắc độ lún tuyến đập công trình thủy điện. Luận án tiến sĩ kỹ thuật, Trường Đại học Mở - Địa chất.

US. Army Corps of Engineers, 2018. *Structural Deformation Surveying*.

Weizhou, junjie Hua, Xiaolin Chang, Chuangbing Zhou, 2011. Settlement analysis of the Shuibuya concrete-face rockfill dam. *Computers and Geotechnics* 38(2): 269-280.

Zhiguo Pang et al, 2023. Deformation Monitoring and Analysis of Reservoir Dams Based on SBAS-InSAR Technology-Banqiao Reservoir. *Remote Sens*, 15(12). <https://doi.org/10.3390/rs15123062>.

ABSTRACT

Establishment of subsidence model of breakwater at Dung Quat oil refinery

Nguyen Thi Kim Thanh^{1*}, Nguyen Ha¹
¹*Hanoi University of Mining and Geology*

Breakwater is an important item ensuring safety of the refinery, but they are susceptible to movement over time due to the impact of factors such as ocean waves, temperature, etc., so they are regularly monitored. Currently, most structures including breakwaters in Vietnam are monitored subsidence by geodetic methods, so the obtained results are only subsidence values of monitoring points. Therefore, to operate the breakwater safely and effectively, it is necessary to make an overall assessment the subsidence trend based on spatial analysis. Analyzing subsidence in space from geodetic data means establishing a subsidence model from mathematical functions. The article has applied this method to establish subsidence model for the breakwater at Dung Quat oil refinery. The achieved result is a probability line of subsidence - established from actual subsidence data of breakwater. The model helps managers monitor projects more intuitively and comprehensively. Simultaneously, it is easy to identify and overcome negative changes of the project, thereby maintaining safe and effective production operations.

Keywords: Subsidence model, deformation monitoring, breakwater, Dung Quat oil refinery.

Nghiên cứu ứng dụng phương pháp phân loại hướng đối tượng và mô hình học máy trong xây dựng bản đồ lớp phủ từ ảnh viễn thám quang học

Khúc Thành Đông^{1*}, Trần Đình Trọng¹, Trần Văn Anh^{2,3}, Đỗ Quang Vinh¹, Hà Trung Khiên¹,
Đào Duy Toàn¹, Trương Xuân Quang⁴

¹Trường Đại học Xây dựng Hà Nội

²Trường Đại học Mỏ - Địa chất

³Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mỏ - Địa chất

⁴Trường Đại học Tài nguyên và Môi trường Hà Nội

TÓM TẮT

Bản đồ lớp phủ bề mặt đất có vai trò quan trọng trong công tác quản lý tài nguyên, giám sát môi trường, hỗ trợ quy hoạch sử dụng đất và phát triển bền vững. Nghiên cứu này kết hợp sử dụng phương pháp phân đoạn ảnh Simple Non-Iterative Clustering (SNIC) và mô hình học máy Gradient Boosting để xây dựng bản đồ lớp phủ khu vực xã Phong Dụ Thượng, huyện Văn Yên, tỉnh Yên Bái theo phân loại hướng đối tượng. Các kênh phổ nhìn thấy, cận hồng ngoại và hồng ngoại sóng ngắn của ảnh viễn thám Sentinel-2 và 203 điểm mẫu thông qua thu thập thực địa và giải đoán từ Google Earth được sử dụng làm dữ liệu đầu vào cho nghiên cứu. Các điểm mẫu tập trung vào sáu loại lớp phủ đặc trưng, gồm: Rừng, thực vật không phải rừng, đất trống, ruộng lúa, mặt nước, và đất xây dựng. Mẫu dữ liệu được chia theo tỷ lệ lần lượt là 70% và 30% để huấn luyện và kiểm chứng mô hình. Nghiên cứu đã xây dựng thành công bản đồ lớp phủ bề mặt xã Phong Dụ Thượng từ ảnh vệ tinh Sentinel-2 vào tháng 02 năm 2019. Đánh giá kết quả phân loại cho thấy bản đồ phân loại đạt độ chính xác cao với độ chính xác tổng thể đạt 0,871 và chỉ số Kappa đạt 0,845.

Từ khóa: Lớp phủ bề mặt, phương pháp Simple Non-Iterative Clustering, mô hình học máy Gradient Boosting.

1. Đặt vấn đề

Bản đồ lớp phủ xây dựng từ ảnh viễn thám là công cụ hữu ích trong quản lý tài nguyên môi trường bởi khả năng cung cấp thông tin chi tiết về lớp phủ trên quy mô lớn (Trần Ngọc Tuấn và nnk., 2016; Hồ Lệ Thu và nnk., 2020). Công cụ này giúp tối ưu hóa quy trình ra quyết định và phát triển các chiến lược quản lý đất đai, tiết kiệm thời gian và nguồn lực đáng kể. Sự tiến bộ trong xây dựng bản đồ lớp phủ hiện nay được góp phần bởi sự phát triển của công nghệ ảnh viễn thám và sự tối ưu của các mô hình học máy (Ngô Đức Anh và nnk., 2016; Jamali, 2021; Ajibola và Cabral, 2024; Saini, 2024; Tran Van Anh và nnk., 2024). Các nghiên cứu hiện nay thường sử dụng các nguồn ảnh vệ tinh phổ biến như MODIS, Landsat, Sentinel-2 kết hợp với phương pháp tích hợp dữ liệu dựa trên điểm ảnh (pixel-based) (Ngô Đức Anh và nnk., 2016; Vũ Thị Phương Thảo và Souksakone Sengchanh, 2022). Ngoài ra, một số khu vực trên thế giới và tại Việt Nam đã áp dụng phương pháp tích hợp dữ liệu dựa trên hướng đối tượng (object-based) cho bài toán này mang lại hiệu quả tính toán cao hơn và giảm các lỗi điểm ảnh nhiễu (pepper noise) (Ajibola và Cabral, 2024). Những năm gần đây, các mô hình học máy được áp dụng rộng rãi vào bài toán phân loại mang lại những kết quả đáng khích lệ (Tran Van Anh và nnk., 2024). Tuy nhiên, chưa có nhiều nghiên cứu đề cập đến việc kết hợp phương pháp phân loại dựa trên đối tượng và mô hình học máy, đặc biệt là các mô hình học máy kết hợp (ensemble models) cho bài toán phân loại lớp phủ từ ảnh viễn thám. Nghiên cứu này sử dụng phương pháp phân loại hướng đối tượng và mô hình học máy Gradient Boosting, là mô hình học máy kết hợp tiên tiến hiện nay để xây dựng bản đồ lớp phủ cho khu vực xã Phong Dụ Thượng, huyện Văn Yên, tỉnh Yên Bái từ ảnh vệ tinh Sentinel-2. Kết quả của nghiên cứu dự kiến sẽ cung cấp những thông tin quan trọng hỗ trợ công tác quản lý tài nguyên và phát triển bền vững tại địa phương.

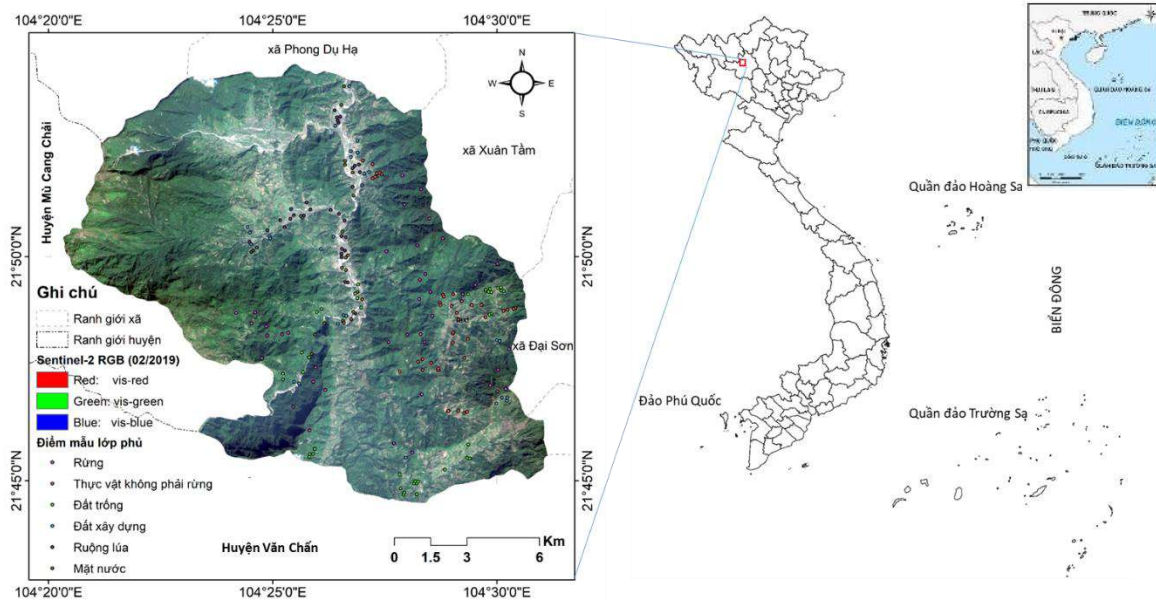
2. Dữ liệu và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Phong Dụ Thượng là xã miền núi nằm tại phía Tây huyện Văn Yên, tỉnh Yên Bái với vị trí địa lý từ 104°20' E đến 21°55' N. Xã Phong Dụ Thượng giáp các xã Nà Hẩu, Đại Sơn, Xuân Tâm ở phía Đông, huyện Mù Cang Chải ở phía Tây, huyện Văn Chấn ở phía Nam, xã Phong Dụ Hạ và địa phận tỉnh Lào Cai ở phía Bắc. Diện tích của Phong Dụ Thượng là 204,95 km², dân số vào năm 2019 là 6.070 người, với mật độ trung bình 30 người/km².

* Tác giả liên hệ
Email: dongkt@huce.edu.vn

Xã Phong Dụ Thượng nổi tiếng với quần thể ruộng bậc thang, nơi có các đồi mâm xôi độc đáo và đồi hoa mua sắc sỡ. Tổng diện tích ruộng bậc thang gần 50 ha, với đỉnh cao nhất đạt độ cao khoảng 1.000 m so với mực nước biển. Xã cũng có tỷ lệ diện tích rừng lớn, đặc biệt là Khu Bảo tồn thiên nhiên Nà Hầu. Ngoài việc canh tác trồng lúa trên ruộng bậc thang, xã Phong Dụ Thượng còn nổi tiếng với các khu vực trồng quế hữu cơ mang lại hiệu quả kinh tế cao. Tuy nhiên, do địa hình phức tạp, nơi đây thường chịu ảnh hưởng nặng nề từ thiên tai, đặc biệt là do những ảnh hưởng của lũ quét và trượt lở đất xảy ra vào mùa mưa hàng năm.



Hình 1. Khu vực nghiên cứu xã Phong Dụ Thượng, huyện Văn Yên, tỉnh Yên Bái

2.2. Dữ liệu

Dữ liệu ảnh viễn thám

Sentinel-2 là một phần của chương trình Copernicus của Liên minh châu Âu, phát triển bởi Cơ quan Vũ trụ châu Âu (ESA), bao gồm hai vệ tinh Sentinel-2A và Sentinel-2B, cung cấp hình ảnh đa phổ với độ phân giải cao. Sentinel-2 có 13 băng tần phổ từ vùng phổ nhìn thấy (VIS), cận hồng ngoại (NIR) đến hồng ngoại sóng ngắn (SWIR), với độ phân giải không gian 10 m, 20 m và 60 m (Bảng 1). Chu kỳ lặp lại là 5 ngày và dải quét rộng 290 km (https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2), giúp cung cấp dữ liệu chi tiết về các đặc điểm bề mặt như đất đai, thảm thực vật và nước, phục vụ cho nhiều ứng dụng quan trọng trong nghiên cứu và quản lý tài nguyên. Bảng 1. Thông tin các kênh phổ của ảnh viễn thám Sentinel-2

STT	Kênh phổ Sentinel-2	Bước sóng trung tâm (µm)	Độ phân giải (m)
1	Kênh 1 - Coastal aerosol	0,443	60
2	Kênh 2 - Blue	0,490	10
3	Kênh 3 - Green	0,560	10
4	Kênh 4 - Red	0,665	10
5	Kênh 5 - Vegetation red edge	0,705	20
6	Kênh 6 - Vegetation red edge	0,740	20
7	Kênh 7 - Vegetation red edge	0,783	20
8	Kênh 8 - NIR	0,842	10
9	Kênh 8A - Vegetation red edge	0,865	20
10	Kênh 9 - Water vapor	0,945	60
11	Kênh 10 - SWIR - Cirrus	1,375	60
12	Kênh 11 - SWIR	1,610	20
13	Kênh 12 - SWIR	2,190	20

Nghiên cứu sử dụng các kênh phổ ở vùng nhìn thấy (kênh 2, 3, 4), cận hồng ngoại (kênh 8) và hồng ngoại sóng ngắn (kênh 11, 12) với độ phân giải không gian 10 m và 20 m vào thời gian tháng 02 năm 2019. Dữ liệu

ảnh của các kênh hồng ngoại sóng ngắn được nội suy và lấy mẫu lại (resampling) về độ phân giải 10 m. Trước khi đưa vào sử dụng, các kênh ảnh được tiền xử lý, gồm: loại bỏ mây thông qua mặt nạ mây QA60 và hiệu chỉnh bức xạ để đưa về giá trị phản xạ.

Dữ liệu mẫu giải đoán:

Việc lựa chọn mẫu giải đoán là công đoạn quan trọng ảnh hưởng đến độ chính xác của mô hình phân loại. Nghiên cứu tiến hành lựa chọn các điểm lấy mẫu từ chuyên thực địa năm 2022 đối với các khu vực dễ tiếp cận và có tính ổn định như khu dân cư, ruộng lúa. Đồng thời, kết hợp sử dụng Google Earth để chọn mẫu cùng với thời điểm chụp của ảnh Sentinel-2. Các điểm lấy mẫu được lựa chọn ngẫu nhiên, được đối chiếu với kết quả giải đoán trên ảnh màu thật của Sentinel-2 để đảm bảo các mẫu được xác định chính xác. Cuối cùng, các mẫu đã thu thập được tổng hợp và biên tập dưới định dạng shapefile trên phần mềm Arcmap 10.8 để tích hợp với dữ liệu phân đoạn ảnh. Số lượng các mẫu và mô tả các loại lớp phủ đặc trưng của khu vực nghiên cứu được thể hiện như trong Bảng 2.

Bảng 2. Bảng phân loại và mô tả các loại lớp phủ đặc trưng của khu vực nghiên cứu

STT	Loại lớp phủ	Ký hiệu	Mô tả	Số lượng mẫu	Tổng diện tích mẫu (m ²)
1	Rừng	FO	Phân bố ở các khu vực núi cao, màu xanh thẫm, đồng nhất.	40	786.991
2	Thực vật không phải rừng	VE	Màu xanh nhạt hoặc xanh không đồng nhất, phân thành các ô thửa lớn.	38	633.681
3	Đất trống	BA	Màu vàng hoặc nâu nhạt, độ sáng cao.	40	380.289
4	Đất xây dựng	BU	Màu trắng, hồng nhạt hoặc đỏ phân bố không đều, theo từng cụm ở các khu vực có địa hình thuận lợi.	29	176.671
5	Ruộng lúa	PA	Màu xanh hoặc nâu đất tùy theo thời điểm, phân thành ô thửa hoặc các đường bậc thang rõ rệt.	35	427.820
6	Mặt nước	WA	Màu xám, đồng nhất, hình dạng tuyến hoặc dạng diện như ao, hồ chứa nước.	21	276.954

2.3. Các thuật toán sử dụng

Thuật toán Simple Non-Iterative Clustering (SNIC):

Simple Non-Iterative Clustering (SNIC) là thuật toán được sử dụng để phân đoạn hình ảnh thành các siêu điểm ảnh (superpixels) dựa trên đồng nhất màu sắc và tính liên tục không gian (Senanayaka và nnk., 2020). SNIC bắt đầu bằng cách khởi tạo các điểm hạt giống phân bố đều trên ảnh. Sau đó, mỗi điểm ảnh được gán nhãn đối tượng theo các điểm hạt giống gần nhất dựa trên khoảng cách không gian và màu sắc. Điểm hạt giống được cập nhật với giá trị trung bình của các điểm ảnh thuộc về nó và quá trình này tiếp tục cho đến khi tất cả các điểm ảnh được gán nhãn.

Ưu điểm của SNIC khi so sánh với các thuật toán phân đoạn khác là đơn giản, hiệu quả và không yêu cầu nhiều bước lặp lại, giúp tăng tốc độ và khả năng mở rộng cho các hình ảnh lớn. SNIC cũng cho phép kiểm soát số lượng và kích thước của các siêu điểm ảnh bằng cách điều chỉnh số lượng điểm hạt giống. Để gán nhãn các điểm ảnh lân cận vào hàng đợi thuật toán SNIC sử dụng hàm khoảng cách theo công thức (1), như sau:

$$\sqrt{\left(\frac{\Delta x}{N}\right)^2 + \left(\frac{\Delta y}{N}\right)^2 + \left(\frac{\Delta L}{M}\right)^2 + \left(\frac{\Delta a}{M}\right)^2 + \left(\frac{\Delta b}{M}\right)^2} \quad (1)$$

trong đó: Δx và Δy là khoảng cách không gian giữa các điểm ảnh.
 ΔL , Δa , và Δb là khoảng cách màu sắc trong không gian màu.
 N và M là các hệ số để cân bằng giữa khoảng cách không gian và khoảng cách màu sắc.

Mô hình Gradient Boosting (GB):

Gradient Boosting là một kỹ thuật học máy kết hợp tiên tiến có khả năng cải thiện hiệu suất và độ chính xác của mô hình dựa trên các mô hình cơ sở yếu hơn (Tran Van Anh và nnk., 2024). Nghiên cứu này sử dụng các mô hình cơ sở dựa trên cây quyết định, thông qua quá trình “boosting” để tạo ra một mô hình mạnh hơn. Quá trình thực hiện gồm các bước chính: khởi tạo mô hình ban đầu, tính toán phần dư (residuals) từ mô hình hiện tại, huấn luyện mô hình con yếu mới để dự đoán phần dư, và cập nhật mô hình dự đoán bằng cách kết hợp mô hình mới. Quá trình này lặp lại cho đến khi đạt tiêu chí dừng như số lượng mô hình con hoặc một ngưỡng lỗi cho phép. Nghiên cứu đã xây dựng mô hình Gradient Boosting với số lượng 80 cây quyết định, mỗi cây có độ sâu tối đa là

3, và học với hệ số học (learning rate) 0.1, đảm bảo mỗi cây quyết định trong số 80 cây sẽ đóng góp 10% vào mô hình tổng.

2.4. Phương pháp đánh giá độ chính xác

Nghiên cứu sử dụng các chỉ số về độ chính xác tổng thể (overall accuracy), hệ số Kappa (Kappa coefficient), và ma trận sai số (confusion matrix) để đánh giá độ chính xác của mô hình phân loại. Độ chính xác tổng thể (Overall Accuracy) là tỷ lệ giữa số lượng dự đoán chính xác và tổng số dự đoán, phản ánh phần trăm mẫu được phân loại đúng. Hệ số Kappa đo lường sự đồng thuận giữa các dự đoán của mô hình và giá trị thực tế, có tính đến khả năng đồng thuận ngẫu nhiên. Hệ số Kappa có giá trị từ 0 đến 1, với giá trị càng cao cho thấy mô hình càng chính xác. Công thức tính hệ số Kappa được thể hiện trong công thức (2):

$$K = \frac{P_o - P_e}{1 - P_e} \quad (2)$$

trong đó: P_o là tỷ lệ dự đoán chính xác và P_e là xác suất dự đoán ngẫu nhiên.

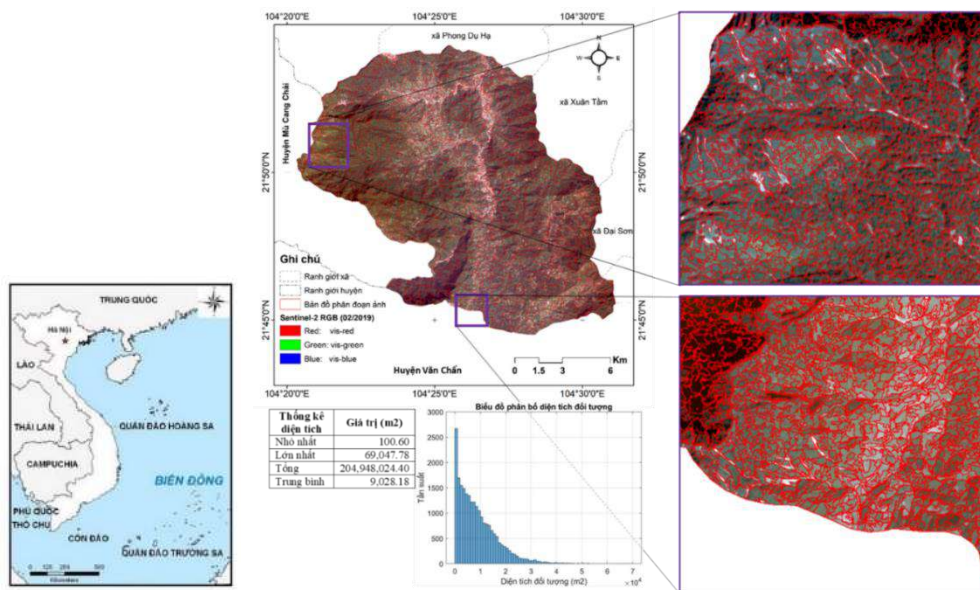
Ma trận sai số là phương pháp biểu diễn trực quan để đánh giá hiệu suất mô hình phân loại, thể hiện số lượng dự đoán đúng và sai cho từng loại lớp phủ. Đây là ma trận vuông với các đường chéo thể hiện số lượng các mẫu được phân loại chính xác, các ô còn lại thể hiện sai số giữa các phân lớp.

3. Kết quả và thảo luận

3.1. Kết quả phân mảnh đối tượng

Dựa vào đặc trưng phản xạ khác nhau của các loại lớp phủ đối với các kênh phổ, nghiên cứu sử dụng đầu vào mô hình phân đoạn là các kênh phổ có độ phân giải không gian 10 m ở bước sóng nhìn thấy và bước sóng cận hồng ngoại của ảnh vệ tinh Sentinel-2. Sau đó, kỹ thuật phân đoạn Simple Non-Iterative Clustering (SNIC) được sử dụng để phân đoạn hình ảnh trên toàn bộ khu vực nghiên cứu. Nghiên cứu đã tiến hành thử nghiệm các bộ tham số khác nhau để lựa chọn ra các giá trị tham số phù hợp của mô hình đối với khu vực nghiên cứu. Các giá trị tham số bao gồm khoảng cách của lưới điểm ban đầu (seed grid) bằng 80 pixel, phương thức kết nối giữa pixel trung tâm với các pixel lân cận (connectivity) bằng 8 được lựa chọn căn cứ vào hai tiêu chí là sự tương đồng về giá trị phản xạ phổ trên cùng một đối tượng và sự khác biệt của các đối tượng liền kề. Quá trình lựa chọn bản đồ phân đoạn phù hợp được thực hiện thông qua đối soát giữa bản đồ phân đoạn và các lớp dữ liệu ảnh (Hình 2).

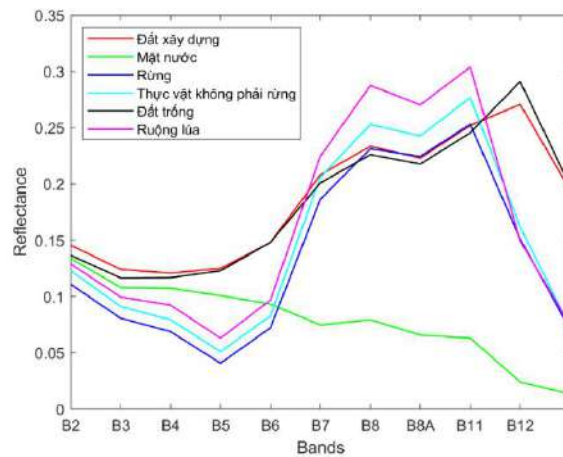
Theo thống kê, nghiên cứu đã phân mảnh bản đồ trên toàn bộ khu vực nghiên cứu với tổng số 22.699 đối tượng với các diện tích khác nhau. Trong đó, diện tích đối tượng nhỏ nhất là 100,60 m² và đối tượng lớn nhất là 69.047,78 m² và các đối tượng có diện tích trung bình là 9.028,18 m². Biểu đồ tần suất cho thấy số lượng đối tượng giảm đều khi diện tích lớn dần, số lượng các đối tượng có diện tích nhỏ hơn 1 ha có số lượng lớn hơn khi so sánh với các nhóm còn lại.



Hình 2. Bản đồ phân đoạn ảnh từ mô hình SNIC và ảnh viễn thám Sentinel-2

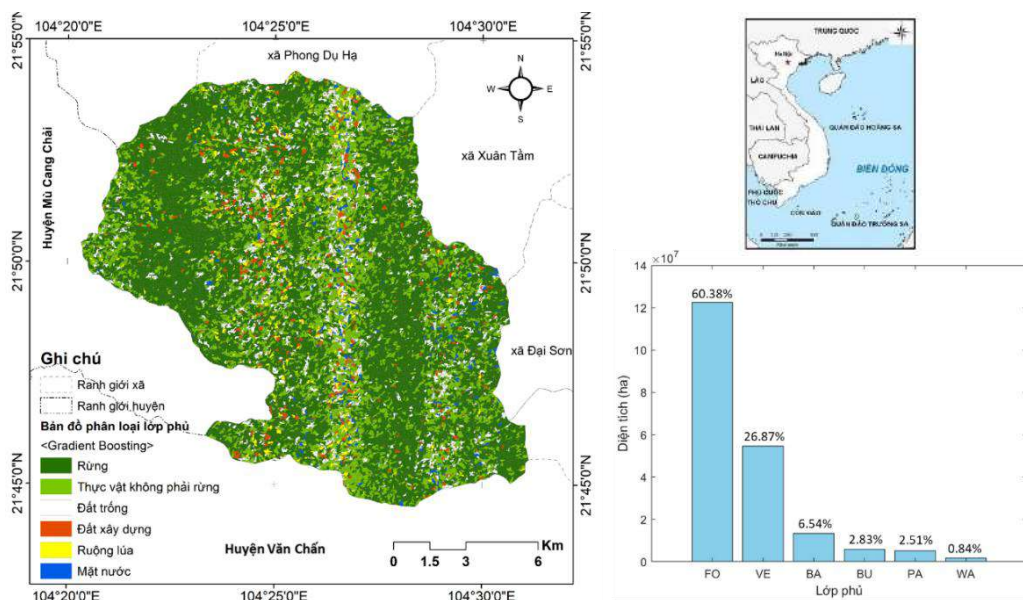
Sau khi thành lập bản đồ phân mảnh, nghiên cứu tiến hành lấy và phân tích đặc trưng phổ các mẫu trên các

kênh có độ phân giải không gian 10 m và 20 m của 06 loại lớp phủ thông qua điểm mẫu thu thập thực địa và thu thập, giải đoán ảnh từ Google Earth (Hình 3). Kết quả cho thấy, đặc trưng phổ có sự tương đồng giữa một số đối tượng. Căn cứ vào sự khác biệt tính chất phổ, có thể chia ra làm ba nhóm chính. Nhóm 1 bao gồm lớp phủ rừng, thực vật không phải rừng, và ruộng lúa có đặc trưng là phản xạ yếu với kênh phổ nhìn thấy nhưng phản xạ mạnh với các kênh cận hồng ngoại và hồng ngoại sóng ngắn. Có thể thấy rằng có sự khác nhau nhất định về giá trị phản xạ của kênh hồng ngoại và hồng ngoại sóng ngắn giữa ba loại lớp phủ này. Đây là tính chất được sử dụng để phân loại ba loại lớp phủ liên quan đến thực vật này. Nhóm 2 bao gồm lớp phủ đất đô thị và đất trống. Biểu đồ phân tích phổ cho thấy có sự tương đồng về phản xạ phổ của hầu hết các kênh. Tuy nhiên, có sự khác nhau nổi bật giữa phản xạ phổ của hai loại lớp phủ này ở kênh phổ SWIR2. Đây là lý do mặc dù kênh phổ hồng ngoại sóng ngắn có độ phân giải không gian là 20 m nhưng vẫn được ưu tiên sử dụng cho mô hình. Nhóm 3 là lớp phủ liên quan đến mặt nước. Kết quả cho thấy có sự khác biệt rõ rệt giữa lớp phủ này với hai nhóm còn lại, có thể phân biệt dễ dàng khi sử dụng kênh sóng nhìn thấy và cận hồng ngoại.



Hình 3. Biểu đồ đặc trưng phổ của các loại lớp phủ với các kênh phổ của ảnh vệ tinh Sentinel-2 tại khu vực nghiên cứu

3.2. Bản đồ phân loại lớp phủ và đánh giá độ chính xác



Hình 4. Bản đồ phân loại lớp phủ bề mặt xã Phong Dụ Thượng, huyện Văn Yên, tỉnh Yên Bái

Bản đồ phân loại lớp phủ xây dựng từ ảnh vệ tinh Sentinel-2 và mô hình Gradient Boosting được thể hiện trong Hình 4. Lớp phủ rừng chiếm diện tích lớn nhất với 12.254,04 ha, tương đương 60,38% tổng diện tích nghiên cứu. Tiếp theo đó là lớp phủ thực vật không phải rừng với 5.453,85 ha, chiếm 26,87%. Các lớp phủ khác như đất trống (BA), đất xây dựng (BU), Ruộng lúa (PA) và mặt nước (WA) chiếm tỷ lệ ít, lần lượt là 6,54%, 2,83%, 2,51%, và 0,84%. Kết quả cho thấy khu vực nghiên cứu có tỷ lệ diện tích rừng và mật độ thực vật cao,

diện tích đất xây dựng và đất canh tác chiếm tỷ lệ thấp và tương đương nhau. Trong nhiều năm trở lại đây, bên cạnh chính sách bảo tồn rừng, người dân và chính quyền địa phương đã có nhiều nỗ lực trong công tác phủ xanh đồi trọc. Tuy nhiên, bản đồ phân loại lớp phủ cho thấy vẫn tồn tại các khu vực đất trống nằm xen kẽ với rừng và thực vật. Ngoài ra, các hiện tượng thiên tai như lũ quét, lũ ống, trượt lở đất xảy ra hằng năm cũng gây tác động không nhỏ đến bề mặt lớp phủ của khu vực này.

Nghiên cứu đã sử dụng tổng số 61 điểm trong tệp kiểm tra để đánh giá hiệu suất mô hình phân loại. Kết quả cho thấy các phân lớp có độ chính xác cao nổi bật là thực vật không phải rừng với tổng số 11 trên 12 mẫu kiểm tra chính xác, tiếp theo đó là lớp phủ rừng với 12 trên 15 mẫu chính xác (Bảng 3). Lớp phủ đất xây dựng cho thấy hiệu suất phân loại thấp nhất trong các nhóm lớp phủ trên với 6 trên 8 mẫu được phân loại chính xác. Đặc trưng của khu vực nghiên cứu có mật độ xây dựng thấp, phân bố không tập trung dẫn đến một số khó khăn trong quá trình chọn mẫu của phân lớp này. Hơn nữa, sự tương đồng giữa đặc trưng phổ của khu vực đất trống và đất xây dựng như thể hiện trong Hình 4 cũng là một nguyên nhân của các sai số. Tuy nhiên, mô hình phân loại lớp phủ của khu vực nghiên cứu đã cung cấp một hiệu suất phân loại tốt với độ chính xác tổng thể đạt 0,871 và hệ số Kappa đạt 0,845.

Bảng 3. Ma trận sai số và tổng hợp thông tin đánh giá độ chính xác mô hình.

Loại lớp phủ	FO	VE	BA	BU	PA	WA	Tổng số
FO	12						12
VE	1	11					12
BA			10	1	1		12
BU		1	2	6			9
PA			1	1	8		10
WA	2					4	6
Tổng số	15	12	13	8	9	4	
Overall Accuracy (OA)= 0,871							
Kappa = 0,845							

4. Kết luận

Nghiên cứu đã xây dựng thành công bản đồ lớp phủ khu vực xã Phong Dụ Thượng, huyện Văn Yên, tỉnh Yên Bái từ ảnh viễn thám Sentinel-2 bằng phương pháp phân loại hướng đối tượng và mô hình học máy. Thông qua sử dụng các kênh phổ của ảnh vệ tinh và các điểm mẫu thu thập, nghiên cứu đã xác định được các khu vực của sáu loại lớp phủ chính bao gồm: rừng, thực vật không phải rừng, đất trống, ruộng lúa, mặt nước và đất xây dựng. Kết quả phân mảnh từ mô hình Simple Non-Iterative Clustering (SNIC) đã phân đoạn ảnh thành 22.699 đối tượng, và mô hình Gradient Boosting đã phân loại lớp phủ bề mặt đạt độ chính xác tổng thể là 0,871, chỉ số Kappa là 0,845. Kết quả phân loại cho thấy phương pháp thực hiện của nghiên cứu là hiệu quả và tin cậy. Kết quả của nghiên cứu cung cấp thông tin hữu ích và quan trọng trong việc thực hiện các công việc liên quan đến giám sát tài nguyên rừng, quy hoạch đất đai và ứng phó với các hiện tượng thiên tai.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Trường Đại học Xây dựng Hà Nội theo đề tài trọng điểm cấp cơ sở, mã số 04-2024/KHXD-TĐ.

Tài liệu tham khảo

- Trần Ngọc Tuấn, Vũ Thị Tuyết, Nguyễn Thanh Nga, Nông Thị Oanh, 2016. Ứng dụng công nghệ viễn thám trong thành lập bản đồ lớp phủ mặt đất theo hướng dẫn của IPCC phục vụ công tác giám sát tài nguyên môi trường và biến đổi khí hậu. *Tạp chí Khoa học Đo đạc và Bản đồ*, trang 39-45.
- Ngô Đức Anh, Đoàn Thị The, Nguyễn Thị Thu Thủy, 2016. Thành lập bản đồ lớp phủ rừng từ ảnh vệ tinh Landsat-8 bằng phương pháp phân loại dựa trên đối tượng kết hợp giải đoán bằng mắt thường. *Hội nghị ứng dụng GIS toàn quốc*, trang 1-11.
- Hồ Lệ Thu, Nguyễn Thanh Hoàn, Lê Minh Hằng, Đặng Xuân Phong, 2020. Xây dựng bản đồ lớp phủ khu vực Tây Nguyên sử dụng dữ liệu ảnh Landsat đa thời gian. *Tạp chí Khoa học - Công nghệ và đổi mới sáng tạo*. Số 3, trang 43-45.
- Vũ Thị Phương Thảo và Souksakone Sengchanh, 2022. 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. *Tạp chí Khoa học Đo đạc và Bản đồ*, trang 26-35.
- Senanayaka, J. B., Morawaliyadda, D. T., Senarath, S. T., Godaliyadda R. I., Ekanayake, M. P. 2020.

Adaptive Centroid Placement Based SNIC for Superpixel Segmentation, *2020 Moratuwa Engineering Research Conference (MERCon)*: 242-247.

Jamali, A. 2021. Land use land cover mapping using advanced machine learning classifiers. *Ekologia*, 40: 286-300.

Ajibola, S. and Cabral, P. 2024. A Review and Meta-analysis of Semantic Segmentation Models in Land Use/ Land Cover Mapping. *Review*: 01-28.

Saini, R. 2024. Land use land cover mapping and snow cover detection in Himalayan region using machine learning and multispectral Sentinel-2 satellite imagery. *International Journal of Information Technology*, 16: 675-686.

Tran, V. A., T. D. Khuc, X. Q. Truong, A. B. Nguyen and T. T. Phi (2024). "Application of potential machine learning models in landslide susceptibility assessment: A case study of Van Yen district, Yen Bai province, Vietnam. *Quaternary Science Advances*, 14: 100181.

ABSTRACT

Application of the object- based classification technique and the machine learning model in mapping land cover utilizing remote sensing optical images

Thanh Dong Khuc^{1*}, Dinh Trong Tran¹, Van Anh Tran^{2,3}, Quang Vinh Do¹, Trung Khien Ha¹,

Toan Duy Dao¹, Xuan Quang Truong⁴

¹Hanoi University of Civil Engineering

²Hanoi University of Mining and Geology

³Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology

⁴Hanoi University of Natural Resources & Environment

Land cover maps play an important role in resource management, environmental monitoring, land use planning, and sustainable development. This study combines the Simple Non-Iterative Clustering (SNIC) segmentation method and the Gradient Boosting machine learning model to create a land cover map for Phong Du Thuong commune, Van Yen district, Yen Bai province, using object- based classification. The input data includes visible, near-infrared, and short-wave infrared spectral bands from Sentinel-2 images, along with 203 sample points collected through field surveys and interpretation on Google Earth. The sample points focus on six characteristic land cover types: forest, non-forest vegetation, bare land, paddy fields, water, and built-up areas. The data samples were divided into 70% for training and 30% for model validation. The study successfully developed a land cover map for Phong Du Thuong commune from Sentinel-2 satellite images in February 2019. The classification results show high accuracy, with an overall accuracy of 0.871 and a Kappa coefficient of 0.845.

Keywords: Land cover maps, the Simple Non-Iterative Clustering method, Gradient Boosting machine learning model.

Xác định lún đất khu vực Cà Mau bằng chuỗi ảnh radar trên nền tảng điện toán đám mây Google Colab

Hà Trung Khiên¹, Trần Văn Anh^{2,3*}, Khúc Thành Đông¹, Nguyễn Đình Huy¹

¹Trường Đại học Xây dựng Hà Nội

²Trường Đại học Mở - Địa chất

³Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

TÓM TẮT

Tình trạng lún đất đang diễn ra phổ biến và phức tạp tại tỉnh Cà Mau, hiện tượng này bắt nguồn từ các yếu tố tự nhiên và các hoạt động của con người. Công nghệ viễn thám radar đã trở thành một công cụ hiệu quả trong việc theo dõi và giám sát lún đất trên một khu vực rộng lớn. Quá trình xử lý ảnh radar để xác định lún đất không chỉ yêu cầu kiến thức chuyên môn và kinh nghiệm, mà còn cần đến sự hỗ trợ của hệ thống phần cứng mạnh mẽ để xử lý dữ liệu đa thời gian. Do vậy, quá trình xử lý sẽ đòi hỏi chi phí cao và mất nhiều thời gian. Nghiên cứu này tập trung đánh giá khả năng ứng dụng nền tảng trực tuyến Google Colab sử dụng kỹ thuật xử lý SBAS (Small Baseline Subset) để xử lý chuỗi 24 ảnh radar khẩu độ tổng hợp từ 01/2022 đến 12/2023 khu vực thành phố Cà Mau và một số khu vực lân cận. Bằng cách so sánh kết quả từ Google Colab với kết quả dữ liệu đo đạc thực địa, nghiên cứu sẽ đánh giá về sự hiệu quả, độ chính xác của kết quả xác định lún bằng kỹ thuật SBAS trên nền tảng Google Colab, kết quả cho độ tương quan $R^2 = 0,71$ và RMSE = 6 mm.

Từ khoá: Lún đất, InSAR, SBAS, Google Colab.

1. Đặt vấn đề

Hiện nay, kỹ thuật InSAR (Interferometric Synthetic Aperture Radar) đã và đang được nghiên cứu trong xác định lún đất bằng nhiều phương pháp như phương pháp giao thoa vi phân D-InSAR (Differential Interferometric Synthetic Aperture Radar), phương pháp giao thoa sử dụng tán xạ cố định PS-InSAR (Permanent Scatterers Interferometric Synthetic Aperture Radar), phương pháp SBAS (Small Baseline Subset) hoặc kết hợp các phương pháp tại nhiều khu vực trên thế giới. Nghiên cứu ứng dụng công nghệ viễn thám radar trong xác định lún đất đầu tiên vào năm 1969, Roger và cộng sự đã thành công trích xuất được thông tin bề mặt sao Kim và Mặt trăng (Roger, 1969). Đến năm 1974, pha địa hình đã được Graham loại bỏ trong phương pháp giao thoa ảnh SAR để xây dựng bản đồ địa hình và xác định độ dịch chuyển bề mặt địa hình (Graham và cộng sự, 1974). Howard và các cộng sự đã lần đầu công bố nghiên cứu chứng minh kỹ thuật D-InSAR bị ảnh hưởng bởi nhiễu do điều kiện khí quyển không ổn định và do điều kiện không gian, thời gian gây ra (Zebker và cộng sự, 1992). Phương pháp PS-InSAR lần đầu được Ferretti và cộng sự đề xuất (Ferretti và cộng sự, 2000). Phương pháp này sử dụng nhiều cặp ảnh giao thoa để xác định các điểm ảnh có độ tương quan cao, từ đó nội suy và xác định biến dạng bề mặt cho cả vùng. Tuy nhiên, phương pháp này đòi hỏi phải có sự xuất hiện của các điểm tán xạ cố định (PS - Persistent Scatterers) trong suốt thời gian thu nhận dữ liệu, đây cũng là hạn chế của phương pháp này khi thực hiện trong các khu vực không có nhiều điểm tán xạ ổn định, chẳng hạn như các khu vực nông thôn, rừng hoặc sa mạc. Với một cách tiếp cận khác, Usai và cộng sự đã lựa chọn để tạo một loạt cặp ảnh giao thoa với đường đáy ảnh nhỏ và thời gian giữa hai lần chụp ngắn, sau đó sử dụng phương pháp bình phương nhỏ nhất để phân tích biến dạng theo thời gian từ đó giảm thiểu ảnh hưởng từ nhiễu phân tán không gian - thời gian. Tuy nhiên, hạn chế trong việc tạo nhiều nhóm dữ liệu đôi khi dẫn đến không thể lập và giải phương trình bình phương nhỏ nhất. Để giải quyết hạn chế này, năm 2002, Berardino và cộng sự đã tiếp tục nghiên cứu áp dụng thuật toán phân tích giá trị suy biến (SVD) và đề xuất phương pháp SBAS trong xử lý chuỗi ảnh SAR (Berardino và cộng sự, 2002). SBAS có thể xử lý các vùng có thay đổi nhanh chóng về điều kiện mặt đất hoặc các thay đổi không liên tục. Từ đó đến nay, kỹ thuật SBAS đã được ứng dụng để xác định biến dạng bề mặt ở nhiều khu vực trên thế giới. Một số nghiên cứu điển hình như nghiên cứu của Yalvac và cộng sự đã sử dụng dữ liệu vệ tinh Sentinel-1A trong giai đoạn từ 2014 đến 2018 để giám sát sự sụt lún ở thành phố Mexico và các tỉnh Aguascalientes bằng kỹ thuật SBAS (Sefa Yalvac, 2020). Nghiên cứu (Bo Xiao và cộng sự, 2022) được Xiao và cộng sự xác định lún đất tại thành phố Côn Minh, Trung Quốc từ tháng 7 năm 2018 đến tháng 11 năm 2020 bằng công nghệ PS-InSAR và SBAS, đồng thời so sánh kết quả của hai phương pháp với kết quả đo thủy chuẩn với sai số trung bình bình phương (RMSE) lần lượt là 6,5 mm/năm và 8,5 mm/năm. Yuejuan Chen và cộng sự đã kết hợp cả ba phương pháp PS-InSAR, D-InSAR và SBAS để xác định các điểm PS (Persistent Scatterers) và DS

* Tác giả liên hệ

Email: tranvananh@humg.edu.vn

(Distributed Scatterers) từ dữ liệu gồm 45 ảnh Sentinel-1B tại khu vực thành phố Thông Liâu - Mông Cổ (Yuejuan Chen và cộng sự, 2024). Kết quả nghiên cứu đã so sánh với kết quả đo bằng GPS cho sai số RMSE là 14 mm.

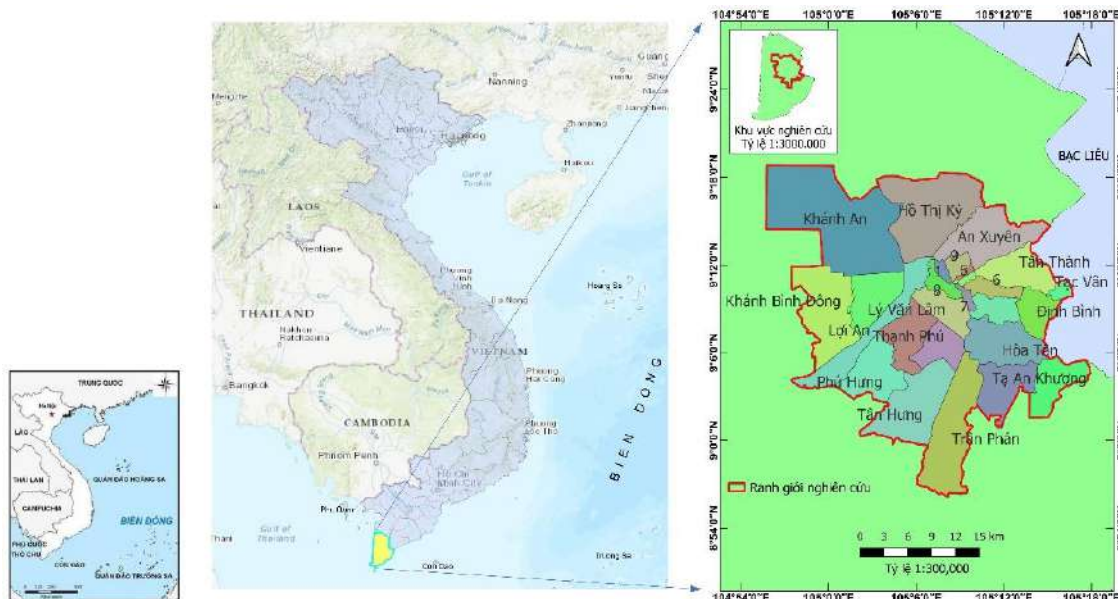
Tại Việt Nam, các phương pháp D-InSAR, PS-InSAR và SBAS cũng đã được nghiên cứu và ứng dụng để xác định lún đất tại một số khu vực. Nghiên cứu của tác giả Trần Văn Anh và cộng sự vào năm 2007 (Tran Van Anh 2007) đã xử lý ba ảnh JERS-1 (band L) bằng phương pháp giao thoa vi phân trong xác định lún tại Hà Nội. Một nghiên cứu khác vào năm 2014 của Đặng Vũ Khắc và cộng sự đã sử dụng ảnh ALOS PalsAR-1 xác định lún do sự phát triển đô thị với kết quả cho thấy độ chính xác cao, có hệ số tương quan R2 đạt 0,98 (Dang Vu Khac, 2014). Năm 2020, Luyen K.B và cộng sự đã sử dụng các phương pháp như PS-InSAR và SBAS trong xử lý ảnh Sentinel-1 để phân tích lún đất tại Hà Nội (Luyen K.B và cộng sự, 2020). Trong phạm vi Đồng bằng sông Cửu Long bao gồm khu vực Cà Mau, nghiên cứu (Erban và cộng sự, 2014) đã ứng dụng công nghệ InSAR để xử lý 121 ảnh từ 2006 đến 2010 tạo ra 78 ảnh giao thoa, sau đó sử dụng phương pháp stacking (chập dữ liệu) để xác định sự dịch chuyển pha trung bình. Phương pháp PS-InSAR đã được ứng dụng để xác định lún đất khu vực Đồng bằng sông Cửu Long trong nghiên cứu (Philip S. J. Minderhoud và cộng sự, 2020) giai đoạn từ 2014 đến 2019.

Các nghiên cứu trên chủ yếu sử dụng các phần mềm chạy trực tiếp trên máy tính như StaMPS/MTI (Stanford Method for Persistent Scatterers/Multi-Temporal InSAR), ESA SNAP (European Space Agency's Sentinel Application Platform), SARscape (Synthetic Aperture Radar Scape),... Khi xử lý bằng các phần mềm này, đòi hỏi hệ thống máy tính phải đảm bảo về cấu hình và bộ nhớ lưu trữ đủ lớn, đặc biệt khi xử lý chuỗi ảnh radar đa thời gian trên một phạm vi lớn. Để giải quyết những hạn chế này, các ứng dụng nền tảng trực tuyến đã được xây dựng và phát triển sử dụng trực tiếp tài nguyên trên máy chủ của các hãng công nghệ lớn như Google, Microsoft,... Điều này giúp tăng hiệu suất và chi phí khi xử lý dữ liệu lớn. Google Colab là một nền tảng trực tuyến của Google đã được ứng dụng nhiều trong xử lý dữ liệu viễn thám với những ưu điểm vượt trội (Johary và cộng sự, 2023). Trong bài báo này, phương pháp SBAS được sử dụng để xử lý chuỗi 24 ảnh radar đa thời gian trên nền tảng trực tuyến Google Colab với mục đích giám sát lún đất khu vực thành phố Cà Mau và các vùng lân cận từ 01/2022 đến 12/2023.

2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1. Giới thiệu khu vực nghiên cứu

Khu vực nghiên cứu gồm toàn bộ diện tích thành phố Cà Mau và một số xã giáp ranh với thành phố Cà Mau thuộc các huyện Đầm Dơi, Cái Nước, Trần Văn Thới, U Minh, Thới Bình nằm ở tỉnh Cà Mau thuộc phía Nam lãnh thổ Việt Nam (Hình 1). Khu vực này địa hình chủ yếu là đồng bằng, với độ cao trung bình từ 0,4 đến 0,6 m trên mực nước biển. Một số khu vực có cao trình khoảng 0,2 m ở vùng thấp và từ 0,8 đến 1,1 m ở các vùng cao hơn. Địa hình có xu hướng dốc nhẹ từ Bắc xuống Nam và từ Đông Bắc xuống Tây Bắc. Trên khu vực có hệ thống sông ngòi và kênh rạch phức tạp, giúp tưới tiêu cho nông nghiệp. Khu vực phía Đông và phía Nam chủ yếu được sử dụng cho nuôi tôm thâm canh và bán thâm canh. Khu vực phía Bắc và phía Tây chủ yếu được sử dụng cho trồng lúa hai vụ, rau màu và nông nghiệp nước ngọt.



Hình 1. Khu vực nghiên cứu

2.2. Dữ liệu sử dụng

Trong bài báo, dữ liệu sử dụng là chuỗi ảnh gồm 24 ảnh của vệ tinh Sentinel-1A từ 01/2022 đến 12/2023 tại khu vực thuộc tỉnh Cà Mau. Dữ liệu ảnh Sentinel-1A là sản phẩm của chương trình quan sát Trái đất Copernicus của Cơ quan Vũ trụ châu Âu (ESA) được trang bị radar khẩu độ tổng hợp băng tần C, có chu kỳ chụp lại là 12 ngày ở chế độ IW (Interferometric Wide Swath) với phân cực VV (Bảng 1). Các tập dữ liệu Sentinel-1A đã được thu thập từ cơ sở dữ liệu vệ tinh Alaska (<https://search.asf.alaska.edu/>). Cùng với đó, nghiên cứu còn sử dụng dữ liệu mô hình số độ cao DEM SRTM do NASA cung cấp với độ phân giải mặt đất 30 m.

Bảng 1. Thông số của ảnh vệ tinh Sentinel-1A trong nghiên cứu

Tham số	Giá trị
Hướng bay	Hướng xuống
Chế độ thu nhận dữ liệu	IW
Định dạng dữ liệu	SLC
Phân cực	VV
Băng tần	C
Bước sóng/cm	5,6
Độ phân giải không gian	20m
Độ rộng dải quét	250km
Số lượng ảnh	24
Thời gian quan trắc	01/2022 - 12/2023

2.3. Cơ sở toán học của phương pháp SBAS

SBAS (Small Baseline Subset) là một phương pháp tiên tiến được sử dụng trong xử lý ảnh radar khẩu độ tổng hợp (SAR) để theo dõi và đo lường sự biến dạng bề mặt Trái đất do Berardino đề xuất năm 2002 (Berardino và cộng sự, 2002). Phương pháp SBAS dựa trên việc phân tích các tập hợp con của ảnh SAR có đường đáy ảnh (baseline) nhỏ để giảm thiểu các lỗi liên quan đến điều kiện quan sát và nhiễu tín hiệu. Phương pháp SBAS có cơ sở toán học như sau:

Giả sử có $N+1$ hình ảnh SAR tại các thời điểm t_0, t_1, \dots, t_N ghi nhận hình ảnh cùng một khu vực, hai hình ảnh thỏa mãn các điều kiện về cơ sở không gian và thời gian được lựa chọn để tạo giao thoa với nhau. Sau khi căn chỉnh chính xác, có thể thu được M cặp ảnh giao thoa:

$$\frac{N+1}{2} \leq M \leq \frac{N \cdot (N+1)}{2} \quad (1)$$

Chọn thời điểm t_0 làm thời điểm tham chiếu, ϕ_i là pha vi phân tại thời điểm t_i ($i = 1, 2, \dots, N$). Trên ảnh giao thoa vi phân thứ i được tạo ra từ cặp ảnh ở 2 thời điểm t_A và t_B , giá trị pha giao thoa của bất kỳ điểm ảnh (r, x) là:

$$\begin{aligned} \delta\phi(r, x) &= \phi(t_A, r, x) - \phi(t_B, r, x) \\ &\approx \delta\phi_{disp,i} + \delta\phi_{topo,i} + \delta\phi_{atm,i} + \delta\phi_{noise,i} \end{aligned} \quad (2)$$

trong đó: $\delta\phi_{disp,i}$ là pha biến dạng bề mặt theo đường ngắm; $\delta\phi_{topo,i}$ là lỗi pha địa hình; $\delta\phi_{atm,i}$ là lỗi pha khí quyển; $\delta\phi_{noise,i}$ là lỗi pha nhiễu. Sau khi đã loại bỏ pha địa hình, pha khí quyển và pha nhiễu công thức (2) trở thành:

$$\delta\phi_i(r, x) \approx \delta\phi_{disp,i} = \frac{4\pi}{\lambda} [d(t_A, r, x) - d(t_B, r, x)] \quad (3)$$

trong đó: λ biểu thị bước sóng; $d(t_A, r, x)$ và $d(t_B, r, x)$ biểu thị các biến dạng tích lũy của bề mặt đất tại các thời điểm t_A và t_B cho phần tử hình ảnh (r, x) so với thời điểm ban đầu t_0 .

Thời gian của các hình ảnh chính và phụ là $IE = [IE_1, \dots, IE_M]$, $IS = [IS_1, \dots, IS_M]$.

Nếu $IE_j > IS_j$ ($j = 1, 2, \dots, M$), thì phương trình của pha giao thoa vi phân là:

$$\delta\phi_j = \phi(t_{IEj}) - \phi(t_{ISj}) \quad (4)$$

Sau đó, mô hình biến dạng tuyến tính của SBAS là:

$$A \cdot \phi = \delta\phi \quad (5)$$

Vì ma trận A có hạng đầy đủ, sử dụng phương pháp bình phương nhỏ nhất để giải ước lượng phương trình (5) ta được:

$$\phi = (AT \cdot A)^{-1} \cdot AT \cdot \delta\phi \quad (6)$$

Quy trình ứng dụng kỹ thuật SBAS trong xử lý chuỗi ảnh radar khẩu độ tổng hợp SAR đa thời gian để giám sát biến dạng bề mặt gồm các bước chính:

(1) Thu thập dữ liệu SAR: Thu thập 24 ảnh SAR của vệ tinh Sentinel-1A từ 01/2022 đến 12/2023.

(2) Tiền xử lý dữ liệu ảnh SAR: Sử dụng các phương pháp hiệu chỉnh hình học để đảm bảo rằng tất cả các ảnh SAR được căn chỉnh chính xác và áp dụng các bộ lọc để lọc nhiễu.

(3) Tạo các cặp ảnh giao thoa: Tính toán ghép các cặp ảnh SAR với khoảng cách thời gian và không gian nhỏ để giảm thiểu các lỗi hình học. Sau đó thực hiện giao thoa vì phân các cặp ảnh đã tạo để xác định sự khác biệt về pha giữa hai cảnh ảnh. Áp dụng các phương pháp để loại bỏ pha nhiễu do thay đổi khí quyển và lỗi hình học.

(4) Giải mở pha: Sử dụng các thuật toán như SNAPHU (Statistical-Cost Network-Flow Algorithm for Phase Unwrapping) để giải mở pha các cặp giao thoa, chuyển đổi từ pha liên tục sang pha tuyệt đối.

(5) Phân tích chuỗi thời gian biến dạng: Sử dụng mô hình số độ cao (DEM) để hiệu chỉnh ảnh hưởng của địa hình đến các cặp giao thoa. Sau đó sử dụng phương pháp SBAS để phân tích chuỗi thời gian của các cặp giao thoa đã giải mở pha, xác định biến dạng bề mặt theo thời gian.

2.4. Nền tảng trực tuyến Google Colab

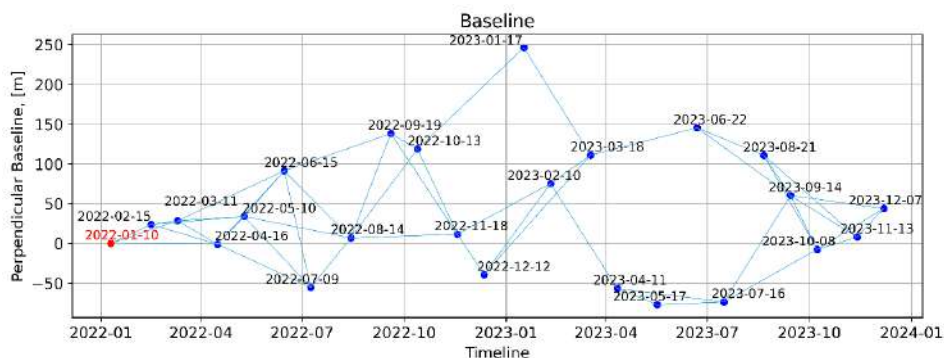
Google Colab (Google Colaboratory) là một nền tảng dịch vụ trực tuyến miễn phí của Google cho phép người dùng có thể biên tập, thực thi mã code ngôn ngữ lập trình Python dưới dạng sổ tay. Do Google Colab chạy và sử dụng tài nguyên trực tuyến của Google, nên khi sử dụng Google Colab, không yêu cầu cấu hình máy tính cao của người sử dụng, chỉ cần một tài khoản Google và kết nối mạng ổn định, người dùng có thể sử dụng miễn phí trên nhiều thiết bị di động. Quá trình xử lý chuỗi ảnh radar là một quá trình phức tạp và tốn nhiều tài nguyên máy tính, việc ứng dụng nền tảng trực tuyến Google Colab sẽ giúp người dùng đảm bảo đủ cấu hình và tài nguyên lưu trữ. Bài báo sử dụng thư viện PyGMTSAR chạy trên nền tảng Google Colab để thu thập, xử lý và phân tích chuỗi ảnh radar bằng phương pháp SBAS. Thư viện PyGMTSAR là một thư viện mã nguồn mở được xây dựng bằng ngôn ngữ lập trình Python.



Hình 2. Nền tảng trực tuyến Google Colab

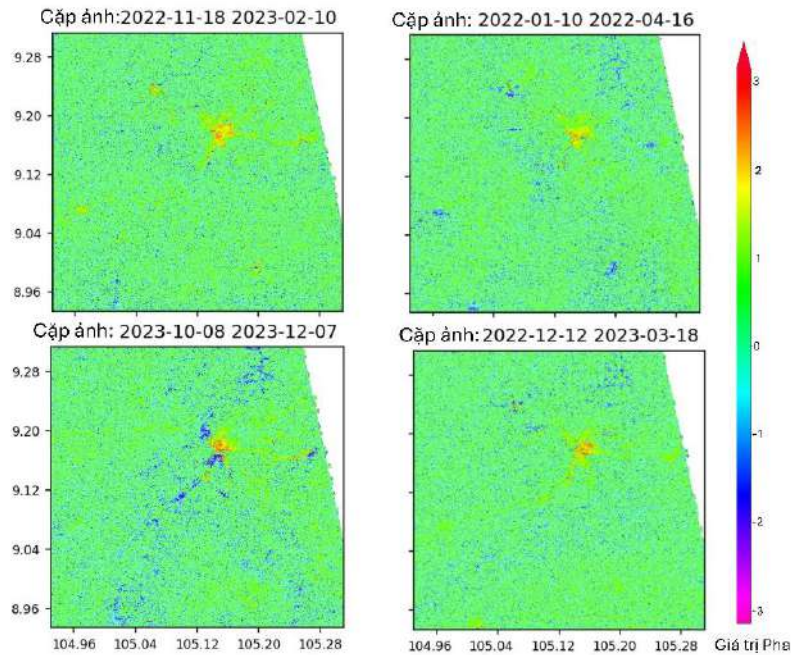
3. Kết quả và thảo luận

Như đã trình bày trong Mục 2.2, bài báo sử dụng dữ liệu gồm 24 cảnh ảnh radar khẩu độ tổng hợp SAR đa thời gian với đường đáy ảnh ngắn (Hình 3).



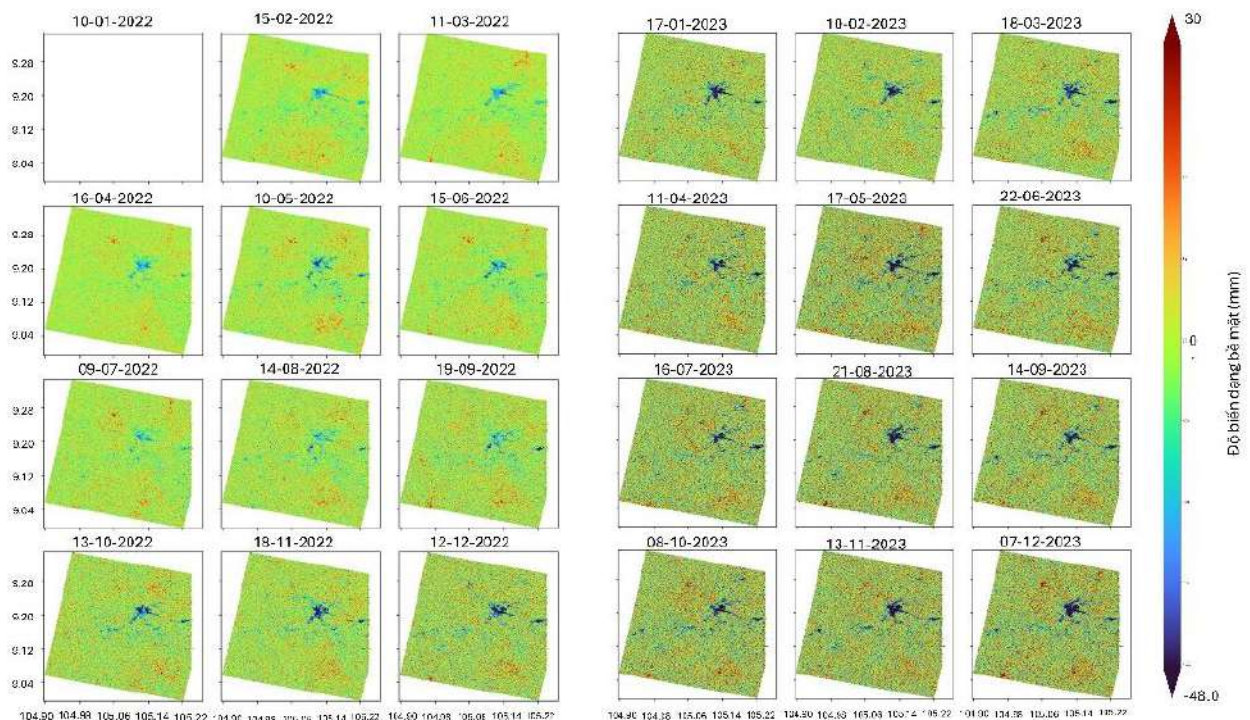
Hình 3. Biểu đồ phân bố đường đáy của tập hợp ảnh Sentinel-1

Sau khi thu thập dữ liệu ảnh thực hiện căn chỉnh với các tệp quỹ đạo chính xác và cắt để khớp chính xác với khu vực nghiên cứu. Quá trình xử lý giao thoa của các cặp ảnh bao gồm việc sử dụng phương pháp lọc Goldstein để loại bỏ nhiễu, giúp tạo ra các ảnh giao thoa rõ ràng khi đã loại bỏ các ảnh hưởng của pha địa hình. Bên cạnh đó và kỹ thuật Minimum Cost Flow (MCF) đã được sử dụng để giải mờ pha. Từ 24 cảnh ảnh đã tạo ra 50 ảnh giao thoa đã được lọc. Hình 4 gồm 4 ảnh giao thoa được chọn với các vân giao thoa rõ ràng, chỉ ra sự biến dạng trong khu vực nghiên cứu.



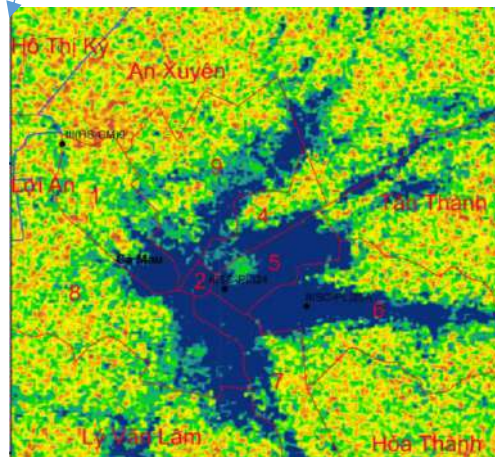
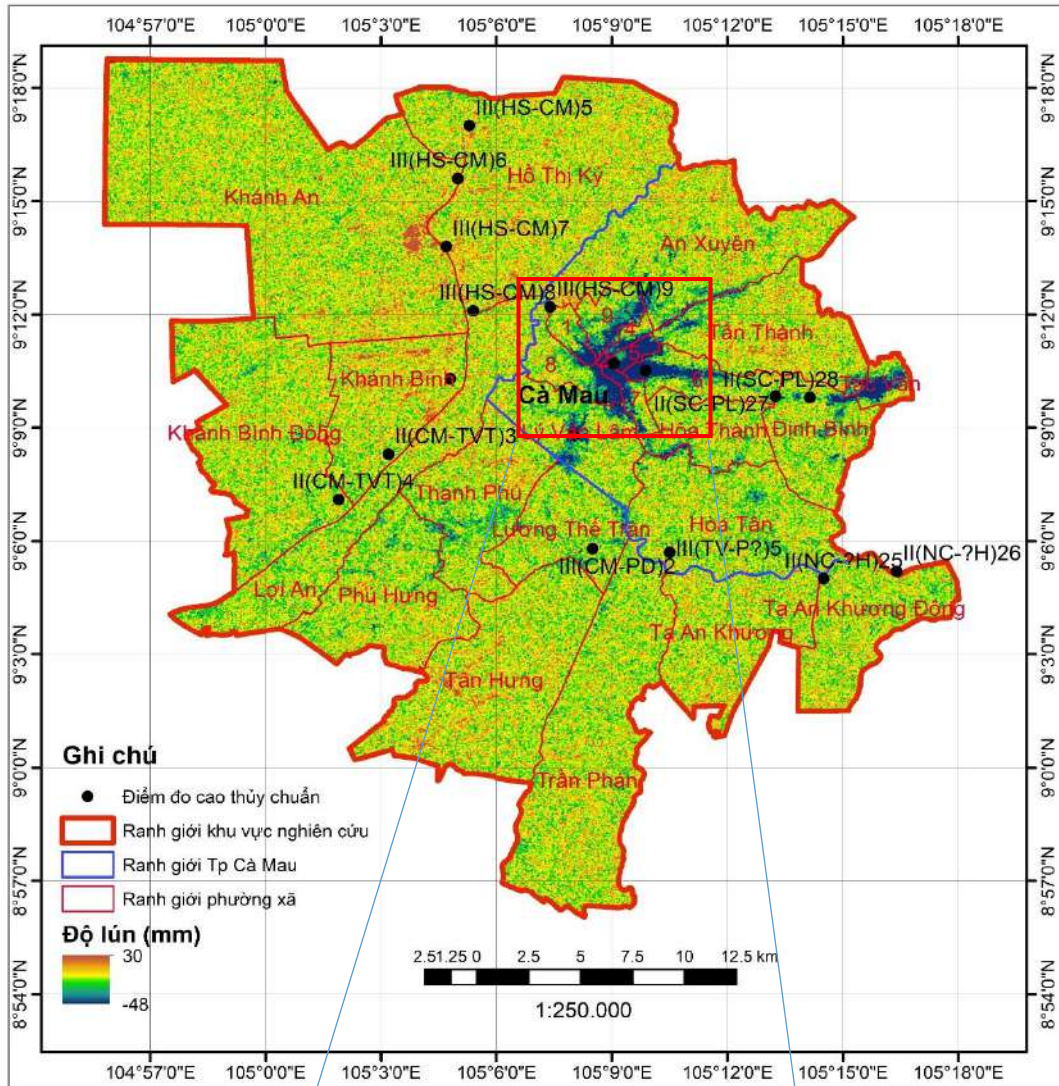
Hình 4. Ảnh giao thoa điển hình

Sau khi sử dụng các kỹ thuật lọc nhiễu để ra được 50 ảnh giao thoa rõ ràng nhất, giải mờ pha bằng thuật toán Minimum Cost Flow (MCF) trên phần mềm SNAPHU đã được cài đặt trong máy chủ Google cuối cùng sử dụng phân tích SBAS để tính toán độ lún theo chuỗi 24 mốc thời gian như Hình 5.

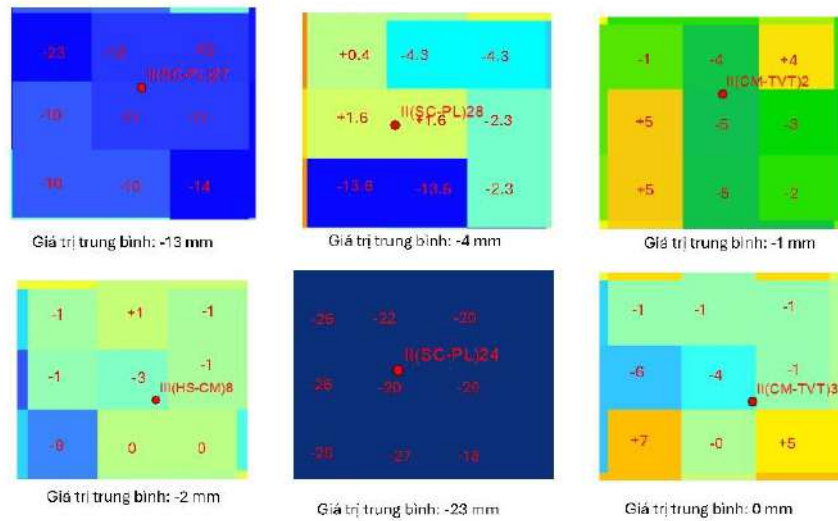


Hình 5. Chuỗi ảnh phân tích biến dạng địa hình bằng SBAS

Hình 6 thể hiện độ lún tổng từ tháng 01/2022 đến tháng 12/2023 sau khi tổng hợp và phân tích bằng phương pháp SBAS. Các vị trí màu xanh lam thể hiện tình trạng lún nhiều nhất trên khu vực thành phố Cà Mau và các vùng lân cận. Từ kết quả trên bản đồ có thể thấy các vị trí lún chủ yếu tập trung ở Phường 2, 5 và các khu vực giáp ranh của hai phường này với các Phường 1, 4, 6, 7, 8, 9. Độ lún lớn nhất trên cả khu vực được ghi nhận bằng SBAS là -48 mm. Độ lún trung bình trên toàn bộ khu vực nghiên cứu là -20 mm.



Hình 6. Bản đồ tổng lún từ 01/2022 đến 12/2023



Hình 7. Tính giá trị lún trung bình của các điểm ảnh xung quanh các điểm thủy chuẩn

Để đánh giá độ tin cậy của kết quả lún xác định bằng phương pháp SBAS chạy trên Google Colab, nghiên cứu đã so sánh với kết quả đo lún thực địa của 13 điểm độ cao đã được xác định độ lún bằng phương pháp đo cao thủy chuẩn (Hình 6). Do ảnh xuất ra với độ phân giải điểm ảnh là 30 m x 30 m, nên để xác định độ lún của các điểm thủy chuẩn tương ứng trên ảnh, nghiên cứu đã lấy trung bình giá trị của 9 điểm ảnh xung quanh điểm thủy chuẩn. Hình 7 thể hiện giá trị của các điểm ảnh xung quanh các điểm thủy chuẩn II(SC-PL)27, II(SC-PL)28, II(CM-TVT)2, III(HS-CM)8, II(CS-PL)24, II(CM-TVT)3. Lần lượt trích xuất các giá trị điểm ảnh xung quanh các điểm thủy chuẩn và tính trung bình các giá trị đó để được giá trị thay đổi địa hình theo hướng tín hiệu radar, áp dụng công thức (7) để tính giá trị dịch chuyển theo phương thẳng đứng.

$$\Delta h = \frac{\Delta LOS}{\cos \theta} \quad (7)$$

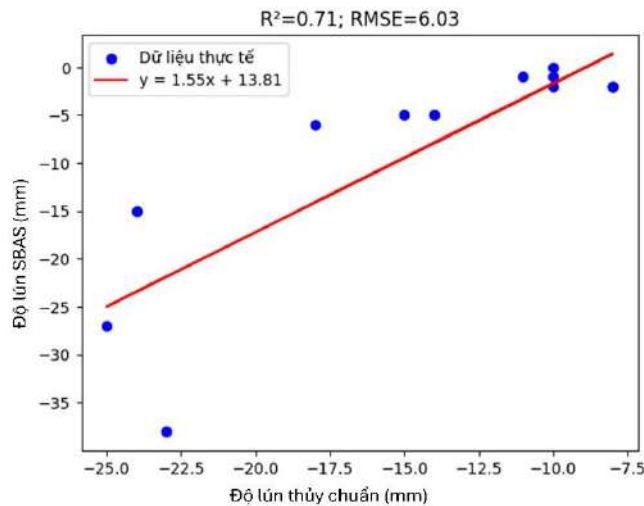
trong đó: Δh là giá trị dịch chuyển theo phương đứng, ΔLOS là giá trị dịch chuyển theo hướng tín hiệu radar và θ là góc tới. Với khu vực nghiên cứu nằm hoàn toàn trong IW1, vậy giá trị góc tới trung bình được lấy là 30,5°.

Sau khi trích xuất và tính toán được giá trị dịch chuyển so sánh với giá trị độ lún bằng đo thủy chuẩn được bảng so sánh kết quả như Bảng 2.

Bảng 2. So sánh độ lún giữa SBAS và thủy chuẩn

STT	Tên điểm	Độ lún SBAS		Độ lún đo thủy chuẩn (mm)	Độ lệch (mm)
		Độ lún theo hướng tín hiệu Radar (mm)	Độ lún theo phương đứng (mm)		
1	II(CM-TVT)3	0	0	-10	10
2	II(CM-TVT)2	-1	-1	-11	10
3	II(NC-H)25	-5	-6	-18	12
4	II(NC-H)26	-2	-2	-8	6
5	II(SC-PL)24	-23	-27	-25	-2
6	II(SC-PL)25A	-33	-38	-23	-15
7	II(SC-PL)27	-13	-15	-24	9
8	II(SC-PL)28	-4	-5	-15	10
9	III(HS-CM)5	-2	-2	-8	6
10	III(HS-CM)6	-2	-2	-8	6
11	III(HS-CM)8	-2	-2	-10	8
12	III(HS-CM)9	-4	-5	-14	9
13	III(TV-P)5	-1	-1	-10	9

Sử dụng mô hình hồi quy tuyến tính để đánh giá sự tương quan giữa kết quả lún xác định bằng SBAS và bằng thủy chuẩn. Kết quả cho thấy $R^2 = 0,71$ và $RMSE = 6 \text{ mm}$ (Hình 8).



Hình 8. Đánh giá tương quan giữa số liệu lún bằng SBAS và đo thủy chuẩn

4. Kết luận

Bài báo đã sử dụng phương pháp SBAS được chạy nền tảng trực tuyến Google Colab với 24 cảnh ảnh trong thời gian xử lý 90 phút, sử dụng 140 GB bộ nhớ đám mây của máy chủ Google. Từ đó có thể thấy sự tiện lợi khi sử dụng nền tảng Google Colab trong xử lý chuỗi nhiều ảnh radar so với việc sử dụng các phần mềm trên máy tính tốn nhiều tài nguyên và yêu cầu cao về cấu hình máy tính.

Kết quả độ lún khu vực thành phố Cà Mau và các xã lân cận cho thấy hiện tượng lún tập trung chủ yếu ở trung tâm thành phố Cà Mau, cụ thể ở Phường 2, 5 và các khu vực giáp ranh của hai phường này với các Phường 1, 4, 6, 7, 8, 9. Độ lún trung bình của toàn khu vực từ 01/2022 đến 12/2023 là -20 mm. So sánh với kết quả đo thực tế bằng phương pháp thủy chuẩn, kết quả xác định độ lún bằng SBAS trên nền tảng Google Colab cho kết quả độ lún với độ tin cậy khá tốt khi với $R^2 = 0,71$ và $RMSE = 6 \text{ mm}$.

Lời cảm ơn

Nhóm nghiên cứu chân thành cảm ơn sự hỗ trợ về mặt tài chính từ đề tài nghiên cứu khoa học công nghệ cấp trường trọng điểm - Trường Đại học Xây dựng Hà Nội, mã số: 05-2024/KHXD-TĐ.

Tài liệu tham khảo

- A. Rogers, R.V.I., 1969. Mapping the surface reflectivity by radar interferometry. *Science* 165, pp. 797-799.
- Graham, L.C., 1974. Synthetic interferometer radar for topographic mapping. *Engrs.* 62, pp. 763-768.
- H.A. Zebker, J.V., 1992. Decorrelation in interferometric radar echoes. *IEEE Trans. Geosci. Rem. Sens.* 30, pp. 950-959.
- A. Ferretti, C.P., F. Rocca., 2000. Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. *IEEE Trans. Geosci. Rem. Sens.* 38, pp. 2202-2212.
- P. Berardino, G.F., R. Lanari, E. Sansosti, 2002. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Trans. Geosci. Rem. Sens.* 40, pp. 2375-2383.
- S., Y., 2020. Validating InSAR-SBAS results by means of different GNSS analysis techniques in medium- and high-grade deformation areas. *Environmental Monitoring and Assessment* 192, pp. 1-12.
- Xiao, B., et al., 2022. The Monitoring and Analysis of Land Subsidence in Kunming (China) Supported by Time Series InSAR. *Sustainability* 14, pp. 12387.
- Chen, Y., et al., 2024. Research on Time Series Monitoring of Surface Deformation in Tongliao Urban Area Based on SBAS-PS-DS-InSAR. *Sensors* 24, pp. 1169.
- Tran Van Anh, S.M., Kiyoshi Shiono and Venkatesh Raghavan, 2007. Accuracy of low relief topographical maps derived from JERS-1 SAR interferometry, experiment in Hanoi, Vietnam. *Journal of Geosciences-OCU-Japan*, 50.
- Dang, V.K., C. Doubre, C. Weber, N. Gourmelen, and F. Masson, 2014. Recent land subsidence caused by the rapid urban development in the Hanoi region (Vietnam) using ALOS InSAR data. *Natural Hazards and Earth System Sciences* 14, pp. 657-674.

Bui, L.K., W. Featherstone, and M. Filmer, 2020. Disruptive influences of residual noise, network configuration and data gaps on InSAR-derived land motion rates using the SBAS technique. *Remote Sensing of Environment*. 247, p. 111941.

Erbas, L., S. Gorelick, and H. Zebker, 2014. Groundwater extraction, land subsidence, and sea-level rise in the Mekong Delta, Vietnam. *Environmental Research Letters* 9, p. 084010.

Philip S. J. Minderhoud, I.H., Jan Kolomaznik, and Olaf Neussner, 2020. Towards unraveling total subsidence of a mega-delta - the potential of new PS InSAR data for the Mekong delta. *Proc. IAHS*. 382, pp. 327-332.

Johary, A.R.F., et al., 2023. Detection of Large-Scale Floods Using Google Earth Engine and Google Colab. *Remote Sensing* 15.

Berardino P, F.G., Lanari R and Sansosti E, 2002. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 40, pp. 2375-2383.

ABSTRACT

Determining land subsidence in the Ca Mau area using a series of radar images on Google Colab Cloud Computing Platform

Ha Trung Khiem¹, Tran Van Anh^{2,3*}, Khuc Thanh Dong¹, Nguyen Dinh Huy¹

¹Hanoi University of Civil Engineering

²Hanoi University of Mining and Geology

³Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology

Land subsidence is a prevalent and complex issue in Ca Mau province, originating from both natural factors and human activities. Radar remote sensing technology has become an effective tool in monitoring and supervising land subsidence over a large area. The process of processing radar images to determine land subsidence not only requires specialized knowledge and experience but also necessitates the support of powerful hardware systems to handle multi-temporal data. Therefore, this process entails high costs and extensive time. This study focuses on evaluating the applicability of the online platform Google Colab using the SBAS processing technique to handle a series of 24 synthetic aperture radar images from January 2022 to December 2023 in the Ca Mau city area and some neighboring regions. By comparing the results from Google Colab with field measurement data, the study will assess the efficiency and accuracy of land subsidence determination results using the SBAS technique on the Google Colab platform, with the results showing a correlation of $R^2 = 0.71$ and RMSE = 6 mm.

Keywords: Land subsidence, InSar, SBAS, Google Colab.

Ứng dụng công nghệ GIS và viễn thám hỗ trợ phân tích, quản lý tài nguyên rừng trong điều kiện biến đổi khí hậu ở tỉnh Đắk Lắk

Nguyễn Sách Thành^{1*}, Trịnh Lê Hùng^{1,3}, Lê Văn Phú¹, Nguyễn Văn Trung^{2,3}

¹Học viện Kỹ thuật quân sự

²Trường Đại học Mở - Địa chất

³Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

TÓM TẮT

Đắk Lắk là một trong những tỉnh có diện tích rừng đứng đầu cả nước, phong phú về hệ sinh thái, đa dạng sinh học. Tuy nhiên, biến đổi khí hậu đang tác động mạnh mẽ đến tài nguyên rừng ở Đắk Lắk, đòi hỏi cần phải có những giải pháp ứng dụng khoa học công nghệ. Hiện nay, công nghệ GIS và viễn thám đóng một vai trò quan trọng trong việc phân tích, quản lý tài nguyên rừng. Chúng cung cấp các công cụ có giá trị để đánh giá, quản lý và bảo tồn tài nguyên rừng, giúp hỗ trợ các hoạt động lâm nghiệp bền vững và nỗ lực bảo tồn đa dạng sinh học. Trong nghiên cứu này, nhóm tác giả sẽ trình bày giải pháp ứng dụng các nền tảng công nghệ GIS và viễn thám mở nhằm hỗ trợ hoạt động phân tích, quản lý của các cơ quan chức năng có liên quan tới lĩnh vực tài nguyên rừng ở tỉnh Đắk Lắk.

Từ khóa: GIS, Viễn thám, Đắk Lắk.

1. Đặt vấn đề

Công nghệ hiện đại nói chung và công nghệ GIS, viễn thám nói riêng đóng vai trò hết sức quan trọng trong quản lý tài nguyên thiên nhiên. GIS là một hệ thống dựa trên máy tính được thiết kế để thu thập, lưu trữ, thao tác, phân tích và trình bày dữ liệu không gian hoặc địa lý. Công nghệ GIS tích hợp nhiều loại dữ liệu khác nhau, chẳng hạn như bản đồ, hình ảnh vệ tinh, ảnh chụp từ trên không và dữ liệu dạng bảng, cho phép người dùng hình dung, hiểu và đưa ra quyết định phù hợp về các mối quan hệ không gian địa lý. Trong khi đó, viễn thám liên quan đến việc sử dụng hình ảnh vệ tinh hoặc ảnh chụp từ trên không để thu thập thông tin về khu rừng mà không cần tiếp xúc vật lý. Việc tích hợp công nghệ GIS và viễn thám thực sự hữu ích trong việc đánh giá tài nguyên rừng và giám sát, bảo vệ rừng, khai thác rừng, phục hồi rừng, bảo tồn và đa dạng sinh học, biến đổi khí hậu, cơ sở dữ liệu không gian về rừng quản lý, chuẩn bị kế hoạch làm việc, bảo tồn môi trường sống của động vật hoang dã, và bảo tồn đất và lưu vực sông (Ankita, 2022),...

Viễn thám kết hợp với GIS giúp hoạt động kiểm kê và phân tích rừng trở nên hiệu quả hơn dựa trên việc cung cấp số liệu thống kê về quần thể với các sai số lấy mẫu thích hợp cho nhiều thuộc tính rừng theo thiết kế mẫu cơ bản (David, 2006). Ngoài ra, Viễn thám và GIS còn được sử dụng kết hợp để quản lý rủi ro và đánh giá cháy rừng, lập bản đồ che phủ rừng (Shruti, 2014),... Năm 2022, Martin đã phác thảo mô hình dạng công thông tin GIS đa chức năng nhằm phục vụ các nhiệm vụ giáo dục, nghiên cứu và quản lý rừng ở Đại học Lâm nghiệp thuộc Đại học Kỹ thuật Zvolen. Đây là đại học ở ở miền Trung Slovakia quản lý 9724 ha rừng cho các hoạt động nghiên cứu và giáo dục. Trong mô hình này, các dạng dữ liệu vector, raster và hình ảnh được lưu trữ trong Microsoft SQL Server 2008 geodatabase. Các dịch vụ web bản đồ và ứng dụng web được xây dựng trên ArcGIS dành cho máy chủ của Esri và các ứng dụng web được phát triển trên nền tảng Microsoft ASP.Net dựa trên kiến trúc C# và Model-View-Controller (Martin, 2022).

Đắk Lắk là một trong 5 tỉnh thuộc vùng Tây Nguyên, có tổng diện tích tự nhiên là 1.312.349,5 ha, trong đó diện tích đất lâm nghiệp chiếm 51,6% tổng diện tích tự nhiên của toàn tỉnh. Tài nguyên rừng Đắk Lắk rất phong phú và đa dạng về các hệ sinh thái các loài động, thực vật. Rừng Đắk Lắk giữ vị trí quan trọng trong đời sống dân sinh kinh tế-xã hội và bảo vệ môi trường sinh thái, bảo vệ đa dạng sinh học, địa chất, cảnh quan thiên nhiên (Bùi Thị Hải Nhung, 2010). Rừng Đắk Lắk có chức năng phòng hộ, bảo vệ nguồn nước, phòng hộ biên giới của tỉnh và là khu vực đầu nguồn, ảnh hưởng dòng chảy các con sông lớn của khu vực duyên hải miền Trung và Đông Nam Bộ như sông Sêrêpôk, Sông Ba, Sông Hinh, sông Đồng Nai,... Rừng Đắk Lắk được phân bố ở khắp các huyện trong tỉnh, đặc biệt là hành lang biên giới của tỉnh giáp Campuchia. Rừng Đắk Lắk phong phú và đa dạng, thường có kết cấu nhiều tầng, đa dạng về cây gỗ, cây bụi và cây thân thảo; có nhiều loại động thực vật quý hiếm phân bố chủ yếu ở Vườn Quốc gia Yok Đôn và các khu bảo tồn thiên nhiên Nam Kar, Ea Sô, Vườn Quốc gia Chư Yang Sin,... có nhiều loại động thực vật quý hiếm ghi trong Sách Đỏ Việt Nam và có loại được ghi trong Sách Đỏ Thế giới.

* Tác giả liên hệ

Email: thanhns.dth@gmail.com

Tuy nhiên, những năm gần đây do việc quản lý chưa tốt làm cho diện tích rừng Đăk Lăk liên tục bị suy giảm nhanh cả số lượng và chất lượng. Đây cũng là nguyên nhân của những thiên tai bất ngờ đã và đang có xu thế mạnh, thường xuyên như lũ lụt, hạn hán. Diện tích rừng tự nhiên bị phá, bị khai thác, lấn chiếm trái phép tập trung nhiều nhất ở các huyện Ea Súp, Buôn Đôn, Krông Bông, Ea Kar, M'Đrăk, Ea H'leo, các Vườn Quốc gia Yok Đôn, Khu Bảo tồn thiên nhiên Ea Sô, Chư Yang Sin,... trong đó, Ea Súp là địa phương mất rừng tự nhiên nhiều nhất. Theo Sở Nông nghiệp và Phát triển nông thôn tỉnh Đăk Lăk, diện tích rừng ở Đăk Lăk năm 2017 có hơn 505.076 ha, giảm hơn 10.533 ha; trong đó, rừng tự nhiên chỉ còn 461.384,5 ha, giảm 10.198 ha so với năm 2016 (Sở Kế hoạch và Đầu tư tỉnh Đăk Lăk, 2019).

Biến đổi khí hậu (BĐKH) là một trong những thách thức nghiêm trọng nhất đối với thế giới trong giai đoạn hiện nay. BĐKH đã tác động tiêu cực lên nhiều lĩnh vực từ hiện tượng nóng lên toàn cầu, nước biển dâng, hạn hán, lũ lụt gia tăng,... gây ra những tổn thất nặng nề cả về kinh tế lẫn tài nguyên môi trường và sức khỏe con người. Những năm gần đây, tác động của BĐKH đã hiện diện trên mọi mặt đời sống tại Đăk Lăk với sự xuất hiện các hiện tượng thời tiết cực đoan như: hạn hán kéo dài và mưa lớn bất thường. Năm 2019, theo Chi cục Kiểm lâm tỉnh Đăk Lăk, tỉnh có diện tích rừng lớn trên 504.000 ha, trong đó trên 200.000 ha rừng được xác định là khu vực trọng điểm có nguy cơ cháy trong mùa khô, thuộc các huyện Ea H'leo, Krông Năng, Ea Kar, Lăk, M'Đrăk, Krông Bông (Trần Lê Kiều Oanh, 2016),...



*Hình 1. Minh họa hiện trường vụ cháy rừng xảy ra tại xã Ea Lai, huyện M'Đrăk, Đăk Lăk
(Nguồn: Tiến Ninh, 2020. Báo Điện tử Đăk Lăk)*

Để phát triển kinh tế-xã hội của tỉnh gắn với việc bảo vệ rừng, tỉnh Đăk Lăk đã ban hành nhiều quy định. Để thực hiện chiến lược phát triển lâm nghiệp tỉnh Đăk Lăk giai đoạn 2006-2020, tỉnh Đăk Lăk tổ chức quy hoạch lại 3 loại rừng và được UBND tỉnh phê duyệt tại Quyết định số 1030/QĐ-UBND ngày 16/5/2007. Năm 2012, Ủy ban nhân dân tỉnh Đăk Lăk đã ra Quyết định số 2332/QĐ-UBND về việc ban hành điều lệ tổ chức và hoạt động của quỹ bảo vệ, phát triển rừng tỉnh Đăk Lăk. Ngày 17/07/2015 UBND tỉnh Đăk Lăk đã ban hành Quyết định số 1860/QĐ-UBND về việc phê duyệt đề cương và dự toán kinh phí điều chỉnh quy hoạch ba loại rừng tỉnh Đăk Lăk giai đoạn 2015-2020 nhằm xác định và cơ cấu lại diện tích đất lâm nghiệp, đất ngoài lâm nghiệp; Xác định lại quy mô diện tích 3 loại rừng phù hợp với tình hình thực tế tại mỗi địa phương. Năm 2019, Ủy ban nhân dân tỉnh Đăk Lăk đã ra Quyết định số 3419/QĐ-UBND về ban hành kế hoạch triển khai thực hiện đề án bảo vệ, khôi phục và phát triển rừng bền vững vùng Tây Nguyên giai đoạn 2016-2030, được phê duyệt tại Quyết định số 297/QĐ-TTG ngày 18/3/2019 của Thủ tướng Chính phủ,...

Trong điều kiện kinh tế - xã hội Việt Nam hiện nay, trước sự phát triển không ngừng của khoa học và công nghệ chủ trương triển khai ứng dụng phần mềm mã nguồn mở của Đảng và Nhà nước thực sự cần thiết, phù hợp với đặc điểm, tình hình của Việt Nam. Ngày 17/7/2006, Thủ tướng Chính phủ ban hành Quyết định số 169/2006/QĐ-TTg quy định về việc đầu tư, mua sắm các sản phẩm công nghệ thông tin của các cơ quan, tổ chức sử dụng nguồn vốn ngân sách nhà nước; trong đó “ưu tiên đầu tư, mua sắm sử dụng các sản phẩm phần mềm mã nguồn mở, đặc biệt là các sản phẩm do các doanh nghiệp trong nước khai thác, cung cấp”. Ngày 15/4/2015, Nghị quyết số 26/NQ-CP của Chính phủ ban hành Chương trình hành động của Chính phủ thực hiện Nghị quyết số 36-NQ/TW ngày 01 tháng 7 năm 2014 của Bộ Chính trị Ban Chấp hành Trung ương Đảng Cộng sản Việt Nam

về đẩy mạnh ứng dụng, phát triển công nghệ thông tin đáp ứng yêu cầu phát triển bền vững và hội nhập quốc tế; trong đó nhiệm vụ phát triển nguồn nhân lực công nghệ thông tin đạt chuẩn quốc tế, đẩy mạnh nghiên cứu, ứng dụng, tiếp thu, làm chủ và sáng tạo công nghệ mới xác định “Tăng cường hoạt động nghiên cứu, tiếp nhận chuyển giao và làm chủ công nghệ mới, phần mềm nguồn mở, công nghệ mới, chuẩn mở”. Ngày nay, với sự ra đời của các hiệp hội địa không gian mở - có sự tham gia liên kết của các trường đại học, các tổ chức phi lợi nhuận, các nền tảng công nghệ GIS, viễn thám mã nguồn mở như QGIS, PostGIS, MapWinGIS, Google EarthEngine,... đã và đang được phát triển ra các sản phẩm phù hợp với nhiều cấp độ, quy mô phát triển khác nhau. Trong dự án phát triển hệ thống thông tin quản lý ngành Lâm nghiệp tại Việt Nam giai đoạn II (FORMIS II) do Chính phủ Việt Nam và Chính phủ Phần Lan tài trợ, được triển khai trong thời gian từ 2013-2018 đã xây dựng phần mềm cập nhật diễn biến rừng. Phần mềm này dựa trên nền tảng phần mềm QGIS và CSDL PostGIS/PostgreSQL. Hệ thống xây dựng các plugin tích hợp vào phần mềm QGIS nhằm phục vụ việc cập nhật CSDL tài nguyên rừng ở các cấp độ quản lý khác nhau (cấp huyện, cấp tỉnh, cấp Trung ương). Năm 2021, Viện Sinh thái rừng và Môi trường đã tổ chức khóa “Tập huấn ứng dụng dữ liệu Landsat và Sentinel giám sát rừng bằng QGIS” cho các cán bộ kỹ thuật của Chi cục Kiểm lâm Hà Nội, Chi cục Kiểm lâm Ninh Bình, Chi cục Kiểm lâm Hòa Bình, Quỹ Bảo vệ và Phát triển rừng tỉnh Hòa Bình và Quỹ Bảo vệ và Phát triển rừng tỉnh Nghệ An. Mục tiêu của khóa tập huấn là giúp các cán bộ kỹ thuật Viện và các cán bộ quản lý, giám sát rừng tại các địa phương nắm được các kiến thức và kỹ năng cơ bản về QGIS, kỹ thuật ứng dụng phần mềm này trong thực tế khảo sát tại hiện trường, xử lý dữ liệu bản đồ trong phòng cũng như trích xuất các dữ liệu từ các loại bản đồ lâm nghiệp phục vụ báo cáo, kiểm tra, giám sát tài nguyên rừng tại các địa phương. Tuy nhiên, đối với tỉnh Đắk Lắk, việc ứng dụng các nền tảng công nghệ GIS và viễn thám cho các hoạt động phân tích, quản lý tài nguyên rừng còn hạn chế. Bởi vậy, trong bài báo này, nhóm tác sẽ trình bày giải pháp ứng dụng công nghệ GIS và viễn thám hỗ trợ phân tích, quản lý tài nguyên rừng trong điều kiện biến đổi khí hậu ở tỉnh Đắk Lắk dựa trên một số nền tảng mã nguồn mở như QGIS, PostGIS, MapWinGIS và Google Earth Engine.

2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1. QGIS

QGIS, ban đầu được gọi là Quantum GIS, là một ứng dụng hệ thống thông tin địa lý mã nguồn mở cho phép người dùng phân tích và trực quan hóa dữ liệu không gian. Năm 2002, QGIS được khởi xướng bởi Gary Sherman khi ông đang tìm cách tạo ra một phần mềm GIS miễn phí, thân thiện với người dùng. QGIS có một cộng đồng mạnh mẽ, đông đảo bao gồm các nhà phát triển, người dùng và những người ủng hộ trên toàn thế giới. QGIS cung cấp các công cụ để tạo bản đồ chi tiết về các khu vực rừng, bao gồm các kiểu thảm thực vật, sự phân bố loài cây và ranh giới rừng. QGIS cung cấp nhiều công cụ phân tích không gian có thể được sử dụng cho các nhiệm vụ quản lý rừng, chẳng hạn như tính toán diện tích rừng, phân tích các mô hình cảnh quan và đánh giá kết nối môi trường sống. QGIS có thể được sử dụng để theo dõi tình trạng rừng bằng cách lập bản đồ và phân tích dữ liệu về tình trạng sâu bệnh, bùng phát dịch bệnh và áp lực môi trường; phân tích chuỗi thời gian giúp theo dõi những thay đổi theo thời gian và xác định xu hướng về sức khỏe rừng (Nicolas, 2018).

2.2. PostGIS

PostGIS là phần mở rộng cơ sở dữ liệu không gian cho PostgreSQL, một trong những hệ thống quản lý cơ sở dữ liệu quan hệ nguồn mở phổ biến nhất. PostGIS bổ sung hỗ trợ cho các đối tượng địa lý vào cơ sở dữ liệu PostgreSQL, cho phép nó được sử dụng làm cơ sở dữ liệu không gian cho các hệ thống thông tin địa lý (GIS). PostGIS được phát triển bởi Research Inc và được phát hành lần đầu vào 5/2001. PostGIS hỗ trợ khả năng thêm các kiểu dữ liệu địa lý nhằm tích hợp trên hệ quản trị cơ sở dữ liệu PostgreSQL. Nó cũng thêm các hàm, toán tử và các chỉ số không gian để áp dụng cho các kiểu dữ liệu đó. Việc bổ sung các hàm, toán tử, các chỉ số và các kiểu dữ liệu làm tăng thêm sức mạnh của lõi của PostgreSQL làm cho nó nhanh hơn và trở thành hệ quản trị cơ sở dữ liệu không gian mạnh mẽ.

Việc sử dụng PostGIS để quản lý rừng mang lại khả năng mạnh mẽ để lưu trữ, truy vấn và phân tích dữ liệu không gian, có thể nâng cao đáng kể các khía cạnh khác nhau của quản lý rừng. PostGIS có thể lưu trữ dữ liệu vector như ranh giới của các khu rừng, vị trí cây riêng lẻ, đường mòn và vùng nước. PostGIS cũng có thể được sử dụng các chức năng để phân tích vùng lân cận, lập mô hình môi trường sống và truy vấn mối quan hệ không gian; thực hiện các hoạt động nối không gian, lớp phủ và đệm để phân tích dữ liệu rừng...

Năm 2013, Evangelos và nhóm nghiên cứu phát triển ứng dụng GIS cho việc lập kế hoạch quản lý lâm nghiệp đô thị. Để phát triển ứng dụng GIS, nhóm nghiên cứu đã sử dụng các nền tảng Geoserver, PostGIS/PostgreSQL, OpenLayers, GDAL, PROJ.4 và Entity Framework 5.0. Hơn nữa ASP.NET MVC4 đã được sử dụng để phát triển trang web của ứng dụng. Giải pháp này có thể hỗ trợ cơ quan chức năng lựa chọn phù hợp các biện pháp quản lý, bảo vệ và sử dụng rừng đô thị bằng cách kết nối thuộc tính với dữ liệu không gian. Ứng dụng này như một công cụ thích hợp cho dịch vụ hành chính đô thị, góp phần quản lý rừng đô thị tốt hơn (Evangelos, 2013).

2.3. MapWinGIS

MapWinGIS là nền tảng phát triển phần mềm GIS mã nguồn mở cho Microsoft Windows. Nó cung cấp một bộ công cụ và chức năng mà các nhà phát triển có thể sử dụng để bổ sung khả năng lập bản đồ và GIS cho các ứng dụng của riêng của họ. MapWinGIS có nguồn gốc từ dự án MapWindow, bắt đầu tại Đại học bang Idaho vào năm 2000. Mục tiêu ban đầu là phát triển một ứng dụng GIS mã nguồn mở, miễn phí dành cho sinh viên và các chuyên gia. Hiện nay, dự án MapWinGIS tiếp tục nhận được thông tin cập nhật và hỗ trợ từ đội ngũ nhà phát triển và cộng tác viên tận tâm. Các phiên bản hiện đại đã cải thiện khả năng tương thích với các hệ điều hành và môi trường phát triển hiện đại, đảm bảo tính phù hợp của nó trong các ứng dụng GIS hiện tại. MapWinGIS đặc biệt hữu ích trong quản lý rừng nhờ khả năng xử lý dữ liệu không gian và thực hiện các hoạt động GIS khác nhau. MapWinGIS có thể được sử dụng trong quản lý rừng; quản lý điều tra rừng, theo dõi sức khỏe rừng, quản lý cháy rừng,...

Năm 2015, Arun và cộng sự đã trình bày nghiên cứu về xây dựng hệ thống thông tin lâm nghiệp xã hội dựa trên nền tảng MapWinGIS và ngôn ngữ lập trình Microsoft Visual Basic. Lâm nghiệp xã hội nhằm mục đích tăng diện tích rừng trồng cho người bình thường nhằm đáp ứng nhu cầu ngày càng tăng về gỗ, nhiên liệu, gỗ, thức ăn gia súc,... Giải pháp đưa ra trong nghiên cứu của Arun có thể hỗ trợ hiển thị, điều hướng, kết xuất dữ liệu không gian địa lý. Các lớp chủ đề khác nhau, địa giới hành chính (có số liệu kinh tế - xã hội) và các lớp cơ sở hạ tầng cùng với dữ liệu vệ tinh Cartosat-1 liên quan đến vòng tròn lâm nghiệp xã hội khu vực Nagpur, Ấn Độ đã được tích hợp vào phần mềm để mô hình hóa đầu ra mong muốn (Arun, 2015).

2.4. Google Earth Engine

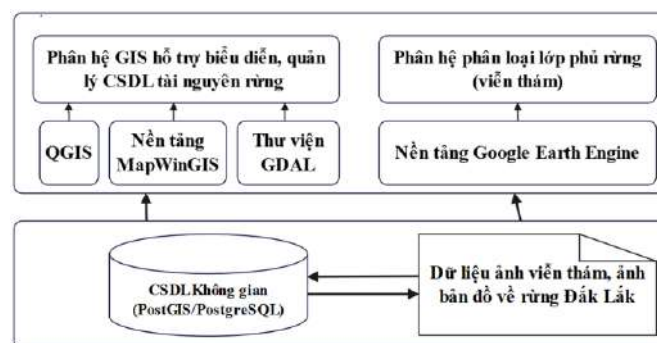
Google Earth Engine (GEE) là một nền tảng dựa trên đám mây được thiết kế để phân tích dữ liệu môi trường ở quy mô hành tinh. Nó kết hợp một danh mục nhiều petabyte gồm hình ảnh vệ tinh và bộ dữ liệu không gian địa lý với các công cụ mạnh mẽ để xử lý và phân tích dữ liệu này. Google Earth Engine được ra mắt chính thức vào năm 2010 tại hội nghị về biến đổi khí hậu COP16 ở Cancun (Mexico). Ban đầu nó được phát triển để hỗ trợ các nhà khoa học, nhà nghiên cứu và nhà hoạch định chính sách theo dõi và hiểu những thay đổi của môi trường. Trong những năm đầu, GEE được sử dụng cho các dự án như lập bản đồ nạn phá rừng ở rừng nhiệt đới Amazon và theo dõi những thay đổi ở các vùng nước (Mojtaba, 2020).

Google liên tục mở rộng danh mục dữ liệu Earth Engine, kết hợp hình ảnh vệ tinh mới, dữ liệu khí hậu và tập dữ liệu kinh tế - xã hội. Công cụ phân tích cải tiến: Nền tảng này giới thiệu các công cụ tiên tiến dành cho máy học, phân tích chuỗi thời gian và tính toán không gian địa lý quy mô lớn. Google Earth Engine (GEE) đã trở thành công cụ quan trọng để quản lý rừng, cung cấp nhiều khả năng giúp giám sát, phân tích và quản lý rừng một cách hiệu quả. GEE có thể phân tích hình ảnh vệ tinh để phát hiện nạn phá rừng và suy thoái rừng gần như theo thời gian thực. Điều này giúp xác định các hoạt động khai thác gỗ bất hợp pháp và các khu vực cần can thiệp. GEE có thể xác định các đám cháy đang hoạt động bằng cách sử dụng dữ liệu vệ tinh nhiệt và theo dõi sự lây lan của chúng trong thời gian gần như thực tế (Yong, 2022).

3. Kết quả và thảo luận

3.1. Mô hình giải pháp

Dưới đây là mô hình giải pháp ứng dụng công nghệ GIS và viễn thám hỗ trợ phân tích, quản lý tài nguyên rừng trong điều kiện biến đổi khí hậu ở tỉnh Đắk Lắk.



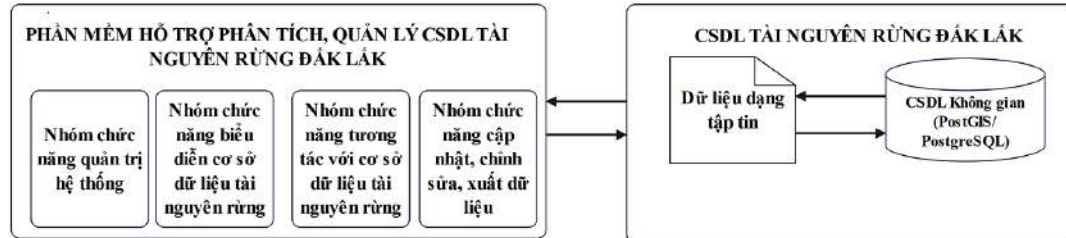
Hình 2. Mô hình giải pháp thám hỗ trợ phân tích, quản lý tài nguyên rừng Đắk Lắk

Mô hình này bao gồm 2 phân hệ:

- Phân hệ GIS hỗ trợ biểu diễn, quản lý CSDL tài nguyên rừng được xây dựng dựa trên các nền tảng QGIS, MapWinGIS, GDAL
- Phân hệ phân loại lớp phủ rừng được xây dựng dựa trên nền tảng Google Earth Engine

a. Thiết kế tổng quan phân hệ GIS hỗ trợ biểu diễn, quản lý CSDL tài nguyên rừng

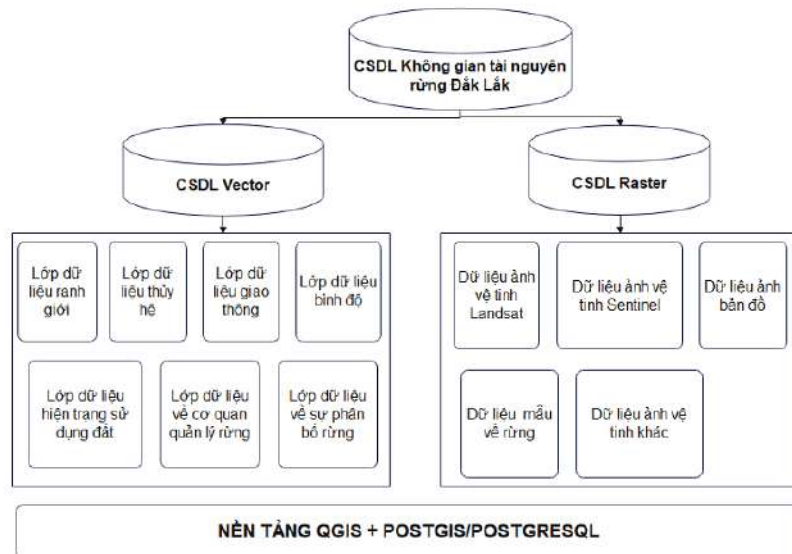
Phân hệ GIS hỗ trợ phân tích, quản lý CSDL tài nguyên rừng bao gồm 4 nhóm chức năng chính: Nhóm chức năng quản trị hệ thống; nhóm chức năng biểu diễn cơ sở dữ liệu tài nguyên rừng, nhóm chức năng tương tác với cơ sở dữ liệu tài nguyên rừng, nhóm chức năng cập nhật, chỉnh sửa, xuất dữ liệu. Phân hệ được phát triển dựa trên các nền tảng GIS mã nguồn mở như MapWinGIS, PostGIS/PostgreSQL và sử dụng các ngôn ngữ lập trình C#, SQL (Hình 3).



Hình 3. Minh họa mô hình các chức năng phân hệ GIS hỗ trợ phân tích, quản lý CSDL tài nguyên rừng

CSDL tài nguyên rừng bao gồm:

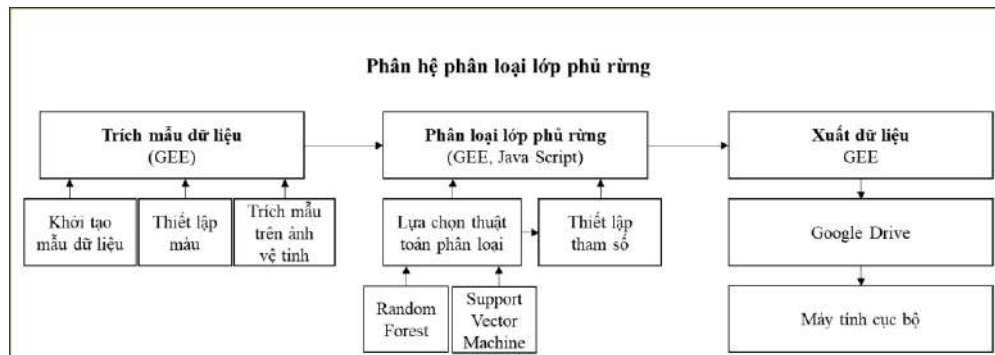
- CSDL vector: Được chuẩn hóa bởi phần mềm QGIS và lưu trữ trên PostGIS, bao gồm các lớp ranh giới, lớp dữ liệu thủy hệ, lớp dữ liệu giao thông, lớp dữ liệu đường bình độ, lớp dữ liệu hiện trạng sử dụng đất, lớp dữ liệu về cơ quan quản lý rừng, lớp dữ liệu về sự phân bố rừng.
- CSDL raster: Được lưu trữ ở dạng cây thư mục, chia thành các danh mục như dữ liệu ảnh vệ tinh Landsat, dữ liệu ảnh vệ tinh Sentinel, dữ liệu mẫu về rừng, dữ liệu ảnh bản đồ và dữ liệu ảnh vệ tinh khác.



Hình 4. Mô hình tổng quan cơ sở dữ liệu không gian tài nguyên rừng Đắc Lắc

b. Thiết kế tổng quan phân hệ phân loại lớp phủ rừng

Hình 5 mô tả chi tiết các chức năng trong phân hệ phân loại lớp phủ rừng. Phân hệ này được chia làm ba chức năng bao gồm: (1) trích mẫu dữ liệu, (2) phân loại lớp phủ rừng và (3) xuất dữ liệu.



Hình 5. Mô hình các chức năng của phân hệ phân loại lớp phủ rừng dựa trên GEE

3.2. Kết quả nghiên cứu

Nhóm tác giả đã xây dựng được chương trình Hệ thống hỗ trợ phân tích, quản lý CSDL tài nguyên rừng Đắk Lắk trong điều kiện biến đổi khí hậu ở tỉnh Đắk Lắk dựa trên việc tích hợp các nền tảng mã nguồn mở như MapWinGIS, PostGIS, GEE. Đây là dạng ứng dụng GIS máy tính để bàn với phân hệ GIS hỗ trợ phân tích, quản lý CSDL tài nguyên rừng và phân hệ phân loại lớp phủ rừng dựa trên GEE chạy ở dạng trực tuyến.

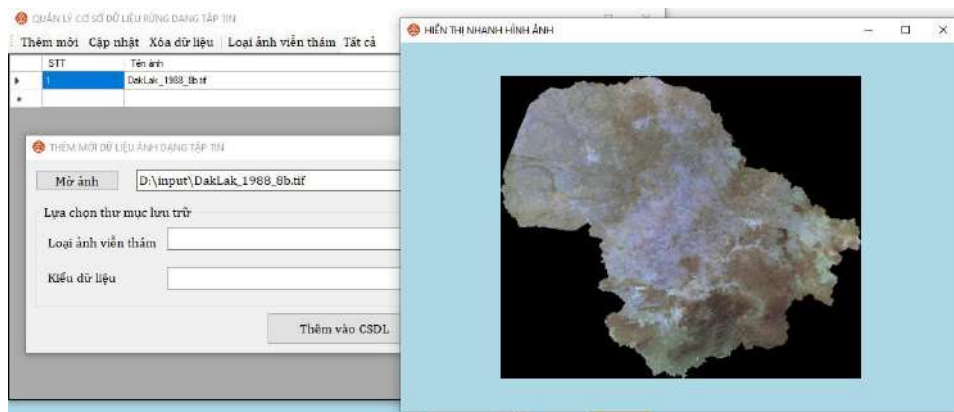
- Đối với phân hệ GIS hỗ trợ phân tích, quản lý CSDL tài nguyên rừng

Phân hệ GIS có thiết kế trực quan theo tiêu chuẩn của một ứng dụng GIS dạng máy tính để bàn với giao diện tiếng Việt. Hệ thống được khởi chạy bởi chức năng đăng nhập, nhằm xác thực tài khoản người sử dụng hệ thống (Hình 6).

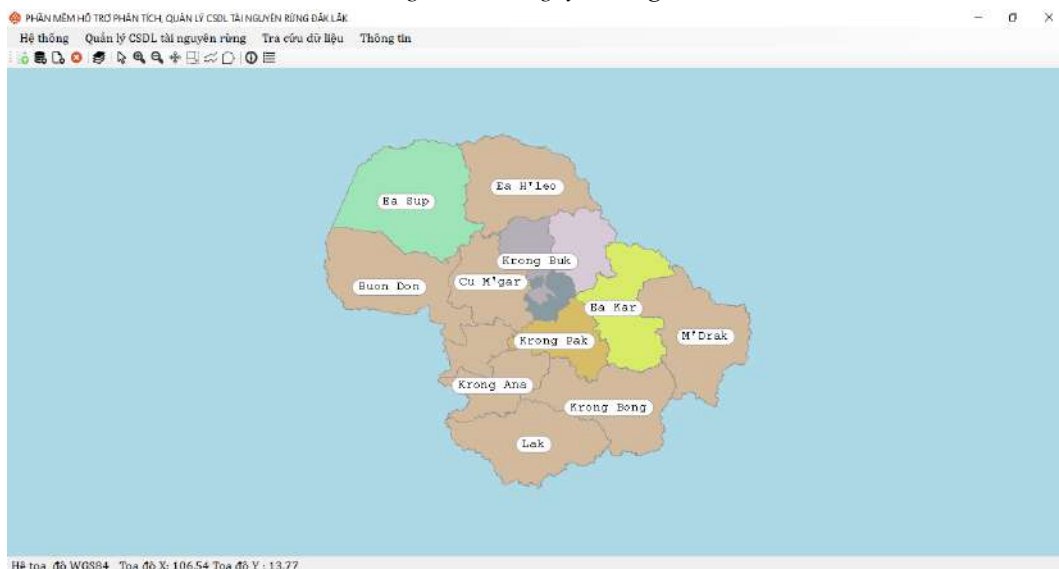


Hình 6. Minh họa giao diện đăng nhập hệ thống

Các chức năng hỗ trợ quá trình quản lý, tìm kiếm các thông tin, dữ liệu về tài nguyên rừng Đắk Lắk được xây dựng trực quan, phù hợp với các nghiệp vụ thường dùng của các cơ quan, đơn vị hoạt động trong ngành lâm nghiệp.



Hình 7. Minh họa giao diện chính các chức năng hỗ trợ quản lý, tìm kiếm các thông tin về tài nguyên rừng Đắk Lắk



Hình 8. Minh họa giao diện chính của phân hệ GIS

- Đối với phân hệ phân loại lớp phủ rừng

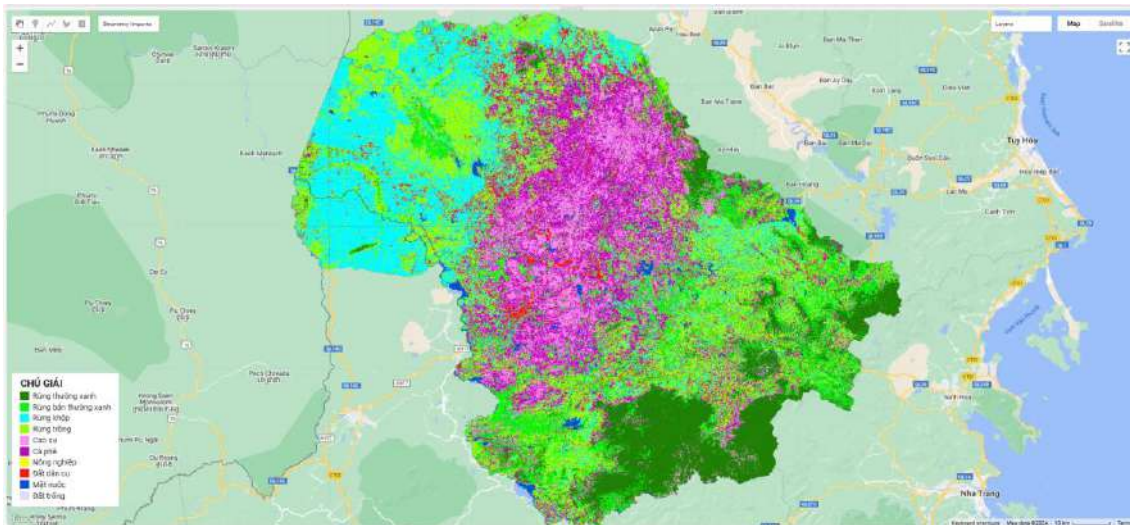
Trước hết, chức năng trích mẫu dữ liệu nhằm mục tiêu xây dựng bộ dữ liệu mẫu, phục vụ quá trình huấn luyện và kiểm tra ở các bước tiếp theo. GEE cung cấp các công cụ về để người dùng khởi tạo và xây dựng mẫu dữ liệu về rừng. Bên cạnh đó, người dùng cần thiết lập các tham số cơ bản cho mẫu dữ liệu như tên đối tượng, màu đối tượng, giá trị điểm mẫu.

Bên cạnh đó, chức năng phân loại lớp phủ rừng được xây dựng bằng ngôn ngữ Java Script trên nền tảng GEE. Người dùng cần lựa chọn thuật toán phân loại trước khi phân loại lớp phủ rừng. Lựa chọn thuật toán phân loại là yếu tố quan trọng trong quá trình xây dựng mô hình học máy. Để đạt được hiệu suất tốt trong việc phân loại dữ liệu, người dùng cần lựa chọn thuật toán và thiết lập các tham số đầu vào tương ứng với thuật toán đó. Trong phân hệ này, các tác giả xây dựng hai thuật toán học máy bao gồm Random Forest (RF) và Support Vector Machine (SVM). Người dùng có thể lựa chọn một thuật toán và thiết lập các tham số đầu vào tối ưu để đảm bảo hình ảnh sau phân loại đạt hiệu quả cao.

Dữ liệu đầu vào của phân hệ là các ảnh viễn thám Landsat 5, Landsat 7, Landsat-8, Landsat 9 và Sentinel-2. Nguồn dữ liệu được lấy trực tiếp trên nền tảng GEE. GEE cung cấp các bộ dữ liệu đã tiền xử lý ảnh bao gồm hiệu chỉnh khí quyển và hiệu chỉnh hình học. Mỗi hình ảnh Landsat có kích thước khoảng 1 GB, thu thập với độ phân giải không gian là 16 ngày. Đồng thời, mỗi ảnh Sentinel-2 có kích thước từ 3 - 5 GB, thu thập với độ phân giải không gian là 5 ngày. Các hình ảnh được lọc mây và cắt theo ranh giới của tỉnh Đắk Lắk.

Để phù hợp với điều kiện tự nhiên ở rừng ở Đắk Lắk và các loại dữ liệu ảnh viễn thám quang học như Landsat, Sentinel-2, các tác giả tiến hành trích xuất mẫu dữ liệu theo 10 đối tượng bao gồm rừng thường xanh, rừng bán thường xanh, rừng khộp, rừng trồng, cao su, cà phê, đất nông nghiệp, đất dân cư, mặt nước và đất trống. Các mẫu dữ liệu được chia thành bộ dữ liệu huấn luyện và bộ dữ liệu kiểm tra theo tỉ lệ 7:3. Kết quả phân loại lớp phủ rừng trên GEE đạt độ chính xác trên 90%.

Ngoài ra, chức năng xuất dữ liệu được xây dựng đảm bảo hình ảnh sau phân loại được lưu trữ về cục bộ. GEE cung cấp liên kết với Google Drive. Điều này cho phép dữ liệu hình ảnh sau phân loại có thể được trích xuất về Drive của tài khoản đang được sử dụng. Cuối cùng, người dùng có thể tải xuống từ Drive về máy tính cục bộ để biên tập và lưu trữ dữ liệu.



Hình 9. Hình ảnh sau phân loại lớp phủ rừng trên GEE

4. Kết luận

Nghiên cứu này đã cho thấy việc ứng dụng các nền tảng công nghệ GIS và viễn thám mở nhằm hỗ trợ các hoạt động phân tích, quản lý tài nguyên rừng ở Đắk Lắk là khả thi, phù hợp với tình hình và điều kiện ở địa phương. Kết quả nghiên cứu này có thể tiếp tục được nghiên cứu và mở rộng không chỉ ở Đắk Lắk mà có thể ở phạm vi rộng hơn như toàn bộ khu vực rừng Tây Nguyên hay các khu vực rừng ở Việt Nam khác. Tuy nhiên, trong thời gian tới, nhóm tác giả sẽ tiếp tục nghiên cứu các giải pháp GIS khác nhằm có thể phát triển một mô hình phân tán, kết hợp với mô hình xử lý dữ liệu ảnh viễn thám dựa trên GEE nhằm nâng cao khả năng sử dụng chung, khả năng chia sẻ và liên kết thông tin dữ liệu được tốt hơn.

Tài liệu tham khảo

Ankita Singh, Ankita Mishra, Hemlata Pant and Jyoti Singh, 2022. GIS Applications in Forest Management.
David C. Chojnacky, Randolph H. Wynne, and Christine, 2006. Mapping forest inventory and analysis data

attributes within the framework of double sampling for stratification design. *Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium*, pp. 237 - 241.

Evangelos Tasoulas, Gregory Varrasa, Ioanninis Tsirogiannisa, Christos Myriounis, 2013. Development of a GIS Application for Urban Forestry Management Planning. *6th International Conference on Information and Communication Technologies in Agriculture, Food and Environment*, pp. 70-80.

Martin Zápotocký and Milan Koreň, 2022. Multipurpose GIS Portal for Forest Management, Research, and Education. *International Journal of Geo-Information*.

Mojtaba Naghdizadegan Jahromi, Maryam Naghdizadegan Jahromi, Babak Zolghadr-Asli, Seyed Kazem Alavipanah, 2020. *Spatial Modeling in Forest Resources Management*. Springer, Cham, pp. 629-649.

Nicolas Baghdadi, Clement Mallet, Mehrez Zribi, 2018. QGIS in Remote Sensing: QGIS and applications in agriculture and forest. Wiley, Vol. 2.

Bùi Thị Hải Nhung, 2010. Đánh giá thực trạng và đề xuất một số giải pháp phát triển lâm nghiệp tỉnh Đắk Lắk. *Tạp chí Khoa học và Công nghệ lâm nghiệp*.

Trần Lê Kiều Oanh, Lê Sỹ Doanh, Lê Nguyên Khang, Mai Ngọc Sơn, Nguyễn Thanh Trà, 2016. Nghiên cứu xây dựng bản đồ phân bố rừng theo nguy cơ cháy tại tỉnh Đắk Lắk. *Tạp chí Khoa học và Công nghệ lâm nghiệp*.

Shruti Kanga, Laxmikant Sharma, Prem Chandra Pandey, Mahendra Singh Nathawat, 2014. GIS Modelling Approach for Forest Fire Risk Assessment and Management. *International Journal of Advancement in Remote Sensing, GIS and Geography*, Vol.2, No.1, pp. 30-44.

Sở Kế hoạch và Đầu tư tỉnh Đắk Lắk, 2019. Báo cáo Đề án hỗ trợ và phát triển doanh nghiệp nhỏ và vừa trên địa bàn tỉnh Đắk Lắk giai đoạn 2019-2020 định hướng đến năm 2030.

Yong Piao, Dong Kun Lee, Sang-Jin Park, Ho Gul Kim, 2022. Forest fire susceptibility assessment using google earth engine in Gangwon-do, Republic of Korea. *Geomatics, Natural Hazards and Risk*, Vol.13, No.1, pp. 432-450.

Tiến Ninh, 2020. Huyện M'Đrăk: Cao điểm phòng, chống cháy rừng. *Báo điện tử Đắk Lắk*, url: <https://baodaklak.vn/channel/3684/202005/huyen-mdrak-cao-diem-phong-chong-chay-rung-5683118/>.

ABSTRACT

Applycation of GIS and remote sensing technology to support analysis and management of forest resources under climate change conditions in Dak Lak province

Thanh Nguyen Sach^{1*}, Hung Le Trinh^{1,3}, Le Van Phu¹, Nguyen Van Trung^{2,3}

¹*Le Quy Don Technical University, Hanoi, Vietnam*

²*Hanoi University of Mining and Geology*

³*Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology*

Dak Lak is one of the provinces with the highest forest area in the country, rich in ecosystems and biodiversity. However, climate change is having a strong impact on forest resources in Dak Lak, requiring solutions that apply science and technology. Currently, GIS and remote sensing technology play an important role in analyzing and managing forest resources. They provide valuable tools for assessing, managing and conserving forest resources, helping to support sustainable forestry practices and biodiversity conservation efforts. In this study, the authors will present solutions for applying open GIS and remote sensing technology platforms to support analysis and management activities of authorities related to the field of forest resources in Dak Lak province.

Keywords: GIS, Remote Sensing, Dak Lak.

Ứng dụng công nghệ địa tin học trong đánh giá hạn hán ở Quảng Nam

Phạm Thị Thanh Hòa^{1,2*}, Trần Hồng Hạnh^{1,2}, Lưu Thị Diệu Chinh³

¹Trường Đại học Mở - Địa chất

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

³Trường Đại học Xây dựng Hà Nội

TÓM TẮT

Nghiên cứu tiến hành đánh giá mức độ hạn hán cho tỉnh Quảng Nam trên cơ sở công nghệ địa tin học. Ở đây, nền tảng điện toán đám mây Google Earth Engine được sử dụng để xử lý, tính toán các chỉ số hạn từ dữ liệu ảnh viễn thám Sentinel-2, giúp thay thế khối lượng lớn công việc thực hiện trên phần mềm thương mại. Công nghệ GIS - hệ thông tin địa lý được sử dụng để trực quan hóa và thành lập bản đồ phân vùng mức độ hạn hán cho tỉnh Quảng Nam. Các chỉ số hạn được sử dụng là LSWI (Land Surface Water Index) - chỉ số nước bề mặt, NDDI (Normalized Difference Drought Index) - chỉ số hạn hán khác biệt chuẩn hóa. Kết quả nghiên cứu cho thấy trên địa bàn tỉnh Quảng Nam trong năm 2023 xảy ra hạn hán với các mức độ khác nhau và không đồng đều. Trong đó hạn nghiêm trọng và hạn nặng tập trung chủ yếu ở phía Đông của tỉnh. Riêng phần phía Tây của tỉnh, do diện tích rừng che phủ lớn nên xảy ra hạn nhẹ, hoặc gần như không hạn. Nhìn chung, cách tiếp cận các chỉ số hạn với công nghệ địa tin học có thể được sử dụng để nhanh chóng thông báo cho chính quyền địa phương về hiện trạng, để từ đó đưa ra các giải pháp thích ứng lâu dài với hạn hán.

Từ khóa: Hạn hán, Quảng Nam, địa tin học, Sentinel-2, chỉ số hạn.

1. Đặt vấn đề

Với sự phát triển không ngừng của công nghệ thông tin, công nghệ địa tin học mang đến những bước đổi mới trong nghiên cứu về trái đất, bao gồm tài nguyên thiên nhiên, môi trường, thiên tai,... Địa tin học đang chứng kiến sự tiến bộ trong công nghệ viễn thám, từ việc phát triển phần mềm xử lý, đến việc không ngừng cải thiện độ phân giải của dữ liệu ảnh vệ tinh. Đồng thời đề tăng khả năng nghiên cứu và đáp ứng nhu cầu sử dụng, nhiều bộ dữ liệu ảnh được cung cấp miễn phí như Modis, Landsat, Sentinel,...

Viễn thám cung cấp một phương tiện hiệu quả để thu thập dữ liệu từ các khu vực địa lý rộng lớn trong thời gian ngắn, kể cả các khu vực không thể tiếp cận hoặc nguy hiểm. Công cụ này cho phép theo dõi những thay đổi về môi trường và tài nguyên thiên nhiên, đồng thời hỗ trợ trong các tình huống khẩn cấp (Navalgund et al., 2007). Tốc độ và phạm vi thu thập dữ liệu có thể giúp ứng phó với thảm họa và cung cấp thông tin theo thời gian thực về mức độ thiệt hại.

Một nhánh khác của công nghệ địa tin học mà không thể thiếu là Hệ thông tin địa lý GIS (Geographic Information System). GIS vẫn được biết đến là một công cụ xử lý đa dạng nguồn thông tin và hỗ trợ ra quyết định. GIS đảm bảo dữ liệu chính xác và phù hợp được tham chiếu theo địa lý. Hiệu quả mà GIS mang lại không còn có gì có thể nghi ngờ, đặc biệt khi kết hợp cùng với công nghệ viễn thám. Đó là việc ứng dụng trong nhiều nghiên cứu về thiên tai như lũ lụt (Al-Tahir và cộng sự, 2014; Irimescu và cộng sự, 2010), đánh giá hạn hán (Huang và cộng sự, 2023), trượt lở (Podolszki & Karlović, 2023),...

Bên cạnh đó, sự bùng nổ của dữ liệu lớn không gian địa lý là xu hướng không thể bỏ qua trong tiến trình phát triển công nghệ. Khi đó, những đổi mới trong phân tích dữ liệu được ghi nhận với việc gia tăng sử dụng nền tảng điện toán đám mây, mang lại khả năng truy cập, phân tích các tập dữ liệu lớn hiệu quả hơn. Điều này tạo điều kiện thuận lợi cho việc cộng tác và xử lý dữ liệu theo thời gian thực trên các vị trí địa lý khác nhau. Google Earth Engine (GEE) là một trong những nền tảng mạnh mẽ cho phép thực hiện các công việc đó. GEE được đánh giá cao trong nhiều lĩnh vực nghiên cứu như nông nghiệp, giám sát thảm thực vật, thiên tai,... (Kumar & Mutanga, 2019; Yu & Gong, 2012).

Trong những năm gần đây, tỉnh Quảng Nam có xu hướng gia tăng tình trạng hạn hán. Điều này dẫn theo các thiệt hại đến sản xuất và sinh hoạt trong khu vực. Đứng trước thách thức với loại hình thiên tai này, cần có các kết quả nghiên cứu nhanh chóng và đảm bảo độ tin cậy, để đánh giá mức độ hạn hán. Tuy nhiên, các nghiên cứu về hạn hán trên địa bàn tỉnh chưa thực sự có nhiều. Điển hình có nghiên cứu của nhóm (Du và cộng sự, 2018) tính một số các chỉ số hạn từ ảnh viễn thám. Hoặc nghiên cứu sử dụng các điểm quan trắc khí tượng có kết hợp công nghệ viễn thám (Hanh Mai và cộng sự, 2022). Các nghiên cứu này mới chỉ sử dụng ảnh có độ phân giải thấp là ảnh Modis có độ phân giải 500 m. Điều này có thể gây hạn chế đến đánh giá tình trạng hạn hán trong khu

* Tác giả liên hệ

Email: phamthithanhhoa@humg.edu.vn

vực. Do đó, với mong muốn thu được kết quả có độ tin cậy cao hơn, nghiên cứu tiến hành tính chỉ số hạn từ dữ liệu ảnh Sentinel-2 có độ phân giải không gian 10-20 m (với các kênh ảnh khác nhau), sử dụng nền tảng GEE. Các chỉ số được lựa chọn sử dụng là LSWI (Land Surface Water Index - chỉ số nước bề mặt) dùng trong đánh giá hạn nông nghiệp (Xiao và cộng sự, 2004), NDDI (Normalized Difference Drought Index - chỉ số hạn hán khác biệt chuẩn hóa), dùng trong đánh giá hạn tổng hợp (Gu và cộng sự, 2007).

2. Khu vực nghiên cứu, dữ liệu sử dụng và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Quảng Nam là một tỉnh nằm ở vùng duyên hải Nam Trung Bộ, giáp tỉnh Thừa Thiên Huế và thành phố Đà Nẵng ở phía Bắc; phía Nam giáp tỉnh Quảng Ngãi và tỉnh Kon Tum; giáp tỉnh Sekong của Lào ở phía Tây; phía Đông giáp Biển Đông. Tỉnh nằm trong tọa độ địa lý khoảng 108026'16"-108044'04" độ kinh Đông và từ 15023'38"-15038'43" độ vĩ Bắc.

Tỉnh chỉ có 2 mùa là mùa mưa (kéo dài từ tháng 10 đến tháng 12) và mùa khô (kéo dài từ tháng 2 đến tháng 8). Còn lại tháng 1, tháng 9 là các tháng chuyển mùa, với đặc trưng là thời tiết hay nhiều loạn và khá nhiều mưa (Cổng thông tin Điện tử tỉnh Quảng Nam).

Quảng Nam là một trong những tỉnh miền Trung chịu mức độ hạn nghiêm trọng, thường xảy ra vào tháng 4 đến tháng 7. Hạn hán còn được xếp hàng vào loại thiên tai gây thiệt hại lớn về kinh tế, môi trường và xã hội trên địa bàn tỉnh Quảng Nam. Theo tài liệu thống kê hạn hán từ năm 1990, hầu như năm nào cũng có hạn hán với các mức độ khác nhau và xu thế không ngừng gia tăng trên địa bàn tỉnh (Vũ Thị Thu Lan, 2011). Tính riêng giai đoạn 1999-2014, thiệt hại do hạn đến lúa là 53.681 ha, 606.000 người trong tình trạng thiếu nước sinh hoạt. Theo (Ủy ban nhân dân tỉnh Quảng Nam, 2010; Viện Khoa học Khí tượng Thủy văn và Môi trường, 2009), do hạn hán kéo dài, 4.841/44.500 ha lúa hè thu ở các huyện đồng bằng bị hư hại. Ngoài ra, có trên 3.000 ha lúa không thể gieo sạ vì khô cằn cùng với 5.000 ha cây trồng thiếu nước tưới và gần 5.000 người dân bị thiếu nước sinh hoạt ở các huyện trung du và miền núi Quảng Nam.



Hình 1. Vị trí khu vực nghiên cứu

2.2. Dữ liệu sử dụng

Nghiên cứu này sử dụng dữ liệu ảnh vệ tinh Sentinel-2 - vệ tinh được nghiên cứu và phát triển bởi Cơ quan Hàng không vũ trụ châu Âu (ESA). Ảnh Sentinel-2 phủ trùm khu vực nghiên cứu vào tháng 5 năm 2023 với độ che phủ mây dưới 30%. Bảng 1 thể hiện thông tin của ảnh Sentinel-2 với cấp xử lý ở mức 2A, bao gồm 12 kênh đa phổ từ kênh nhìn thấy đến hồng ngoại sóng ngắn, tương ứng từ B1 đến B12. Ngoài ra, còn có ba dải QA, trong đó một (QA60) là dải bitmask có thông tin về mặt nạ đám mây.



Hình 2. Ảnh Sentinel-2 tỉnh Quảng Nam (tổ hợp màu tự nhiên)

Bảng 1. Thông tin kênh đa phổ của ảnh Sentinel-2 Level 2A

Kênh	Tên kênh ảnh	Độ phân giải	Bước sóng
B1	Aerosols (Sol khí)	60 m	443,9 nm (S2A)/442,3 nm (S2B)
B2	Blue (Xanh lam)	10 m	496,6 nm (S2A)/492,1nm (S2B)
B3	Green (Xanh lục)	10 m	560 nm (S2A)/559 nm (S2B)
B4	Red (Đỏ)	10 m	664,5 nm (S2A)/665 nm (S2B)
B5	Red Edge 1 (Kênh rìa đỏ 1)	20 m	703,9 nm (S2A)/703,8 nm (S2B)
B6	Red Edge 2 (Kênh rìa đỏ 2)	20 m	740,2 nm (S2A)/739,1nm (S2B)
B7	Red Edge 3 (Kênh rìa đỏ 3)	20 m	782,5 nm (S2A)/779,7 nm (S2B)
B8	NIR (cận hồng ngoại)	10 m	835,1 nm (S2A)/833 nm (S2B)
B8A	Red Edge 4 (Kênh rìa đỏ 4)	20 m	864,8 nm (S2A)/864 nm (S2B)
B9	Water vapor (Hơi nước)	60 m	945 nm (S2A)/943,2 nm (S2B)
B11	SWIR 1 (Hồng ngoại sóng ngắn 1)	20 m	1613,7 nm (S2A)/1610,4 nm (S2B)
B12	SWIR 2 (Hồng ngoại sóng ngắn 2)	20 m	2202,4 nm (S2A)/2185,7 nm (S2B)

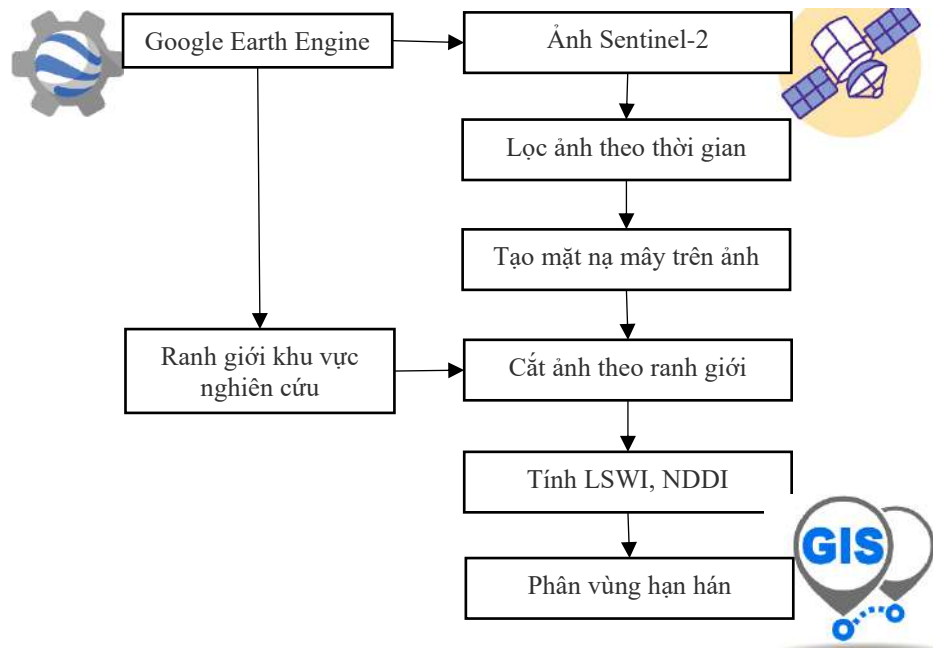
2.3. Phương pháp nghiên cứu

2.3.1. Sơ đồ quy trình

Nghiên cứu thực hiện trên nền tảng điện toán đám mây GEE, với việc sử dụng ảnh Sentinel-2. Ảnh Sentinel-2 được chọn theo thời gian, tạo mặt nạ mây, sau đó cắt ảnh theo ranh giới tỉnh Quảng Nam.

Hai chỉ số hạn là LSWI và NDDI được tính từ các kênh ảnh Sentinel-2.

Sơ đồ quy trình được thể hiện trong Hình 3.



Hình 3. Sơ đồ quy trình nghiên cứu hạn hán sử dụng công nghệ địa tin học

2.3.2. Các chỉ số hạn

Chỉ số LSWI thể hiện mức độ thay đổi hàm lượng nước của lớp phủ bề mặt. LSWI là một trong những chỉ số để đánh giá mức độ hạn hán của lớp phủ thực vật (R. Bajgain và cộng sự, 2017; K. và cộng sự, 2011). Chỉ số LSWI càng nhỏ mức độ hạn hán càng nghiêm trọng.

Chỉ số hạn NDDI được tính toán liên quan đến các chỉ số NDVI (Normalized Difference Vegetation Index - Chỉ số thực vật khác biệt chuẩn hóa) và NDWI (Normalized Difference Water Index - Chỉ số nước khác biệt chuẩn hóa), tương ứng nhạy cảm với sự hiện diện của thảm thực vật và của nước trong một khu vực.

Bảng 2. Công thức tính các chỉ số trong nghiên cứu

Chỉ số	Tên đầy đủ chỉ số	Công thức	Ý nghĩa
LSWI	Chỉ số nước bề mặt (Land Surface Water Index)	$LSWI = \frac{NIR - SWIR1}{NIR + SWIR1} \quad (1a)$ $LSWI_{Sentinel\ 2} = \frac{B8 - B11}{B8 + B11} \quad (1b)$	Chỉ số hạn nông nghiệp
NDDI	Chỉ số hạn hán khác biệt chuẩn hóa (Normalized Difference Drought Index)	$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI} \quad (2)$	Chỉ số hạn tổng hợp
NDVI	Chỉ số thực vật khác biệt chuẩn hóa (Normalized Difference Vegetation Index)	$NDVI = \frac{NIR - RED}{NIR + RED} \quad (3a)$ $NDVI_{Sentinel\ 2} = \frac{B8 - B4}{B8 + B4} \quad (3b)$	Chỉ số cần tính trong công thức (2)
NDWI	Chỉ số nước khác biệt chuẩn hóa (Normalized Difference Water Index)	$NDWI = \frac{NIR - SWIR}{NIR + SWIR} \quad (4a)$ $NDWI_{Sentinel\ 2} = \frac{B8A - B11}{B8A + B11} \quad (4b)$	Chỉ số cần tính trong công thức (2)

trong đó:

NIR: kênh cận hồng ngoại; SWIR: kênh hồng ngoại sóng ngắn;

NDVI: Chỉ số thực vật khác biệt chuẩn hóa (Tucker, 1979) với giá trị trong khoảng từ -1 đến 1. Giá trị NDVI gần đến 1 cho chỉ thị thực vật khỏe mạnh, NDVI thấp cho nhận biết tình trạng thực vật suy giảm;

NDWI: Chỉ số nước khác biệt chuẩn hóa (Gao, 1996), nằm trong khoảng từ -1 đến 1. Giá trị NDWI cao cho thấy đủ độ ẩm, trong khi giá trị thấp cho thấy tình trạng thiếu nước.

Trong Bảng 2, công thức tính các chỉ số hạn được thể hiện trong công thức (1b) và (2) khi áp dụng tính cho

ảnh Sentinel-2. Các chỉ số khác như NDVI, NDWI được tính phục vụ cho công thức (2).

Các thông số B4, B8, B8A, B11 trong các công thức (1b), (3b), (4b) tương ứng với các kênh ảnh được ghi trong Bảng 1.

Việc phân chia mức độ hạn hán cho hai chỉ số LSWI và NDDI được thể hiện trong Bảng 3 và Bảng 4.

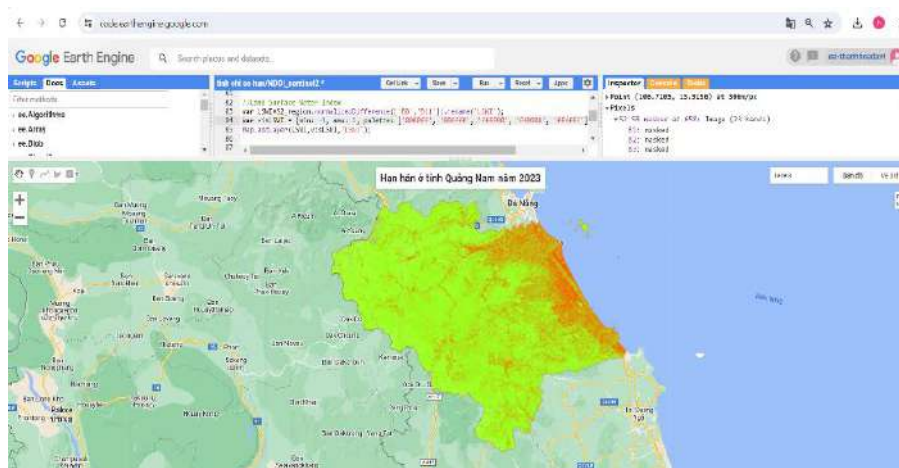
Bảng 3. Bảng phân chia mức độ hạn hán với chỉ số LSWI (Rajen Bajgain và cộng sự, 2015)

STT	Phân lớp	Mức độ
1	$LSWI \leq -0,1$	Hạn nghiêm trọng
2	$-0,1 < LSWI \leq 0$	Hạn nặng
3	$0 < LSWI \leq 0,1$	Hạn trung bình
4	$LSWI > 0,1$	Không hạn

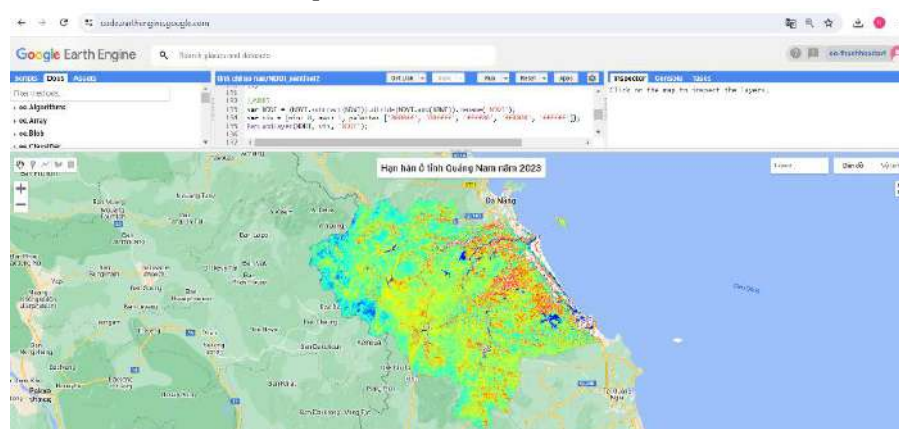
Bảng 4. Bảng phân chia mức độ hạn hán với chỉ số NDDI (Nepal và cộng sự, 2021)

STT	Phân lớp	Mức độ
1	$NDDI > 3$	Hạn nghiêm trọng
2	$1,25 < NDDI \leq 3$	Hạn nặng
3	$0,7 < NDDI \leq 1,25$	Hạn trung bình
4	$-2 \leq NDDI \leq 0,7$	Hạn nhẹ
5	$NDDI < -2$	Không hạn

3. Kết quả và thảo luận



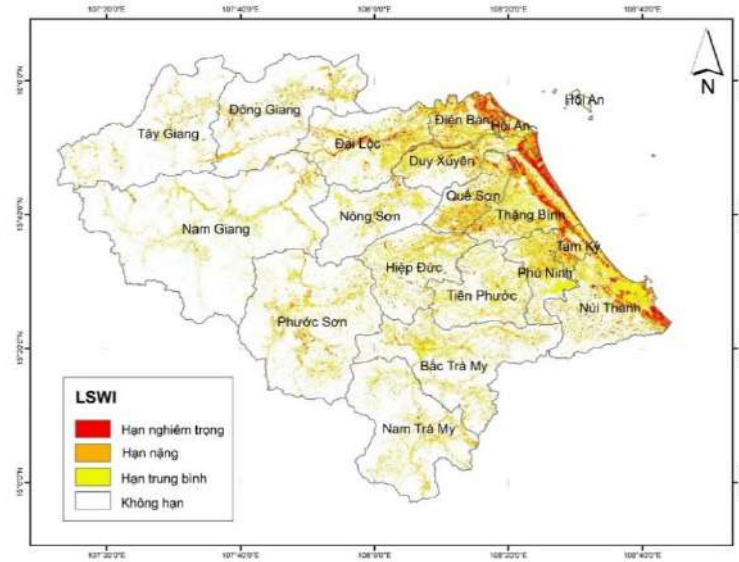
Hình 4. Kết quả chỉ số hạn LSWI trên GEE



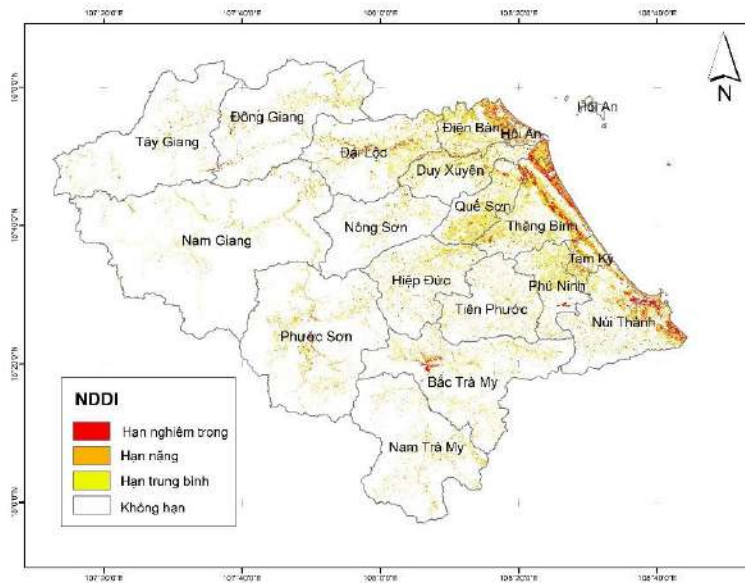
Hình 5. Kết quả chỉ số hạn NDDI trên GEE

Các kết quả thu được là các ảnh chỉ số được tính toán và hiển thị trên Google Earth Engine: LSWI, NDVI, NDWI, NDDI. Hình 4 và Hình 5 thể hiện kết quả ảnh LSWI và NDDI trên nền tảng GEE. Để thu được các kết quả này, việc lập trình được tiến hành trực tiếp trên Code Editor - môi trường tương tác để phát triển các ứng dụng Earth Engine, trên cơ sở các hàm và thuật toán được định nghĩa sẵn và phát triển các code riêng cho bài toán.

Sau đó ảnh chỉ số hạn được tải về và trình bày trên phần mềm ArcGIS với các mức độ phân chia vùng hạn theo Bảng 3 và Bảng 4, với các mức: Không hạn, hạn trung bình, hạn nặng, hạn nghiêm trọng.



Hình 6. Kết quả tình trạng hạn hán tỉnh Quảng Nam dùng chỉ số LSWI (5/2023)



Hình 7. Kết quả tình trạng hạn hán tỉnh Quảng Nam dùng chỉ số NDDI (5/2023)

Kết quả tính toán LSWI, NDDI dựa trên ảnh Sentinel-2 năm 2023 thể hiện trên Hình 6 và Hình 7 cho thấy hạn hán gần như xảy ra trên địa bàn tỉnh, phân bố không đồng đều với các mức hạn từ trung bình cho đến nghiêm trọng. Trong đó, vùng hạn nghiêm trọng (màu đỏ) tập trung ven biển cồn cát phía Đông của tỉnh, hoàn toàn tương đồng kết quả phân ngưỡng hạn giữa hai chỉ số: $LSWI \leq -0,1$ (Rajen Bajgain và cộng sự, 2015) và $NDDI > 3$ (Nepal và cộng sự, 2021). Các huyện thuộc khu vực phía Đông cũng có tình trạng hạn hán cao hơn (tương ứng màu vàng đến màu đỏ), điển hình tại khu vực thuộc hạ lưu sông Vu Gia - Thu Bồn gồm các huyện Đại Lộc, huyện Duy Xuyên, thị xã Điện Bàn, thành phố Hội An và các khu vực khác thuộc huyện Quế Sơn, huyện Thăng Bình, thành phố Tam Kỳ. Tại đây, mực nước tại các con sông và các hồ chứa cũng đã giảm đến mức khô hạn do không có mưa và nhiệt độ cao trong thời gian dài. Riêng phần phía Tây của tỉnh, do có diện tích thực phủ che phủ lớn (vùng có thảm thực vật rừng lá rộng) với giá trị NDVI cao nên mức độ hạn hán có xu hướng thấp hơn so với các vùng còn lại. Đó là các huyện Nam Giang, Phước Sơn, Tây Giang. Khi đó, chỉ số $LSWI > 0,1$ và $NDDI < 0,7$, tương ứng hạn nhẹ hoặc không hạn. Nhìn chung, thời điểm hạn hán ở Quảng Nam thường rơi vào thời kỳ từ tháng 4 đến tháng 7. Do đó với thời gian nghiên cứu là tháng 5, gần như huyện nào cũng ghi nhận tình trạng hạn hán.

Khi so sánh kết quả tình trạng hạn hán, quan sát thấy rằng bản đồ NDDI và LSWI thể hiện điều kiện hạn hán có sự chênh lệch. Theo đó, diện tích hạn năm 2023 là 120.574,64 ha (chiếm khoảng 10,9%) với chỉ số NDDI và 238.025 ha (chiếm khoảng 21,6%) với chỉ số LSWI so với tổng diện tích toàn tỉnh. Lý do cho kết quả khác biệt này là NDDI xem xét cả điều kiện thực vật NDVI, thay vì chỉ xem xét đến điều kiện mặt nước NDWI như chỉ số LSWI. Vùng có giá trị NDWI cao cho biết diện tích bề mặt nước nên mức độ hạn hán nghiêm trọng sẽ giảm. Ngược lại NDWI thấp và xem xét điều kiện NDVI thấp (nhưng không âm) cho ra kết quả giá trị NDDI cao, nghĩa là bề mặt của khu vực rất ít hoặc không phải là mặt nước, kèm theo điều kiện thực vật giảm, do đó mức độ nghiêm trọng của hạn hán cao. Đồng thời giá trị NDVI cũng là một phương tiện hiệu quả để xác định các bất thường về tình trạng thực vật, chẳng hạn như sự suy giảm rõ rệt về sức khỏe thực vật với các giá trị thấp (nhưng không âm) có thể suy đoán về tình trạng hạn hán (Sruthi & Aslam, 2015).

Kết quả nghiên cứu được so sánh trực quan với nghiên cứu trước đó của (Du và cộng sự, 2018), cho thấy các khu vực hạn hán tương đồng với nhau ở vị trí thuộc hạ lưu sông Vu Gia - Thu Bồn, cũng như nhận định thời điểm hạn hán là tháng 4, tháng 5 với nền nhiệt cao và lượng mưa ít. Tuy nhiên so sánh này mang tính chất tương đối, bởi thời điểm so sánh khác nhau, dữ liệu và các chỉ số không giống nhau. Vì thế trong các nghiên cứu tiếp theo về hạn hán ở tỉnh Quảng Nam, nhóm nghiên cứu sẽ bổ sung thêm các số liệu đánh giá, đồng thời đưa thêm các dữ liệu đầu vào khác để đánh giá tình trạng hạn hán, cũng như ảnh hưởng của hạn hán đến hiện trạng thảm thực vật tại đây.

4. Kết luận

Lĩnh vực địa tin học đang phát triển trong kỷ nguyên khoa học kỹ thuật, với những đổi mới mới hứa hẹn sẽ mang đến đột phá trong phân tích và biểu diễn dữ liệu không gian. Vì thế việc ứng dụng công nghệ này trong nghiên cứu hạn hán đang được nhiều nước trên thế giới, cũng như Việt Nam quan tâm.

Nghiên cứu đã sử dụng các chỉ số hạn từ ảnh viễn thám Sentinel-2 để xác định mức độ nghiêm trọng và phạm vi không gian hạn ở tỉnh Quảng Nam. Các chỉ số hạn hán là sự bổ sung tốt cho đánh giá hạn hán và tìm hiểu lịch sử hạn hán. Sự kết hợp công nghệ viễn thám cùng khả năng xử lý ảnh thực hiện trên nền tảng điện toán đám mây và công nghệ GIS phân tích dữ liệu, mang lại hiệu quả về mặt thời gian, giảm tải việc lưu trữ dữ liệu, giúp trực quan hóa kết quả.

Đây là nghiên cứu ban đầu về tình trạng hạn hán trên địa bàn Tỉnh, với việc sử dụng ảnh có độ phân giải cao hơn so với các nghiên cứu trước đó. Tuy nhiên, để có thể mang lại kết quả cao hơn, cần có thêm đa dạng các chỉ số hạn hán, cũng như cần bổ sung nguồn dữ liệu ảnh hưởng đến hạn hán như lượng mưa, nhiệt độ, địa hình, lớp phủ,...

Tài liệu tham khảo

- Al-Tahir, R., Saeed, I., & Mahabir, R., 2014. *Applications of Remote Sensing and GIS Technologies in Flood Risk Management*. In book: Flooding and Climate Change: Sectorial Impacts and Adaptation Strategies for the Caribbean Region.
- Bajgain, R., Xiao, X., Basara, J., Wagle, P., Zhou, Y., Zhang, Y., & Mahan, H., 2017. *Assessing agricultural drought in summer over Oklahoma Mesonet sites using the water-related vegetation index from MODIS*. Int J Biometeorol, 61(2), 377-390. Doi:10.1007/s00484-016-1218-8.
- Bajgain, R., Xiao, X., Wagle, P., Basara, J., & Zhou, Y., 2015. *Sensitivity analysis of vegetation indices to drought over two tallgrass prairie sites*. ISPRS Journal of Photogrammetry and Remote Sensing, 108, 151-160. Doi:https://doi.org/10.1016/j.isprsjprs.2015.07.004.
- Du, T., Bui, D., Nguyen, M., & Lee, H., 2018. *Satellite-Based, Multi-Indices for Evaluation of Agricultural Droughts in a Highly Dynamic Tropical Catchment, Central Vietnam*. Water, 10. Doi:10.3390/w10050659.
- Gao, B.-c., 1996. *NDWI-A normalized difference water index for remote sensing of vegetation liquid water from space*. Remote Sensing of Environment, 58(3), 257-266. Doi:https://doi.org/10.1016/S0034-4257(96)00067-3.
- Gu, Y., Brown, J., Verdin, J., & Wardlow, B., 2007. *A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States*. Geophysical Research Letters, 34, 6. Doi:10.1029/2006GL029127.
- Hanh Mai, B., Dang, T., Tu, Tâm, N., Van, D., The Anh, L., Thi, H., & Duyen, T., 2022. *Effects of Drought Levels on Soil Microbial Respiration and Biomass in Correspondence with Land use Types*. VNU Journal of Science: Earth and Environmental Sciences, 38(1), 31-44. Doi:10.25073/2588-1094/vnuees.4731.
- Huang, Y., Liu, Y., Shi, R., & Ren, H., 2023. *Application of Remote Sensing and GIS in Drought and Flood Assessment and Monitoring*. Water, 15(3). Doi:10.3390/w15030541.
- Irimescu, A., Craciunescu, V., Stancalie, G., & Nertan, A., 2010. *Remote sensing and GIS techniques for flood monitoring and damage assessment. Study case in Romania*.

- K, C., Seshasai, M., & Behera, G., 2011. Assessment of Early Season Agricultural Drought through Land Surface Water Index (LSWI) and Soil Water Balance Model (Vol. XXXVIII-8/W20).
- Kumar, L., & Mutanga, O., 2019. *Google Earth Engine Applications*. Remote Sensing, 11, 591. Doi:10.3390/rs11050591.
- Navalgund, R., V, J., & Roy, P., 2007. *Remote sensing applications: An overview*. Current science, Vol. 93.
- Nepal, S., Tripathi, S., & Adhikari, H., 2021. Geospatial approach to the risk assessment of climate-induced disasters (drought and erosion) and impacts on out-migration in Nepal. International Journal of Disaster Risk Reduction, 59, 102241. Doi:https://doi.org/10.1016/j.ijdr.2021.102241.
- Podolszki, L., & Karlović, I. (2023). Remote Sensing and GIS in Landslide Management: An Example from the Kravarsko Area, Croatia. *Remote Sensing*, 15(23). doi:10.3390/rs15235519.
- Sruthi, S., & Aslam, M. A., 2015. Agricultural Drought Analysis Using the NDVI and Land Surface Temperature Data; a Case Study of Raichur District. Aquatic Procedia, 4, 1258-1264. doi:10.1016/j.aqpro.2015.02.164.
- Tucker, C. J., 1979. *Red and photographic infrared linear combinations for monitoring vegetation*. Remote Sensing of Environment, 8(2), 127-150. Doi:https://doi.org/10.1016/0034-4257(79)90013-0.
- Ủy ban nhân dân tỉnh Quảng Nam, 2010. Kế hoạch thực hiện Đề án Nâng cao nhận thức cộng đồng và quản lý rủi ro thiên tai dựa vào cộng đồng trên địa bàn tỉnh Quảng Nam đến năm 2020. Báo cáo lưu trữ UBND tỉnh Quảng Nam, Tam Kỳ.
- Viện Khoa học Khí tượng Thủy văn và Môi trường, 2009. *Đề án xây dựng bản đồ hạn hán và mức độ thiếu nước sinh hoạt ở Nam Trung Bộ và Tây Nguyên*. Báo cáo lưu trữ tại Viện Khoa học Khí tượng thủy văn và môi trường, Hà Nội.
- Vũ Thị Thu Lan, 2011. Chuyên đề Đề xuất các giải pháp phòng tránh và giảm thiểu thiên tai lũ lụt, hạn hán tỉnh Quảng Nam, Dự án P1-08-VIE. Viện địa lý, Viện Khoa học và Công nghệ Việt Nam.
- Xiao, X., Hollinger, D., Aber, J., Goltz, M., Davidson, E., Zhang, Q., & Moore, B., 2004. *Satellite-based modeling of gross primary production in an evergreen needleleaf forest*. Remote Sensing of Environment, 89, 519-534. Doi:10.1016/j.rse.2003.11.008
- Yu, L., & Gong, P., 2012. *Google Earth as a virtual globe tool for Earth science applications at the global scale: progress and perspectives*. International Journal of Remote Sensing, 33(12), 3966-3986. Doi:10.1080/01431161.2011.636081

ABSTRACT

Application of geoinformatics technology for drought assessment in Quang Nam province

Pham Thi Thanh Hoa^{1,2*}, Tran Hong Hanh^{1,2}, Luu Thi Dieu Chinh³

¹Hanoi University of Mining and Geology

²Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology

³Hanoi University of Civil Engineering

This study assessed the level of drought in Quang Nam province based on geoinformatics technology. Google Earth Engine - a cloud computing platform - was used to calculate drought indicators from Sentinel-2 images, replacing a large amount of image processing in commercial software. GIS technology - geographic information system - was used to visualize and create drought maps for Quang Nam province. The drought indices used such as LSWI (Land Surface Water Index), NDDI (Normalized Difference Drought Index). The research results showed that droughts occurred in all districts of Quang Nam in 2023 with different levels and uneven distribution. Severe and extreme drought areas were mainly concentrated in the East of the province. In the West, due to the large forest, there was mild drought or no drought. In general, the approach using drought index and geoinformatics technology can be effective in quickly informing local authorities about the drought situation, thereby providing long-term adaptation solutions to drought.

Keywords: Drought, Quang Nam, geoinformatics, Sentinel-2, drought index.

Nghiên cứu ứng dụng mô hình LSTM (Long Short-Term Memory) trong dự báo độ mặn từ chuỗi dữ liệu đa biến tại trạm Đại Ngãi, Sóc Trăng

Bùi Duy Quỳnh^{1*}, Hà Thị Hằng¹, Lưu Thị Diệu Chinh², Trần Xuân Trường^{3,4}

¹Khoa Cầu đường - Trường Đại học Xây dựng Hà Nội

²Khoa Công trình thủy - Trường Đại học Xây dựng Hà Nội

³Khoa Trắc địa- Bản đồ và Quản lý đất đai - Trường Đại học Mỏ - Địa chất

⁴Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES) - Trường Đại học Mỏ - Địa chất

TÓM TẮT

Xâm nhập mặn đang đặt ra những thách thức đáng kể đối với các khu vực ven biển trên toàn thế giới. Các mô hình dự đoán xâm nhập mặn có thể cung cấp thông tin quý giá trong chủ động lên kế hoạch phòng chống, giảm thiểu tác động và ảnh hưởng của xâm nhập mặn. Nghiên cứu này ứng dụng mô hình học sâu mạng bộ nhớ dài-ngắn LSTM (Long Short-Term Memory) - một dạng đặc biệt của mạng nơ-ron tái phát (RNN-Recurrent Neural Network) - để dự báo độ mặn dựa trên việc phân tích chuỗi dữ liệu đa biến, bao gồm: dữ liệu độ mặn, mực nước và lượng mưa tại trạm Đại Ngãi từ năm 2011 đến năm 2022. Độ tin cậy của mô hình dự báo độ mặn được đánh giá dựa trên các chỉ số R², MAE và RMSE cho các tập dữ liệu kiểm chứng và dữ liệu kiểm tra. Kết quả đánh giá cho thấy mô hình dự báo độ mặn có độ tin cậy cao, với giá trị R² > 95,8, giá trị MAE dao động từ 0,1 g/L đến 0,29 g/L và giá trị RMSE cũng rất nhỏ, thay đổi trong phạm vi từ 0,34 g/L đến 0,54 g/L. Kết quả thực nghiệm này cho thấy việc áp dụng mô hình học sâu trên cơ sở sử dụng chuỗi dữ liệu đa biến trong dự báo độ mặn có ý nghĩa rất lớn trong việc cảnh báo sớm cũng như lên kế hoạch phòng tránh, thực thi các giải pháp ngăn mặn lâu dài. Ngoài ra, kết quả thực nghiệm này cũng có thể áp dụng hiệu quả cho các trạm quan trắc độ mặn khác trong khu vực.

Từ khóa: Dự báo độ mặn, chuỗi dữ liệu đa biến, LSTM, trạm Đại Ngãi.

1. Đặt vấn đề

Xâm nhập mặn là hiện tượng xảy ra khi nước biển xâm nhập vào các khu vực nước ngọt thông qua hệ thống sông, suối (Ippen và cộng sự, 1961) và thường xảy ra ở các khu vực ven biển và đồng bằng nơi có sự giao thoa giữa nước biển và nước ngọt (Mahmuduzzaman và cộng sự, 2014). Đồng bằng sông Cửu Long (ĐBSCL) được coi là một trong những vùng bị ảnh hưởng nặng nề trên thế giới trước các thiên tai như hạn hán, xâm nhập mặn (Xiao và cộng sự, 2021). Hiện nay, với sự gia tăng của mực nước biển dâng do biến đổi khí hậu (BĐKH), việc trữ nước tại các đập thủy điện đầu nguồn, sự hạ thấp đáy sông do khai thác cát, sự gia tăng sử dụng nước nội vùng đã và đang làm cho tình trạng xâm nhập mặn ngày càng trở nên nghiêm trọng, nước biển lấn sâu vào nội đồng, gây ảnh hưởng lớn đến đời sống sinh hoạt và sản xuất tại khu vực ĐBSCL (Duong Tran Anh và cộng sự, 2018; Doan Van Binh và cộng sự, 2020; Nguyen Huu Thoi, 2001). Để dự báo xâm nhập mặn, một số nghiên cứu thường sử dụng bộ mô hình MIKE và thu được kết quả tương đối tốt (Đặng Hoàng Lam và các cộng sự, 2022; Đặng Văn Dũng và các cộng sự, 2018; Đoàn Quang Trí, 2016); tuy nhiên, các mô hình này thường yêu cầu số lượng dữ liệu đầu vào lớn và chi tiết, như dữ liệu địa hình, mặt cắt sông, các công trình thủy lợi, hệ số nhám,... để kiểm định mô hình và cần nhiều thời gian tính toán.

Ngày nay, các mô hình trí tuệ nhân tạo đã cung cấp thêm một cách tiếp cận mới trong việc xây dựng các mô hình dự báo độ mặn dựa trên chuỗi số liệu theo thời gian và đạt được độ chính xác cũng như độ tin cậy cao. Các mô hình trí tuệ đã từng được áp dụng nhiều trong dự báo độ mặn và cho kết quả tốt, như mô hình trung bình Bayes (BMA) (Kairong Lin và các cộng sự, 2019; Nguyen, Thu Thi Minh và các cộng sự, 2023), ARIMA (Tran Thanh Thai và các cộng sự, 2021), rừng ngẫu nhiên (RF), máy vector hỗ trợ (SVM), mạng nơ-ron Elman (ENN) (Pengyu Lu và các cộng sự, 2021) hoặc các mô hình kết hợp Boruta - mạng nơ-ron nhân tạo (B-ANN), Boruta - hồi quy vector hỗ trợ (B-SVR) (Abbasi, Mahdi và các cộng sự, 2022). Song song với các mô hình học máy truyền thống, mô hình học sâu (Deep Learning) mà điển hình là mạng nơ-ron hồi quy (RNN- Recurrent Neural Network) với phiên bản mở rộng là mạng bộ nhớ dài ngắn LSTM (Long Short-Term Memory) đã được sử dụng trong các bài toán về dự báo chuỗi thời gian, như dự đoán lưu lượng xả nước hàng ngày trong lưu vực (Gauch, M. và các cộng sự, 2021), dự báo độ mặn (Nguyễn Công Thành và Nguyễn Tiên Giang, 2022), đánh giá sự dịch chuyển của cầu (Thai Ha Vu, Ngoc Quang Vu và Nguyen Van Thieu, 2024), hay dự báo độ cao sóng (Trần

* Tác giả liên hệ

Email: quynhbd@huce.edu.vn

Hồng Thái và các cộng sự, 2022). Hầu hết các nghiên cứu này đều khẳng định tính hiệu quả của mô hình LSTM trong dự báo, cho kết quả khả quan nhờ khả năng ghi nhớ các bước và không bị ảnh hưởng nhiều khi số liệu đầu vào bị thiếu. Trong một nghiên cứu dự báo độ mặn mới đây nhất của Thành và cộng sự (Nguyễn Công Thành và Nguyễn Tiền Giang, 2022) vào năm 2023, nghiên cứu này chỉ sử dụng một biến dữ liệu đầu vào duy nhất, đó là dữ liệu độ mặn cao nhất trong ngày tại trạm Đại Ngãi trong 6 tháng mùa khô (từ tháng 1 đến tháng 6) năm 2002 đến năm 2021 để dự báo độ mặn. Kết quả của nghiên cứu này cho thấy mô hình LSTM mặc dù có thể xử lý và dự báo chuỗi dữ liệu theo chuỗi thời gian khá tốt, nhưng mô hình không dự đoán được các giá trị cực trị trong quá trình huấn luyện. Trong khi đó, ưu điểm nổi trội của mô hình LSTM là xử lý dữ liệu đa biến theo chuỗi thời gian (S. Alhirmizy và B. Qader, 2019). Đối với dự báo xâm nhập mặn, mô hình LSTM có thể xử lý tốt dữ liệu đa biến, nơi mà nhiều biến số có thể ảnh hưởng qua lại lẫn nhau theo thời gian. Bằng cách này, LSTM không chỉ phân tích từng biến số riêng lẻ mà còn học được các mối quan hệ phức tạp giữa chúng, qua đó, giúp cải thiện độ chính xác của mô hình (Zhang, Xiaoying và các cộng sự, 2023).

Chính vì vậy, nghiên cứu này đề xuất và trình bày kết quả ứng dụng mô hình bộ nhớ dài-ngắn LSTM trong dự báo độ mặn tại trạm đo mặn Đại Ngãi, tỉnh Sóc Trăng dựa trên chuỗi dữ liệu đa biến, đa thời gian trong quá khứ, bao gồm độ mặn, mực nước, lượng mưa. Trên cơ sở đó đánh giá hiệu quả, độ tin cậy và khả năng áp dụng rộng rãi mô hình này cho các trạm quan trắc độ mặn khác trong khu vực.

2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Trạm thủy văn Đại Ngãi ($9^{\circ}44'4,60''N$, $106^{\circ}4'27,03''E$) là một trong các trạm quan trắc khí tượng của tỉnh Sóc Trăng, nằm ở khu vực cửa Trần Đề của nhánh Sông Hậu, thuộc thị trấn Đại Ngãi, huyện Long Phú, tỉnh Sóc Trăng (Hình 1). Hiện nay, tại trạm Đại Ngãi quan trắc nhiều thông tin thủy văn như mực nước, độ mặn, đo mưa tự động,... phục vụ cung cấp thông tin khí tượng thủy văn cho khu vực cũng như Ủy ban hội sông Mê Kông vì sự phát triển bền vững của khu vực này.



Hình 1. Vị trí trạm thủy văn Đại Ngãi, tỉnh Sóc Trăng

2.2. Số liệu sử dụng

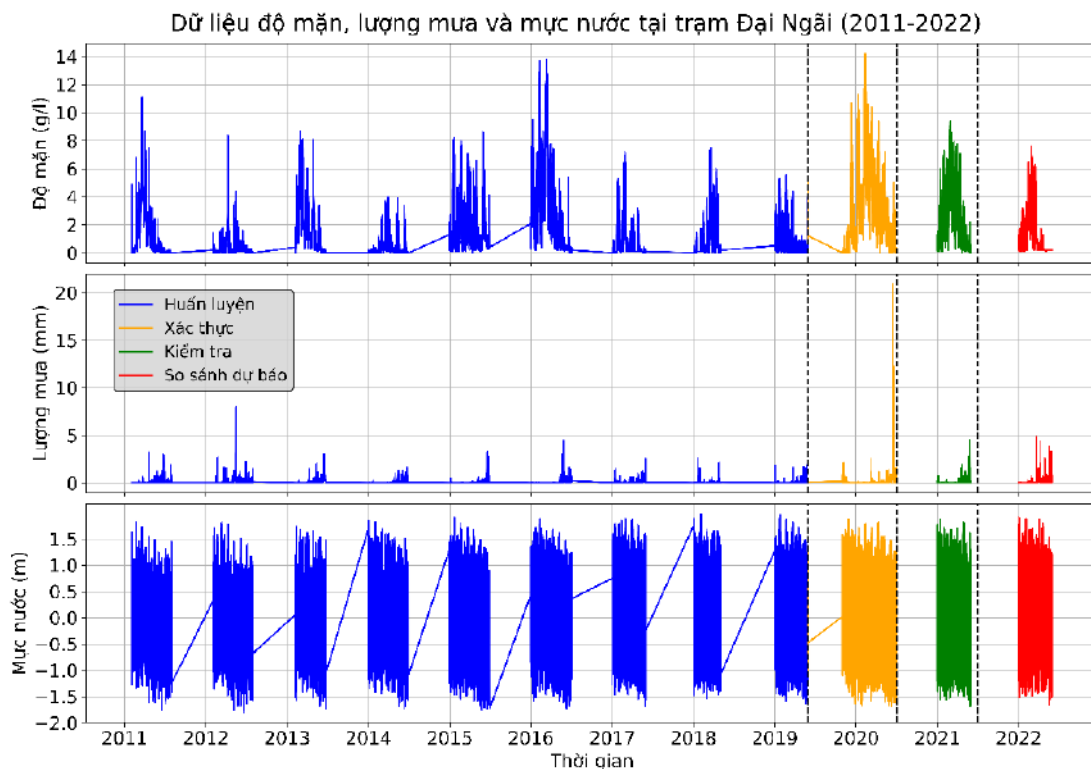
Trong nghiên cứu này, số liệu được sử dụng gồm: Số liệu quan trắc mặn (g/L), số liệu mực nước (cm) được đo liên tục (2 h/lần) trong ngày tại trạm Đại Ngãi trong 12 năm (2011-2022) ở các tháng mùa khô (từ tháng 01 đến tháng 6); số liệu lượng mưa thu thập từ dữ liệu vệ tinh và được lấy từ trang web của dự án POWER (Dự đoán tài nguyên năng lượng toàn cầu) của NASA, <https://power.larc.nasa.gov/data-access-viewer/>.

Toàn bộ dữ liệu được chia thành 4 phần: 70% cho tập huấn luyện (training), 10% cho tập xác thực (validation), 10% cho tập kiểm tra (testing) và 10% cho tập so sánh dự báo (prediction) được thể hiện trong Hình 2. Số liệu từ 01/01/2011 tới 31/05/2019 dùng để huấn luyện mô hình, số liệu từ 01/11/2019 đến 30/6/2020 dùng để kiểm chứng, số liệu từ 01/01/2021 đến 31/5/2021 dùng để kiểm tra và phần còn lại từ 01/01/2022 đến 31/5/2022 dùng để kiểm tra kết quả dự báo trong tương lai.

2.3. Phương pháp nghiên cứu

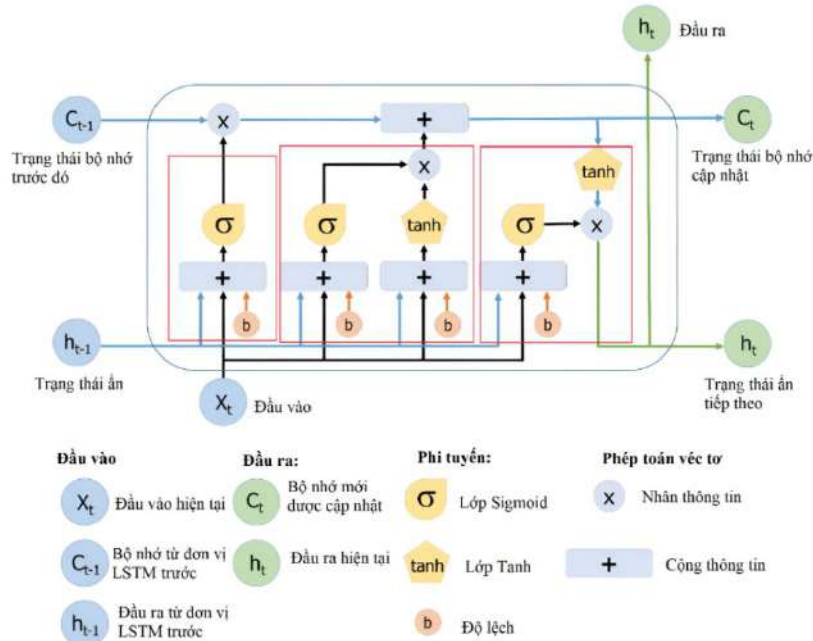
a. Mạng LSTM

Mô hình mạng LSTM được xây dựng dựa trên mạng nơron hồi quy RNN do Hochreiter và Schmidhuber đề xuất năm 1997 (Hochreiter, S. và Schmidhuber, J., 1997) và ngày càng được cải tiến (Yao, Kaisheng và các cộng sự, 2015) với mục đích khắc phục các thiếu sót của RNN bao gồm sự phụ thuộc lâu dài trong RNN và hiện tượng gradient biến mất và bùng nổ. Về lý thuyết, RNN có khả năng tự xử lý các phụ thuộc theo thời gian (temporal dependencies) bằng việc sử dụng bộ nhớ ngắn hạn và dựa trên việc xác định các tham số một cách hiệu quả. Tuy nhiên, trong thực tế RNN không thể tự giải quyết các phụ thuộc theo thời gian khi chuỗi số liệu có các phụ thuộc xa (long-term dependencies) (Bengio, Samy và Bengio, Yoshua, 2000).



Hình 2. Dữ liệu thủy văn tại trạm Đại Ngãi từ 2011-2022 (Nhóm tác giả thu thập)

LSTM có cấu trúc giống như RNN, nhưng cấu trúc bên trong thì lại phức tạp hơn nhằm khắc phục các vấn đề của RNN (Hình 3). Có ba cổng trong LSTM với các chức năng khác nhau: cổng đầu vào (kiểm soát mỗi quan hệ trạng thái và dữ liệu đầu vào), cổng đầu ra (kiểm soát quan hệ trạng thái và đầu ra) và cổng quên (kiểm soát thang thời gian bộ nhớ).



Hình 3. Cấu trúc của mạng LSTM

Công quen sử dụng hàm sigmoid áp dụng cho dữ liệu đầu vào hiện tại và trạng thái ẩn trước đó để quyết định sự đóng góp của chúng ở các trạng thái tiếp theo, được biểu diễn qua công thức (1). Đây chính là công có vai trò chia khóa để giải quyết các vấn đề tồn tại của mô hình RNN truyền thống. Công vào được trộn giữa hàm sigmoid (xác định các giá trị cần cập nhật) và hàm tanh (tạo ra các giá trị mới) để áp dụng cho trạng thái đầu vào và trạng thái ẩn để tính toán trạng thái hiện tại của của ô LSTM, biểu diễn qua công thức (2) và (3). Đầu ra của hai công cuối cùng được sử dụng cho tính trạng thái ẩn tiếp theo của ô, biểu diễn qua công thức (4) và (5).

$$f_t = s(W_f \cdot [h_{t-1}, X_t] + b_f) \quad (1)$$

$$i_t = s(W_i \cdot [h_{t-1}, X_t] + b_i) \quad (2)$$

$$C_t = f_t \times c_{t-1} + i_t \times \tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \quad (3)$$

$$o_t = s(W_o \cdot [h_{t-1}, X_t] + b_o) \quad (4)$$

$$h_t = o_t \times \tanh(C_t) \quad (5)$$

b. Thiết lập mô hình

Dữ liệu gồm độ mặn, lượng mưa và mực nước đo 2 h/lần tại trạm Đại Ngãi, được sử dụng làm dữ liệu đa biến đầu vào cho mô hình LSTM. Trong nghiên cứu này, mô hình LSTM áp dụng độ dài chuỗi đầu vào trong quá khứ kí hiệu $w = 84$ (tương ứng với dữ liệu 7 ngày), $w = 168$ (14 ngày) và $w = 360$ (30 ngày) để đánh giá cho 2 h tiếp theo, được kí hiệu là $w84$, $w168$ và $w360$. Ngoài ra, mô hình LSTM còn bị ảnh hưởng của các tham số đầu vào, vì vậy, việc lựa chọn các tham số này rất quan trọng. Do đó, trong nghiên cứu này đã sử dụng phương pháp tìm kiếm ngẫu nhiên (Random Search) dựa trên tập dữ liệu huấn luyện và tập dữ liệu kiểm chứng để tối ưu hóa các tham số cho từng loại kích thước cửa sổ trượt. Random search sẽ áp dụng các bộ thông số khác nhau của các mô hình được thiết lập, qua đó tìm được bộ thông số tối ưu nhất thông qua bộ dữ liệu huấn luyện và kiểm chứng. Quá trình thực hiện tìm kiếm cũng như chạy mô hình được thực hiện trên nền tảng Google Colab bằng ngôn ngữ Python kết hợp các thư viện có sẵn scikit-learn. Các siêu tham số này được trình bày trong Bảng 1 dưới đây.

Bảng 1. Các siêu tham số của mô hình LSTM

Siêu tham số	Phạm vi	Mô hình		
		w = 84 (7 ngày)	w = 168 (14 ngày)	w = 360 (30 ngày)
Số đơn vị ẩn LSTM	[16,32,64,128,256]	16	32	128
Dropout	[0,1, 0,2, 0,3, 0,5]	0,5	0,3	0,5
Batch size	[16,32,64,128,256]	256	256	256
Epochs	[10, 50, 100, 500]	50	50	50

c. Các chỉ số đánh giá độ tin cậy của mô hình

Để đánh giá độ tin cậy của mô hình dự báo LSTM, nghiên cứu này sử dụng một số chỉ số sau, bao gồm: hệ số xác định R^2 (Di Bucchianico, Alessandro, 2008; Wright, Sewall, 1921), sai số trung phương (RMSE) (Nevitt, Jonathan và Hancock, Gregory R., 2000), độ lệch tuyệt đối (MAE) (Sammur, Claude và Webb, Geoffrey I., 2010), giá trị phương sai được giải thích Ex_VAR (explained variance) (Nevitt, Jonathan và Hancock, Gregory R., 2000). Các giá trị RMSE và MAE đạt giá trị tối ưu khi bằng 0; giá trị R^2 và giá trị Ex_VAR đạt tối ưu khi bằng 100%. Các chỉ số này được tính theo công thức sau:

$$R^2 = \left(1 - \frac{\sum_{i=1}^n (y_{tt,i} - y_{db,i})^2}{\sum_{i=1}^n (y_{tt,i} - \bar{y}_{tt})^2}\right) \times 100 \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{tt,i} - y_{db,i})^2} \quad (7)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{db,i} - y_{tt,i}| \quad (8)$$

$$Ex_VAR = 100 \times \left(1 - \frac{var(y_{tt,i} - y_{db,i})}{var(y_{tt,i})}\right) \quad (9)$$

3. Kết quả và thảo luận

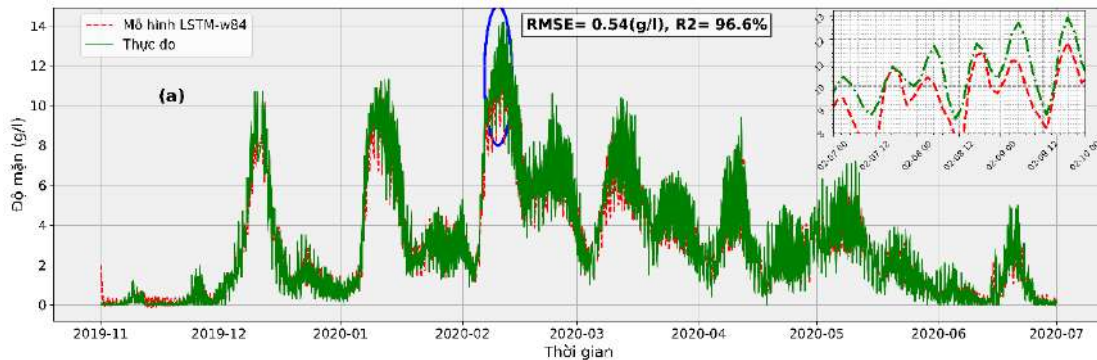
Với 3 mô hình LSTM có các khoảng dữ liệu đầu vào khác nhau (w84, w168 và w360) được huấn luyện với bộ tham số trong Bảng 1, sẽ tiến hành đánh giá các mô hình này dựa trên bộ dữ liệu trong tệp kiểm chứng và kiểm tra. Các chỉ số đánh giá mô hình được thông kê trong Bảng 2.

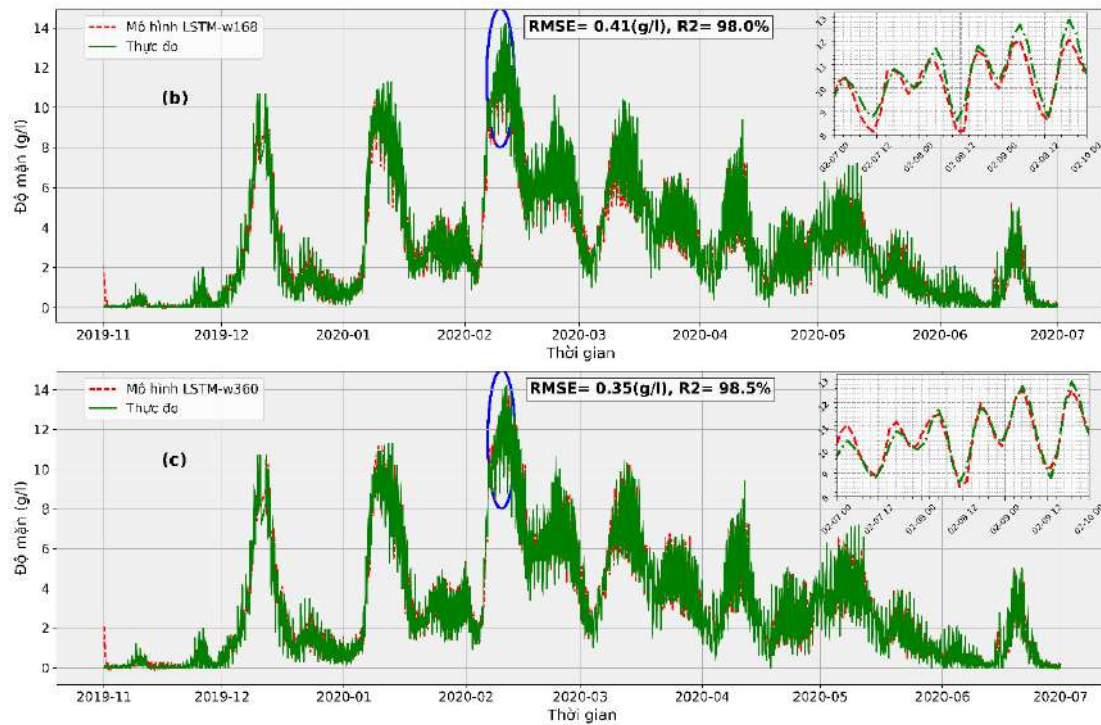
Dựa trên các chỉ số trong Bảng 2 cho thấy, các mô hình có độ chính xác rất cao với hệ số xác định R^2 đều lớn hơn 95% ở cả tệp kiểm chứng và tệp kiểm tra. Các chỉ số như RMSE và MAE đều nhỏ hơn 1 g/L.

Bảng 2. Các chỉ số đánh giá mô hình LSTM

Chỉ số	Mô hình w84		Mô hình w168		Mô hình w360	
	Kiểm chứng	Kiểm tra	Kiểm chứng	Kiểm tra	Kiểm chứng	Kiểm tra
RMSE (g/l)	0,54	0,43	0,41	0,34	0,35	0,39
MAE (g/L)	0,29	0,19	0,17	0,12	0,12	0,10
R^2 (%)	96,58	95,82	98,02	97,29	98,48	97,77
Var_ex (%)	96,77	95,96	98,05	97,35	98,48	97,72

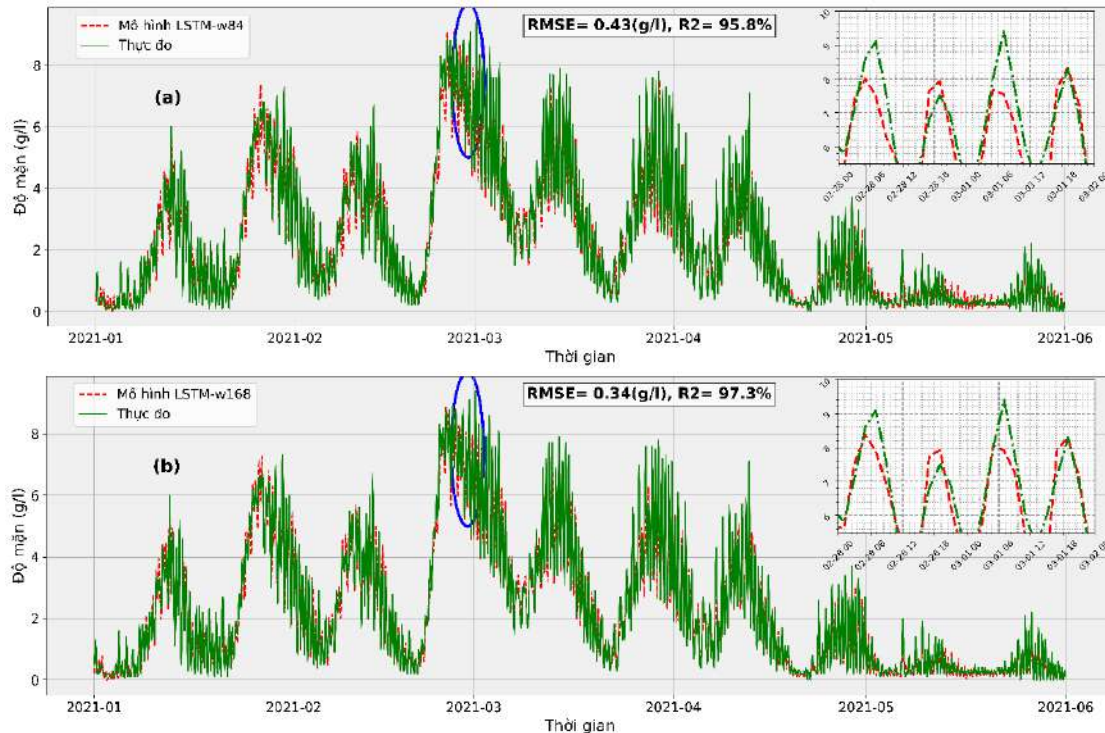
Hình 4 trình bày kết quả của các tệp kiểm chứng với các khoảng dữ liệu đầu vào khác nhau được dự báo bằng mô hình LSTM. Mặc dù, chuỗi dữ liệu độ mặn có sự biến đổi phức tạp, tuy nhiên, các kết quả cho thấy có sự tương đồng rất cao giữa giá trị thực đo và giá trị dự báo từ mô hình, kể cả tại các vị trí đỉnh mặn của chuỗi dữ liệu. Có thể thấy, mô hình LSTM-w360 có giá trị dự báo tại các vị trí đỉnh mặn sát với giá trị thực đo nhất trong 3 mô hình (khu vực phóng to Hình 4).

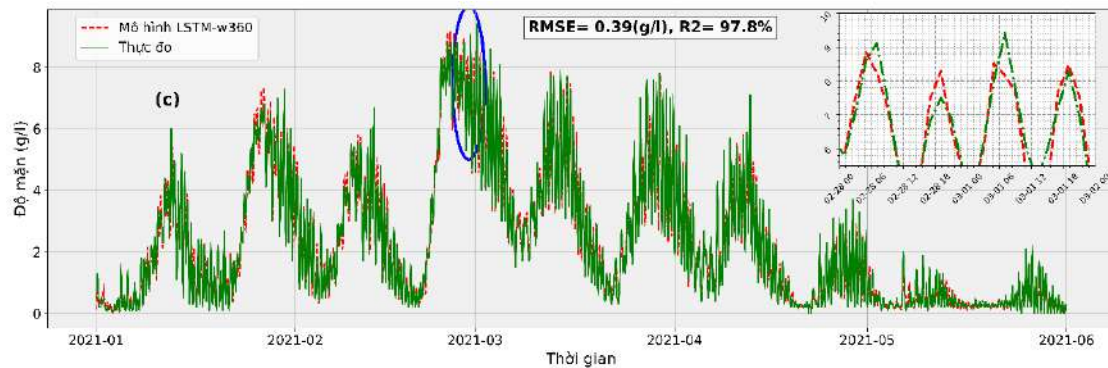




Hình 4. Kết quả các mô hình với các khoảng dữ liệu w84, w168 và w360 của tập kiểm chứng

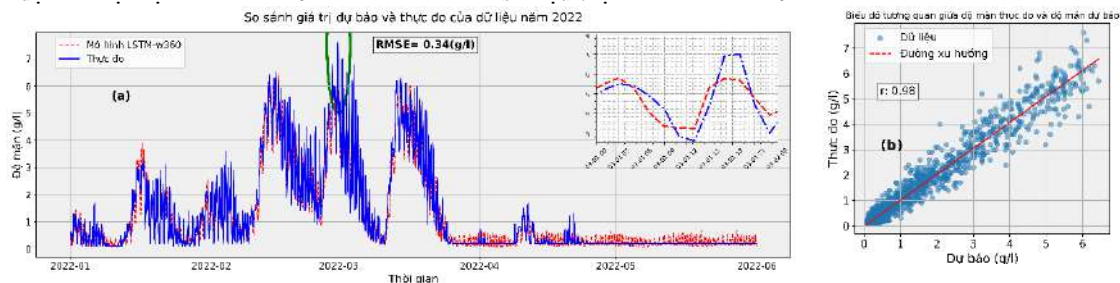
Tương tự, việc so sánh giữa kết quả của mô hình và dữ liệu thực tế đối với tập dữ liệu kiểm tra thể hiện trong Hình 5. Các kết quả cho thấy có sự tương đồng rất cao giữa giá trị thực đo và giá trị dự báo từ mô hình, kể cả tại các vị trí đỉnh mặn của chuỗi dữ liệu. Có thể thấy, mô hình LSTM-w360 có giá trị dự báo tại các vị trí đỉnh mặn sát với giá trị thực đo nhất trong các mô hình LSTM-w84, LSTM-w168 và LSTM-w360 (khu vực phóng to trong Hình 5).





Hình 5. Kết quả các mô hình với các khoảng dữ liệu w84, w168 và w360 của tập kiểm tra

Việc các mô hình LSTM có thể dự báo được trong tương lai dựa trên dữ liệu lịch sử, trong nghiên cứu đã tiến hành dự báo độ mặn cho năm 2022 và so sánh với số liệu thực đo của năm 2022.



Hình 6. Kết quả giữa mô hình dự báo và dữ liệu thực đo năm 2022

Với kết quả mô hình LSTM-w360 cho độ chính xác cao nhất nên được sử dụng làm mô hình dự báo độ mặn cho năm 2022. Kết quả dự báo được thể hiện trong Hình 6, cho thấy độ chính xác dự báo rất cao với sai số RMSE = 0,34 g/L (Hình 6a) và hệ số tương quan $r = 0,977$ (Hình 6b).

4. Kết luận

Các kết quả đạt được trong nghiên cứu này cho thấy mô hình LSTM có thể xử lý hiệu quả chuỗi số liệu đa biến, đa thời gian (độ mặn, mực nước và lượng mưa) trong dự báo độ mặn. Nghiên cứu này đã sử dụng mô hình LSTM với kích thước các chuỗi dữ liệu đầu vào khác nhau, bao gồm LSTM-w84, LSTM-w168, LSTM-w360 để dự báo độ mặn tại trạm Đại Ngãi, Sóc Trăng. Trong đó, mô hình LSTM-w360 có hiệu suất dự báo tốt nhất so với các mô hình còn lại (LSTM-w84 và LSTM-w168). Tuy nhiên, các mô hình này cần được thử nghiệm mở rộng hơn trong thời gian tới với sự tham gia của các yếu tố khác như gió, nhiệt độ, lưu lượng dòng chảy,... cũng như áp dụng dự báo độ mặn tại các trạm quan trắc khác trong khu vực để kiểm chứng độ chính xác và tính hiệu quả của mô hình LSTM này.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi đề tài khoa học và công nghệ cấp bộ “Nghiên cứu phương pháp đánh giá rủi ro xâm nhập mặn có xét tới ảnh hưởng của biến đổi khí hậu” mã số CT.2022.01.XDA.04.

Tài liệu tham khảo

- Di Bucchianico, Alessandro (2008). "Coefficient of Determination (R^2)", *Encyclopedia of Statistics in Quality and Reliability*, <https://doi.org/10.1002/9780470061572.eqr173>.
- Duong Tran Anh và cộng sự (2018). "Simulating future flows and salinity intrusion using combined one- and two-dimensional hydrodynamic modelling-the case of Hau River, Vietnamese Mekong Delta", *Water (Switzerland)*, 10(7).
- Nevitt, Jonathan & Hancock, Gregory R. (2000). "Improving the root mean square error of approximation for nonnormal conditions in structural equation modeling", *The Journal of Experimental Educational*, tr. 251-268.
- Sammut, Claude & Webb, Geoffrey I. (2010). "Mean Absolute Error", *Encyclopedia of Machine Learning*, Springer Reference, tr. 652.
- Wright, Sewall (1921). "Correlation and causation", *Journal of Agricultural Research*, XX(7), tr. 557-585.
- Abbasi, Mahdi et al. (2022). "Long-term streamflow prediction using hybrid svr-ann based on bayesian model averaging", *Journal of Hydrologic Engineering*, 27(11).

- Bengio, Samy & Bengio, Yoshua, IEEE Transactions on Neural Networks (2000). "Taking on the curse of dimensionality in joint distributions using neural networks", 11(3), tr. 550-557.
- Đặng Hoàng Lam và cộng sự (2022). "Xây dựng mô hình MIKE 11 phục vụ công tác dự báo thủy văn và xâm nhập mặn tỉnh Bến Tre", *Tạp chí Khí tượng thủy văn*, 740(1), tr. 38-49.
- Đặng Văn Dũng và cộng sự (2018). "Khai thác mô hình MIKE 11 trong dự báo, cảnh báo xâm nhập mặn vùng Đồng bằng sông Cửu Long", *Tạp chí Khí tượng thủy văn*, 693, tr. 48-58.
- Đoàn Quang Trí (2016). "Ứng dụng mô hình Mike 11 mô phỏng và tính toán xâm nhập mặn cho khu vực Nam bộ", *Tạp chí Khí tượng Thủy văn*, 11, tr. 39-46.
- Doan Van Binh et al. (2020). "Long-term alterations of flow regimes of the Mekong River and adaptation strategies for the Vietnamese Mekong Delta", *Journal of Hydrology: Regional Studies*, 32, tr. 100742.
- Gauch, M. et al. (2021). "Rainfall-runoff prediction at multiple timescales with a single Long Short-Term Memory network", *Hydrol. Earth Syst. Sci.*, 25(4), tr. 2045-2062.
- Hochreiter, S. và Schmidhuber, J. (1997). "Long Short-Term Memory", *Neural Computation*, 9(8), tr. 1735-1780.
- Ippen, Arthur T & Harleman, Donald RF (1961). *One-dimensional analysis of salinity intrusion in estuaries*, Waterways Experiment Station.
- Kairong Lin et al. (2019). "Modeling saltwater intrusion using an integrated Bayesian model averaging method in the Pearl River Delta", *Journal of Hydroinformatics*, 21(6), tr. 1147-1162.
- Mahmuduzzaman, Mohammed et al. (2014). "Causes of salinity intrusion in coastal belt of Bangladesh", 4(4A), tr. 8-13.
- Nguyễn Công Thành và Nguyễn Tiền Giang (2022). "Xây dựng mô hình máy học LSTM (Long Short Term Memory) phục vụ công tác dự báo mặn tại trạm đo mặn Đại Ngãi", *Tạp chí Khí tượng Thủy văn*, 740(1).
- Nguyen Huu-Thoi và Ashim Das Gupta (2001). "Assessment of Water Resources and Salinity Intrusion in the Mekong Delta", *Water International*, 26(1), tr. 86-95.
- Nguyen, Thu Thi Minh và cộng sự. (2023), "Mapping of Top-Soil Salinity Zoning in the Coastal Area of Ben Tre province, Vietnam", *International Journal of Intelligent Systems Applications in Engineering*, 11(4s).
- Pengyu Lu et al. (2021). "An integrated framework of input determination for ensemble forecasts of monthly estuarine saltwater intrusion", *Journal of Hydrology*, 598.
- S. Alhirmizy & B. Qader (2019). Multivariate Time Series Forecasting with LSTM for Madrid, Spain pollution, *2019 International Conference on Computing and Information Science and Technology and Their Applications (ICCISTA)*, tr. 1-5.
- Thai Ha Vu, Ngoc Quang Vu & Nguyen Van Thieu (2024). "Chapter 24 - Spatial prediction of bridge displacement using deep learning models: A case study at Co Luy bridge", *Applications of Artificial Intelligence in Mining and Geotechnical Engineering*, Elsevier, tr. 437-461.
- Trần Hồng Thái và cộng sự (2022). "Xây dựng mô hình mạng nơron hồi quy dự báo độ cao sóng có nghĩa tại trạm Cồn Cỏ, Quảng Trị, Việt Nam", *Tạp chí Khí tượng thủy văn*.
- Tran Thanh Thai et al. (2021). "Performance evaluation of Auto-Regressive Integrated Moving Average models for forecasting saltwater intrusion into Mekong river estuaries of Vietnam", *Vietnam Journal of Earth Sciences*, 43(4), tr. 428-443.
- Xiao, Han et al. (2021). "Saltwater intrusion into groundwater systems in the Mekong Delta and links to global change", 12(3), tr. 342-352.
- Yao, Kaisheng et al. (2015). "Depth-gated recurrent neural networks", *arXiv preprint arXiv:1503.03790*, 9, tr. 98.
- Zhang, Xiaoying et al. (2023). "Advance prediction of coastal groundwater levels with temporal convolutional and long short-term memory networks", *Hydrology and Earth System Sciences*, 27(1), tr. 83-96.

ABSTRACT

Research on the application of the LSTM (Long Short-Term Memory) model in salinity forecasting from multivariate data series at Dai Ngai station, Soc Trang province

Bui Duy Quynh^{1*}, Ha Thi Hang¹, Luu Thi Dieu Chinh², Tran Xuan Truong^{3,4}

¹*Faculty of Bridges and Roads, Hanoi University of Civil Engineering*

²*Faculty of Hydraulic Engineering, Hanoi University of Civil Engineering*

³*Faculty of Geodesy and Land Management, Hanoi University of Mining and Geology*

⁴*Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology*

Saltwater intrusion is posing significant challenges for coastal areas worldwide. Predictive models of saltwater intrusion can provide valuable information for proactive planning, mitigating impacts, and reducing the effects of saltwater intrusion. This study applies the Long Short-Term Memory (LSTM) deep learning model, a special type of Recurrent Neural Network (RNN), to forecast salinity based on multivariate time series analysis, including salinity data, water level, and rainfall at the Dai Ngai station from 2011 to 2022. The reliability of the salinity prediction model is assessed based on the R^2 , MAE, and RMSE indicators for validation and test datasets. The evaluation results show that the salinity prediction model is highly reliable, with an R^2 value > 95.8 , MAE values ranging from 0.1 g/L to 0.29 g/L, and RMSE values also being very small, varying from 0.34 g/L to 0.54 g/L. These experimental results indicate that applying deep learning models using time series data in salinity prediction is highly significant for early warning, planning, and implementing long-term saltwater intrusion prevention measures. These experimental results can also be effectively utilized at other regional salinity monitoring stations.

Keywords: Salinity forecasting, multivariate data series, LSTM, Dai Ngai station.

Kiểm định giả thuyết thống kê trong bình sai lưới quan trắc chuyển dịch ngang đập công trình thủy điện

Trần Trung Anh^{1,2*}, Nguyễn Quang Hà³

¹Khoa Trắc địa - Bản đồ và Quản lý Đất đai, Trường Đại học Mở - Địa chất

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

³Công ty Thủy điện Bản Vẽ

TÓM TẮT

Trong nghiên cứu này, nhóm tác giả đã áp dụng kiểm định giả thuyết thống kê Chi - bình phương (χ^2) để đánh giá hiệu quả của quá trình bình sai trong lưới quan trắc chuyển dịch ngang đập thủy điện. Sử dụng phần mềm ApnetPro, nhóm tác giả đã xử lý và phân tích dữ liệu đo quan trắc bằng máy toàn đạc điện tử (đo 286 góc đo, 194 cạnh đo được tính chuyển về mặt chiếu 150 m của hệ công trình) gồm 6 điểm lưới cơ sở, 38 điểm quan trắc trên đập thủy điện Bản Vẽ. Trọng số đầu vào theo độ chính xác máy toàn đạc theo lý lịch của máy có sai số đo góc là 0,5", sai số đo cạnh là 0,6 mm+1 ppm, với giá trị $\chi^2 = 1348.487$, thì mức kiểm định 95% không đạt kỳ vọng. Khi điều chỉnh trọng số theo sai số đo góc 1", sai số đo cạnh 1 mm+1 ppm, thì $\chi^2 = 420.394$ và đạt kiểm định giả thuyết thống kê. Kết quả bình sai cho thấy sự cải thiện đáng kể về độ chính xác (sai số tương hỗ cạnh yếu nhất từ 1/67740 xuống còn 1/81400), với sai số sau bình sai có xu hướng nhỏ hơn và phân phối đều hơn so với trước khi bình sai (sai số trung phương trọng số đơn vị từ 1.8270 hạ xuống còn 1.0201). Giá trị χ^2 tính toán được từ dữ liệu quan trắc sau bình sai nằm trong khoảng giá trị tới hạn từ bảng phân phối Chi - bình phương (với kiểm định ở mức ý nghĩa 95%), chỉ ra rằng không có sự khác biệt đáng kể giữa sai số quan trắc giả thuyết và kỳ vọng. Điều này khẳng định rằng quá trình bình sai đã được thực hiện một cách hiệu quả, góp phần vào việc nâng cao chất lượng tọa độ bình sai và phản ánh trung thực về chất lượng trị đo. Nghiên cứu mở ra hướng tiếp tục tối ưu hóa phương pháp bình sai và kiểm định thống kê, đồng thời áp dụng chúng trong các điều kiện quan trắc khác nhau.

Từ khóa: Lưới quan trắc chuyển dịch ngang, kiểm định giả thuyết thống kê, chi - bình phương, đập thủy điện.

1. Đặt vấn đề

Trong quản lý, vận hành các công trình thủy điện, việc theo dõi và đánh giá chính xác chuyển dịch ngang của đập là một yếu tố quan trọng cần theo dõi nhằm đảm bảo an toàn, ổn định. Lưới quan trắc chuyển dịch đập được thiết lập nhằm mục đích thu thập dữ liệu chính xác về các chuyển dịch và biến dạng của đập, từ đó phát hiện kịp thời các dấu hiệu bất thường có thể dẫn đến sự cố. Trong bối cảnh này, việc bình sai lưới quan trắc không chỉ giúp cải thiện độ chính xác của dữ liệu đo đạc mà còn góp phần vào việc đánh giá sự chuyển dịch khách quan chính xác (Phạm Quốc Khánh, 2021).

Tuy nhiên, một vấn đề đặt ra là làm thế nào để xác định liệu quá trình bình sai đã được thực hiện một cách hiệu quả hay chưa, điều này đòi hỏi phải có một phương pháp đánh giá khách quan và khoa học. Thông thường, sai số đo đạc trước bình sai được lấy từ thông số của máy đo từ đó làm cơ sở để tính trọng số của trị đo và kết quả sau bình sai thường bị ảnh hưởng bởi thông số này. Để đánh giá có sự phù hợp của sai số đo trước bình sai của dây trị đo với sai số nhận được sau bình sai cần bài toán kiểm định thống kê. Kiểm định giả thiết thống kê Chi - bình phương xuất hiện như một công cụ mạnh mẽ, cho phép chúng ta so sánh giữa sai số đo đạc trước và sau khi bình sai, từ đó đánh giá xem liệu có sự cải thiện đáng kể nào không đồng thời trả lời về sự phù hợp của sai số đo trước bình sai tương ứng với chất lượng trị đo (Charles D. Ghilani, 2017).

Các phần mềm bình sai ở nước ta thường bỏ qua công tác đánh giá kiểm định thống kê sau bình sai. Nghiên cứu này nhằm mục đích áp dụng kiểm định Chi - bình phương để kiểm tra giả thiết liệu quá trình bình sai lưới quan trắc chuyển dịch ngang tại đập thủy điện có thực sự mang lại kết quả chính xác hơn hay không, việc này được tự động hóa dưới sự xử lý của phần mềm ApnetPro. Bằng cách này, nhóm tác giả hy vọng sẽ cung cấp một cái nhìn sâu sắc hơn về việc áp dụng các phương pháp thống kê trong trắc địa, đồng thời góp phần nâng cao chất lượng quan trắc chuyển dịch ngang đập của công trình thủy điện.

2. Cơ sở của phương pháp nghiên cứu

2.1. Cơ sở của kiểm định thống kê

* Tác giả liên hệ

Email: trantrunganh@humg.edu.vn

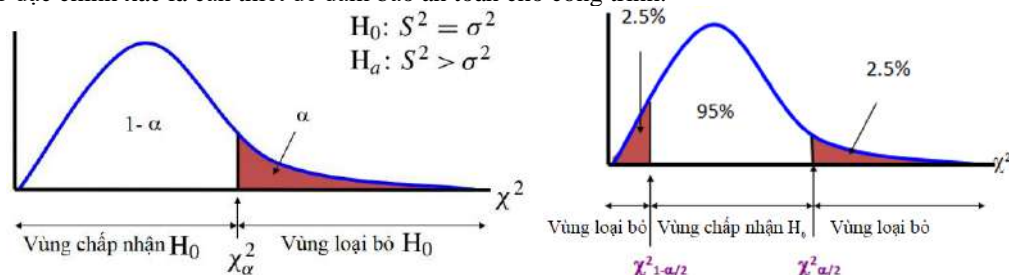
Cơ sở khoa học của kiểm định thống kê trong xử lý lưới trắc địa dựa trên việc sử dụng các phép thử thống kê để đánh giá sự chính xác và độ tin cậy của dữ liệu đo đạc. Trong đó, kiểm định Chi - bình phương (χ^2) là một công cụ thống kê mạnh mẽ, cho phép chúng ta so sánh xác suất của dãy trị đo với xác suất kỳ vọng dựa trên một phân phối nhất định. Phân phối Chi - bình phương là một phân phối xác suất được sử dụng rộng rãi trong các bài toán kiểm định giả thiết. Đặc điểm của phân phối này là nó chỉ có một tham số, đó là số bậc tự do (hay trị đo thừa), và hình dạng của phân phối sẽ thay đổi tùy thuộc vào giá trị của trị đo thừa. Phân phối này được sử dụng để mô tả tổng bình phương của các biến ngẫu nhiên độc lập, chuẩn hóa theo phân phối chuẩn (Charles D. Ghilani, 2017).

Bảng 1. Phân phối Chi - bình phương

	Kiểm định một đuôi	Kiểm định hai đuôi	
Giả thuyết ban đầu	$H_0 : S^2 = \sigma^2$	$H_0 : S^2 = \sigma^2$	(1)
Giả thuyết thay thế	$H_a : S^2 > \sigma^2$	$H_a : S^2 \neq \sigma^2$	(2)
Kiểm định thống kê	$\chi^2 = \frac{vS^2}{\sigma^2}$	$\chi^2 = \frac{vS^2}{\sigma^2}$	(3)
Vùng bác bỏ: Giả thuyết H_0 bị bác bỏ khi tiêu chí bác bỏ được thỏa mãn.	$\chi^2 > \chi^2_\alpha$	$\begin{cases} \chi^2 > \chi^2_{\alpha/2} \\ \chi^2 < \chi^2_{1-\alpha/2} \end{cases}$	(4)

trong đó: H_0 là giả thuyết ban đầu của S;
 H_a là giả thuyết thay thế của S;
 σ là giá trị kỳ vọng;
 v là số trị đo thừa;
 α là mức xác suất của kiểm định.

Kiểm định giả thiết thống kê chi bình phương được sử dụng để so sánh giữa tần suất quan sát và tần suất kỳ vọng. Nếu giá trị Chi - bình phương tính toán từ dữ liệu nằm ngoài khoảng giá trị tới hạn tìm được trong bảng phân phối chi bình phương, chúng ta có cơ sở để bác bỏ giả thuyết ban đầu và chấp nhận giả thuyết thay thế. Trong trắc địa, kiểm định Chi - bình phương có thể được áp dụng để đánh giá sự phân phối của sai số đo đạc. Khi bình sai lưới quan trắc chuyển dịch ngang đập thủy điện, sai số đo đạc của máy được đưa vào giả thuyết là giá trị kiểm định của máy. Kết quả bình sai lưới cần đánh giá sai số này có phù hợp với chất lượng trị đo hay không. Người dùng có thể điều chỉnh sai số đo đạc giả thuyết ban đầu, nếu điều chỉnh không phù hợp (hạ thấp quá hoặc nâng cao quá giá trị) thì kiểm định Chi - bình phương sẽ trả lời. Điều này đặc biệt quan trọng trong việc đánh giá chất lượng của lưới quan trắc chuyển dịch ngang trả đúng về chất lượng trị đo đạt được của máy đo, nơi mà việc đo đạc chính xác là cần thiết để đảm bảo an toàn cho công trình.



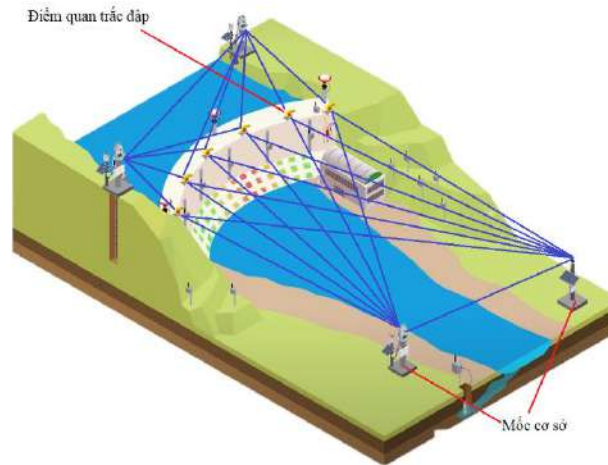
Hình 1. Phân phối Chi bình phương cho kiểm định 1 đuôi và 2 đuôi (Ảnh: financetrain.com)

Vùng chấp nhận giả thuyết ban đầu được biểu thị trong Hình 1 cho kiểm định 1 đuôi và 2 đuôi. Ở đây giá trị χ^2 giới hạn được tra từ bảng χ^2 dựa vào số trị đo thừa và mức xác suất tương ứng.

Mặc dù kiểm định chi bình phương là một công cụ hữu ích, nhưng nó cũng có những giới hạn. Các giả định cơ bản của kiểm định, như kích thước mẫu và tần suất kỳ vọng, cần được đáp ứng để kết quả kiểm định là hợp lệ. Ngoài ra, kiểm định này không cung cấp thông tin về mức độ của sự khác biệt hoặc về hướng của mối quan hệ. Nghiên cứu này được thiết kế để đánh giá hiệu quả của quá trình bình sai trong lưới quan trắc chuyển dịch ngang đập thủy điện thông qua việc áp dụng kiểm định giả thiết thống kê Chi - bình phương.

2.2. Phương pháp nghiên cứu

Lưới quan trắc chuyển dịch ngang đập (Hình 2) là lưới đo góc, cạnh với độ chính xác cao với trạm máy và gương đều có định tâm bắt buộc. Từ các điểm cơ sở đã phân tích độ ổn định, tiến hành đo bằng toàn đạc điện tử tới các điểm quan trắc chuyển dịch gấn trên đập thủy điện.



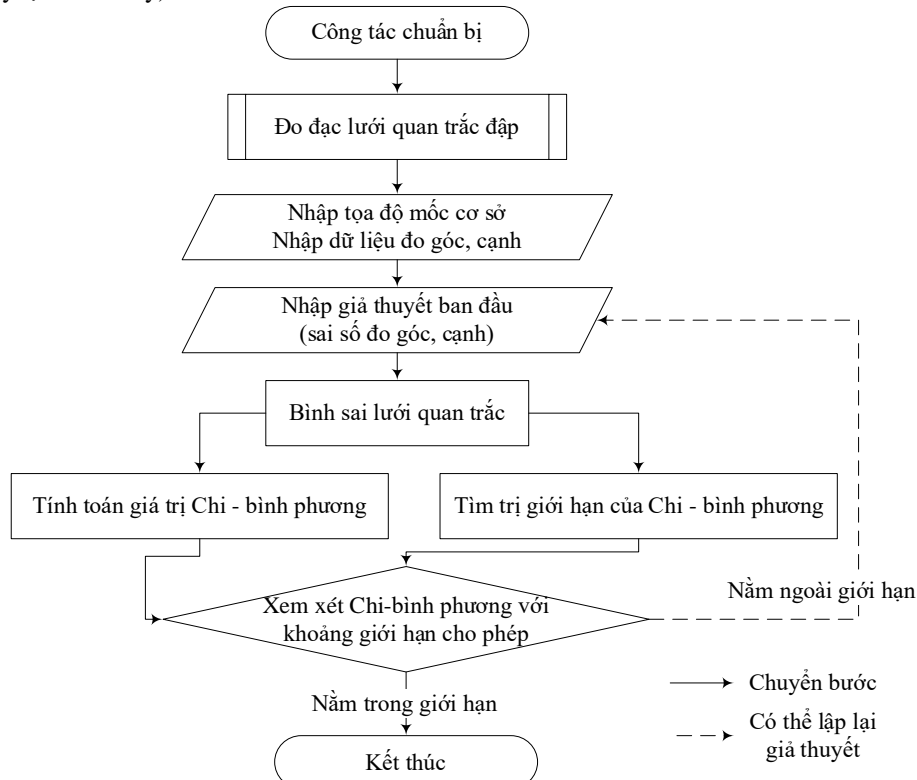
Hình 2. Lưới quan trắc đập thủy điện (ảnh Hexagon - Leica, chỉnh bởi Trần Trung Anh)

Sau khi đo đạc hiện trường, tiến hành xử lý trị đo góc, cạnh ở các vòng đo. Một việc quan trọng là cải chính trị đo cạnh về mặt chiếu công trình (công thức 5).

$$D_{mc} = D_{do} - \frac{H_p - H_{mc}}{R} D_{do} \quad (5)$$

trong đó: H_{mc} là độ cao mặt chiếu của hệ công trình (ở thực nghiệm lấy bằng 150 m);
 H_p là độ cao của điểm đứng máy đo tại mốc cơ sở;
 D_{do} là cạnh đo trực tiếp của máy toàn đạc điện tử;
 D_{mc} là cạnh đo được tính chuyển về mặt chiếu công trình;
 $R=6378000$ m là bán kính Trái đất.

Các bước xử lý lưới được trình bày trong quy trình ở Hình 3. Nhập tọa độ điểm gốc cơ sở đã xử lý và phân tích ổn định (Trần Trung Anh, Nguyễn Quang Hà, 2023), nhập trị đo góc, cạnh, cần thiết phải nhập giả thuyết thống kê ban đầu (đồng thời nhập sai số tiên nghiệm của góc, cạnh, giá trị này ở lần xử lý đầu tiên thường lấy theo thông số lý lịch của máy).



Hình 3. Quy trình kiểm định giả thuyết thống kê Chi - bình phương cho lưới quan trắc đập

Tiến hành bình sai lưới quan trắc: Ở bước này cần bình sai theo phương pháp phụ thuộc, tính tọa độ gần đúng của điểm quan trắc, cải chính trị đo cạnh về mặt chiếu công trình, lựa chọn phương pháp kiểm định (1 đuôi hoặc 2 đuôi).

Sau bình sai, tiến hành tính giá trị χ^2 trên cơ sở công thức (3) dựa vào trị đo thừa, sai số trung phương trọng số đơn vị nhận được sau bình sai, và sai số trung phương tiên nghiệm.

Tra bảng phân phối Chi - bình phương để xác định giá trị $\chi^2_{\alpha, v}$ giới hạn dựa vào trị đo thừa, mức kiểm định xác suất tin cậy thường chọn 95% với $\alpha = 0,05$ (Charles D. Ghilani 2017). Giá trị này còn phụ thuộc vào kiểm định 1 đuôi hoặc 2 đuôi ở công thức (4).

Trên cơ sở χ^2 tính toán sau bình sai, so sánh với $\chi^2_{\alpha, v}$ giới hạn để đưa ra quyết định là giả thuyết thống kê có phù hợp với dãy trị đo hay không. Ở đây có 2 trường hợp: trường hợp thỏa mãn điều kiện (4) tức là giả thuyết thống kê ban đầu được chấp nhận thì kết thúc xử lý lưới; trường hợp không thỏa mãn điều kiện (4) chúng ta phải thay đổi sai số đo góc, đo cạnh và tiến hành bình sai lại lần 2. Quá trình này có thể lặp lại nhiều lần để kết quả kiểm định giả thuyết thống kê được chấp nhận.

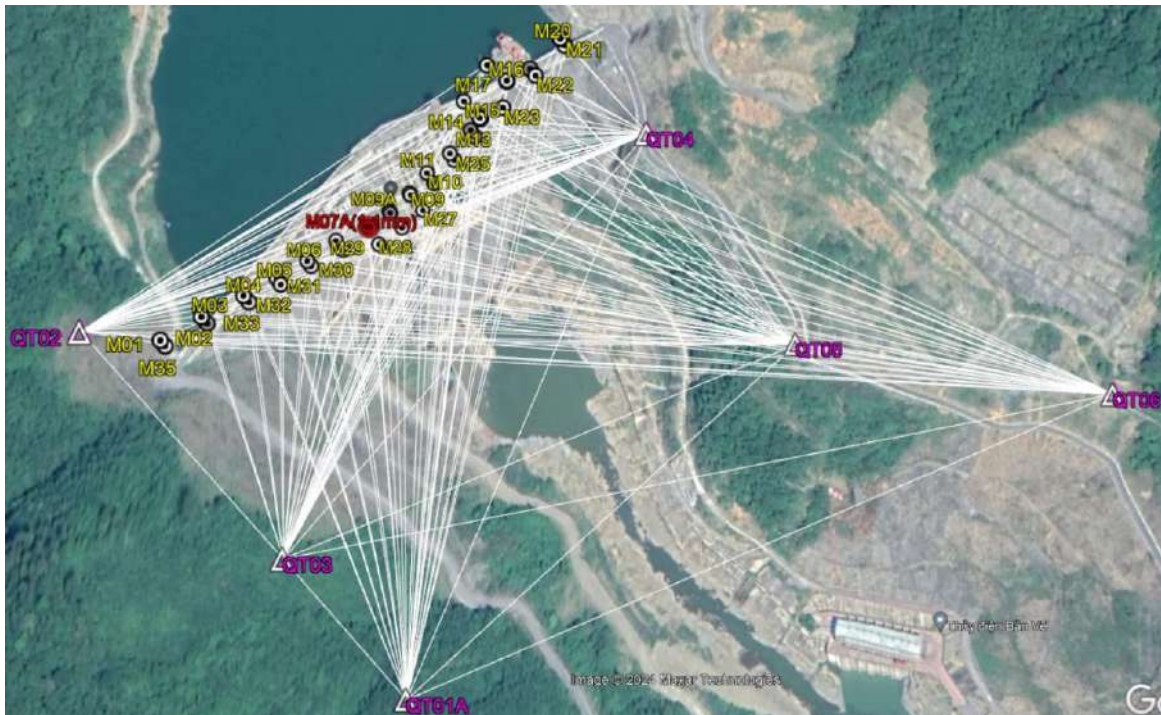
Trong quá trình phân tích dữ liệu, nhóm tác giả đã sử dụng phần mềm ApnetPro, một công cụ tốt trong xử lý bình sai lưới trắc địa có kiểm định thống kê. Phần mềm này xử lý nhanh chóng, cung cấp các modul để tính một số phương pháp kiểm định thống kê phục vụ bài toán bình sai trong xử lý số liệu quan trắc công trình.

3. Thực nghiệm

Khu vực thực nghiệm là thủy điện Bản Vẽ được xây dựng trên dòng sông Nậm Non tại xã Yên Na, huyện Tương Dương, tỉnh Nghệ An. Đây là thủy điện lớn nhất khu vực Bắc miền Trung với công suất lắp máy 320 MW, sản lượng điện trung bình năm 1.084 triệu kWh, mực nước dâng bình thường 200 m và có dung tích hồ chứa 1,83 tỷ m³ (EVN-GENCO1, 2017).

Tọa độ các điểm mốc cơ sở bao gồm cả độ cao đã được xử lý được lấy từ tài liệu đã công bố (Trần Trung Anh, Nguyễn Quang Hà, 2023).

Lưới quan trắc chuyển dịch ngang đập gồm 6 điểm mốc cơ sở, 38 điểm quan trắc gắn trên thành đập, các điểm gắn liền để định tâm (Hình 3).



Hình 4. Lưới quan trắc đập thủy điện Bản Vẽ

Đo góc cạnh bằng máy toàn đạc điện tử kèm phụ kiện gương tiêu chuẩn, máy và gương đều được định tâm bắt buộc. Máy toàn đạc điện tử Leica có thông số công bố độ chính xác đo góc 0,5"; độ chính xác đo cạnh đạt 0,6 mm + 1 ppm x D; (D là chiều dài cạnh được tính bằng km). Các thông số của máy đo là tiên nghiệm giả thuyết ban đầu của lần tính thứ 1.



Hình 4. Đo lưới quan trắc đập thủy điện Bản Vẽ (ảnh Nguyễn Quang Hà)

Tiến hành đo đặc hiện trường và xử lý trạm máy với 286 góc đo, 194 cạnh đo được trình bày sơ lược ở Bảng 2 (EVN-PECC1, 2022).

Bảng 2. Trị đo lưới quan trắc đập thủy điện Bản Vẽ

TT	Cạnh đo			Góc đo					
	Đầu	Cuối	Đo (m)	Trái	Giữa	Phải	Độ	Phút	Giây
1	QT04	M22	117.5766	QT01A	QT04	M32	46	20	43,1
2	QT01A	M15	551.4738	QT03	QT05	M31	31	59	58,1
3	QT05	M09	409.1877	QT05	QT01A	M21	325	26	51,3
...
66	QT03	M23	450.1947	QT03	QT05	M09A	44	34	59,7
67	QT06	M23	620.8871	QT04	QT03	M02	297	52	22,7
68	QT06	M27	676.9140	QT04	QT03	M13	341	38	37,2
...
128	QT02	M33	122.5105	QT01A	QT06	M11	43	43	31,8
129	QT02	M29	239.5238	QT01A	QT06	M12	45	59	11,6
130	QT02	M16	424.9572	QT01A	QT06	M13	47	07	08,3
...

Với đồ hình lưới quan trắc chuyển dịch ngang đập thủy điện Bản vẽ như trên, số trị đo thừa là 404, chọn mức kiểm định xác suất tin cậy 95% với $\alpha = 0,05$. Kiểm định thống kê Chi - bình phương được thực hiện thông qua phần mềm ApnetPro của tác giả Trần Trung Anh với việc sử dụng thuật toán kiểm định giả thuyết thống kê và tự động tra bảng Chi - bình phương được tích hợp sẵn. Điều này không chỉ giúp tăng cường độ chính xác của kết quả kiểm định mà còn giảm thiểu khả năng sai sót do tính toán thủ công.

Bảng 3. Kiểm định giả thuyết thống kê Chi - bình phương lưới quan trắc đập thủy điện Bản Vẽ

Chỉ số	Lần tính 1	Lần tính 2	Lần tính 3	Ghi chú
Giả thuyết ban đầu	$S^2 = \sigma^2 = 1$	$S^2 = \sigma^2 = 1$	$S^2 = \sigma^2 = 1$	
Sai số đo góc	0,5''	1''	2''	
Sai số đo cạnh	0,6 mm+1 ppm	1 mm+1 ppm	2 mm+1 ppm	
Giả thuyết thay thế	$S^2 = 3.33784$	$S^2 = 1.040579$	$S^2 = 0.294689$	
Chi - bình phương (χ^2)	1348.487	420.394	119.055	
Giá trị giới hạn chấp nhận của (χ^2) cho kiểm định 1 đuôi	$\chi^2_{\alpha} = 451.865$			$v = 404$ $\alpha = 0,05$
Khoảng giới hạn chấp nhận của (χ^2) cho kiểm định 2 đuôi	$[\chi^2_{1-\alpha/2} = 350.205; \chi^2_{\alpha/2} = 461.582]$			
Sai số trung phương vị trí điểm yếu nhất	1,2 mm	1,1 mm	1,1 mm	
Sai số tương hỗ cạnh yếu nhất	1/67740	1/81400	1/86272	
Quyết định theo kiểm định 1 đuôi	Loại bỏ	Chấp nhận	Chấp nhận	
Quyết định theo kiểm định 2 đuôi	Loại bỏ	Chấp nhàn	Loại bỏ	

Nhóm tác giả tiến hành xử lý với cả 2 phương án kiểm định giả thuyết thống kê 1 đuôi và 2 đuôi với Chi - bình phương giới hạn. Số lần tính là 3 lần, kết quả được trình bày trong Bảng 3.

Từ kết quả ở Bảng 3, nhóm tác giả nhận thấy với mức giả thuyết thống kê ban đầu $S^2 = \sigma^2 = 1$, thì tiên nghiệm theo thông số máy đo của nhà sản xuất ở lần tính 1, nhận được $\chi^2 = 1348.487$, thì kiểm định giả thuyết thống kê không đạt ở cả 2 phương án 1 đuôi và 2 đuôi. Cho nên, nhóm tác giả đã điều chỉnh thông số này ở lần lặp thứ 2 với sai số đo góc 1", cạnh 1 mm + 1 ppm, nhận được $\chi^2 = 420.394$ kiểm định giả thuyết thống kê đạt trong vùng chấp nhận cho cả 2 phương án. Để khách quan, nhóm tác giả tiếp tục thử nghiệm thay đổi thông số sai số đo góc lên 2", cạnh 2 mm+1 ppm thì nhận được $\chi^2 = 119.055$ giảm xuống, kiểm định giả thuyết thống kê theo phương án 1 đuôi thì đạt, phương án 2 đuôi thì bị loại bỏ. Nhận thấy sai số tương hỗ vị trí điểm đều tăng qua các lần tính 2 và 3, riêng lần 3 giá trị Chi - bình phương giảm nhưng sai số tương hỗ cạnh lại tăng, điều này là bất hợp lý vì độ tin cậy của lần 3 không tốt bằng lần 2. Do vậy, kết quả được chấp nhận ở lần tính thứ 2 là hợp lý.

4. Kết luận

Nghiên cứu này đã áp dụng kiểm định giả thiết thống kê Chi - bình phương để đánh giá hiệu quả của quá trình bình sai trong lưới quan trắc chuyển dịch ngang đập thủy điện. Các kết quả thu được từ phân tích dữ liệu đã cung cấp bằng chứng vững chắc về sự cải thiện đáng kể trong độ chính xác của dữ liệu quan trắc đồng thời với độ tin cậy sau khi bình sai. Quá trình bình sai cùng kiểm định giả thuyết thống kê đã được chứng minh là hiệu quả thông qua sự phân phối đồng nhất của giả thuyết trước bình sai và kết quả sau bình sai. Điều này góp phần vào việc nâng cao chất lượng xử lý dữ liệu quan trắc, từ đó tăng cường độ tin cậy trong việc theo dõi và đánh giá an toàn đập thủy điện. Quy trình xử lý tự động gồm cả bình sai và kiểm định giả thuyết thống kê bằng phần mềm ApnetPro đảm bảo sự thống nhất và tránh sai sót. Nghiên cứu khẳng định rằng kiểm định giả thiết thống kê Chi - bình phương là một công cụ hữu ích trong việc đánh giá và cải thiện quy trình bình sai lưới quan trắc chuyển dịch ngang đập thủy điện. Việc áp dụng phương pháp này góp phần vào việc đảm bảo an toàn và tăng cường hiệu quả quản lý rủi ro cho các công trình thủy điện.

Kiến nghị đơn vị quản lý vận hành công trình thủy điện áp dụng quy trình và phần mềm xử lý số liệu ApnetPro có kiểm định thống kê để tăng năng suất lao động, đảm bảo độ chính xác khi xử lý lưới quan trắc chuyển dịch ngang đập.

Tài liệu tham khảo

Trần Trung Anh, Nguyễn Quang Hà (2023). Xây dựng lưới tam giác thủy công trên cơ sở kết hợp tính chuyển tọa độ và bình sai tự do, Hội nghị toàn quốc Trắc địa công trình vì sự phát triển bền vững ESSD2023, Nhà xuất bản Khoa học tự nhiên và Công nghệ, trang 147-154.

Benzao Tao, bản dịch tiếng Việt của Phan Văn Hiến và nnk. (2017). Bình sai lưới tự do và phân tích biến dạng, Nhà xuất bản Tài nguyên Môi trường và Bản đồ Việt Nam.

Charles D. Ghilani (2017). Adjustment Computations: Spatial Data Analysis, 6th Edition, John Wiley & Sons.

Đặng Nam Chinh (Chủ biên), 2015. Giáo trình Bình sai lưới Trắc địa, Nhà xuất bản Khoa học và Kỹ thuật.

EVN (2005). Quyết định số 4389/CV-EVN-TĐ ngày 26/8/2005 của Tổng Công ty Điện lực Việt Nam về xây dựng lưới tam giác thủy công, lưới thủy chuẩn thủy công phục vụ thi công và quản lý vận hành các công trình thủy điện.

EVN-GENCO1 (2011). Giới thiệu Công ty Thủy điện Bản Vẽ, website: <https://banvehpc.com/c3/vi-VN/gioi-thieu/Gioi-thieu-chung-2-2288>, truy cập ngày 30/6/2024.

EVN-PECC1 (2022). Công ty CP Tư vấn Điện 1, Báo cáo kỹ thuật Nhà máy thủy điện Bản Vẽ - Giai đoạn vận hành, Quan trắc biến dạng đập dâng, đập tràn và cửa nhận nước chu kỳ 16.

Phạm Quốc Khánh (2021). Ứng dụng kiểm định thống kê xác định điểm không ổn định trong lưới cơ sở quan trắc chuyển dịch biến dạng công trình, Tạp chí Khoa học kỹ thuật Mỏ - Địa chất, tập 2, kỳ 62, trang 35-41.

Phạm Quốc Khánh, Nguyễn Thị Kim Thanh (2021). Ứng dụng phương pháp ước lượng vững theo phương sai hậu nghiệm phát hiện sai số thô trong lưới khống chế trắc địa, Tạp chí Khoa học kỹ thuật mỏ - Địa chất, tập 2, kỳ 62, trang 36-42.

ABSTRACT

Statistical hypothesis testing for horizontal monitoring network adjustment in a hydroelectric dam

Tran Trung Anh^{1,2*}, Nguyen Quang Ha³

¹*Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology*

²*Geomatics in Earth Sciences Research Group (GES), Hanoi University of Mining and Geology*

³*EVN-GENCO1*

In this study, we applied the Chi-square statistical test (χ^2) to evaluate the effectiveness of the adjustment process in the transverse displacement monitoring grid of a hydroelectric dam. Using the ApnetPro software, we processed and analyzed measurement data obtained from electronic theodolites (286 angle measurements and 194 distance measurements transformed to a 150 m reference plane of the dam structure). The input weights based on the theodolite accuracy were 0,5" for angle measurement error and 0,6 mm+1 ppm for distance measurement error. The calculated χ^2 value was 1348.487, which did not meet the expected significance level at 95%. However, after adjusting the weights to 1" for angle measurement error and 1 mm+1ppm for distance measurement error, the χ^2 value became 420.394, meeting the statistical hypothesis testing criteria. The results of the adjustment process showed significant improvement in accuracy (weakest equivalent distance error reduced from 1/67740 to 1/81400). The post-adjustment errors tended to be smaller and more evenly distributed compared to pre-adjustment errors (mean square error of unit weight decreased from 1.8270 to 1.0201). The calculated χ^2 value from the post-adjustment measurement data fell within the critical range from the Chi-square distribution table (at a 95% confidence level), indicating no significant difference between the theoretical measurement error and the expected error. This confirms the effective implementation of the adjustment process, contributing to improved coordinate quality and reflecting measurement accuracy. The study suggests further optimization of the adjustment method and statistical testing, applying them to various monitoring conditions.

Keywords: Horizontal monitoring network, statistical hypothesis testing, Chi - square, hydroelectric dam.

Phân tích, lọc và hiển thị dữ liệu đám mây điểm LiDAR với thư viện Pandas và Matplotlib của Python

Nguyễn Thị Hữu Phương^{1*}, Phạm Thị Hải Vân¹, Đào Thị Hồng Thắm¹

¹ Trường Đại học Mở - Địa chất

TÓM TẮT

Thư viện Pandas và Matplotlib của Python là hàm thư viện có thể mạnh rất lớn khi đây là thư viện mã nguồn mở, có độ mạnh và linh hoạt trong xử lý dữ liệu, nhất là với những bộ dữ liệu lớn và rất lớn. Thư viện Pandas được thiết kế để dễ dàng làm việc được với dữ liệu có cấu trúc (như dữ liệu dạng bảng, đa chiều, không đồng nhất) và dữ liệu chuỗi thời gian. Pandas và Matplotlib có thể dễ dàng hiển thị và trực quan dữ liệu sau khi đã xử lý với những hàm và phương thức được cung cấp nhằm giúp cho người dùng dễ dàng xử lý trên dữ liệu của mình. Với dữ liệu LiDAR, Pandas và Matplotlib cung cấp một cách mạnh mẽ các phương thức lọc, phân tích và hiển thị trực quan dữ liệu dạng điểm. Trong bài báo này, nhóm tác giả giới thiệu về các hàm và phương thức để thao tác trên bộ dữ liệu đám mây điểm LiDAR với khoảng 2 triệu bản ghi. Kết quả thực nghiệm đã thể hiện rõ ưu điểm của thư viện Pandas và Matplotlib trong lọc, phân tích và hiển thị dữ liệu LiDAR.

Từ khóa: Pandas, Matplotlib, trực quan hóa dữ liệu, phân tích dữ liệu LiDAR, lọc dữ liệu LiDAR, LiDAR.

1. Mở đầu

Dữ liệu đóng vai trò rất quan trọng với nhiều hệ thống thông tin trong thế giới số hiện nay. Những lợi ích mà dữ liệu mang lại là rất lớn, giúp chúng ta có cái nhìn chi tiết và hiểu rõ về vấn đề. Tuy nhiên, chúng ta cũng phải đối mặt với không ít những thách thức mà dữ liệu đặt ra. Dữ liệu được thu từ nhiều nguồn với nhiều công nghệ và phương pháp khác nhau. Dữ liệu khi thu thập sẽ có nhiều, bị khuyết thiếu và nhiều những vấn đề khác. Dữ liệu với hàng triệu, trăm triệu bản ghi và nhiều thuộc tính, yêu cầu cấp thiết cần đặt ra đó chính là xử lý, phân tích dữ liệu để có thể hiểu dữ liệu một cách dễ dàng. Việc hiểu dữ liệu là điều không dễ dàng, nhất là với dữ liệu lớn, có cấu trúc phức tạp vì khi dữ liệu được áp dụng cho các bài toán khác nhau sẽ có đặc trưng, kỹ thuật và phương pháp phân tích khác nhau.

LiDAR là công nghệ viễn thám chủ động mới, với khả năng thu thập được dữ liệu trên bề mặt Trái đất, LiDAR ngày càng được sử dụng nhiều và rộng rãi trong các lĩnh vực không chỉ riêng Trắc địa - Bản đồ. Dữ liệu thu được chính từ LiDAR là đám mây điểm, là tập hợp các điểm được tạo ra bởi cảm biến. Những đám mây điểm có thể rất dày đặc và chứa hàng tỷ điểm, cho phép tạo ra các mô hình 3D có độ chi tiết cao về một khu vực. Chúng được sử dụng trong các ứng dụng như giám sát môi trường, lập bản đồ, quy hoạch đô thị, giám sát cơ sở hạ tầng,...(Yellow Scan, 2024). Đám mây điểm LiDAR được tạo ra bằng cách sử dụng các xung laser để đo khoảng cách đến các đối tượng và tạo ra các điểm biểu diễn vật thể trên bề mặt Trái đất từ những xung phản hồi. Đám mây điểm cung cấp dữ liệu vị trí có giá trị cho các ứng dụng khác nhau. Các điểm riêng lẻ trong đám mây điểm chứa thông tin như tọa độ 3D (X, Y và Z. Ngoài ra, chúng có thể được bổ sung các thuộc tính như số trả về, góc quét, hướng quét, mật độ điểm, giá trị màu RGB và dấu thời gian (Yellow Scan, 2024). Từ những đám mây điểm với hàng triệu, hàng tỷ điểm và số lượng thuộc tính lớn, sẽ cung cấp thông tin có giá trị cho chúng ta biết thêm về mật độ của vật thể hoặc thành phần vật chất.

Tuy nhiên, đám mây điểm thô chưa thể sử dụng được ngay sau khi thu thập, chúng cần phải được xử lý trước khi đưa vào các ứng dụng cụ thể. Quá trình xử lý đám mây điểm bao gồm nhiều nước và nhiều công đoạn như loại bỏ nhiễu, lọc, phân đoạn, phân cụm và mô hình hóa. Có nhiều ngôn ngữ lập trình, phương pháp, công cụ và mô hình được các nhà khoa học sử dụng để xử lý, phân tích và trực quan dữ liệu đám mây điểm LiDAR. Trong (Wei Yao, Wenbing Tao và cộng sự, 2023) đã chỉ ra phân tích hoặc xử lý đám mây điểm LiDAR, bao gồm đăng ký hoặc ánh xạ đám mây điểm, lọc, phân đoạn và phân loại, lập mô hình 3D và trực quan hóa, là điều kiện tiên quyết cơ bản để ứng dụng có hiệu quả các đám mây điểm LiDAR cho các bài toán cụ thể. Các thuật toán đa dạng đã được cung cấp dưới dạng phương pháp tiếp cận dựa trên dữ liệu, dựa trên mô hình hoặc kết hợp để phân tích và khám phá các đám mây điểm LiDAR. Với các kỹ thuật mới nhất trong học máy và học sâu được nhóm tác giả sử dụng giúp trích xuất thông tin ngữ nghĩa từ các đám mây điểm LiDAR theo cách thông minh và hiệu quả hơn,

* Tác giả liên hệ

Email: nguyenthihuuphuong@humg.edu.vn

nhanh chóng và chính xác hơn, đồng thời mở rộng hơn nữa phạm vi ứng dụng của đám mây điểm LiDAR. Nghiên cứu (Hadi AliAkbarpour, Yuan Li, 2024), nhóm tác giả đã chỉ ra sự phức tạp về mặt địa lý của một số địa hình cụ thể và sự thay đổi nhanh chóng trong môi trường đô thị vẫn đòi hỏi các giải pháp hiệu quả và hiệu quả hơn cho mô hình hóa bề mặt kỹ thuật số và tái tạo cảnh 3D. Do đó, cần quá trình phân tích, xử lý dữ liệu để tìm hiểu dữ liệu rõ hơn.

Mặc dù quy trình chụp là điểm khởi đầu cho nhiều ứng dụng, nhưng việc xử lý các đám mây điểm là điều cần thiết để trực quan hóa, làm phong phú, phân tích, định lượng, đánh giá, mô hình hóa và hiểu đối tượng hoặc cảnh được đo. Một quy trình xử lý thường bao gồm nhiều giai đoạn, chẳng hạn như định hướng đám mây điểm, đồng đăng ký, kiểm soát chất lượng, trích xuất tính năng, phân đoạn và phân loại ngữ nghĩa, phát hiện và nhận dạng đối tượng, phát hiện thay đổi và mô hình hóa đối tượng. Nghiên cứu của nhóm tác giả (Johannes Otepka và cộng sự, 2022) đã chỉ ra rằng phương pháp, thuật toán và cấu trúc dữ liệu tiên tiến của học máy cùng với các quy trình xử lý toàn diện cho các ứng dụng hoặc cảm biến cụ thể. Chính vì những lý do đó đặt ra yêu cầu cần phải có hướng giải quyết nhanh chóng để xử lý đám mây điểm LiDAR và đưa đến tay người sử dụng hay người dùng có thể sử dụng trực tiếp với những dạng dữ liệu dễ xử lý và thao tác hơn như excel, csv,... thay vì file .las định dạng ban đầu. Trong bài báo này, nhóm tác giả sử dụng hai thư viện mạnh của Python là Pandas và Matplotlib để xử lý, phân tích và trực quan hóa dữ liệu đám mây điểm LiDAR. Sự lựa chọn đối với ngôn ngữ lập trình Python mang lại nhiều điểm mạnh to lớn như: Python là một ngôn ngữ đủ đơn giản để mọi người có thể nhanh chóng hiểu được và làm quen được với nó, Python đủ mạnh để xử lý các tập dữ liệu lớn và nó cho phép tận dụng các thư viện máy học mạnh mẽ được cung cấp bởi cộng đồng rộng lớn.

2. Nội dung nghiên cứu

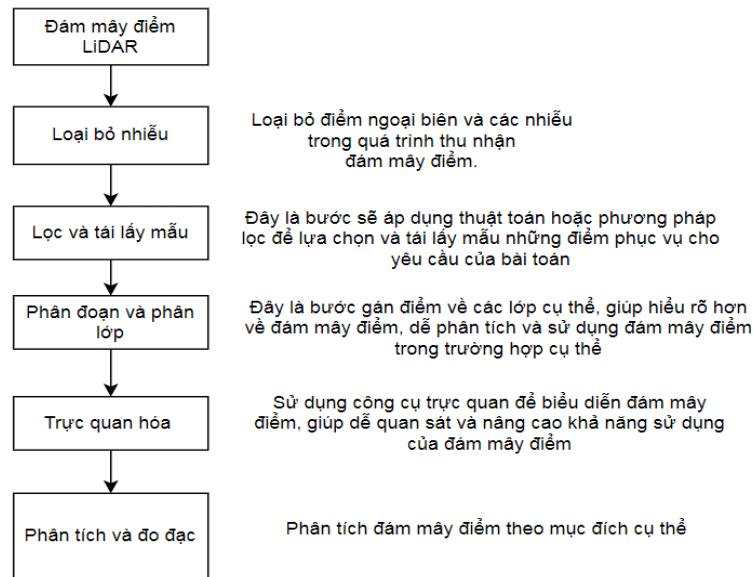
2.1. Phân tích, lọc và trực quan hóa dữ liệu LiDAR

Phân tích dữ liệu là quá trình kiểm tra, làm sạch, chuyển đổi và mô hình hóa dữ liệu với mục tiêu khám phá thông tin hữu ích, đưa ra kết luận và hỗ trợ ra quyết định (H. Taderhdost, 2022). Phân tích dữ liệu có nhiều khía cạnh, nhiều hướng tiếp cận và nhiều phương pháp thực hiện và được áp dụng cho nhiều ngành, lĩnh vực khoa học khác nhau. Phân tích dữ liệu là một quá trình thu thập dữ liệu thô và sau đó chuyển đổi nó thành thông tin hữu ích cho việc ra quyết định của người dùng (Stephan Kudyba, 2014). Theo (John W. Tukey, 2020) các quy trình phân tích dữ liệu, các cách lập kế hoạch thu thập dữ liệu để giúp việc phân tích dữ liệu dễ dàng hơn, chính xác hơn hoặc chính xác hơn và tất cả các máy móc và kết quả thống kê (toán học) áp dụng để phân tích dữ liệu. Quy trình phân tích dữ liệu bao gồm 6 bước đó là: định nghĩa bài toán, thu thập dữ liệu, làm sạch dữ liệu, phân tích dữ liệu, trực quan hóa dữ liệu và cuối cùng là biểu diễn dữ liệu. Trong đó, lọc dữ liệu là quá trình chọn một phần nhỏ hơn trong tập dữ liệu và sử dụng tập hợp con đó để xem hoặc phân tích. Đây là quá trình sắp xếp thông qua một tập dữ liệu lớn để xác định các tập hợp thông tin cụ thể dựa trên các tiêu chí đã xác định. Điều này rất quan trọng để tìm kiếm dữ liệu quan trọng, loại bỏ thông tin không cần thiết và cải thiện chất lượng tổng thể của dữ liệu. Lọc dữ liệu nếu được thực hiện một cách chính xác có thể giúp cho quá trình phân tích và xử lý dữ liệu được nhanh chóng, chính xác và đơn giản hơn. Lọc dữ liệu chọn và hiển thị một tập hợp con dữ liệu dựa trên các tiêu chí cụ thể. Phương pháp lọc dữ liệu có thể khác nhau tùy theo ngữ cảnh, ta có thể lựa chọn phương pháp và công cụ khác nhau (F. Khan, 2024).

Dữ liệu sau khi phân tích và lọc cần được trực quan hóa để dễ hiểu, dễ quan sát và dễ thao tác hơn. Trực quan hóa dữ liệu là sự thể hiện dữ liệu thông qua việc sử dụng đồ họa phổ biến, chẳng hạn như biểu đồ, sơ đồ, đồ họa thông tin và thậm chí cả hình ảnh động. Những màn hình hiển thị thông tin trực quan này truyền đạt các mối quan hệ dữ liệu phức tạp và thông tin chuyên sâu dựa trên dữ liệu theo cách dễ hiểu (IBM, 2024). Trực quan hóa dữ liệu được sử dụng trong nhiều trường hợp, nhiều ngành và nhiều lĩnh vực với nhiều mục đích khác nhau.

Với dữ liệu đám mây điểm LiDAR, do số lượng điểm thu thập rất lớn, trong đó có một số dữ liệu được thu thập thêm ngoài giá trị tọa độ, độ cao như giá trị màu của điểm, thứ tự tia phản xạ, cường độ phản xạ,... nhưng những dữ liệu đó sẽ không cung cấp cho chúng ta bất kỳ thông tin nào nếu chúng ta không phân tích và trích xuất thông tin vào những ứng dụng cụ thể. Việc xử lý và phân tích dữ liệu LiDAR thường được thực hiện với các bước như trong Hình 1:

Quy trình xử lý và phân tích đám mây điểm LiDAR rất phức tạp bao gồm nhiều bước, nhiều công đoạn cần thực thi. Chính vì thế, khi công nghệ đám mây điểm phát triển, các kỹ thuật xử lý yêu cầu ngày càng phức tạp hơn, do đó, việc kết hợp học sâu, học nhất đa cảm biến và xử lý dựa trên đám mây là hoàn toàn cần thiết. Những tiến bộ này giúp chúng ta trong sự hiểu biết phức tạp hơn nữa về thế giới xung quanh chúng ta, tạo điều kiện cho các ứng dụng thú vị trong xe tự hành, xây dựng thành phố thông minh, bảo tồn di sản văn hóa, nhận dạng và nhiều hơn thế nữa.



Hình 5. Quy trình xử lý và phân tích dữ liệu LiDAR

2.2. Thư viện Pandas và Matplotlib của Python

Pandas là thư viện thường được sử dụng để xử lý phân tích dữ liệu. Nó đơn giản hóa việc tải dữ liệu từ các nguồn bên ngoài như tệp văn bản và cơ sở dữ liệu, cũng như cung cấp các cách phân tích và thao tác dữ liệu sau khi được tải vào máy tính một cách đơn giản, thuận tiện và dễ dàng. Các tính năng được cung cấp trong Pandas tự động hóa và đơn giản hóa rất nhiều tác vụ phổ biến cần nhiều dòng mã để viết bằng ngôn ngữ Python cơ bản (Declan V., 2024). Pandas là một thư viện Python cực kỳ phổ biến và vẫn đang phát triển, được sử dụng trong nhiều lĩnh vực từ khoa học môi trường và khí hậu, cho đến khoa học xã hội, ngôn ngữ học, sinh học cũng như một số ứng dụng trong ngành như phân tích dữ liệu, giao dịch tài chính, và nhiều người khác (M-E.OGBEMI, 2024). Pandas có hai kiểu cấu trúc dữ liệu cơ bản là Dataframe (Dữ liệu khối) và Series (Dữ liệu chuỗi). Trong đó, Dataframe được tổ chức dạng bảng, với hàng và cột, mỗi dòng sẽ lưu trữ được kiểu dữ liệu đơn, như dữ liệu dạng số, chuỗi, logic,... Kiểu dữ liệu dạng chuỗi (Series) được tổ chức dạng mảng một chiều, mỗi phần tử trong mảng sẽ được gán một chỉ mục (index), ta sẽ sử dụng chỉ mục để truy xuất vào đối tượng.

Để sử dụng được thư viện Pandas ta cần sử dụng câu lệnh cài đặt như sau: `pip install pandas`. Sau khi cài đặt, ta có thể làm việc với các kiểu dữ liệu như Series và Dataframe được hiển thị trong Pandas:

```
>>> import pandas as pd
>>> my_series = pd.Series([4.6, 2.1, -4.0, 3.0])
>>> print(my_series)
0    4.6
1    2.1
2   -4.0
3    3.0
dtype: float64
```

Hình 6. Kiểu dữ liệu series của pandas

```
>>> scottish_hills = {'Ben Nevis': (1345, 56.79685, -5.003508),
...                  'Ben Macdui': (1309, 57.070453, -3.668262),
...                  'Braeriach': (1296, 57.078628, -3.728024),
...                  'Cairn Toul': (1291, 57.054611, -3.71042),
...                  'Sgòr an Lochain Uaine': (1258, 57.057999, -3.725416)}
>>>
>>> dataframe = pd.DataFrame(scottish_hills)
>>> print(dataframe)
   Ben Nevis  Ben Macdui  Braeriach  Cairn Toul  Sgòr an Lochain Uaine
0  1345.000000  1309.000000  1296.000000  1291.000000  1258.000000
1    56.796850    57.070453    57.078628    57.054611    57.057999
2   -5.003508   -3.668262   -3.728024   -3.710420   -3.725416
```

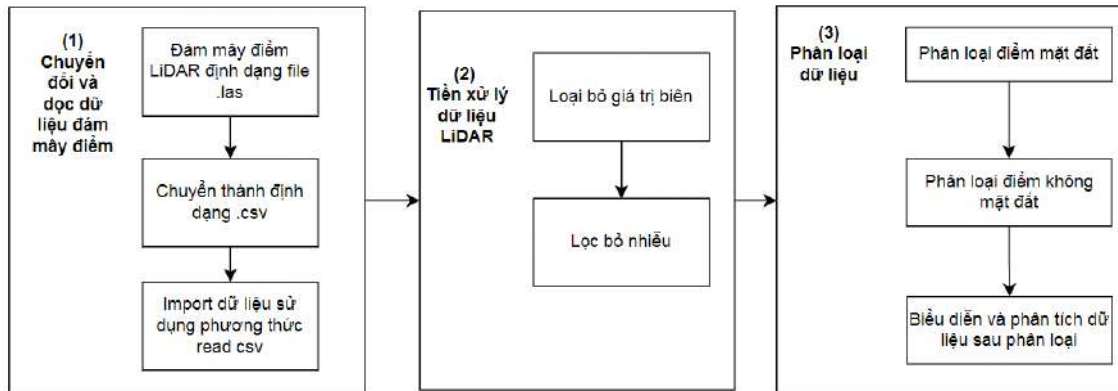
Hình 7. Kiểu dữ liệu Dataframe trong pandas

Matplotlib là một thư viện Python phổ biến để hiển thị dữ liệu và tạo các biểu đồ tĩnh, hoạt hình và tương tác. Chương trình này cho phép bạn vẽ đồ họa hấp dẫn và giàu thông tin như biểu đồ đường, biểu đồ phân tán, biểu đồ và biểu đồ thanh. Matplotlib có khả năng tùy biến cao và linh hoạt, khiến nó trở thành lựa chọn ưu tiên của các nhà phân tích dữ liệu và nhà khoa học làm việc trong các lĩnh vực như tài chính, khoa học, kỹ thuật và khoa học xã hội (J. Field, 2009). Các loại biểu đồ Matplotlib hỗ trợ cho người dùng gồm: Bar, Pie, Box plot,

Histogram, Line, Subplots, Scatter, Contour. Matplotlib là một lựa chọn phù hợp cho các yêu cầu cần trực quan hóa dữ liệu khác nhau, bao gồm phân tích dữ liệu, vẽ sơ đồ và tạo các biểu đồ chất lượng. Nó vượt trội trong các bài toán mà người dùng yêu cầu quyền kiểm soát chi tiết đối với việc tùy chỉnh kịch bản trực quan hóa và cần tạo các hình ảnh trực quan phức tạp hoặc chuyên biệt.

2.3. Phân tích, lọc và trực quan dữ liệu LiDAR của Pandas và Matplotlib

Quy trình phân tích, lọc và trực quan dữ liệu đám mây điểm LiDAR với Pandas và Matplotlib được nhóm tác giả thực hiện trong bài báo như trong Hình 4 dưới đây:



Hình 8. Quy trình phân tích, lọc và trực quan hóa dữ liệu được sử dụng trong bài báo

Quy trình phân tích, lọc và trực quan hóa dữ liệu đám mây điểm LiDAR được thực hiện với bốn bước gồm:

a) Bước 1: Chuyển đổi và đọc dữ liệu đám mây điểm LiDAR:

Tại bước này, nhóm tác giả chuyển đổi định dạng đám mây điểm từ .las thành định dạng .csv, đây là định dạng dữ liệu được phân cách với nhau bằng dấu phẩy, giống với kiểu dữ liệu dataframe của Pandas.

Để chuyển đổi định dạng .las sang .csv, nhóm tác giả sử dụng phương thức `to_csv` của thư viện Lasio. Lasio là thư viện được sử dụng để đọc và ghi dữ liệu, siêu dữ liệu của các tập tin LAS.

Đọc file .las và chuyển đổi định dạng sang csv sử dụng câu lệnh:

```
>>> import lasio.examples
>>> from sys import stdout
>>> las = lasio.examples.open('sample.las')
>>> las.to_csv(stdout)
```

Hình 9. Câu lệnh đọc và chuyển định dạng với thư viện Lasio

```
>>> rd_las = lp.read(r'E:\NCS\ThuNghiemMCC\Output.las')
>>> print(rd_las)
<LasData(1.1, point fmt: <PointFormat(0, 0 bytes of extra dims)>, 2253292 points, 3 vlrs)>
```

Hình 10. Hiển thị metadata của file .las

Sau khi chuyển đổi thành công định dạng, sử dụng phương thức `read_csv` của Pandas để đọc dữ liệu trước khi phân tích và xử lý dữ liệu đám mây điểm.

```
>>> pd.read_csv(r"C:\Users\nguyen\Downloads\Fm.csv", usecols=[1,2,3,4])
stdin:1: DtypeWarning: Columns (1,2,3,4) have mixed types. Specify dtype option on import or set low_memory=False.
   Unnamed: 1  Unnamed: 2  Unnamed: 3  Unnamed: 4
0          0           1           2           3
1    225450    22513    224590    224849
2      NaN      NaN      NaN      NaN
3          0           1           2           3
4    230593           5    223031           1
...
943939  4366533.54    2271.18         2.0         4.0
943940  4366533.68    2271.2         2.0         4.0
943941  4366533.82    2271.15         2.0         4.0
943942  4366533.96    2271.15         2.0         4.0
943943  4366534.1     2271.16         NaN         NaN
[343944 rows x 4 columns]
```

Hình 11. Đọc file csv

b) Tiền xử lý dữ liệu:

Trong bước tiền xử lý dữ liệu, cần thực hiện loại bỏ biên và những điểm nhiễu có trong đám mây điểm. Với hàng triệu bản ghi được thu thập, nhiễu xuất hiện trong đám mây điểm là không tránh khỏi. Nhóm tác giả đã thực hiện lọc bỏ điểm biên và giá trị biên theo các bước như sau:

- Bước 1: Thu thập mọi điểm thứ n.

- Bước 2: Gán mã nhị phân cho những điểm được chọn. Hiển thị các điểm được chọn và điểm không được chọn theo mã nhị phân này.
- Bước 3: Xác định số lượng điểm ít điểm lân cận hơn so với khoảng cách trung bình của đám mây điểm sử dụng hai tham số đầu vào: số điểm lân cận được dùng để tính khoảng cách trung bình và ngưỡng được xác định dựa trên độ lệch chuẩn của khoảng cách trung bình trên đám mây điểm. Trong bước thử nghiệm này, nhóm tác giả lựa chọn số điểm lân cận là 20, tỉ lệ lệch chuẩn là 2.0, việc lựa chọn này dựa trên mật độ điểm và giá trị độ lệch chuẩn của đám mây điểm để làm ngưỡng cho bước loại bỏ điểm biên và nhiễu.
- Loại bỏ những điểm xác định ở Bước 3

```
print("Statistical outlier removal")
cl, ind = voxel_down_pcd.remove_statistical_outlier(nb_neighbors=20,
                                                    std_ratio=2.0)
display_inlier_outlier(voxel_down_pcd, ind)
```

Statistical outlier removal
Showing outliers (red) and inliers (gray):

Hình 12. Loại bỏ điểm biên sử dụng thống kê số lượng điểm có ít điểm lân cận

c) Phân loại dữ liệu

Đây là bước gán nhãn cho các điểm, độ chính xác của quá trình phân loại dữ liệu sẽ quyết định đến độ chính xác của bài toán ứng dụng. Phân loại đám mây điểm được chia thành hai bước với phân loại điểm mặt đất và điểm không mặt đất được thực hiện như sau:

- Tìm các điểm là tia phản xạ đơn có độ cao z lớn hơn ngưỡng trung bình độ cao của các điểm → Gán điểm không mặt đất,
- Lặp lại quá trình tìm kiếm và tìm kiếm được các điểm lân cận với những điểm tìm kiếm với ngưỡng là độ cao trung bình,
- Phân loại lại điểm nếu độ cao của điểm cao hơn so với trung bình độ cao,
- Gán nhãn điểm theo độ cao với các điểm mặt đất, thực vật.

```
in_file = laspy.file.File(r"points.las", mode = "r")

point_records = in_file.points

#find all points that are single-return vegetation points (class 3, 4, or 5)

#note: num_returns is not included in the point_records above. So I am
#creating new arrays here.
class_arr = in_file.raw_classification
return_arr = in_file.num_returns

single_veg_pts = np.where(np.logical_and(np.logical_or(class_arr == 3,
                                                         class_arr == 4,
                                                         class_arr == 5), return_arr == 2))

#get the filtered points' records
single_veg_pt_records = in_file.points[single_veg_pts]
```

Hình 13. Phân loại đám mây điểm LiDAR

Sau khi gán nhãn và phân loại các điểm vào các lớp chuyên biệt, thực hiện phân tích dữ liệu với Pandas và trực quan hóa dữ liệu với thư viện Matplotlib. Phân tích dữ liệu sử dụng các phương thức sau:

```
>>> dt.describe()
```

	4366595.37	2253.23	2	3
count	3.439360e+05	343936.000000	343935.0	343935.000000
mean	4.366621e+06	2271.774274	2.0	2.532647
std	5.750620e+01	22.564076	0.0	1.435856
min	4.366530e+06	2211.880000	2.0	0.000000
25%	4.366572e+06	2253.410000	2.0	2.000000
50%	4.366618e+06	2270.950000	2.0	3.000000
75%	4.366667e+06	2288.110000	2.0	4.000000
max	4.366736e+06	2319.840000	2.0	4.000000

Hình 14. Mô tả dữ liệu với phương thức describe

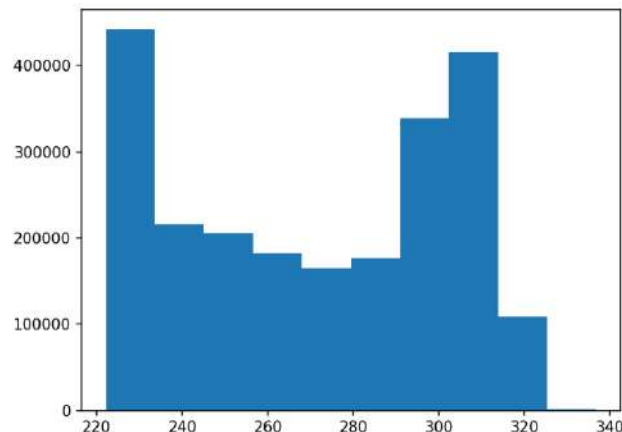
```
>>> dt.corr()
      4366595.37    2253.23    2         3
4366595.37    1.000000   -0.090842  NaN   -0.174847
2253.23      -0.090842   1.000000  NaN   -0.221658
2          NaN         NaN   NaN    NaN
3          -0.174847   -0.221658  NaN   1.000000

>>> dt.std()
      4366595.37    57.506196
2253.23         22.564076
2              0.000000
3              1.435856
dtype: float64
```

Hình 15. Độ tương quan và độ lệch chuẩn dữ liệu với phương thức corr

[illegible]

Hình 16. Phân tích thuộc tính với phương thức vlrs

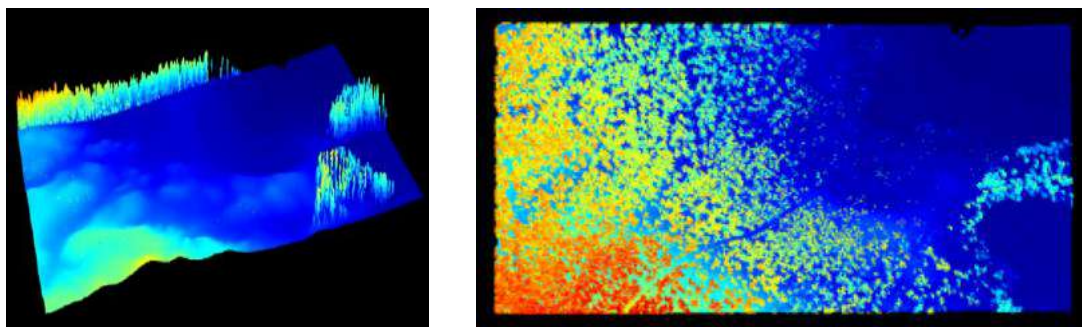


Hình 17. Biểu đồ histogram thống kê phân bố độ cao trong đám mây điểm

3. Thử nghiệm

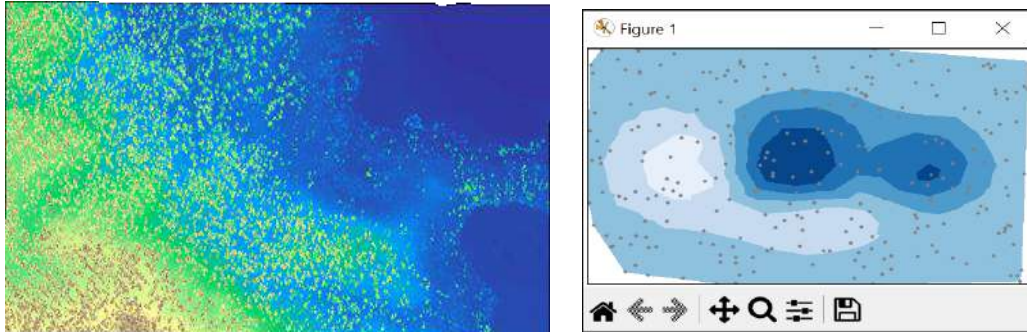
Dữ liệu thử nghiệm được nhóm tác giả sử dụng trong bài báo là dữ liệu đám mây điểm được thu nhận tại khu vực Bắc Ninh với 2.253.292 điểm. Các thuộc tính của điểm của thể hiện trong Bảng 1. Đây là khu vực có mật độ dân cư đông, có diện tích khoảng 10 km², mật độ dân cư cao, có đường giao thông và nhiều cây. Giá trị tọa độ $X_{\min} = 501674.204076$; $Y_{\min} = 3757299.628317$; $X_{\max} = 503228.707749$; $Y_{\max} = 3758184.08866$. Mật độ điểm là 2.98 pt/m².

Đám mây điểm sau khi loại bỏ nhiễu và điểm biên được hiển thị như sau:

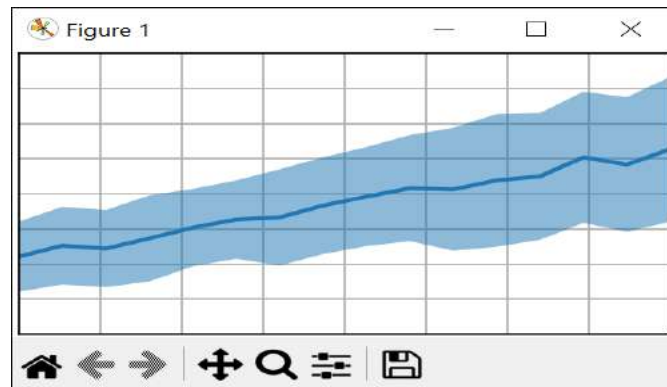


Hình 18. Đám mây điểm được hiển thị theo độ cao sau khi loại bỏ nhiễu

Kết quả phân loại điểm mặt đất và điểm không mặt đất được nhóm tác giả thành lập mô hình TIN độ phân giải 1 m và biểu đồ đường bình độ được thể hiện trong Hình 18, 19.



Hình 19. Mô hình TIN và đường bình độ của khu vực



Hình 20. Đồ thị biểu diễn giá trị trung vị của độ cao điểm

Qua kết quả thực nghiệm với các thuật toán lọc, phân loại điểm được nhóm tác giả sử dụng, đám mây điểm sau phân loại được trực quan hóa và thể hiện được độ tương quan cũng như cấu trúc của dữ liệu.

4. Kết luận

Qua kết quả thử nghiệm sử dụng thư viện Pandas và Matplotlib của Python trong phân tích, lọc và trực quan hóa dữ liệu đám mây điểm LiDAR có thể thấy đây là hai thư viện dễ sử dụng, dễ cài đặt, có độ mạnh trong phân tích và biểu diễn dữ liệu LiDAR với các hàm được phát triển bởi một cộng đồng rộng lớn, các hàm phân tích dữ liệu ngắn gọn và thực hiện trực quan với nhiều loại biểu đồ thay vì phải cài đặt những phần mềm chuyên biệt có thể mất phí và khó cài đặt. Đồng thời, hai thư viện cung cấp các công cụ mạnh mẽ để giúp người sử dụng có thể dễ dàng xử lý dữ liệu và trực quan dữ liệu. Cùng với sự lớn mạnh của cộng đồng sử dụng ngôn ngữ Python, đây sẽ là điểm mạnh của Pandas và Matplotlib so với ngôn ngữ và công cụ khác.

Trong bài báo, nhóm tác giả đã sử dụng Pandas và Matplotlib thực hiện loại bỏ nhiễu và điểm biên, phân loại và trực quan hóa được dữ liệu đám mây điểm LiDAR. Với các loại biểu đồ trực quan và quy trình xử lý, phân tích dữ liệu đã đề xuất, xử lý dữ liệu thô LiDAR không còn là một quy trình phức tạp, khó thực hiện.

Tài liệu tham khảo

- Yellow Scan, "LiDAR Point Cloud Basics," Yellow Scan, 24 5 2024. Available: <https://www.yellowscan.com/knowledge/lidar-point-cloud-basics/>, truy cập 8/7/2024.
- Wei Yao, Wenbing Tao, et al., 2023. "Machine Learning for LiDAR Point Cloud Analysis", Remote Sensing.
- Hadi AliAkbarpour, Yuan Li, 2024. "Processing for Digital Surface Modelling and 3D Scene Reconstruction", Remote Sensing.
- Johannes Otepka, Martin Weinmann, Di Wang, Kourosh Khoshelham, 2022. "Point Cloud Processing in Remote Sensing Technology", Remote Sensing.
- H. Taderhdost, 2022. "Different type of Data analysis; data analysis methods and technique in research project", International Journal of Academic Research in Management, pp. 1-9.
- Stephan Kudyba, "Transforming unstructured data into useful information," Big Data, Mining, and Analytics, pp. 227 - 246, 2014.
- John W. Tukey, 2020. "Jonh Turkey The future of data analysis," The Annals of Mathematical Statistics, 33(1).
- F. Khan, "Data Filtering: A Comprehensive Guide to Techniques, Benefits, and Best Practices," Astera, 10 5

2024. Available: <https://www.astera.com/type/blog/data-filtering/>, truy cập 10/7/2024.

IBM, “What is data visualization?,” IBM, 2024. Available: <https://www.ibm.com/topics/data-visualization#:~:text=Data%20visualization%20is%20the%20representation,that%20is%20easy%20to%20understand>, truy cập 10/7/2024.

Declan. V, “PYTHON DATA ANALYSIS WITH PANDAS AND MATPLOTLIB,” Coding club. Available: <https://ourcodingclub.github.io/tutorials/pandas-python-intro/>, truy cập 11/7/2024.

M.-E. OGBEMI, “Data Visualization with Matplotlib - a Step by Step Guide,” freeCodeCamp, 24/4/2023. [Trực tuyến]. Available: <https://www.freecodecamp.org/news/introduction-to-data-vizualization-using-matplotlib/>, truy cập 11/7/2024.

J. Field, 2009. “Dividing listening into its component,” Listening in the Language Classroom, pp. 96-109.

ABSTRACT

Analyze, filter, and visualize LiDAR point cloud data with the python libraries pandas and matplotlib

Nguyen Thi Huu Phuong^{1*}, Pham Thi Hai Van¹, Dao Thi Hong Tham¹

¹*Hanoi University of Mining and Geology*

Python's Pandas and Matplotlib libraries are the strongest functional libraries as they are open source library code, with speed and performance in data processing, best with large and very large data sets. Pandas library is designed to easily work with structured data (such as tabular, multidimensional, heterogeneous) and time series data. Pandas and Matplotlib can easily display and visualize data after processing the functions and methods provided to help users easily process their data. With LiDAR data, Pandas and Matplotlib provide powerful methods for filtering, analyzing, and visualizing data patterns. In this article, the authors introduce functions and methods to manipulate the LiDAR point cloud data set with about 2 million records. The resulting implementation was able to determine the priority of the Pandas and Matplotlib libraries during the filtering, analysis and visualization of LiDAR data.

Keyword: Pandas, Matplotlib, data visualization, LiDAR data analysis, LiDAR data filtering, LiDAR.

Nghiên cứu xác định sự biến động độ sâu ven đảo xa bờ sử dụng dữ liệu viễn thám Sentinel-2 trên nền tảng Google Earth Engine

Nguyễn Như Hùng^{1*}, Lê Văn Phú¹, Nguyễn Văn Sơn¹

¹ Học viện Kỹ thuật quân sự

TÓM TẮT

Sự biến động độ sâu ven đảo xa bờ có thể ảnh hưởng đến hệ sinh thái biển, quy hoạch và phát triển kinh tế ven biển, cũng như các hoạt động hàng hải. Việc theo dõi và đánh giá sự biến động này giúp bảo vệ môi trường, quản lý tài nguyên biển và định hướng phát triển bền vững. Nghiên cứu xác định sự biến động độ sâu ven đảo xa bờ sử dụng dữ liệu viễn thám trên nền tảng Google Earth Engine (GEE) là một hướng nghiên cứu rất hiệu quả nhờ khả năng xử lý dữ liệu lớn và cung cấp các công cụ phân tích mạnh mẽ. Trong nghiên cứu này xác định sự thay đổi độ sâu ven đảo xa bờ bằng dữ liệu ảnh viễn thám được thu thập tại hai thời điểm lần lượt là 27/3/2020 và 20/9/2020. Kết quả thu được giá trị tương quan giữa giá trị đo độ sâu thực tế ngoài địa và giá độ sâu được xác định từ ảnh vệ tinh là 0,87 và 0,93. Giá trị RMSE thu được tương ứng lần lượt là 0,06 m và 0,04 m. Bên cạnh đó, giá trị biến thiên độ sâu lớn nhất là 8,26 m. Từ kết quả tính toán độ sâu ảnh vệ tinh và cao độ thủy triều tại thời điểm chụp ảnh tiến hành hiệu chỉnh đồng bộ độ sâu và so sánh, đánh giá sự thay đổi. Kết quả của phương pháp bước đầu cho thấy tính tự động, nhanh và hiệu quả trong phục vụ đánh giá sự thay đổi độ sâu trên vùng biển Việt Nam, đặc biệt là ở những khu vực ven bờ, ven quần đảo. Đồng thời, phương pháp còn hỗ trợ công tác theo dõi, cập nhật thường xuyên.

Từ khóa: Độ sâu, Sentinel-2, Trường Sa, Google Earth Engine.

1. Đặt vấn đề

Đánh giá biến động độ sâu biển (biến động địa hình đáy biển) là một phần quan trọng trong nghiên cứu hải dương học, quản lý tài nguyên biển và bảo vệ môi trường. Bên cạnh đó, đáy biển chứa nhiều tài nguyên quan trọng như dầu khí, khoáng sản, và các nguồn sinh vật biển. Việc theo dõi biến động địa hình giúp quản lý và khai thác các tài nguyên này hiệu quả và bền vững.

Với nền khoa học ngày càng phát triển, các phương pháp ước tính độ sâu trên biển đã có nhiều sự thay đổi lớn. Các phương pháp đo sâu truyền thống như đơn tia, đa tia trên các phương tiện đường thủy hoặc LiDAR có phạm vi không gian hạn chế (Zhao và nnk., 2013). Các phương pháp này có độ chính xác cao tuy nhiên cần phải đo đạc trực tiếp, tốn nhiều thời gian, chi phí và công sức khi được triển khai. Do đó, việc sử dụng nguồn dữ liệu miễn phí, đáp ứng nhu cầu về không gian và thời gian là điều cần thiết (Traganos và nnk., 2018).

Do nhược điểm của các phương pháp đo truyền thống, phương pháp xác định độ sâu sử dụng dữ liệu vệ tinh đã được phát triển và cho thấy tiềm năng của nó. Các nghiên cứu trước đây đã áp dụng các cảm biến vệ tinh có độ phân giải thấp nhằm xây dựng các mô hình đối tượng nước. Trong nghiên cứu (Balsamo và nnk., 2010; Tedesco và nnk., 2011), các tác giả đã sử dụng dữ liệu vệ tinh MODIS (Moderate Resolution Imaging Spectroradiometer) nhằm xác định độ sâu khu vực hồ trên toàn cầu. Kết quả nghiên cứu được áp dụng nhằm xây dựng mô hình hồ trong dự báo thời tiết, xác định tốc độ bay hơi và dự báo nhiệt độ bề mặt chính xác. Bên cạnh đó, nghiên cứu (Barale và nnk., 2008) đã sử dụng nguồn dữ liệu SeaWiFS (1998-2003) với mục đích xác định vị trí, theo dõi các loại thực vật dưới nước và dị thường trong lưu vực Địa Trung Hải.

Các cảm biến có độ phân giải không gian thấp có thể gây nhiễu bởi các điểm ảnh phân tách nước và đất liền dọc theo bờ biển. Do đó, một số hình ảnh vệ tinh quan sát mặt đất với độ phân giải không gian và thời gian nhỏ đã được ứng dụng trong xây dựng mô hình độ sâu. Nhằm xác định chất lượng nước, thành phần chất nền và lập bản đồ độ sâu một cách chính xác, các nghiên cứu đã sử dụng nguồn dữ liệu đa phổ như CASI-2, Ocean PHILLS (Dekker và nnk., 2011; Brando và nnk., 2009; Kerr và nnk., 2018). Ngoài ra, nghiên cứu (Traganos và nnk., 2018) sử dụng nguồn hình ảnh viễn thám Sentinel-2, dữ liệu được thu thập tại khu vực biển Aegean và Đông Địa Trung Hải với độ phân giải cao phục vụ lập bản đồ độ sâu, phân tích kinh tế và sinh thái ở các rìa ven biển. Bên cạnh đó, trong nghiên cứu (Li và nnk., 2018), hình ảnh vệ tinh Landsat-8 được sử dụng để xác định độ phân xạ của đáy biển, đồng thời nghiên cứu hệ sinh thái dưới nước.

Trong thời gian gần đây, nhiều nền tảng điện toán đám mây được phát triển hỗ trợ người dùng xử lý nguồn dữ liệu vệ tinh và các dữ liệu bản đồ khác. Google Earth Engine (GEE) là một nền tảng điện toán đám mây mạnh mẽ cung cấp khả năng truy cập dễ dàng vào bộ dữ liệu vệ tinh với thời gian tính toán nhanh hỗ trợ ước tính độ

* Tác giả liên hệ:

Email: nguyennhuhung@lqdtu.edu.vn

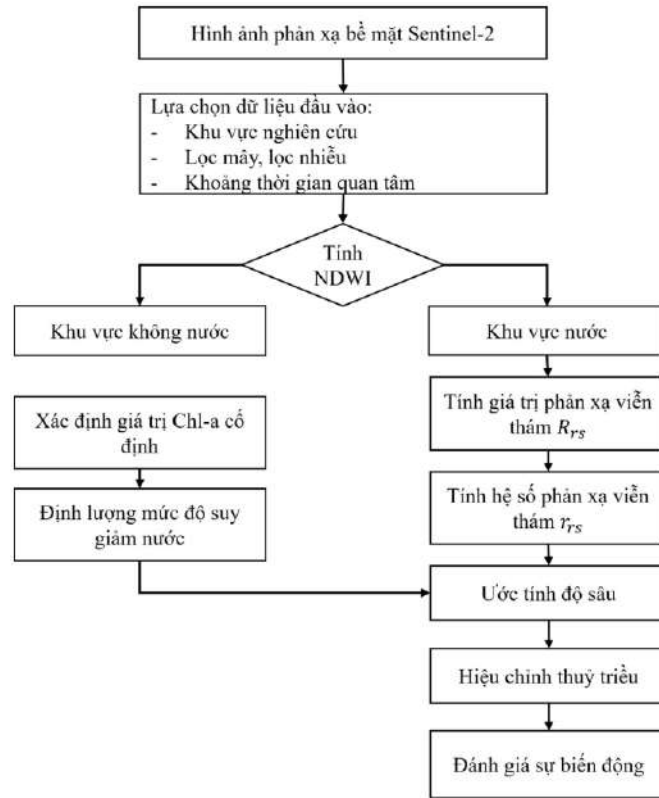
sâu. GEE đã được sử dụng trong nhiều sản phẩm với quy mô toàn cầu bao gồm độ che phủ đất, thay đổi rừng, diện tích mặt nước và sử dụng đất đô thị (Traganos và nnk., 2018; Li và nnk., 2021).

Tuy nhiên, các nghiên cứu hiện nay chủ yếu nghiên cứu về phương pháp xác định độ sâu, chưa có nhiều nghiên cứu về ứng dụng dữ liệu viễn thám để đánh giá sự thay đổi độ sâu địa hình theo thời gian. Vì vậy trong nghiên cứu này tác giả nghiên cứu phương pháp xác định sự thay đổi cao độ địa hình ven đảo nổi tại các đảo xa bờ, dữ liệu ảnh vệ tinh thực nghiệm được sử dụng ở hai mùa gió Tây Nam và Đông Nam.

2. Phương pháp nghiên cứu

Quy trình xử lý dữ liệu viễn thám trong nghiên cứu được trình bày cụ thể trên Hình 1. Trước hết, nghiên cứu tiến hành lọc bộ dữ liệu hình ảnh phản xạ bề mặt từ dữ liệu Sentinel-2 dựa trên các tham số như khu vực nghiên cứu, độ che phủ của mây, khoảng thời gian quan tâm. Nghiên cứu lựa chọn hình ảnh có độ che phủ của mây trên khu vực nghiên cứu là tối thiểu. Điều này làm giảm nhiễu do mây và bóng mây gây ra trong quá trình thu nhận thông tin bề mặt của cảm biến.

Tiếp theo, nghiên cứu tiến hành lọc các đối tượng không phải là nước trên hình ảnh phản xạ, sử dụng các giá trị ngưỡng của kênh ảnh để lọc các khu vực nước có độ đục cao, sóng và phản chiếu của ánh sáng mặt trời. Giá trị được sử dụng để lọc các đối tượng phù hợp với từng khu vực nghiên cứu. Đối với khu vực nghiên cứu, các giá trị được sử dụng là $\rho_{GREEN} > 0,01$, $\rho_{RED} < 0,1$ và $\rho_{NIR} < 0,03$ (Hùng & Phú, 2023).



Hình 1. Quy trình đánh giá sự biến động độ sâu từ dữ liệu vệ tinh

Bên cạnh đó, nghiên cứu tiếp tục lọc các đối tượng không phải là nước bằng cách sử dụng chỉ số nước chênh lệch chuẩn hoá (NDWI) được trình bày trong công thức (1):

$$NDWI = \frac{\rho_{GREEN} - \rho_{NIR}}{\rho_{GREEN} + \rho_{NIR}} \quad (1)$$

Dữ liệu hình ảnh sau khi lọc bỏ các đối tượng không nước được trích xuất phục vụ tính toán độ sâu tự động dựa trên thuật toán đo độ sâu (Li và nnk., 2019). Quy trình xử lý dữ liệu được thực hiện trên nền tảng GEE. Đầu tiên, nghiên cứu tính toán giá trị phản xạ viễn thám R_{rs} từ dữ liệu phản xạ bề mặt ρ_{λ} (Vanhellemont, 2019). Công thức tính toán giá trị phản xạ viễn thám R_{rs} được trình bày như công thức (2):

$$R_{rs}(\lambda) = \frac{\rho_{\lambda}}{\pi} \quad (2)$$

Tiếp theo, nghiên cứu loại bỏ sự ảnh hưởng giữa hai bề mặt không khí và nước thông qua hệ số phản xạ viển thám dưới bề mặt được trình bày ở công thức (3) (Lee và nnk., 2013):

$$r_{rs}(\lambda) = \frac{R_{rs}(\lambda)}{0.52 + 1.7 * R_{rs}(\lambda)} \quad (3)$$

Sau khi thu được giá trị hệ số phản xạ viển thám dưới bề mặt, nghiên cứu tính toán giá trị nồng độ Chlorophyll-a (*Chl-a*) với mục đích khởi tạo tham số ước tính độ sâu (Li và nnk., 2019). Nghiên cứu sử dụng giá trị *Chl-a* cố định (*Chl-a* = 0,5 mg.m⁻³) (Li và nnk., 2021). Các tham số ước tính độ sâu (m_0 và m_1) lần lượt được tính theo công thức (4) và (5):

$$m_0 = 52.073 * e^{(0.957 * Chl-a)} \quad (4)$$

$$m_1 = 50.156 * e^{(0.957 * Chl-a)} \quad (5)$$

Cuối cùng, nghiên cứu ước tính độ sâu nước bằng cách định lượng các mức độ suy giảm khác nhau giữa kênh ảnh màu lam và màu lục (Stumpf và nnk., 2003):

$$Depth = m_0 \frac{\ln(1000 * r_{rs,blue})}{\ln(1000 * r_{rs,green})} - m_1 \quad (6)$$

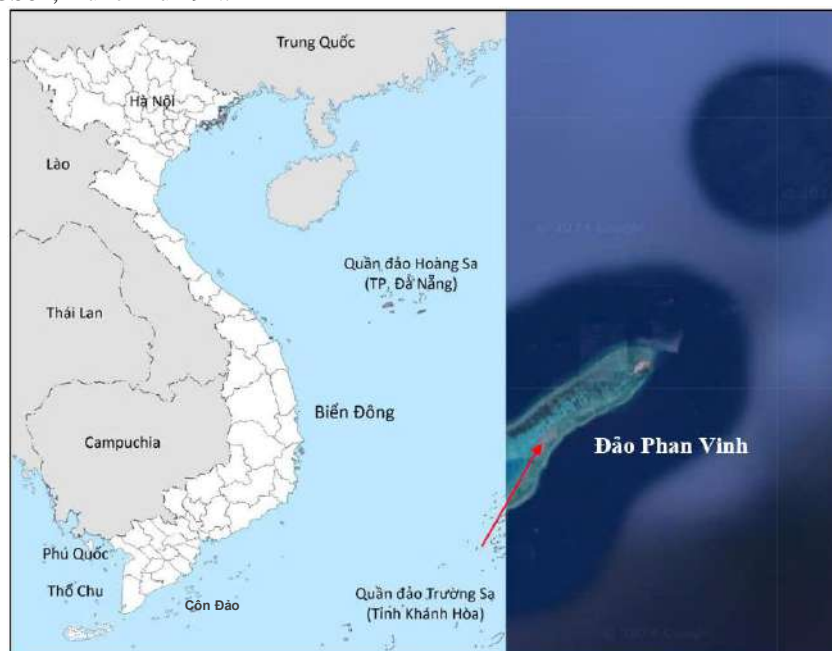
Sau khi thu được độ sâu được tính toán cho các khoảng thời gian quan tâm, các tác giả tiến hành hiệu chỉnh thủy triều. Điều này đảm bảo độ sâu ở các thời điểm khác nhau có cùng mặt thủy chuẩn. Bên cạnh đó, nghiên cứu sử dụng các điểm đo thực tế nhằm hiệu chỉnh độ sâu tại hai thời điểm khác nhau. Cuối cùng, thuật toán trừ ảnh được sử dụng để đánh giá sự biến động độ sâu. Bằng cách lấy hai hình ảnh độ sâu và trừ chúng với từng vị trí điểm ảnh tương ứng, hình ảnh hiển thị sự khác biệt độ sâu giữa hai thời điểm thu nhận ảnh.

3. Khu vực và dữ liệu thực nghiệm

3.1. Khu vực và dữ liệu ảnh vệ tinh khu vực nghiên cứu

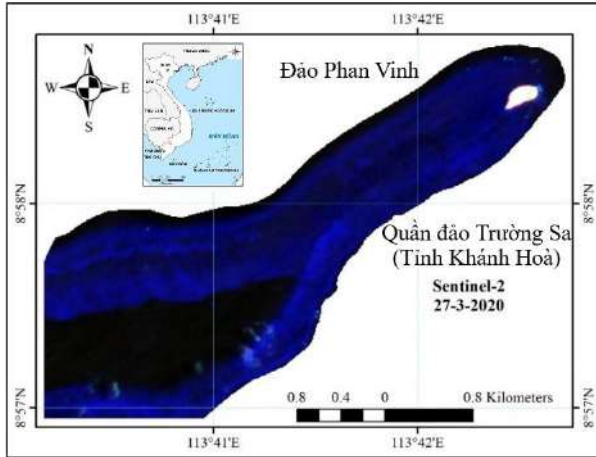
Khu vực nghiên cứu xác định sự biến động độ sâu là một phần của đảo nổi thuộc Huyện đảo Trường Sa, Tỉnh Khánh Hòa. Khu vực này nằm theo trục Đông Bắc - Tây Nam dài khoảng 9 km và rộng khoảng 1,8 km, ở giữa có một vịnh biển kín với độ sâu từ 3 m đến 20 m. Các đảo xa bờ có hệ sinh thái phong phú và vị trí đặc biệt quan trọng về quốc phòng an ninh (Hình 2).

Nghiên cứu xây dựng bản đồ độ sâu dựa trên bộ dữ liệu phản xạ bề mặt Sentinel-2 có sẵn trên nền tảng điện toán đám mây Google Earth Engine. Dữ liệu sử dụng trong nghiên cứu đã xử lý mức 2A với phép chiếu UTM, hệ quy chiếu WGS84, múi chiếu 49N.

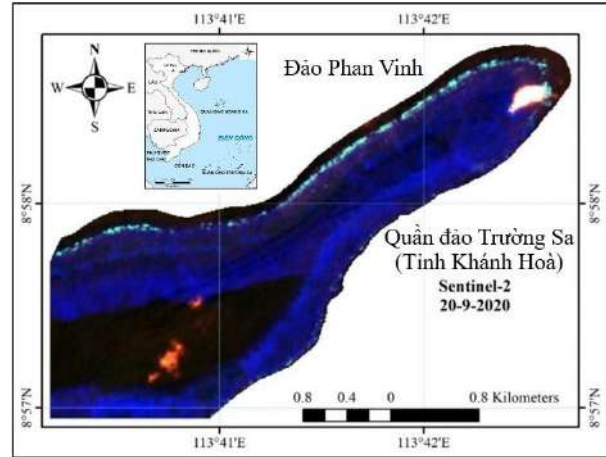


Hình 2. Khu vực nghiên cứu

Nghiên cứu tiến hành thu thập dữ liệu phản xạ bề mặt của vệ tinh Sentinel-2 khu vực đảo Phan Vinh thuộc huyện đảo Trường Sa, Tỉnh Khánh Hoà. Hai hình ảnh được thu thập lần lượt ngày 27-3-2020 và 20-9-2020. Dữ liệu vệ tinh sau khi lọc mây, lọc nhiễu và lọc các đối tượng không nước được thể hiện trong Hình 3 và Hình 4.



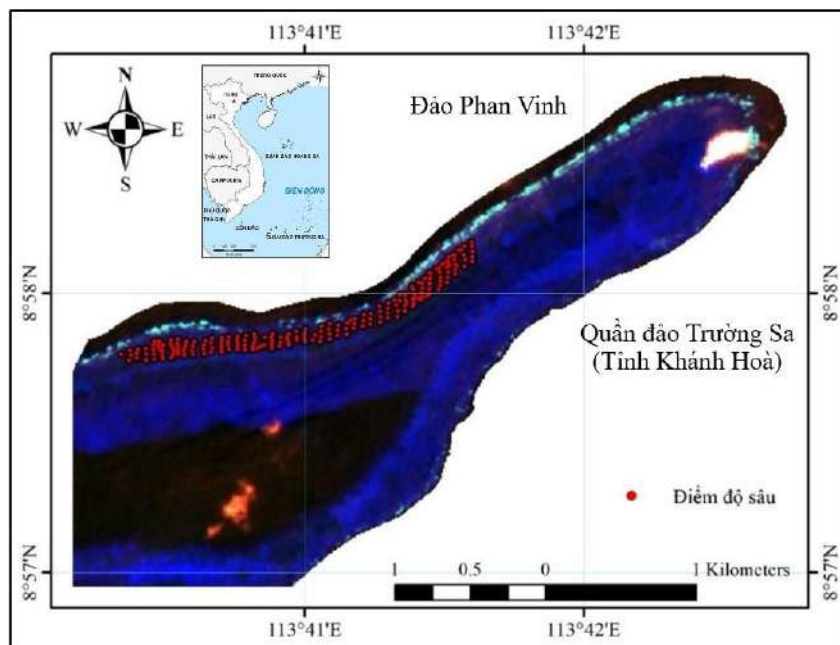
Hình 3. Hình ảnh Sentinel-2 khu vực nghiên cứu chụp ngày 27/3/2020



Hình 4. Hình ảnh Sentinel-2 khu vực nghiên cứu chụp ngày 20/9/2020

3.2. Dữ liệu đo thực địa

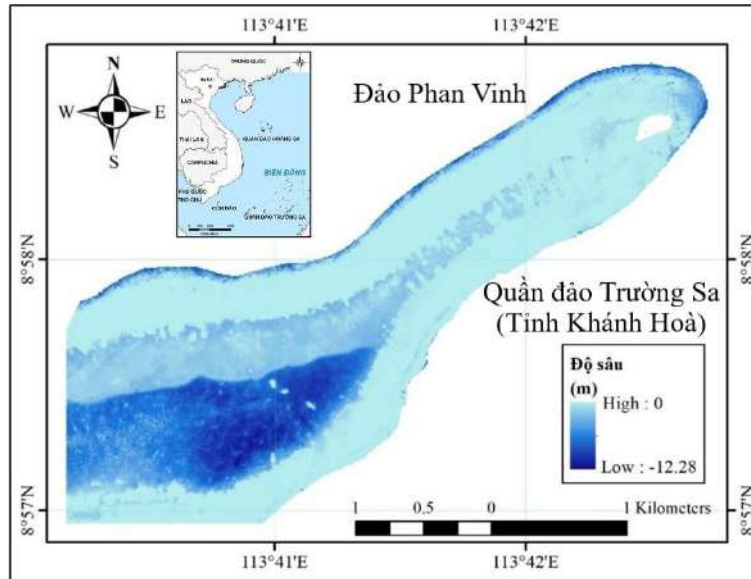
Hình 5 minh họa vị trí đo đạc tại thực địa, bao gồm 210 điểm. Để thu thập dữ liệu tại hiện trường sử dụng phương pháp đo GNSS (RTK) kết hợp đo sâu hồi âm đơn tia Hidrobox đã được đồng bộ với hệ quy chiếu dữ liệu ảnh thực nghiệm. Các điểm đo này có đặc điểm nằm trên vùng biển có địa chất đáy biển cứng, không bị cát và đối tượng khác che phủ nên có độ sâu không thay đổi trong thời gian nghiên cứu. Đồng thời, giá trị độ sâu của các điểm đo thực địa là từ 0,5 m đến 1,3 m.



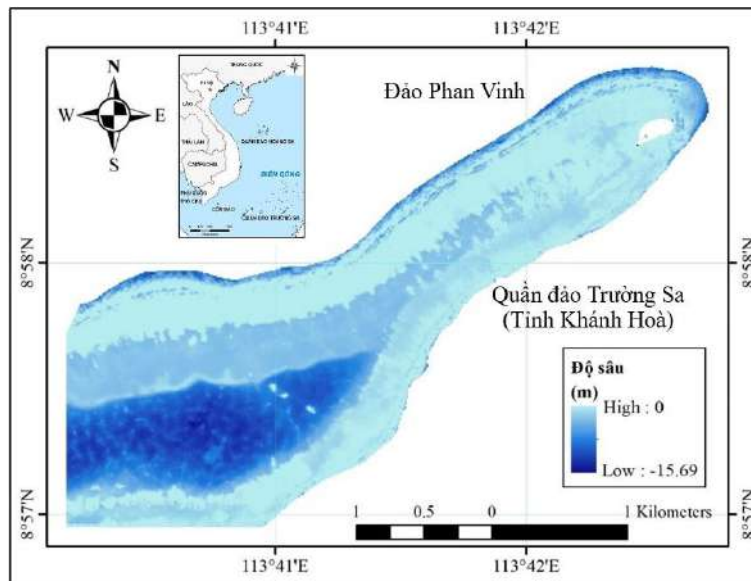
Hình 5. Các điểm đo thực địa được sử dụng cho mục đích hiệu chỉnh

4. Kết quả nghiên cứu và thảo luận

Tiếp theo, các hình ảnh phản xạ bề mặt Sentinel-2 được sử dụng để tính toán độ sâu. Thuật toán được tính toán độ sâu được xây dựng theo ngôn ngữ JavaScript trên nền tảng GEE với các tham số mô hình được trình bày trong Phần 2. Các bản đồ độ sâu lần lượt được trình bày trong Hình 6 và Hình 7.

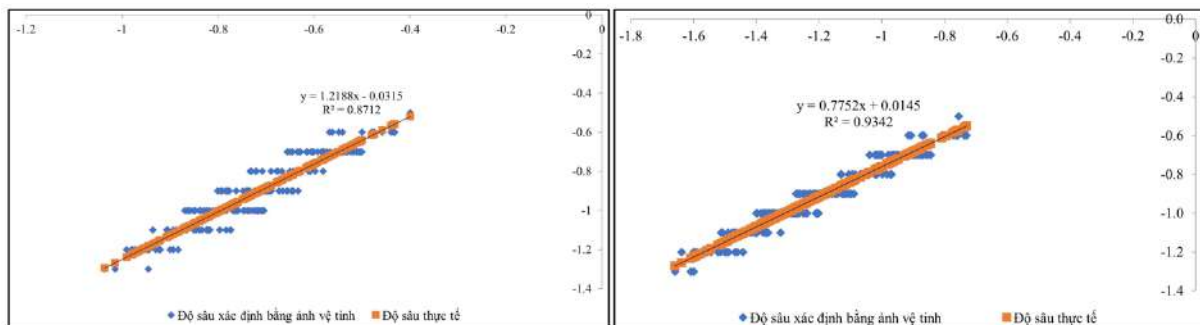


Hình 6. Bản đồ độ sâu được tính bằng dữ liệu vệ tinh ngày 27/3/2020



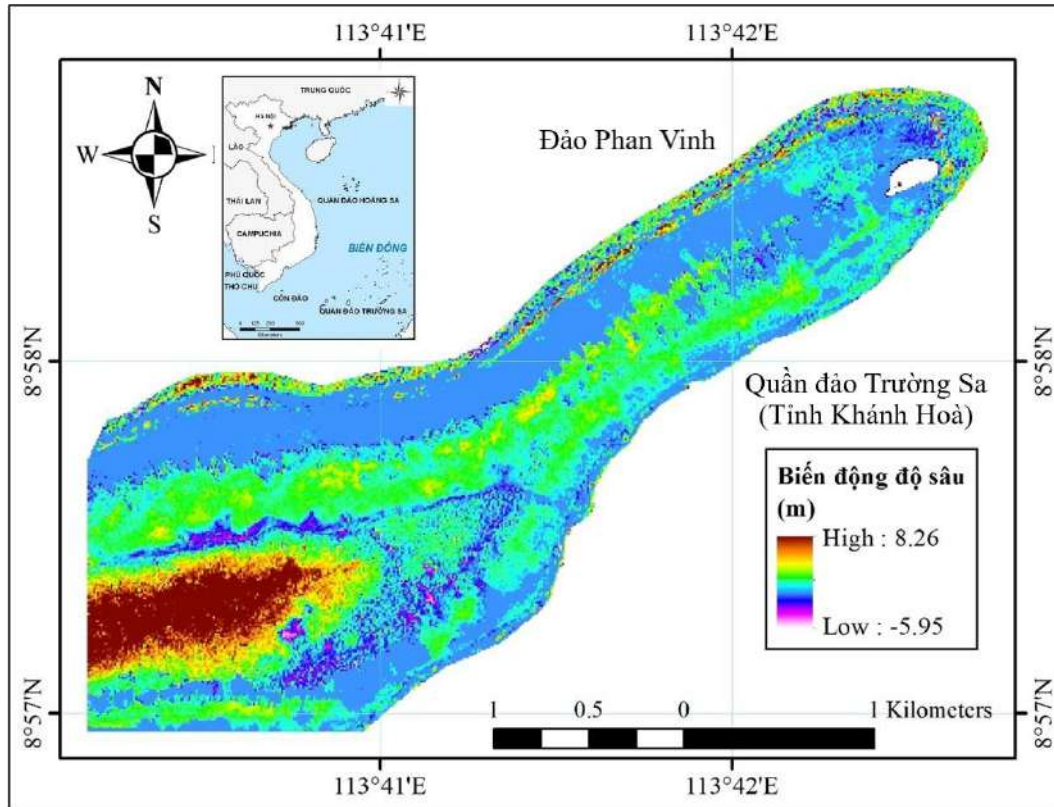
Hình 7. Bản đồ độ sâu được tính bằng dữ liệu vệ tinh ngày 20/9/2020

Kết quả tính toán độ sâu từ ảnh vệ tinh được kiểm tra với 210 điểm đo đặc thực tế (Hình 8) cho ta thấy tính tương quan cao giữa giá trị đo độ sâu thực tế ngoài địa và giá độ sâu được xác định từ ảnh vệ tinh (với R^2 của thời điểm chụp ảnh là 0,87 và 0,93). Bên cạnh đó, giá trị RMSE thu được tương ứng với dữ liệu ngày 27/3/2020 và ngày 20/9/2020 lần lượt là 0,06 m và 0,04 m.



Hình 8. Hàm tương quan giữa độ sâu được xác định bằng dữ liệu vệ tinh và độ sâu thực tế tại thực địa

Thủy triều đo được tại khu vực nghiên cứu trong thời điểm thu nhận ảnh ngày 27/3/2020 là 1,2 m, ngày 20/9/2020 là 1,3 m. Như vậy, sự khác biệt về thủy triều giữa hai thời điểm là 0,1 m. Bên cạnh đó, các điểm đo thực tế được sử dụng để hiệu chỉnh độ sâu. Vị trí các điểm đo được hiển thị trong Hình 5.



Hình 9. Bản đồ biến động độ sâu khu vực nghiên cứu

Hình 9 trình bày sự thay đổi độ sâu khu vực nghiên cứu tại hai thời điểm khác nhau. Kết quả cho thấy khu vực nghiên cứu có sự dịch chuyển mạnh của các đối tượng đáy biển bao gồm các rạn san hô, thảm cỏ biển và đặc biệt là cát biển.

Khu vực biến động lớn nhất tập trung ở giữa lòng hồ. Nơi có sự biến thiên độ sâu lớn lên đến +8,26 m. Điều này cho thấy, độ sâu ở trong lòng hồ tại thời điểm tháng 9/2020 đã thấp hơn so với độ sâu tại thời điểm tháng 3/2020. Khu vực lòng hồ có giá trị biến động độ sâu lớn ngoài nguyên nhân sự lắng đọng của cát san hô thì sự phát triển của cỏ và tảo biển cũng ảnh hưởng chủ yếu đến sự thay đổi lớn của giá trị độ sâu.

Bên cạnh đó, cát biển và rạn san hô có sự dịch chuyển sang hai bên lòng hồ, tạo thành hình vành khuyên. Vành khuyên này có sự biến thiên độ sâu lớn nhất là -5,95 m. Kết quả đó đồng nghĩa với việc cát biển đã được bồi tụ tại vị trí vành khuyên này.

Ngoài ra, địa hình tự nhiên của lòng hồ có thể ảnh hưởng đến cách cát phân bố dẫn tới độ sâu của lòng hồ thay đổi. Dựa trên kết quả độ sâu tại Hình 6 và Hình 7, lòng hồ có độ dốc cao. Điều này khiến lòng hồ có thể dễ dàng bị xói mòn và cát bị đẩy đi (Mỹ, 1995).

5. Kết luận

Nghiên cứu xác định sự biến động độ sâu khu vực ven đảo xa bờ sử dụng dữ liệu viễn thám trên nền tảng điện toán đám mây Google Earth Engine. Trước hết, các tác giả sử dụng hai hình ảnh Sentinel-2 tại hai thời điểm khác nhau thuộc khu vực nghiên cứu. Tiếp theo, thuật toán Stump được sử dụng để tính toán độ sâu tự động. Đồng thời, nghiên cứu sử dụng dữ liệu thủy triều và điểm đo thực tế để hiệu chỉnh độ sâu được tính toán từ dữ liệu vệ tinh. Từ độ sâu tính toán và hiệu chỉnh cao triều thời điểm chụp ảnh được tích hợp bằng phần mềm ArcGIS để đánh giá sự thay đổi địa hình giữa hai thời điểm. Kết quả của phương pháp bước đầu cho thấy tính tự động, nhanh và hiệu quả trong phục vụ đánh giá sự thay đổi độ sâu trên vùng biển Việt Nam, đặc biệt là ở những khu vực ven bờ, ven quần đảo. Việc đánh giá biến động độ sâu biển không chỉ mang lại lợi ích trực tiếp mà còn gián tiếp hỗ trợ nhiều lĩnh vực khác nhau, từ kinh tế, quốc phòng, môi trường, đến giáo dục và nghiên cứu khoa học, góp phần vào sự phát triển bền vững của đất nước. Tuy nhiên, do điều kiện thực nghiệm, dữ liệu đo và quan trắc cao độ thủy triều chưa được nhiều chu kỳ, nên cần nghiên cứu tiếp để hoàn thiện tiếp quy trình xác định biến động độ sâu từ dữ liệu ảnh vệ tinh.

Tài liệu tham khảo

- Zhao, J.; Barnes, B.; Melo, N.; English, D.; Lapointe, B.; Muller-Karger, F.; Schaeffer, B.; Hu, C. 2013. Assessment of Satellite-Derived Diffuse Attenuation Coefficients and Euphotic Depths in South Florida Coastal Waters. *Remote Sens. Environ.*, 131, 38-50.
- Traganos, D.; Poursanidis, D.; Aggarwal, B.; Chrysoulakis, N.; Reinartz, P. 2018. Estimating Satellite-Derived Bathymetry (SDB) with the Google Earth Engine and Sentinel-2. *Remote Sens.*, 10, 859.
- Balsamo, G.; Dutra, E.; Stepanenko, V.M.; Viterbo, P.; Miranda, P.; Mironov, D. 2010. Deriving an Effective Lake Depth from Satellite Lake Surface Temperature Data: A Feasibility Study with MODIS Data; ECMWF: Reading, UK.
- Barale, V.; Jaquet, J.-M.; Ndiaye, M. 2008. Algal Blooming Patterns and Anomalies in the Mediterranean Sea as Derived from the SeaWiFS Data Set (1998-2003). *Remote Sens. Environ.*, 112, 3300-3313.
- Tedesco, M.; Steiner, N. 2011. In-Situ Multispectral and Bathymetric Measurements over a Supraglacial Lake in Western Greenland Using a Remotely Controlled Watercraft. *Cryosphere*, 5, 445-452.
- Dekker, A.G.; Phinn, S.R.; Anstee, J.; Bissett, P.; Brando, V.E.; Casey, B.; Fearn, P.; Hedley, J.; Klonowski, W.; Lee, Z.P. 2011. Intercomparison of Shallow Water Bathymetry, Hydro-optics, and Benthos Mapping Techniques in Australian and Caribbean Coastal Environments. *Limnol. Oceanogr. Methods*, 9, 396-425.
- Brando, V.E.; Anstee, J.M.; Wettle, M.; Dekker, A.G.; Phinn, S.R.; Roelfsema, C. 2009. A Physics Based Retrieval and Quality Assessment of Bathymetry from Suboptimal Hyperspectral Data. *Remote Sens. Environ.*, 113, 755-770.
- Kerr, J.M.; Purkis, S. 2018. An Algorithm for Optically-Deriving Water Depth from Multispectral Imagery in Coral Reef Landscapes in the Absence of Ground-Truth Data. *Remote Sens. Environ.*, 210, 307-324.
- Li, J.; Yu, Q.; Tian, Y.Q.; Becker, B.L.; Siqueira, P.; Torbick, N. 2018. Spatio-Temporal Variations of CDOM in Shallow Inland Waters from a Semi-Analytical Inversion of Landsat-8. *Remote Sens. Environ.*, 218, 189-200.
- Li, J.; Knapp, D.E.; Schill, S.R.; Roelfsema, C.; Phinn, S.; Silman, M.; Mascaro, J.; Asner, G.P. 2019. Adaptive Bathymetry Estimation for Shallow Coastal Waters Using Planet Dove Satellites. *Remote Sens. Environ.*, 232, 111302.
- Vanhellemont, Q. 2019. Daily Metre-Scale Mapping of Water Turbidity Using CubeSat Imagery. *Opt. Express*, 27, A1372-A1399.
- Lee, Z.; Weidemann, A.; Arnone, R. 2013. Combined Effect of Reduced Band Number and Increased Bandwidth on Shallow Water Remote Sensing: The Case of Worldview 2. *IEEE Trans. Geosci. Remote Sens.*, 51, 2577-2586.
- Li, J.; Knapp, D.E.; Lyons, M.; Roelfsema, C.; Phinn, S.; Schill, S.R.; Asner, G.P. 2021. Automated Global Shallow Water Bathymetry Mapping Using Google Earth Engine. *Remote Sens.*, 13, 1469.
- Stumpf, R.P.; Holderied, K.; Sinclair, M. 2003. Determination of Water Depth with High-resolution Satellite Imagery over Variable Bottom Types. *Limnol. Oceanogr.*, 48, 547-556.
- Nguyen, N. H., Le, V. P. (2023). Bathymetric mapping of Co To island area, Quang Ninh province using sentinel-2 and Landsat-8 images. *Journal of Science and Technique-Section on Special Construction Engineering*, 6(01).
- Nguyễn, Q. M. (1995). Ảnh hưởng của yếu tố địa hình đến xói mòn đất ở Việt Nam. *VNU Journal of Science: Natural Sciences and Technology*, 11(1).

ABSTRACT

Research to determine depth variations along offshore islands using Sentinel-2 remote sensing data on the Google Earth Engine platform

Nguyen Nhu Hung^{1*}, Le Van Phu¹, Nguyen Van Son¹

¹*Le Quy Don Technical University, Hanoi, Vietnam*

Variations in depth along offshore islands can affect marine ecosystems, coastal planning and economic development, as well as maritime activities. Monitoring and evaluating these fluctuations helps protect the environment, manage marine resources and guide sustainable development. Research to determine depth fluctuations along offshore islands using remote sensing data on the Google Earth Engine (GEE) platform is a very effective research direction thanks to its ability to process large data and provide powerful analysis tools. In this study, the deep changes along the offshore island are determined using remote sensing image data collected at two times, March 27, 2020 and September 20, 2020, respectively. The results show that the correlation values between the actual depth measurement in the field and the depth value determined from satellite images are 0.87 and 0.93. The corresponding RMSE values are 0.06 m and 0.04 m, respectively. In addition, the maximum depth variation value is 8.26 m. From the results of calculating the satellite image depth and tidal height at the time the image was taken, synchronize the depth correction and compare and evaluate the changes. The initial results of the method show that it is automatic, fast and effective in assessing depth changes in the Vietnamese and Southern waters, especially in coastal and archipelagic areas. At the same time, the method also supports regular monitoring and updating.

Keywords: Bathymetric, Sentinel-2, Google Earth Engine.

Nghiên cứu một số thuật toán huấn luyện mạng nơron nhân tạo trong dự báo độ lún công trình thủy điện

Phạm Quốc Khánh^{1,*}

¹Trường Đại học Mở - Địa chất

TÓM TẮT

Mạng nơron nhân tạo (Artificial Neural Network - ANN) đã và đang được ứng dụng rộng rãi trong ngành trắc địa, trong xây dựng mô hình chuyển dịch và dự báo thì mạng ANN Back-Propagation (BP) được ứng dụng nhiều hơn cả. Để huấn luyện mạng BP có rất nhiều thuật toán khác nhau, mỗi thuật toán lại có ưu, nhược điểm riêng trong quá trình huấn luyện mẫu. Bài báo sử dụng 02 tập mẫu là số liệu quan trắc độ lún của công trình thủy điện Hòa Bình với số mẫu và lớp đầu vào khác nhau, huấn luyện các tập mẫu này bằng 06 và 03 thuật toán như thuật toán Levenberg-Marquardt (Trainlm), tốc độ học thay đổi (Traingdx), thuật toán Quasi-Newton,... rồi so sánh kết quả dự báo với số liệu đo thực tế. Kết quả cho thấy với tập mẫu ít (17 mẫu) và không có nhiều lớp đầu vào (chỉ có lớp thời gian và độ lún đo bằng máy thủy chuẩn hạng 2) thì giá trị dự báo (01 giá trị) là như nhau với các thuật toán khác nhau. Cũng với tập mẫu này khi sử dụng một thuật toán thì trong các lần huấn luyện, mặc dù các chỉ tiêu đánh giá như hiệu suất huấn luyện, trạng thái huấn luyện và phân tích tương quan khác nhau nhưng vẫn cho kết quả dự báo giống nhau. Khi giá trị tương quan R càng tiến sát 1 thì giá trị mô hình dự báo (không phải giá trị dự báo) và giá trị đo càng gần với nhau. Với tập mẫu phức tạp hơn (20 mẫu) với nhiều lớp đầu vào (thời gian, mực nước và độ lún) thì các thuật toán này cho kết quả dự báo khác nhau, chứng tỏ số tập mẫu và số lớp đầu vào có ảnh hưởng đến kết quả học của thuật toán. Như vậy, khi xây dựng mô hình chuyển dịch và dự báo cần làm phép thử với một số thuật toán để có được kết quả dự báo tốt hơn và sát với giá trị thực tế hơn.

Từ khóa: ANN, BP, dự báo chuyển dịch.

1. Đặt vấn đề

Dự báo chuyển dịch các công trình lớn, trọng điểm quốc gia như đập thủy điện, cầu lớn, nhà cao tầng và siêu cao tầng, trượt lở đất đá,... là một mắt xích vô cùng quan trọng trong quy trình quan trắc chuyển dịch công trình. Trên thế giới, đây là phần không thể thiếu của quá trình quan trắc nhưng ở nước ta chưa có quy định nào về vấn đề này. Nghiên cứu về dự báo gần đây ở nước ta vẫn là các phương pháp tương đối kinh điển là sử dụng các hàm toán học để mô phỏng chuyển dịch, các phương pháp này mô phỏng được cả quá trình chuyển dịch chứ không mô phỏng xu thế chuyển dịch của từng chu kỳ quan trắc. Phương pháp tự hồi quy, lọc Kanmal, BP nhân tạo (Nguyễn Đình Thúc, 2000; Phạm Quốc Khánh, Nguyễn Quốc Long, 2015; Phạm Quốc Khánh, Nguyễn Việt Hà, 2/2015; Cao Đình Trọng và nnk., 2011) cũng đã được nghiên cứu nhưng chưa phổ biến do ở Việt Nam số lượng mẫu quan trắc hạn chế.

Ứng dụng mạng nơron nhân tạo (ANN) trên thế giới có nhiều nghiên cứu ví dụ như dự báo động đất, dự báo độ lún công trình, dự báo trọng lực khu vực đập thủy điện, dự báo trượt lở (S. Narayanakumar, K.Raja, 2016; George Pantazis, Eleni-Georgia Alevizakou, 2013; Xiaoyu Wang et, 2017; Ke Fuyang, Li Yayun, 2014) với nhiều thuật toán dự báo của ANN khác nhau. Các nghiên cứu này đều nhận xét mạng ANN-BP (nơron nhân tạo lan truyền ngược) được áp dụng nhiều hơn cả. Tuy nhiên, mạng ANN-BP có rất nhiều thuật toán huấn luyện mẫu, các thuật toán này hiện nay đã được xây dựng thành các modul trên nhiều phần mềm phổ biến như Matlab. Tuy nhiên, chưa thấy có đánh giá về việc nên sử dụng thuật toán nào là tối ưu và vượt trội hơn cả khi xây dựng mô hình chuyển dịch và dự báo biến dạng công trình. Chính vì thế, bài báo này dựa trên mạng ANN-BP để dự báo chuyển dịch với các thuật toán khác nhau là TrainLM, TrainGDX, TrainRP, TrainSCG, TrainBFG, TrainOSS (J. Zupan, J.Gasteiger. 1993; Prashant Rao, 2016; Supported and Compatible Compilers - Release 2016a, 2016); xây dựng mô hình chuyển dịch và dự báo độ lún công trình thủy điện Hòa Bình theo các thuật toán này, sau đó tiến hành so sánh kết quả dự báo với giá trị thực tế để xác định xem thuật toán nào có hiệu quả hơn. Kết quả cho thấy, với một lượng mẫu không nhiều và lớp đầu vào ít thì các thuật toán này cho kết quả dự báo như nhau. Nhưng giá trị dự báo sẽ khác nhau khi lớp đầu vào nhiều hơn dù số lượng mẫu gần tương tự. Do đó, để dự báo được giá trị gần sát với thực tế thì khi thực hiện tính toán cần thiết phải làm các phép thử với các thuật toán này rồi mới ra quyết định nên chọn thuật toán nào ứng dụng trong bài toán mà chúng ta đang cần.

* Tác giả liên hệ

Email: phamquockhanh@hmg.edu.vn

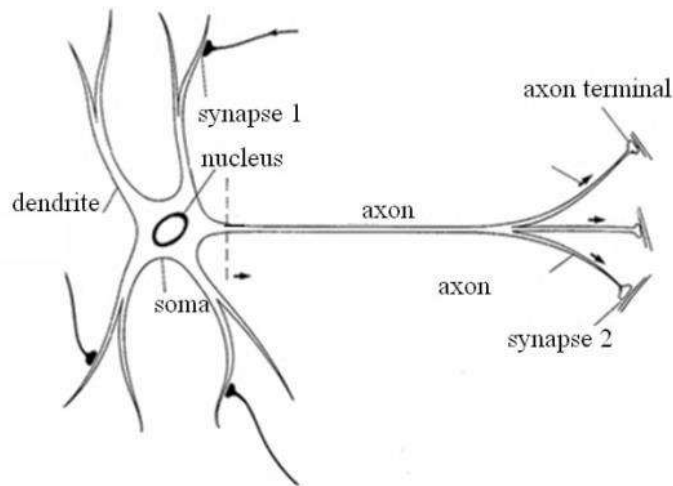
2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1. Cơ sở lý thuyết của mạng nơron nhân tạo

2.1.1. Nơron thần kinh và cấu trúc của nó

Các tế bào thần kinh trong não được gọi là các nơron thần kinh. Mỗi nơron liên kết với hàng nghìn nơron khác. Nơron là đơn vị cơ bản mà bộ não dùng để xử lý thông tin, ước tính rằng trong vỏ não con người có xấp xỉ 10 tỷ nơron.

Cấu trúc một nơron thần kinh với các thành phần chính được minh họa trong như Hình 1.



Hình 1. Cấu trúc của một nơron thần kinh

trong đó:

- Soma là thân của nơron.
- Synapse là kết nối giữa một nơron và nơron khác, hay còn gọi là khớp thần kinh. Thông tin đi ra từ một nơron qua Axon của nó sau đó được truyền qua một Synapse đến nơron nhận.
- Dendrite là các dây mảnh gắn liền với soma (tua gai thần kinh), chúng truyền dữ liệu dưới dạng xung điện thế đến cho soma xử lý. Bên trong soma các dữ liệu đó được tổng hợp lại, có thể xem gần đúng sự tổng hợp ấy như là một phép lấy tổng tất cả các dữ liệu mà nơron nhận được.
- Axon là một các loại tín hiệu khác cũng được gắn với soma. Khác với dendrites, các axon có khả năng phát các xung điện thế, chúng là dây dẫn tín hiệu từ nơron đi các nơi khác. Các axon có một giá trị ngưỡng threshold, chỉ khi nào điện thế trong soma vượt quá một threshold nào đó thì axon mới phát một xung điện thế, còn nếu không thì nó ở trạng thái nghỉ.

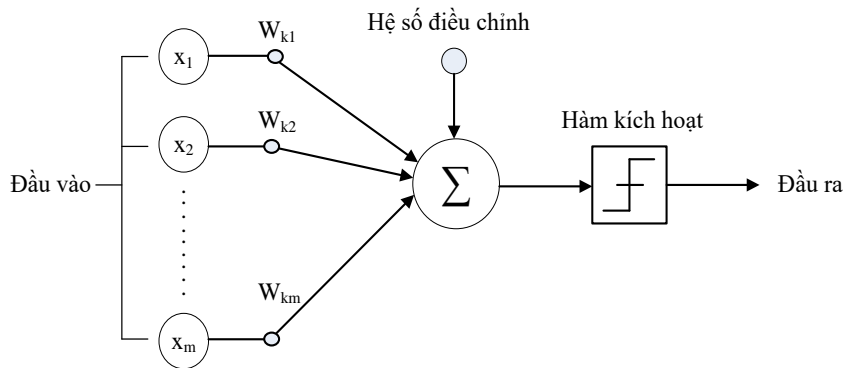
2.1.2. Mạng nơron sinh học và mạng nơron nhân tạo

Theo nghĩa sinh học, mạng nơron là một tập hợp các dây thần kinh kết nối với nhau. Hiện nay, thuật ngữ này còn dùng để chỉ mạng nơron nhân tạo, cấu thành từ các nơron nhân tạo. Do đó thuật ngữ “mạng nơron” xác định hai khái niệm phân biệt: mạng nơron sinh học và mạng nơron nhân tạo (Hu Wusheng, 2006).

a. Mạng nơron sinh học là một mạng lưới các nơron có kết nối hoặc có liên quan về mặt chức năng trực thuộc hệ thần kinh ngoại biên hay hệ thần kinh trung ương. Thông thường, một mạng nơron bao gồm một hoặc nhiều nhóm các nơron được kết nối vật lý với nhau hoặc có liên quan với nhau về chức năng. Một nơron đơn có thể được nối với nhiều nơron khác và tổng số nơron và kết nối trong một mạng có thể là một giá trị cực kỳ lớn. Các kết nối (khớp thần kinh) từ các axon tới các tế bào tua gai thần kinh (dendrite). Ngoài tín hiệu điện, còn có các dạng tín hiệu khác phát sinh từ việc khuếch tán các chất dẫn truyền xung động thần kinh (neurotransmitter), chúng có ảnh hưởng đối với tín hiệu điện. Do vậy mạng nơron vô cùng phức tạp.

b. Mạng nơron nhân tạo mô phỏng các tế bào thần kinh và các khớp nối thần kinh của bộ não con người, trong mạng nơron nhân tạo cũng có các thành phần có vai trò tương tự các thành phần trong nơron thần kinh. Hai thành phần chính cấu tạo nên mạng nơron nhân tạo là các nơron (mô phỏng các tế bào thần kinh) và các synapse (mô phỏng các khớp nối thần kinh).

Mô hình mạng nơron nhân tạo được mô tả như Hình 2:



Hình 2. Mô hình mạng neuron nhân tạo

Trong một mô hình neuron bao gồm 3 thành phần chính đó là:

- Một tập hợp các synapse hay các kết nối, mà mỗi một trong chúng được đặc trưng bởi một trọng số của riêng nó. Tức là một tín hiệu x_j tại ngõ vào của synapse j nối với neuron k sẽ được nhân với trọng số synapse w_{kj} . Ở đó k là chỉ số của neuron tại ngõ ra của synapse đang xét, còn j chỉ điểm ngõ vào của synapse. Các trọng số synapse của một neuron nhân tạo có thể nhận cả các giá trị âm và các giá trị dương.
- Một bộ cộng để tính tổng các tín hiệu ngõ vào của neuron, đã được nhân với các trọng số synapse tương ứng; phép toán được mô tả ở đây tạo nên một bộ tổ hợp tuyến tính.
- Một hàm kích hoạt (activation function) để giới hạn biên độ ngõ ra của neuron. Hàm kích hoạt cũng được xem xét như là một hàm nén; nó nén (giới hạn) phạm vi biên độ cho phép của tín hiệu ngõ ra trong một khoảng giá trị hữu hạn. Mô hình neuron trong Hình 2 còn bao gồm một hệ số hiệu chỉnh tác động từ bên ngoài, b_k . Hệ số hiệu chỉnh b_k có tác dụng tăng lên hoặc giảm đi ngõ vào thực của hàm kích hoạt, tùy theo nó dương hay âm.

Thể hiện bằng công thức toán học, một neuron k được mô tả như sau:

$$u_k = \sum_{j=0}^m w_{kj} x_j \quad (1)$$

và:

$$y_k = \varphi(u_k + b_k) \quad (2)$$

trong đó: x_1, x_2, \dots, x_m là các tín hiệu đầu vào; $w_{k1}, w_{k2}, \dots, w_{km}$ là các trọng số synapse của neuron k ; u_k là đầu ra bộ tổ hợp tuyến tính tương ứng; b_k là hệ số hiệu chỉnh.

Hệ số hiệu chỉnh b_k là một tham số ngoài của neuron nhân tạo k , từ đó công thức (1) và (2) có thể viết lại một cách tương đương là:

$$v_k = \sum_{j=0}^m w_{kj} x_j \quad (3)$$

và:

$$y_k = \varphi(v_k) \quad (4)$$

Trong công thức (3), chúng ta đã thêm một synapse mới với đầu vào (Bias) của nó là:

$$x_0 = +1 \quad (5)$$

Và trọng số là:

$$w_{k0} = b_k \quad (6)$$

Như vậy, mô hình của neuron k sẽ được mô phỏng lại như Hình 3. Trong Hình 3, nhiệm vụ của hệ số hiệu chỉnh là thực hiện hai việc: Một là, thêm một tín hiệu đầu vào cố định là 1; hai là, thêm một trọng số synapse mới bằng giá trị của hệ số b_k .

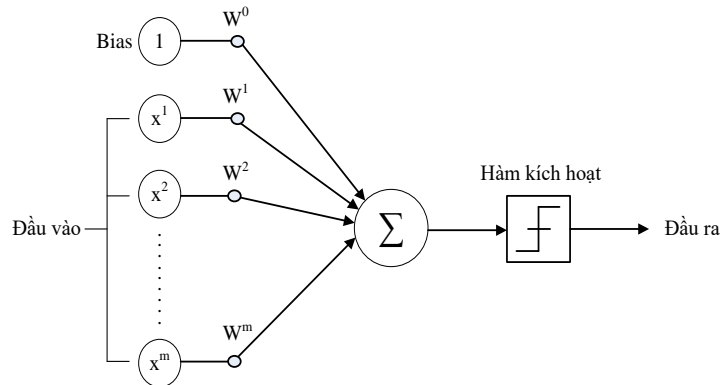
Mô hình này trên thực tế được gọi là một “Perceptron”. Perceptron, một phát minh của Rosenblatt [1962] là một trong những mô hình neuron sớm nhất. Một Perceptron mô hình một neuron bằng cách tính tổng các trọng số của ngõ vào (input). Nếu tổng này lớn hơn một giá trị ngưỡng (threshold) thì cho giá trị ngõ ra (output) bằng 1 ngược lại cho giá trị ngõ ra (output) bằng 0.

Ngoài ra, trong mô hình mạng neuron còn có giá trị ngưỡng và hàm kích hoạt:

Giá trị ngưỡng “threshold” là một giá trị có thể điều chỉnh được. Các ngõ vào (x_1, x_2, \dots, x_m) và các trọng số (w_1, w_2, \dots, w_m) trong Perceptron là các giá trị thực, có thể âm (-) hoặc dương (+). Bản thân Perceptron bao gồm các trọng số, bộ xử lý tổng, hàm kích hoạt và một bộ xử lý ngưỡng có thể điều chỉnh được gọi là “Bias”. Bias có

thể được xem là xu hướng (một xu hướng theo một cách cư xử cụ thể) của Perceptron để kích khởi bất chấp ngõ vào của nó. Mạng mô hình Perceptron kích khởi khi tổng trọng số lớn hơn 0.

Hàm kích hoạt, ký hiệu bởi $\varphi(v)$, xác định đầu ra của nơron, hàm kích hoạt bao gồm hàm ngưỡng và hàm sigmoid.



Hình 3. Mô hình mạng nơron nhân tạo với giá trị Bias

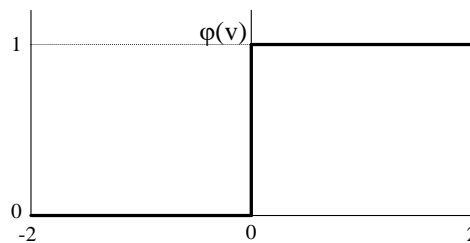
Hàm ngưỡng được biểu diễn dưới dạng toán học là:

$$\varphi(v) = \begin{cases} 1 & \text{nếu } v \geq 0 \\ 0 & \text{nếu } v < 0 \end{cases} \quad (7)$$

Khi đó đầu ra của bộ tổ hợp tuyến tính theo hàm ngưỡng là

$$v_k = \sum_{j=1}^m w_{kj} x_j + b_k \quad (8)$$

Và đồ thị của nó là

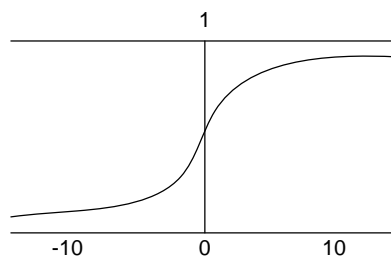


Hình 4. Hàm ngưỡng

Hàm sigmoid là dạng chung nhất của hàm kích hoạt được sử dụng trong cấu trúc mạng nơron nhân tạo. Nó là một hàm tăng và hàm này thể hiện sự trung gian giữa tuyến tính và phi tuyến. Một ví dụ của hàm này là hàm logistics, xác định như sau:

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (9)$$

trong đó: a là tham số độ dốc của hàm sigmoid. Bằng việc biến đổi tham số a, chúng ta thu được các hàm sigmoid với các độ dốc khác nhau. Đồ thị hàm sigmoid như Hình 5.

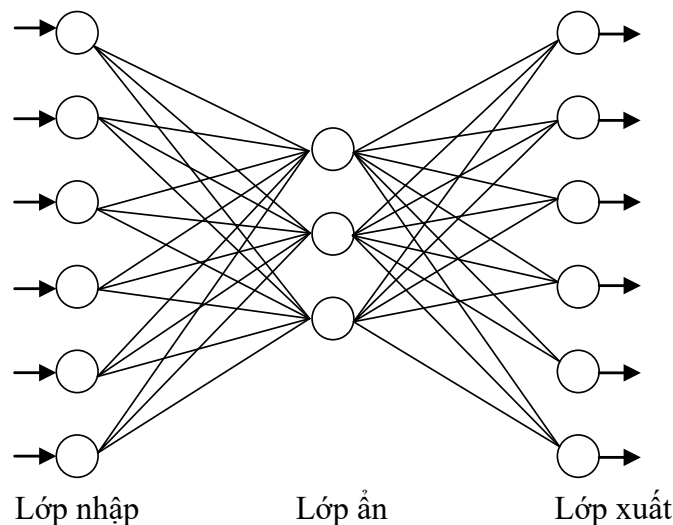


Hình 5. Hàm logistic

2.1.3. Mạng neuron lan truyền ngược (Back-Propagation)

Mạng lan truyền ngược là một hàm phi tuyến có thể xấp xỉ gần đúng nhất một hàm đích được cho qua một số mẫu trong tập mẫu.

Một mạng lan truyền tổng quát là một mạng có n ($n > 2$) lớp: lớp thứ nhất gọi là lớp nhập, lớp thứ n là lớp xuất và $n-2$ lớp ẩn. Phần lớn các mạng chỉ có một biến phụ thuộc nhưng với mô hình tả trên mạng neuron cho phép số nút trong lớp nhập, lớp xuất và lớp ẩn là tùy ý. Số nút trong lớp nhập và lớp xuất phụ thuộc vào yêu cầu bài toán, số nút ẩn do người xây dựng mô hình quyết định. Hình 6 minh họa một mạng lan truyền với một lớp ẩn, mạng này thường được sử dụng xây dựng mô hình biến dạng và dự báo chuyển dịch công trình (Hu Wusheng, 2006).



Hình 6. Mạng lan truyền với một lớp ẩn

Mạng Back-Propagation thực hiện dựa trên quy luật học Widrow-Hoff tổng quát hóa cho mạng đa lớp và các hàm truyền phi tuyến khác nhau. Mạng có ngưỡng, một lớp sigmoid và một lớp tuyến tính ngõ ra có thể mô phỏng bất kỳ hàm nào với số mẫu rời rạc hữu hạn.

Thuật toán huấn luyện Back-Propagation chuẩn là gradient descent, chính là quy luật học Widrow-Hoff. Khái niệm Back-Propagation mô tả cách tính gradient trong mạng đa lớp phi tuyến. Thuật toán cơ bản có một số thay đổi dựa trên các kỹ thuật tối ưu hóa chuẩn như là các phương pháp “conjugate gradient” và “Newton”.

Gọi ANN-BP là mạng lan truyền ngược vì quá trình luyện mạng được bắt đầu với các giá trị trọng số tùy ý, có thể là ngẫu nhiên và tiến hành lặp đi lặp lại. Mỗi lần lặp gọi là một thế hệ. Trong mỗi thế hệ, mạng hiệu chỉnh các trọng số sao cho sai số giảm dần. Sai số là độ lệch giữa kết quả tính được và kết quả đích. Tiến trình điều chỉnh nhiều lần giúp cho trọng số dần đạt các giá trị tối ưu. Để cập nhật trọng số trong mỗi thế hệ, mạng phải xử lý tất cả các mẫu trong tập mẫu.

Như vậy, đối với từng mẫu, mạng sẽ thực hiện các thao tác sau đây: (1) Trước tiên mạng thực hiện quá trình lan truyền tiến, nghĩa là mạng ánh xạ các biến nhập của mẫu hiện hành thành các giá trị xuất. (2) Sai số được tính dựa trên giá trị tính được và giá trị đích. Trên cơ sở sai số tính toán được, mạng sẽ cập nhật lại các trọng số theo nguyên tắc lan truyền ngược sai số - gọi là lan truyền ngược. Như vậy để học mỗi mẫu, mạng thực hiện hai bước: lan truyền tiến - thực hiện ánh xạ và lan truyền ngược - cập nhật trọng số (Zhang Zhenglu et al., 2007).

Mạng Back-Propagation được huấn luyện chính xác sẽ cho đáp ứng hợp lý khi đưa ngõ vào chưa từng được huấn luyện. Thông thường tín hiệu mới vào có ngõ ra tương tự với ngõ ra chính xác của tín hiệu vào đã được huấn luyện giống với ngõ vào mới này. Do tính chất tổng quát hóa này, ta có thể huấn luyện mạng dựa trên các cặp vào/ra đại diện mà vẫn cho kết quả tốt đối với các tín hiệu chưa được huấn luyện. Có rất nhiều thuật toán huấn luyện mạng BP và sẽ được nghiên cứu trong phần tiếp theo của bài báo này.

2.2. Phương pháp nghiên cứu

Dựa vào lý thuyết của mạng BP và đã được xây dựng thành modul về trên phần mềm Matlab 2016a. Để xây dựng mô hình và dự báo chuyển dịch công trình cần có dữ liệu đầu vào, dữ liệu học và dữ liệu đầu ra. Quá trình huấn luyện mẫu có các thuật toán khác nhau, chúng khác nhau về cách tính toán, dung lượng bộ nhớ,... và được tóm tắt trong Bảng 1 (Prashant Rao, 2016; Supported and Compatible Compilers - Release 2016a, 2016).

Bảng 1. Một số thuật toán huấn luyện mạng BP

HÀM	MÔ TẢ
Traingd	Gradient Descent cơ bản. Đáp ứng chậm, sử dụng trong chế độ incremental.
Traingdm	Gradient Descent có quán tính. Thường nhanh hơn Traingd , có thể sử dụng trong chế độ nấc.
Traingdx	Tốc độ học thay đổi. Huấn luyện nhanh hơn Traingd nhưng chỉ có thể sử dụng trong chế độ batch.
Trainrp	Resilient Back-Propagation. Đây là thuật toán huấn luyện chế độ batch đơn giản, có độ hội tụ nhanh và ít tốn bộ nhớ.
Traincgf	Thuật toán Conjugate Gradient Fletcher - Reeves. Ít tốn bộ nhớ nhất trong các thuật toán conjugate gradient.
Traincgp	Thuật toán Conjugate Gradient Polak - Ribière. Yêu cầu bộ nhớ nhiều hơn Traincgf , đối với một số bài toán có độ hội tụ nhanh.
Traincgb	Thuật toán conjugate gradient Powell - Beale. Yêu cầu bộ nhớ lớn hơn Traincgp , tốc độ hội tụ nhanh hơn.
Trainscg	Thuật toán Scaled Conjugate Gradient. Đây là thuật toán Conjugate Gradient duy nhất không yêu cầu định hướng tìm.
Trainbfg	Phương pháp Quasi - Newton BFGS. Yêu cầu lưu trữ ma trận Hessian và có nhiều phép tính hơn các thuật toán Conjugate Gradient trong mỗi vòng lặp, nhưng thường hội tụ sau một vài phép lặp.
Trainoss	Phương pháp One Step Secant. Đây là phương pháp kết hợp giữa phương pháp Conjugate Gradient và phương pháp Quasi - Newton.
Trainlm	Thuật toán Levenberg - Marquardt. Thuật toán huấn luyện nhanh đối với mạng có kích thước vừa phải, có tính chất giảm bộ nhớ khi tập huấn luyện lớn.
Trainbr	Thuật toán chuẩn tổng quát hóa Bayesian. Là giải thuật huấn luyện Levenberg - Marquardt được sửa đổi, tạo ra mạng tốt hơn, giảm sự phức tạp để mạng hội tụ.

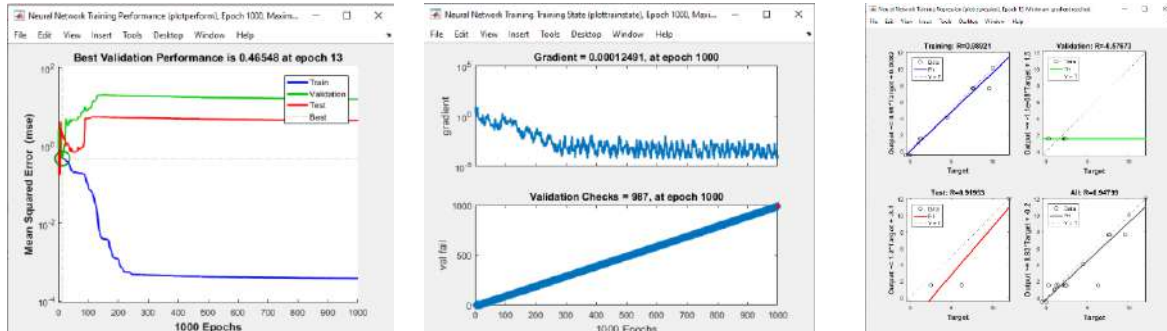
Để có thể so sánh và đánh giá xem trong các thuật toán này thuật toán nào có thể áp dụng tốt hơn cho công tác xây dựng mô hình và dự báo chuyển dịch. Nhóm tác giả đã sử dụng hai bộ dữ liệu quan trắc độ lún của thủy điện Hòa Bình để tiến hành làm thực nghiệm.

- Thứ nhất là số liệu đo lún của điểm PV-01 thuộc hệ thống quan trắc trên cơ đập thủy điện Hòa Bình với 18 chu kỳ, thời gian quan trắc một tháng một lần, số liệu quan trắc đầu vào chỉ có 2 thành phần là thời gian và kết quả đo lún như Bảng 2.

Bảng 2. Số liệu đo lún điểm PVM-1 thủy điện Hòa Bình

Chu kỳ (tháng)	Độ lún (mm)	Chu kỳ (tháng)	Độ lún (mm)
1	0	10	2,5
2	-0,5	11	4,5
3	1	12	7,6
4	6,3	13	7,7
5	2,3	14	9,6
6	2,1	15	7,7
7	1,4	16	10,1
8	1,2	17	12,1
9	0,3	18	13,3

Sử dụng 17 chu kỳ đầu để xây dựng mô hình, dự báo cho chu kỳ thứ 18 với các thuật toán huấn luyện (lấy ngẫu nhiên) của mạng BP là TrainLM, TrainGDX, TrainRP, TrainSCG, TrainBFG, TrainOSS. Thực hiện khi huấn luyện mạng với các điều kiện như nhau là: mạng BP có 3 lớp gồm lớp đầu vào, lớp ẩn và lớp đầu ra, trong đó lớp ẩn gồm 6 nơron. Trong quá trình huấn luyện, qua mỗi lần tính đều có thể xem kết quả huấn luyện một cách trực quan thông qua hiệu suất, trạng thái và hệ số hồi quy trong từng lần huấn luyện như được trình bày trong Hình 8.



a. Hiệu suất huấn luyện

b. Trạng thái huấn luyện

c. Phân tích hồi quy

Hình 8. Kết quả huấn luyện qua từng lần huấn luyện mạng

Tiếp đó, nhóm tác giả sử dụng thuật toán huấn luyện là TrainBFG và tiến hành so sánh kết quả dự báo qua từng lần huấn luyện để xem kết quả có khác nhau không?

- Thứ 2 là sử dụng số liệu đo lún khi mực nước hồ thay đổi với tập mẫu đầu vào có 3 thành phần là thời gian, mực nước hồ và độ lún theo mực nước hồ (Nguyễn Thị Kim Thanh, 2021). Tập mẫu xây dựng mô hình gồm 20 chu kỳ như Bảng 3, dự báo cho chu kỳ thứ 21 (thời gian: 32 tháng; mực nước: 104,68 m; độ lún: -32,00 mm). Sử dụng 3 thuật toán huấn luyện của mạng BP là TrainLM, TrainGDX, TrainRP để so sánh kết quả dự báo với nhau.

Bảng 3. Số liệu quan trắc mực nước và độ lún đập thủy điện Hòa Bình

Thời gian (T)	0	1	4	5	7	8	10	13	14	16
Mực nước (m)	0	102,71	101,82	89,62	109,8	117,2	116,41	101,86	109,91	101,85
Độ lún (mm)	0	-4,1	-7,4	-11,9	-7,4	-3,5	-4,1	-14	-13,4	-18,4
Thời gian (T)	17	18	19	20	22	25	27	28	29	31
Mực nước (m)	86,84	104,09	107,59	115,4	116,72	114,32	111,6	109,11	85,83	104,68
Độ lún (mm)	-24,8	-21,1	-19,5	-16	-15,6	-20,7	-23,1	-24,1	-35,6	-32

3. Kết quả và thảo luận

3.1. Với thực nghiệm số 1

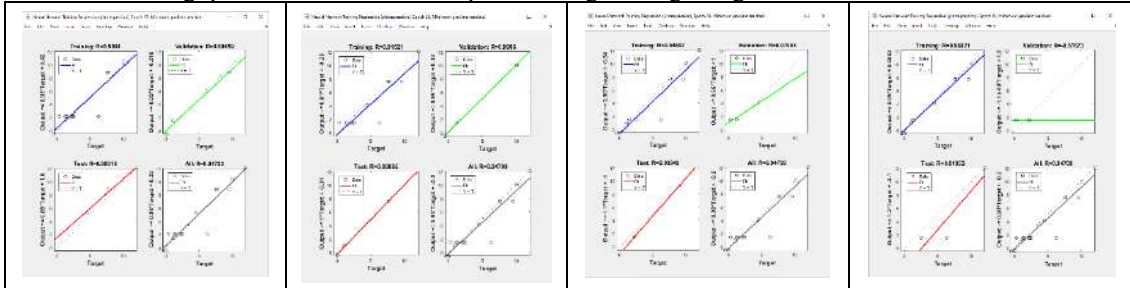
Quy trình huấn luyện của các thuật toán là như nhau, chỉ khác ở phần lựa chọn thuật toán. Kết quả dự báo được lấy sau khi huấn luyện nhiều lần, khi hệ số hồi quy $R > 0,9$ thì cho là đạt để lấy kết quả, kết quả dự báo được trình bày trong Bảng 4.

Bảng 4. Kết quả dự báo của 6 thuật toán huấn luyện

Thuật toán huấn luyện	TrainLM	TrainGDX	TrainRP	TrainSCG	TrainBFG	TrainOSS
Dự báo (mm)	12,10	12,03	12,10	12,08	12,10	11,70
Độ lệch (mm)	1,2	1,27	1,2	1,22	1,2	1,6

Nhận xét: Kết quả dự báo của các thuật toán trên đối với tập mẫu đo lún thủy điện Hòa Bình sai khác nhau không nhiều, kết quả huấn luyện đều có hệ số tương quan tốt.

Sử dụng thuật toán huấn luyện TrainBFG và huấn luyện với 4 lần, sau mỗi lần đều lấy kết quả dự báo, kết quả phân tích hệ số tương quan như Hình 9 và kết quả dự báo ghi trong Bảng 5.



Hình 9. Kết quả phân tích hệ số tương quan qua từng lần huấn luyện

Bảng 5. Kết quả dự báo sau từng lần huấn luyện của thuật toán TrainBFG

Thuật toán huấn luyện TrainBFG	Lần 1	Lần 2	Lần 3	Lần 4
Dự báo (mm)	12,1	12,1	12,1	12,1

Nhận xét: Qua 4 lần huấn luyện, mặc dù kết quả huấn luyện (hiệu suất, trạng thái, hệ số hồi quy) trong từng lần là khác nhau nhưng giá trị dự báo là giống nhau. Như vậy, chúng ta thấy rằng sau mỗi lần huấn luyện giá trị của mô hình là khác nhau, thể hiện rõ nhất là hệ số tương quan.

3.2. Với thực nghiệm số 2

Thực nghiệm này có thành phần số liệu đầu vào nhiều hơn 1 lớp, kết quả tính toán dự báo của 3 thuật toán huấn luyện khi có hệ số tương quan $R > 0,9$ trình bày trong Bảng 6.

Bảng 6. Kết quả dự báo với 3 thuật toán huấn luyện

Thuật toán	Dự báo (mm)	Số liệu đo (mm)	Độ lệch (mm)
TrainLM	-30,62	-32,00	1,38
TrainGDX	-31,82	-32,00	0,18
TrainRP	-32,82	-32,00	-0,82

Nhận xét:

- Kết quả dự báo theo từng thuật toán huấn luyện đối với tập mẫu này là không giống nhau;
- Giá trị dự báo cũng tương đối sát với giá trị thực đo nhưng lệch nhau khá nhiều với từng thuật toán.

4. Kết luận

Qua tìm hiểu và nghiên cứu một số thuật toán huấn luyện mạng trí tuệ nhân tạo, thực nghiệm với số liệu đo lún xấp xỉ 20 chu kỳ của thủy điện Hòa Bình và dự báo cho chu kỳ tiếp theo, nhóm tác giả có một số kết luận ban đầu như sau:

- Với tập mẫu ít lớp đầu vào thì các thuật toán huấn luyện cho kết quả đầu ra như nhau, ngược lại với tập mẫu nhiều lớp đầu vào, sẽ có kết quả khác nhau;
- Giá trị tương quan của thuật toán huấn luyện sẽ tương ứng với độ chính xác của mô hình. Giá trị này càng gần 1 thì mô hình càng gần với giá trị tập mẫu đầu vào.
- Trên thực tế, để ứng dụng tốt mạng nơron nhân tạo để xây dựng mô hình và làm công tác dự báo thì số tập mẫu càng nhiều càng tốt, số lượng lớp đầu vào sẽ làm tăng mức độ học của thuật toán, từ đó sẽ cho kết quả tốt hơn và là cơ sở cho người dùng lựa chọn thuật toán khi huấn luyện mạng BP.

Tài liệu tham khảo

Hu Wusheng, 2006. *Lý thuyết và ứng dụng mạng thần kinh nhân tạo trong công trình*. Nhà xuất bản Trắc Hội Bắc Kinh, Trung Quốc (tiếng Trung Quốc).

Phạm Quốc Khánh, Nguyễn Quốc Long, 2015. Nghiên cứu ứng dụng phương pháp lọc Kalman dự báo chuyển dịch đứng bề mặt. *Tạp chí Công nghiệp mở*, số 5, 23-25.

- Phạm Quốc Khánh, Nguyễn Việt Hà, 2015. Ứng dụng phương pháp tự hồi quy trong dự báo lún công trình. *Tạp chí Công nghiệp mỏ*, số 1, 57-60.
- S. Narayanakumar, K.Raja, 2016. A BP artificial neural network model for earthquake magnitude prediction in Himalayas, India, *Scientific Research Publishing*.
- George Pantazis, Eleni-Georgia Alevizakou, 2013. The use of artificial neural networks in predicting vertical displacements of structures, *International Journal of applied Science and Technology*.
- Nguyễn Thị Kim Thanh, 2021. Nghiên cứu giải pháp nâng cao hiệu quả xử lý số liệu quan trắc độ lún tuyến đập công trình thủy điện. Luận án tiến sĩ kỹ thuật. Đại học Mỏ - Địa chất.
- Nguyễn Đình Thúc, 2000. Trí tuệ nhân tạo - Mạng nơron - Phương pháp và ứng dụng. Nhà xuất bản Giáo dục.
- Cao Đình Trọng và nnk., 2011. Ứng dụng mạng nơron trong dự báo độ lớn (magnitude) động đất khu vực Tây Bắc Việt Nam, *Tạp chí Các khoa học về Trái đất*.
- Ke Fuyang, Li Yayun, 2014. The forecasting method of landslides based on improved BP neural network, *Geotechnical Investigation and Surveying*.
- Xiaoyu Wang et al., 2017. Study on MPGA-BP of Gravity Dam Deformation Prediction, *Mathematical Problems in Engineering*.
- J. Zupan, J.Gasteiger, 1993. *Neural Networks for chemists: an introduction*, John Wiley & Sons, Inc.
- Zhang Zhenglu, 2007. Deformation monitoring analysis and prediction for engineering construction. Surveying and Mapping Publishing House, pp. 174-209.
- Prashant Rao, 2016, <https://www.mathworks.com/content/dam/mathworks/mathworks-dot-com/images/events/matlabexpo/in/2016/whats-new-in-matlab-simulink-r2016a.pdf>.
- Supported and Compatible Compilers - Release 2016a, 2016, https://www.mathworks.com/content/dam/mathworks/mathworks-dot-com/support/sysreq/files/SystemRequirements-Release2016a_SupportedCompilers.pdf.

ABSTRACT

Research on some artificial intelligence training algorithms in predicting structural displacement

Pham Quoc Khanh^{1*}

¹Hanoi University of Mining and Geology

Artificial neural network (ANN) has been widely applied in geodesy, ANN Back-Propagation (BP) is used more in modelling of displacement and prediction. To train the BP network, there are many different algorithms, each algorithm has its own advantages and disadvantages. The article used 02 sample sets of subsidence monitoring data of the Hoa Binh hydropower project with different amount of samples and input data which were trained by 06 and 03 algorithms such as Levenberg-Marquardt (Trainlm), variable learning rate (Traingdx), Quasi-Newton, ... Then compare the forecasting results with actual measurements. The results showed that the smaller sample set (17 samples) without many input layers (only time layer and subsidence layer measured by class 2 levelling equipment), the prediction value (one value) is the same when using different algorithms. With the same sample set, when using an algorithm, during the training sessions, although the evaluation criteria such as training performance, training status and correlation analysis are different, the prediction results are still similar. With the bigger sample set (20 samples) with many input layers (time, water level, subsidence), using different algorithms create different prediction results. This proves that a number of samples and input layers affect learning results of algorithm. Therefore, in modelling displacement prediction, it is necessary to test some algorithms to get better forecasting results and closer to the actual value.

Keywords: ANN, BP, displacement prediction.

Công nghệ máy bay không người lái và thuật toán bán tự động trong việc xác định hệ thống khe nứt chính trên mỏ đá ốp lát

Phạm Văn Việt^{1,2*}, Nguyễn Anh Tuấn^{1,2,3}, Trần Trung Anh^{1,3}, Phạm Văn Hòa^{1,2}

¹Trường Đại học Mở - Địa chất

²Nhóm Nghiên cứu mạnh ISRM, Trường Đại học Mở - Địa chất

³Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

TÓM TẮT

Các mỏ đá ốp lát ngày càng được mở rộng khai thác do giá trị kinh tế đem lại lớn hơn so với khai thác đá làm vật liệu xây dựng thông thường. Nhưng công tác khai thác đá ốp lát phụ thuộc rất nhiều vào việc xác định sự phân bố khe nứt trong khối đá, đặc biệt là hệ khe nứt chính trong mỏ, làm cơ sở đưa ra hướng khai thác, công nghệ khai thác và quy mô khai thác phù hợp với điều kiện mỏ. Hiện nay, việc xác định khe nứt mới thu thập bằng phương pháp thủ công sử dụng địa bàn và thước đo gây khó khăn cho việc đánh giá tổng thể khe nứt trên một khu vực do số lượng khe nứt thu thập ít, một số khu vực không thể tiếp cận đo, rủi ro đá rơi. Do đó, bài báo đã đề xuất sử dụng công nghệ máy bay không người lái và thuật toán bán tự động trong việc xác định hệ thống khe nứt chính cho mỏ đá ốp lát. Bài báo đã phân tích sử dụng dữ liệu đám mây điểm từ thu thập bằng công nghệ máy bay không người lái kết hợp với phân tích đám mây điểm bằng thuật toán bán tự động được viết mã trên phần mềm Matlab từ việc hình thành mặt khe nứt từ đám mây điểm trên số lượng điểm, thống kê các mặt khe nứt, phân nhóm khe nứt dựa trên giá trị góc hợp bởi các hệ khe nứt cùng với góc phân tán khe nứt đưa ra để tìm kiếm các mặt khe nứt, phân nhóm xác định hệ khe nứt chính và xác định sự phân bố khoảng các khe nứt trong hệ khe nứt. Bài báo đã áp dụng phân tích thử cho dữ liệu đám mây điểm tường và mái một phần của điện thái hòa (cung đình Huế) với độ chính xác về góc phương vị và góc dốc 0,7-4%, sai số về khoảng cách 2%, tiến hành áp dụng xác định hệ khe nứt chính cho mỏ đá ốp lát Nam Đông, tỉnh Thừa Thiên Huế và đã xác định được 7 hướng khe nứt chính trong mỏ và phân bố khoảng cách giữa chúng.

Từ khóa: Mỏ đá ốp lát, UAV, hệ khe nứt chính, mỏ đá ốp lát Nam Đông.

1. Giới thiệu

Mỏ đá ốp lát là loại đá tự nhiên bao gồm các nhóm đá như magma, trầm tích, biến chất không chứa các khe nứt được khai thác và chế biến tới các kích thước, hình dạng, màu sắc và độ bóng khác nhau. Đá ốp lát có nhiều ứng dụng trong thực tế như làm đá ốp tường, vỉa hè, bếp ăn, làm tượng,... Đá ốp lát phải đảm bảo về thạch học, khoáng vật và cơ học, đặc biệt là không chứa nứt nẻ trong đá (Morales Demarco, Oyhançabal, Stein & Siegesmund, 2013; Mosch, Nikolayew, Ewiak, & Siegesmund, 2011; Taboada, Vaamonde, & Saavedra, 1999).

Trên thế giới, yêu cầu cho đá khối đang tăng do các hoạt động xây dựng sử dụng đá tự nhiên. Cụ thể, khối lượng đá khối cho hoạt động ốp lát, các ứng dụng khác, trang nghệ thuật và vật liệu xây dựng chiếm tương ứng 35%, 25%, 20%, 16% và 9%. Do đó, các mỏ đá khối như đá gabro, đá hoa cương và đá granit tăng về quy mô và công suất. Các nước trên thế giới sản xuất nhiều đá ốp lát phải kể đến Trung Quốc, Ấn Độ, Thổ Nhĩ Kỳ, Brazil, Ý, Iran, Tây Ban Nha và Ai Cập. Hiệp hội xây dựng thế giới chỉ ra rằng thế giới sẽ sử dụng đá làm ốp lát chiếm nhiều hơn 30% và tăng 2,5 lần suốt 15 năm qua và xu thế sẽ tăng cao do yêu cầu cao và công nghệ đầu tư hiện đại và giá cả cạnh tranh với gạch ốp ceramic (Stonedead, 2021).

Qua tất cả các tài liệu đánh giá chỉ ra rõ ràng rằng nhân tố ảnh hưởng tới khả năng thu hồi đá ốp lát bao gồm khe nứt, cơ lý, độ bóng, kích thước. Một trong chúng, khe nứt là nhân tố chính trong giảm thu hồi đá ốp lát đã được chỉ ra bởi một vài tác giả. (Palmström, 2001; Sousa, 2010; Sousa, Oliveira & Alves, 2016). Hiện nay việc tính toán thu hồi đá ốp lát thông qua phân bố các hệ khe nứt chính được đo bằng phương pháp khảo sát theo đường và khảo sát theo diện tích trên bề mặt, trong lỗ khoan và các moong khai thác thử. Nhưng việc xác định hệ khe nứt chính vẫn còn khó khăn vì dữ liệu thu thập khe nứt còn ít mà mới chỉ dựa vào việc xác định thủ công thông qua đo bằng địa bàn, hạn chế trong điều kiện địa hình phức tạp không thể tiếp cận để đo trực tiếp (Phạm Văn Việt và nnk., 2020). Các thuật toán phân nhóm hệ khe nứt vẫn chỉ dựa vào việc phân nhóm các khe nứt đơn lẻ được thu thập bằng thủ công do số lượng khe nứt hạn chế như phương pháp phân nhóm hệ khe nứt theo thuật toán fuzzy đã được sử dụng trong phần mềm Dips của hãng rocscience (Hammah & Curran, 1998) và thuật toán nhóm các hệ khe nứt dựa trên hệ số phân tán khe nứt trong hệ (Nguyen, Merrien-Soukatchoff, Vinches & Gasc-Barbier, 2016). Việc thu thập khe nứt mang tính tự động, trên diện tích

* Tác giả liên hệ

Email: phamvanviet@humg.edu.vn

lớn nhằm đánh giá chính xác các hệ khe nứt chính trên mô từ dữ liệu khe nứt đủ lớn, phân bố diện rộng là vấn đề cần phải giải quyết.

Hiện nay, với sự phát triển của công nghệ máy bay không người lái UAV (Unmanned Aerial Vehicle) kết hợp với công nghệ định vị vệ tinh theo thời gian thực GNSS (Global Navigation Satellite System) có thể tiếp đo vẽ bề mặt địa hình mọi vị trí với độ chính xác cao và độ chi tiết lớn, có thể thu thập dữ liệu đám mây điểm PCL (Point Cloud). Từ dữ liệu PCL thuật toán bán tự động sẽ nhận diện ra các mặt khe nứt chứa dữ liệu điểm, đưa ra các hệ khe nứt chính, khoảng cách giữa các khe nứt và sự phân bố của các khe nứt trong không gian. Đây là một thực tế quan trọng trong đánh giá nứt nẻ của các mỏ đá ôp lát hiện nay, khắc phục các hạn chế của phương pháp thủ công và phù hợp với xu thế của thế giới.

2. Phương pháp

Xác định khe nứt từ dữ liệu PCL trải qua 3 giai đoạn gồm (tính toán hình thành các mặt phẳng, phân tích thống kê các điểm trên mặt phẳng, phân tích phân nhóm các khe nứt).

2.1. Tính toán hình thành các mặt phẳng

2.1.1. Tìm kiếm các điểm gần nhất

Thuật toán thực hiện tìm kiếm các điểm bên cạnh sử dụng 2 cách tiếp cận khác nhau: cố định khoảng cách hoặc cố định số lượng điểm bên cạnh. Một vài lỗi có thể tăng khi sử dụng cách cố định khoảng cách do sự không đồng nhất của mật độ đám mây điểm (Lato, Diederichs & Hutchinson, 2010). Trong thuật toán ưu tiên lựa chọn cố định số lượng các điểm sử dụng trong nghiên cứu.

Hàm knnsearch trong Matlab sử dụng một thuật toán cao cấp nhanh và hiệu quả tìm các điểm bên cạnh (Friedman, Bentley & Finkel, 1977). Các điểm gần nhất được tính toán bằng hàm knnsearch và khoảng cách Euclid. Tiếp theo, mỗi điểm trong tập điểm ban đầu P_i , một tập Q_i các điểm gần nhau hình thành. Như vậy, xác định k điểm gần nhau nhất cho mỗi điểm trong tập 3D đám mây điểm. Bước tiếp theo kiểm tra xem tập với $k + 1$ điểm có hình thành một mặt phẳng hay không.

2.1.2. Kiểm tra cùng mặt phẳng

Trên thực tế phương pháp xem xét mọi điểm và các điểm gần nhất như là một đối tượng mặt phẳng. Như tập Q_i được hình thành qua bước trên có hình thành mặt phẳng không. Quá trình thử được thực hiện tính toán phương vị. Nếu tập điểm Q_i là mặt phẳng, quá trình sẽ tiếp tục. Ngược lại, tập ghép Q_i sẽ bị loại bỏ.

Quá trình kiểm tra cùng mặt phẳng dựa trên phân tích thành phần chính. Hàm matlab princomp cho phép thực hiện phân tích xác định các giá trị riêng ($\lambda_1, \lambda_2, \lambda_3$) và vector (V_1, V_2, V_3). Phần biến chiếm bởi k thành phần H_k xác định theo công thức (Rencher & Christensen, 2012):

$$H_k = \frac{\sum_i^k l_i}{\sum_1^3 l_i} \quad (1)$$

Giả định rằng phần đánh giá được xác định trong mặt phẳng (π) ở không gian R^3 , sẽ có 2 chiều có thể giải thích phần chính của dữ liệu. Kích thước thứ 3 sẽ giải thích sự thể hiện lỗi trong dữ liệu. Nếu bề mặt không phẳng hoặc lỗi do đó sẽ đủ liên quan, kích thước thứ 3 sẽ yêu cầu mức độ quan trọng tới 2 chiều kia.

Để biết 1 tập hợp điểm có là mặt phẳng hoặc không, các thông số dung sai (η) được xác định theo công thức sau:

$$h = \frac{l_3}{l_1 + l_2 + l_3} \quad (2)$$

Thông số dung sai (η_{\max}) là độ lệch cho phép lớn nhất của một tập điểm, sao cho mặt phẳng tập điểm xem xét hợp lý như một mặt phẳng. Giá trị η_{\max} hình thành thông qua phân tích độ nhạy với dữ liệu thực dưới các điều kiện kiểm tra nhất định. Chấp nhận rằng nếu một tập các thành phần chính có 80% hoặc nhiều hơn sẽ thể hiện dữ liệu hợp lý (Rencher & Christensen, 2012). Vậy đề xuất giá trị dung sai lớn nhất η_{\max} là 20%, khi các trường hợp có $\eta > \eta_{\max}$ thì tập điểm sẽ không đủ điều kiện hình thành mặt phẳng và bị loại bỏ.

2.1.3. Điều chỉnh mặt và tính toán vector pháp tuyến (PCA)

Khi phát hiện những tập hợp điểm của các điểm cùng mặt phẳng, bước tiếp theo là tính toán mặt phẳng phù hợp nhất với những tập điểm này. Công thức giải tích thể hiện trong công thức dưới, trong đó A, B, C là 3 thành phần của vector pháp tuyến đơn vị (PCA) với mặt phẳng và giá trị D đưa khoảng cách vuông góc từ gốc tới mặt phẳng.

$$Ax + By + Cz + D = 0 \quad [A, B, C, D] \in R \quad (3)$$

Khi PCA được tính toán trong bước trước, mặt phẳng xác định rõ ràng theo vector $V_3 = (A, B, C)$.

Tóm lại, trong phần này tính toán xác định hướng của các tập hợp điểm. Bước tiếp là tính toán các hướng đại diện nhất của các điểm 3D và k các điểm gần nhất.

2.2. Phân tích thống kê các điểm khe nứt lên mặt phẳng

Phương pháp tiếp sau là dựa trên các vector pháp tuyến song song kết hợp với các điểm, sẽ xem xét các điểm Q_i kết hợp với một điểm (P_i) mà thuộc một khe nứt xác định bởi một mặt phẳng chưa biết π . Nếu α là mặt phẳng hợp nhất với mặt phẳng Q_i thì hướng của mặt α và mặt π được cho là gần. Tương tự, xem xét một tập khác Q_j với cùng kích thước với Q_i kết hợp với 1 điểm P_j là thành phần của P . Tập hợp điểm còn là một phần mặt π và có thể tính toán mặt phẳng phù hợp nhất β và 3 mặt (π, α, β) có một hướng gần.

Phân tích thống kê bằng đồ họa stereo graphic của các điểm nhận mặt phẳng để xác định các hệ khe nứt chính:

- Tính toán vector pháp tuyến cho mỗi mặt và chuyển đổi đưa về biểu diễn đồ họa nổi stereographic (Ramsay, Huber & Lisle, 2000);
- Tính toán mật độ điểm nhân mặt phẳng cho mỗi vùng trên đồ họa;
- Tính điểm lớn nhất.

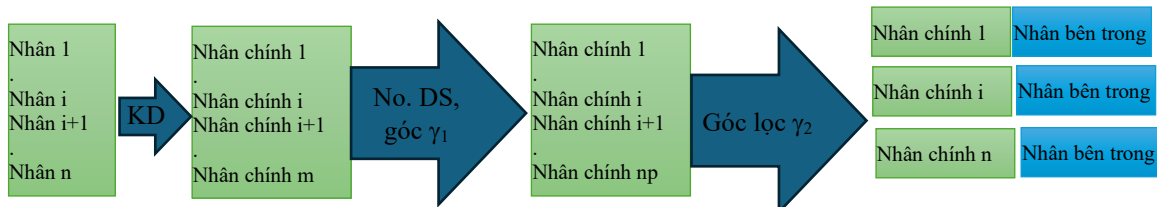
2.2.1. Ước tính mật độ (KDE)

Sử dụng ước tính mật độ theo phương pháp Kernel (KDE) là cách ước tính mật độ thống kê không thông số của các biến ngẫu nhiên. Nghiên cứu đã chứng minh rằng để ước tính mật độ không thông số nhiều biến thì sử dụng KDE tốt hơn sử dụng đồ thị tần số truyền thống (Silverman, 2018). Phương pháp thực hiện bằng sử dụng hàm matlab kde2d bởi Gaussian kernel (Botev, Grotowski & Kroese, 2010). Thực hiện trên một mã code cho phép tính toán tự động chiều rộng của kernel và tính toán mật độ của chúng.

2.2.2. Xác định bán tự động các hệ khe nứt

Trong bước này, phương pháp gán một hướng chính tới mọi điểm trong đám mây điểm. Khi phương pháp phát hiện không đại diện bởi hướng chính chúng sẽ không được gán Sơ đồ tóm tắt các bước thể hiện trong hình bên dưới. Bình thường phân tích hàm mật độ thể hiện nhiều điểm nhất, nhưng chỉ vài là nhân chính, bởi vì sự tồn tại của các lỗi đọc và các điểm đơn của các mặt cong đưa ra sự phân tán các điểm nhân mặt phẳng. Vậy, 2 yêu cầu cần gán để giám sát cho phép xác định các điểm như các nhân mặt phẳng chính:

- Điều kiện số 1 (Nón tìm kiếm): Lựa chọn một giá trị nhất định (γ_1) sao cho góc hình thành bởi hai vector pháp tuyến chính phải lớn hơn giá trị này.
- Điều kiện 2 (Lọc số điểm nhân chính): Do có số lượng lớn nhất các hệ khe nứt (n_p) hình thành. Hệ thống lọc và chấp nhận số lượng mặt hệ khe nứt chính có mật độ cao hơn.



Hình 1. Sơ đồ trình tự xác định hệ khe nứt chính.

Bước tiếp theo bao gồm chia đám mây điểm bằng gán các nhân cho mỗi điểm của đám mây điểm theo các hệ khe nứt gần nhất đã được xác định. Đối với mọi điểm, tìm kiếm các khe nứt mà góc nhỏ nhất giữa các vector pháp tuyến của điểm nhân và vector pháp tuyến của khe nứt chính được gán. Một giới hạn sau đó được xác định để giới hạn theo giá trị cho phép (γ_2).

Ở phần này, xác định hướng của hệ khe nứt chính và các điểm nhân xung quanh. Hướng tương ứng được gán cho mỗi điểm nhân phụ thuộc vào điểm và hướng xung quanh của chúng. Sau đó loại bỏ các điểm nhân có hướng khác với hướng điểm nhân chính (góc lớn hơn góc γ_2). Khi chúng ta khai thác các điểm thuộc các hệ khe nứt, tiếp theo xác định các phân nhóm trong các hệ khe nứt chính.

2.3. Phân tích các nhóm khe nứt

2.3.1. Phân nhóm hệ khe nứt

Đối với phân nhóm các hệ dữ liệu 3D thuật toán quét dựa trên mật độ với độ nhiễu (Density-Based Scan Algorithm with Noise (DBSCAN) (Ester, Kriegel, Sander & Xu, 1996). Thuật toán phân nhóm cho phân loại các lớp trong cơ sở dữ liệu đã được chứng minh trong các ứng dụng hiệu quả xử lý các đám mây điểm LiDAR (Tonini & Abellan, 2014). Thuật toán yêu cầu các thông số đầu vào như:

- ϵ là khoảng cách lớn nhất giữa 2 điểm để xem xét chúng như lân cận.
- Min-pts là số lượng nhỏ nhất các điểm lân cận điểm p để xem xét q như là một điểm lõi. Sự khác nhau lớn về mật độ có thể ảnh hưởng đến phương pháp tự động (Ester et al., 1996; Lato et al., 2010). Ester et al. (1996)

gợi ý thiết lập thông số min-pts là 4. Thêm rằng, giá trị ε nên được xem xét khoảng cách của điểm lân cận thứ 4 tới mỗi hệ khe nứt. Phương pháp tính toán tất cả khoảng cách lân cận thứ 4 cho tất cả các điểm được chấp nhận của mặt chính. Giá trị ε được xem xét như giá trị trung bình, lớn nhất hoặc giá trị đại diện thống kê khác. Giá trị lớn nhất không hợp lý bởi vì sự tồn tại của giá trị nhiễu. Khi phương pháp thực hiện với các dữ liệu thô, số lượng các khoảng cách đủ cao để xem xét sự phân phối chuẩn. Vậy giá trị ε đề xuất là trung bình các khoảng cách lân cận thứ 4 cộng với 2 độ lệch chuẩn.

Phân tích các nhóm thực tế có thể tìm một số lượng lớn các nhóm nhỏ. Người sử dụng chỉ ưu tiên các nhóm lớn. Vậy phương pháp đề xuất tùy chọn loại bỏ các nhóm nhỏ bằng lựa chọn giới hạn số lượng các điểm trên một nhóm. Thực vậy, chỉ các nhóm có kích thước bằng số điểm cân bằng hoặc cao hơn số điểm giới hạn trên một nhóm được thể hiện.

Ở mục này, tính toán các nhóm không gian của mỗi hệ khe nứt. Bước tiếp theo tính toán công thức mặt phẳng của các khe nứt này để xác định thông số toán học các mặt phẳng khe nứt.

2.3.2. Tạo ra mặt khe nứt trong hệ (PCA)

Sinh mặt phẳng thực hiện như sau:

- Một tập hợp các điểm thuộc một hệ khe nứt I - R_i
- Một tập hợp các điểm thành viên của một nhóm j thay thế một tập hợp $R_i - R_{ij}$
- Xác định mặt phù hợp nhất R_{ij} mà công thức mặt phẳng có thể được xác định theo:

$$A_{ij}x + B_{ij}y + C_{ij}z + D_{ij} = 0 \quad (4)$$

Phân tích thành phần chính (PCA) tới k điểm P_{ijk} có tọa độ $(x_{ij}^k, y_{ij}^k, z_{ij}^k)$ và là thành viên của tập hợp không rỗng R_{ij} đặt được vector chính V_3 , Các thông số A_{ij} , B_{ij} , C_{ij} được tính toán theo công thức (4). Tính toán những thông số này sử dụng vector pháp tuyến của các hệ khe nứt chính. Giá trị độc lập D_{ij} của công thức mặt phẳng được tính theo phương pháp bình phương tối thiểu theo công thức:

$$D_{ij} = -\frac{A_{ij}}{n} \sum_{k=1}^n x_{ij}^k - \frac{B_{ij}}{n} \sum_{k=1}^n y_{ij}^k - \frac{C_{ij}}{n} \sum_{k=1}^n z_{ij}^k \quad (5)$$

trong đó: n là kích thước nhóm R_{ij} .

2.3.3. Kiểm tra lỗi

Khi công thức mặt R_{ij} được tính toán, rất dễ kiểm tra chất lượng của fitting dữ liệu. Đưa ra rằng $\overline{V_3}=1$, lỗi trùng khớp er_{ij} được xác định như khoảng cách mặt:

$$er_{ij}^k = Ax_{ij}^k + Bx_{ij}^k + Cx_{ij}^k + D \quad (6)$$

Do đó, các lỗi er_{ij} kết hợp với nhóm R_{ij} có thể được xác định bởi công thức (6). Tập hợp er_{ij} phải thỏa mãn 2 đặc trưng: giá trị của modun er_{ij} phải nhỏ nhất (điều này sẽ thỏa mãn như công thức được tính toán bởi phương pháp bình phương tối thiểu và giá trị độ lệch chuẩn $\sigma(er_{ij})$ phải thấp hợp lý.

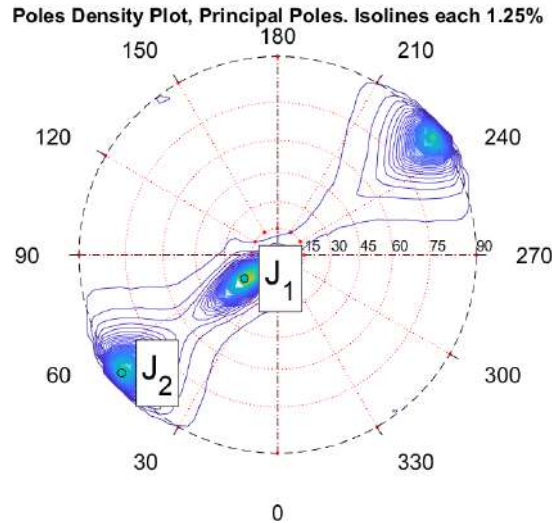
3. Xây dựng thử nghiệm đánh giá xác định hệ mặt

Dữ liệu tiến hành thử nghiệm là bộ dữ liệu đám mây điểm của một phần tường và mái của Điện Thái Hòa, thuộc khuôn viên Hoàng thành ở thành phố Huế, tỉnh Thừa Thiên Huế (Hình 2).

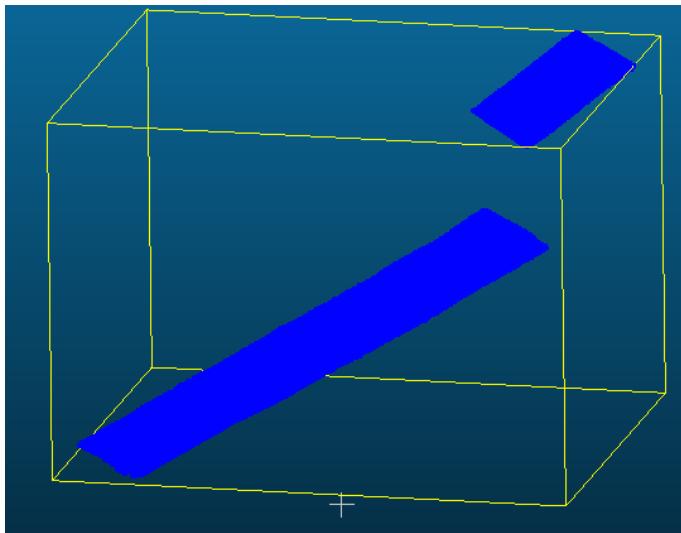


Hình 2. Dữ liệu khảo sát một phần Điện Thái Hòa (phần khung màu đỏ) (a), dữ liệu đám mây điểm (b).

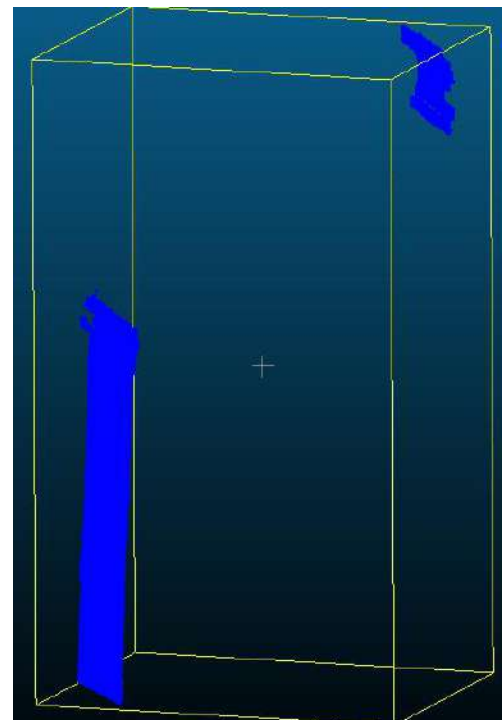
Dữ liệu dưới dạng đám mây điểm gồm có 2 hệ mặt phẳng mái nhà (mái trên và mái dưới) và hệ mặt phẳng tường (tường chân móng và tường trên mái) được đưa vào phần mềm lập trình matlab chạy chương trình đã được viết code dựa trên ngôn ngữ lập trình matlab dạng mã nguồn mở để nhận diện hệ mặt phẳng tường và hệ mặt phẳng mái nhà. Qua quá trình chạy đã xác định được 2 hệ mặt phẳng (gồm tường và mái nhà), thể hiện cụ thể trên Hình 3. Đã phân tích ra được các hệ mặt tường và hệ mặt mái thể hiện cụ thể trên Hình 4 và Hình 5. Giá trị thu được từ kết quả phân tích hệ mặt trên phần mềm và đo trực tiếp thu được dữ liệu Bảng 1.



Hình 3. Sự phân bố hệ mặt phẳng trên đồ họa nổi gồm 2 hệ mặt phẳng.



Hình 4. Hệ mặt mái nhà (gồm mái dưới và mái trên)



Hình 5. Hệ mặt tường nhà (tường chân móng và tường trên mái)

Bảng 1. Bảng so sánh giữa giá trị đo thực và giá trị thu được từ phần mềm.

TT	Thành phần đánh giá	Đo thực tế			Giá trị phân tích phần mềm		
		Góc phương vị, độ	Góc dốc, độ	Khoảng cách, m	Góc phương vị, độ	Góc dốc, độ	Khoảng cách, m
1	Hệ mặt mái nhà	55	24	1	54,6	23,1	0,98
2	Hệ mặt tường	233	90	3	232,8	89,14	3,06

Qua việc so sánh giữa giá trị đo và giá trị thu được từ phần mềm xác định hệ bề mặt thấy rằng, thuật toán tính toán nhận biết các hệ bề mặt đảm bảo độ tin cậy thông qua sai số về góc phương vị là 0,7%, sai số về góc dốc không quá 4%, sai số về khoảng cách 2%, có thể triển khai áp dụng trong việc xác định các hệ khe nứt trong các mỏ đá nói chung và các mỏ đá ốp lát nói riêng đảm bảo độ tin cậy trong thu thập dữ liệu.

4. Nghiên cứu triển khai áp dụng công nghệ UAV, thuật toán bán tự động trong việc xác định hệ khe nứt chính cho mỏ đá ốp lát

Nghiên cứu tiến hành xác định các hệ khe nứt chính tại mỏ đá ốp lát Nam Đông, tỉnh Thừa Thiên Huế. Công tác thu thập dữ liệu đám mây điểm được thực hiện bởi định vị GNSS RTK và máy bay không người lái UAV DJI Phantom 4 pro v2, với thực hiện định vị đo 3 điểm GCP và 2 điểm CP (Hình 6), thực hiện bay bằng phần mềm DJI pilot, xây dựng mô hình bằng phần mềm Pix4Dmapper tạo ra mô hình đám mây điểm Hình 7.

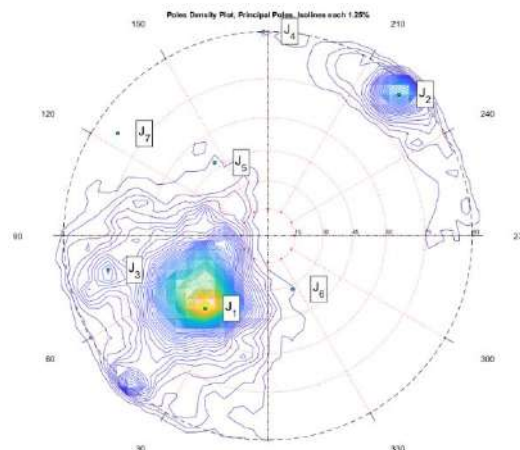


Hình 6. Các điểm GCP và CP trên mỏ để định vị và kiểm tra

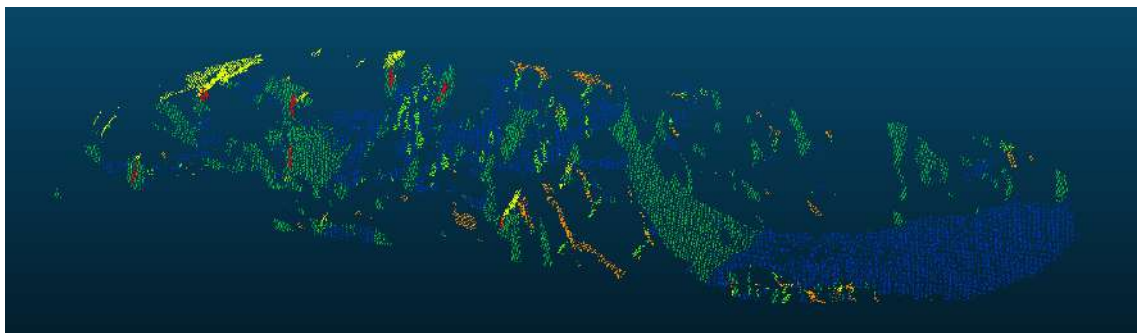


Hình 7. Đám mây điểm khu vực mỏ xác định khe nứt.

Dữ liệu đám mây điểm khu vực xác định khe nứt được đưa vào phần mềm lập trình matlab chạy chương trình đã được viết code dựa trên ngôn ngữ lập trình matlab với các thông số đầu vào số lượng điểm đưa nội suy mặt phẳng là 30, sai số hình thành mặt phẳng là 20%, góc hợp tối thiểu giữa các hệ khe nứt là 30 và số lượng hệ khe nứt sơ bộ đưa ra là 7 hệ khe nứt, góc hình thành hệ khe nứt là 30°. Qua quá trình phân tích trên phần mềm Matlab với code thuật toán đưa ra xác định sự phân bố các hệ khe nứt trên đồ họa nổi Hình 8 và sự phân bố các hệ khe nứt trên Hình 9, phân bố các khoảng cách giữa các khe nứt và số lượng khe nứt thể hiện ở Hình 10 và bảng tổng hợp các thông số hệ thống khe nứt thể hiện Bảng 1.



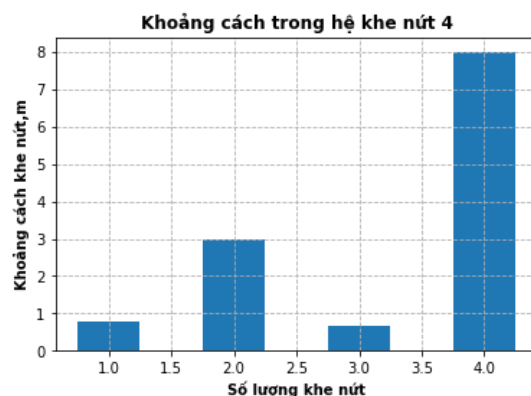
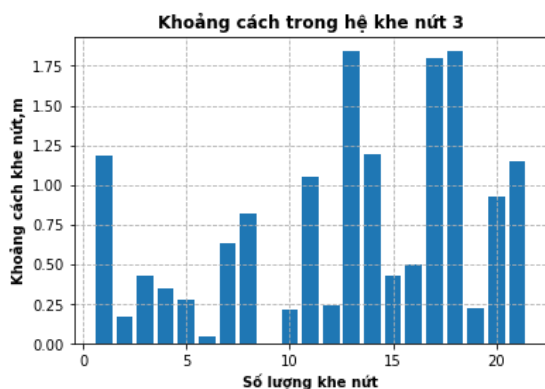
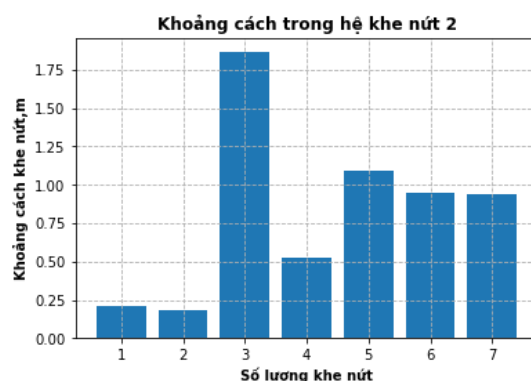
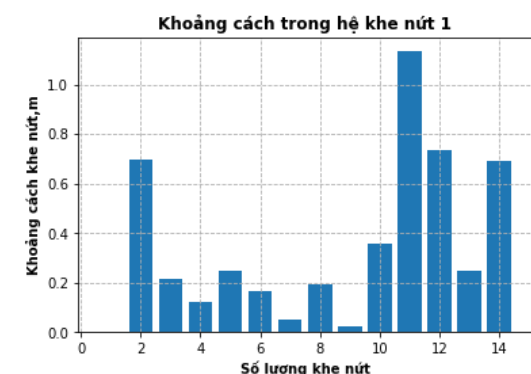
Hình 8. Sự phân bố các hệ khe nứt trên mỏ đá ốp lát Nam Đông

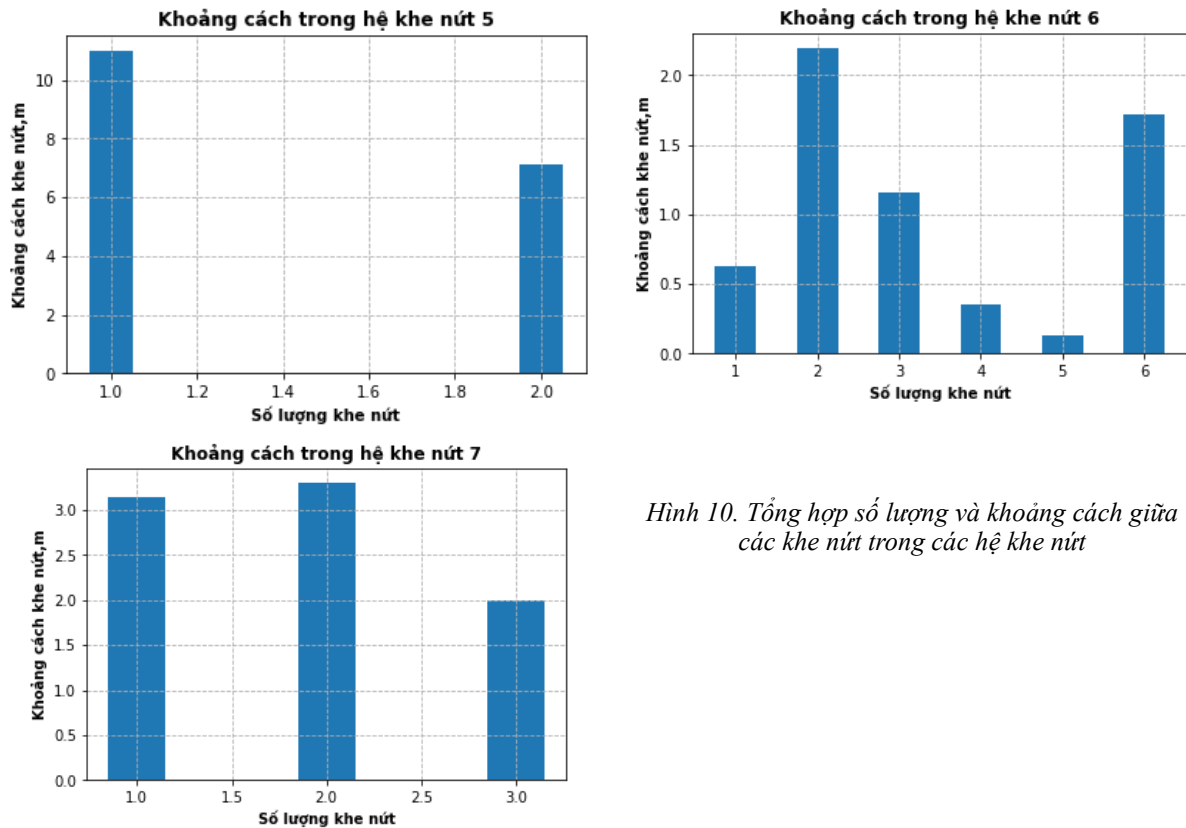


Hình 9. Sự phân bố các hệ khe nứt khu vực khảo sát

Bảng 2. Phân bố thể nằm của các hệ khe nứt trên mỏ đá ốp lát

TT	Hệ khe nứt	Góc phương vị hướng dốc, độ	Góc dốc, độ
1	Hệ khe nứt 1	40,7	50,5
2	Hệ khe nứt 2	222,8	86,5
3	Hệ khe nứt 3	78	77
4	Hệ khe nứt 4	178	90
5	Hệ khe nứt 5	144	48
6	Hệ khe nứt 6	335	32
7	Hệ khe nứt 7	124	83





Hình 10. Tổng hợp số lượng và khoảng cách giữa các khe nứt trong các hệ khe nứt

5. Kết luận

Khe nứt đóng vai trò quan trọng trong việc xác định khả năng thu hồi đá ốp lát cho các mỏ. Giải pháp thu thập thông tin dữ liệu về khe nứt từ các dữ liệu đám mây điểm bề mặt địa hình khu vực mỏ đá ốp lát để xác định các thông số hệ thống khe nứt là hữu ích khi kết hợp với thuật toán bán tự động xác định các hệ khe nứt từ các mặt đá trên bề mặt địa hình bằng phân tích đám mây điểm trên bề mặt địa hình. Qua phân tích thực nghiệm dữ liệu mây điểm tường và mái Điện Thái Hòa so với đo thực tế đảm bảo đủ độ tin cậy để thực hiện phân tích hệ thống khe nứt trên các mỏ đá nói chung và mỏ đá ốp lát nói riêng. Việc thu thập dữ liệu từ đám mây điểm có thể khắc phục được việc thu thập dữ liệu ít, không phân bố đều, không tiếp cận đo những khu vực cao, hiểm trở bằng thủ công.

Tài liệu tham khảo

- Botev, Z. I., Grotowski, J. F., & Kroese, D. P. (2010). Kernel density estimation via diffusion. *Annals of Statistics*, 38(5), 2916-2957. <https://doi.org/10.1214/10-AOS799>.
- Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings - 2nd International Conference on Knowledge Discovery and Data Mining, KDD 1996*, 96(34), 226-231.
- Friedman, J. H., Bentley, J. L., & Finkel, R. A. (1977). An Algorithm for Finding Best Matches in Logarithmic Expected Time. *ACM Transactions on Mathematical Software (TOMS)*, 3(3), 209-226. <https://doi.org/10.1145/355744.355745>.
- Hammah, R. E., & Curran, J. H. (1998). Fuzzy cluster algorithm for the automatic identification of joint sets. *International Journal of Rock Mechanics and Mining Sciences*, 35(7), 889-905. [https://doi.org/10.1016/S0148-9062\(98\)00011-4](https://doi.org/10.1016/S0148-9062(98)00011-4).
- Lato, M. J., Diederichs, M. S., & Hutchinson, D. J. (2010). Bias correction for view-limited lidar scanning of rock outcrops for structural characterization. *Rock Mechanics and Rock Engineering*, 43(5), 615-625. <https://doi.org/10.1007/s00603-010-0086-5>.
- Morales Demarco, M., Oyhançabal, P., Stein, K. J., & Siegesmund, S. (2013). Granitic dimensional stones in Uruguay: Evaluation and assessment of potential resources. *Environmental Earth Sciences*, 69(4), 1397-1438. <https://doi.org/10.1007/s12665-012-2027-y>.

- Mosch, S., Nikolayew, D., Ewiak, O., & Siegesmund, S. (2011). Optimized extraction of dimension stone blocks. *Environmental Earth Sciences*, 63(7), 1911-1924. <https://doi.org/10.1007/s12665-010-0825-7>.
- Nguyen, A. T., Merrien-Soukatchoff, V., Vinches, M., & Gasc-Barbier, M. (2016). Grouping discontinuities in representative sets: influence on the stability analysis of slope cuts. *Bulletin of Engineering Geology and the Environment*, 75(4), 1429-1444. <https://doi.org/10.1007/s10064-015-0822-x>.
- Palmström, A. (2001). Measurement and characterizations of rock mass jointing. *In-Situ Characterization of Rocks - Chapter 2, January 2001*, 1-40.
- Phạm Văn Việt, Nguyễn Anh Tuấn, Phạm Văn Hòa, Trần Đình Bảo, N. D. P. (2020). Nghiên cứu khả năng ứng dụng sóng siêu âm trong việc xác định khe nứt cho khối đá ốp lát. *Tạp chí Khoa học Kỹ thuật mỏ-Địa Chất*, 61(5), 98-106. <https://doi.org/10.46326/JMES.KTTLT2020.08>.
- Ramsay, J. G., Huber, M. I., & Lisle, R. J. (2000). The techniques of modern structural geology: Applications of continuum mechanics in structural geology (Vol. 3). Elsevier.
- Rencher, A. C., & Christensen, W. F. (2012). Wiley series in probability and statistics. *Methods of Multivariate Analysis*, 759-768.
- Silverman, B. W. (2018). Density estimation: For statistics and data analysis. In: *Density Estimation: For Statistics and Data Analysis*. Routledge. <https://doi.org/10.1201/9781315140919>.
- Sousa, L. M. O. (2010). Evaluation of joints in granitic outcrops for dimension stone exploitation. *Quarterly Journal of Engineering Geology and Hydrogeology*, 43(1), 85-94. <https://doi.org/10.1144/1470-9236/08-076>.
- Sousa, L. M. O., Oliveira, A. S., & Alves, I. M. C. (2016). Influence of fracture system on the exploitation of building stones: the case of the Mondim de Basto granite (north Portugal). *Environmental Earth Sciences*, 75(1), 1-16. <https://doi.org/10.1007/s12665-015-4824-6>.
- Stonedead (2021). *The worldwide stone trade business*. <http://www.stonedead.com/stone-trading-around-the-world>.
- Taboada, J., Vaamonde, A., & Saavedra, A. (1999). Evaluation of the quality of a granite quarry. *Engineering Geology*, 53(1), 1-11. [https://doi.org/10.1016/S0013-7952\(98\)00074-X](https://doi.org/10.1016/S0013-7952(98)00074-X).
- Tonini, M., & Abellan, A. (2014). Rockfall detection from terrestrial lidar point clouds: A clustering approach using R. *Journal of Spatial Information Science*, 8(1), 95-110. <https://doi.org/10.5311/JOSIS.2014.8.123>.

ABSTRACT

Unmanned Aerial Vehicle Technology and Semi-automatic Algorithm in Identifying Main Joint Sets for Dimension Stone Quarries

Pham Van Viet^{1,2*}, Nguyen Anh Tuan^{1,2,3}, Tran Trung Anh^{1,3}, Pham Van Hoa^{1,2}

¹Hanoi University of Mining and Geology, Hanoi

²Innovations for Sustainable and Responsible Mining (ISRM) Research Group,
Hanoi University of Mining and Geology, Hanoi

³Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology

Dimension stone quarries have been strongly expanded due to bringing back huge economic value from dimension stones in comparison to quarrying them to produce common construction materials. However, quarrying dimension stone depends really more on defining the distribution of joints in rockmass, especially joint sets, contributing to choosing mining direction, mining technologies and suitable output scale for each mine condition. Nowadays, collecting joints has just been carried out by manual ways with using geological compass, length tape, causing more difficulty to analyze the whole joints in a particular area due to small numbers of joints, some untouched areas, rockfall risk. Therefore, the paper suggests the application of Unmanned Aerial Vehicle (UAV) technology and semi-automatic algorithm identify main joint sets for dimension stone quarries. The study analyzed point cloud data, collected by the UAV technology, with semi-automatic algorithm coded on Matlab software for joint plane formation from point cloud, joint plane statistics, joint plane classification based on the angles between joint sets and the disperse angle to find out joint planes, cluster the planes into joint sets and calculate spacing distribution of each joint set. The algorithm was experimented with a point cloud file of a part of wall and roof of Thai Hoa Palace (Nguyen dynasty palace) with reliable exactness. The knowledge of the study was applied to Nam Dong dimension stone quarry in Thua Thien Hue province. The results of application found out seven main joint sets and their spacing distribution.

Keywords: Dimension stone, UAV, main joint sets, Nam Dong.

Phương trình LiDAR tính tọa độ điểm đo trong công nghệ bay quét laser hàng không

Trần Trung Anh^{1,2*}, Trần Hồng Hạnh^{1,2}, Lưu Hải Âu³, Quách Mạnh Tuấn⁴

¹Khoa Trắc địa - Bản đồ và Quản lý đất đai, Trường Đại học Mở - Địa chất

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

³Viện Khoa học Đo đạc và Bản đồ

⁴Công ty CP Thương mại và Xây dựng QT Miền Bắc

TÓM TẮT

Nghiên cứu này đã áp dụng phương trình LiDAR tính tọa độ điểm đo mặt đất trong công nghệ bay quét laser hàng không. Các kết quả thu được từ cơ sở toán học, hình học để xây dựng phương trình LiDAR, phân tích tính toán dữ liệu bay quét laser hàng không theo số liệu tựa trên vùng thực nghiệm. Từ đó nhận thấy rằng: trong phương trình LiDAR có những thông số cố định tham gia tính toán toàn bộ dữ liệu như thông số kiểm định các giá trị độ lệch (tọa độ, góc xoay) giữa IMU và bộ phát Laser; có thông số dùng chung cho thời điểm phát chùm tia laser như: tọa độ và góc xoay của IMU; thông số riêng của từng tia quét laser là góc xoay để đồng bộ tọa độ về hệ bộ phát laser, còn khoảng cách thì một tia laser phát ra có thể thu được nhiều lớp phản xạ khác nhau theo tầng dữ liệu mặt đất (có thể từ 1 đến 3 hoặc nhiều hơn tùy theo khả năng kỹ thuật của thiết bị quét và đối tượng mặt đất) dẫn đến có nhiều khoảng cách, từ đó có thể tính được tọa độ điểm đo ở nhiều lớp khác nhau. Quy trình của phương trình LiDAR đảm bảo sự thống nhất và tránh sai sót. Thực nghiệm tại phường Tăng Tiến, thị xã Việt Yên, tỉnh Bắc Giang, bay quét laser hàng không với máy bay DJI M350 và bộ quét Zenmuse L2. Các dữ liệu kiểm chứng phương trình LiDAR tựa trên vùng thực nghiệm này cho thấy các công thức trong phương trình LiDAR đảm bảo sự tường minh và chính xác.

Từ khóa: Phương trình LiDAR, quét laser hàng không, điểm đo, hệ thống định vị toàn cầu, bộ đo quán tính.

1. Đặt vấn đề

Hệ thống LiDAR bao gồm một thiết bị quét đo khoảng cách laser và một hệ thống định vị toàn cầu GNSS (Global Navigation Satellite System) cùng bộ định hướng POS (Position and Orientation System) hay hệ thống đạo hàng quán tính INS (Inertial Navigation System), định vị vi phân DGPS (Differential Global Positioning System), bộ đo quán tính IMU (Inertial Measurement Unit). Khi lắp đặt hệ thống LiDAR trên phương tiện bay (có người lái hoặc không người lái) để thực hiện bay quét địa hình được gọi là công nghệ bay quét laser hàng không. Máy quét laser đo khoảng cách từ cảm biến đến mặt đất, trong khi định vị GPS/IMU tích hợp cung cấp thông tin về vị trí và trạng thái của thiết bị (Habib và nnk., 2010). Tọa độ của các điểm LiDAR là kết quả của việc kết hợp các phép đo xuất phát từ từng thành phần hệ thống của nó, cũng như các thông số lắp đặt liên quan đến các thành phần đó. Các nghiên cứu trên thế giới (Vaughn và nnk., 1996; Schenk, 2001; El-Sheimy, 2005) về phương trình LiDAR là cơ sở, tuy nhiên chưa có sự thống nhất về ma trận xoay. Trong nghiên cứu sự thay đổi tỷ số LiDAR theo khoảng cách, tác giả (Bùi Thị Thanh Lan và Đồng Thị Linh, 2016) trình bày phương trình LiDAR nhấn mạnh về ảnh hưởng của khí quyển đến khoảng cách đo laser là một thành tố của phương trình LiDAR. Nghiên cứu khác về LiDAR (Nguyễn Hữu Phương, 2017; Nguyễn Hữu Phương và nnk., 2022; Trần Trung Anh và nnk., 2022; Bùi Ngọc Quý và nnk., 2023) trình bày về phân loại đám mây điểm LiDAR và ứng dụng, chưa đi sâu vào phương trình LiDAR. Nghiên cứu của Nguyễn Trường Xuân và nnk. (2017) trình bày về LiDAR trên thiết bị di động trên mặt đất, phạm vi nhỏ, có đề cập đến công thức LiDAR ở dạng góc nhỏ chưa phải công thức gốc.

Một vấn đề đặt ra là các hệ thống xử lý tức thời của công nghệ LiDAR thường như một hộp đen đóng kín, các tệp dữ liệu đo được mã hóa, kể cả các phần mềm xử lý dữ liệu cũng được tự động hóa gần như toàn bộ (DJI, 2022, 2023). Để làm chủ được công nghệ LiDAR cần nắm rõ bản chất toán học, để từ đó có thể giúp người sử dụng có tính chủ động, các thao tác tại thực địa để hạn chế các sai số và sai sót. Từ công tác thiết kế bay quét LiDAR đến xử lý dữ liệu sản phẩm (Trần Trung Anh và nnk., 2022). Trong nghiên cứu này, trình bày về phương trình LiDAR trong tính toán tọa độ điểm đo của công nghệ quét laser hàng không ở dạng công thức gốc, đầy đủ và thống nhất về ma trận xoay. Nghiên cứu làm rõ cơ sở hình học, toán học làm quy trình, thuật toán để giúp làm chủ công tác bay quét LiDAR tại thực địa cũng như xây dựng ứng dụng xử lý LiDAR sau này.

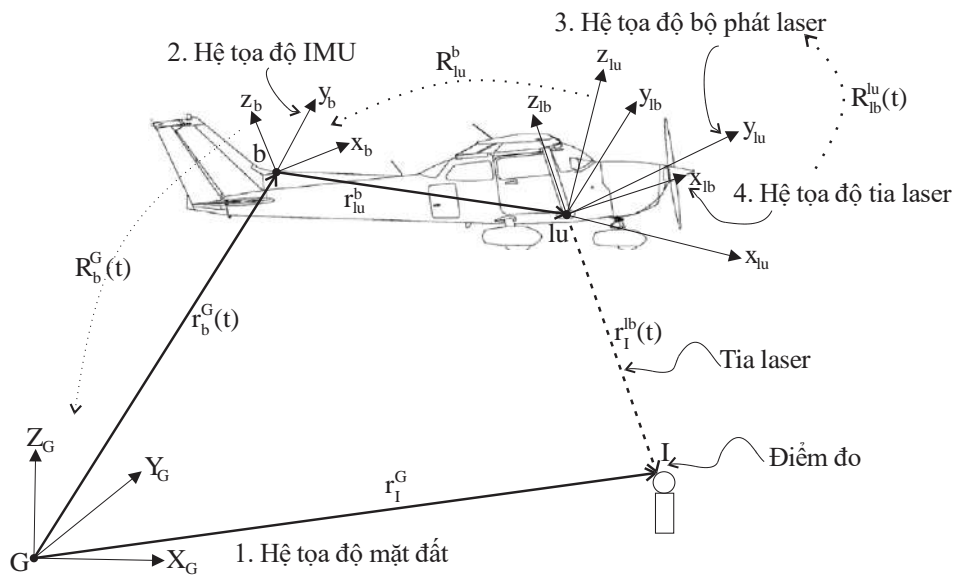
* Tác giả liên hệ

Email: trantrunganh@humg.edu.vn

2. Cơ sở của phương pháp nghiên cứu

2.1. Cơ sở toán học của phương trình LiDAR

Để xây dựng phương trình LiDAR, cần xuất phát từ cơ sở hình học của công nghệ bay quét LiDAR (Ayman Habib, 2008), hình vẽ về mối quan hệ vector được thể hiện trong Hình 1. Mối quan hệ giữa các phép đo hệ thống và các tham số được thể hiện trong phương trình LiDAR ở công thức 6 (Vaughn và nnk., 1996; Schenk, 2001; El-Sheimy, 2005). Như có thể thấy trong Hình 1, vị trí của điểm đo laser I, có thể được tính toán thông qua tổng của các vector sau khi áp dụng các phép quay thích hợp (Habib và nnk., 2010). Trong phương trình LiDAR, những thông số cố định như độ lệch tâm, lệch góc giữa IMU và bộ quét laser cần kiểm định hiệu chuẩn trước khi sử dụng (Mostafa, 2001; Gonçalves và nnk., 2011; Li và nnk., 2019). Những thông số cần đo/tính và dùng chung tại một thời điểm quét của chùm tia laser như vị trí và góc xoay của IMU so với hệ tọa độ mặt đất. Những thông số góc quét dùng chung của 1 tia quét laser cho các lớp phản xạ khác nhau (tầng dữ liệu mặt đất).



Hình 1. Phương trình LiDAR theo tính tổng vector (Vẽ bởi: Trần Trung Anh)

Trong Hình 1:

$r_I^G = \begin{bmatrix} X_I^G \\ Y_I^G \\ Z_I^G \end{bmatrix}$ là tọa độ điểm đo I trong hệ tọa độ mặt đất $G.X_G Y_G Z_G$;

$r_b^G(t) = \begin{bmatrix} X_b^G \\ Y_b^G \\ Z_b^G \end{bmatrix}_{(t)}$ là tọa độ mặt đất của gốc hệ tọa độ IMU tại thời điểm t;

$$R_b^G(t) = \begin{bmatrix} \cos(\alpha)\cos(\beta) & \cos(\alpha)\sin(\beta)\sin(\gamma) - \sin(\alpha)\cos(\gamma) & \cos(\alpha)\sin(\beta)\cos(\gamma) + \sin(\alpha)\sin(\gamma) \\ \sin(\alpha)\cos(\beta) & \sin(\alpha)\sin(\beta)\sin(\gamma) + \cos(\alpha)\cos(\gamma) & \sin(\alpha)\sin(\beta)\cos(\gamma) - \cos(\alpha)\sin(\gamma) \\ -\sin(\beta) & \cos(\beta)\sin(\gamma) & \cos(\beta)\cos(\gamma) \end{bmatrix}_{(t)} \quad (1)$$

là ma trận xoay giữa hệ tọa độ mặt đất và hệ tọa độ IMU của các góc xoay (α, β, γ) tại thời điểm t;

$r_{lu}^b = \begin{bmatrix} x_{lu}^b \\ y_{lu}^b \\ z_{lu}^b \end{bmatrix}$ là vector độ lệch giữa bộ phát laser và hệ tọa độ IMU (độ lệch cánh tay đòn);

$$R_{lu}^b = \begin{bmatrix} \cos(\phi)\cos(\kappa) & -\cos(\phi)\sin(\kappa) & -\sin(\phi) \\ \cos(\omega)\sin(\kappa) + \sin(\omega)\sin(\phi)\cos(\kappa) & \cos(\omega)\cos(\kappa) - \sin(\omega)\sin(\phi)\sin(\kappa) & -\sin(\omega)\cos(\phi) \\ \sin(\omega)\sin(\kappa) - \cos(\omega)\sin(\phi)\cos(\kappa) & \sin(\omega)\cos(\kappa) + \cos(\omega)\sin(\phi)\sin(\kappa) & \cos(\omega)\cos(\phi) \end{bmatrix} \quad (2)$$

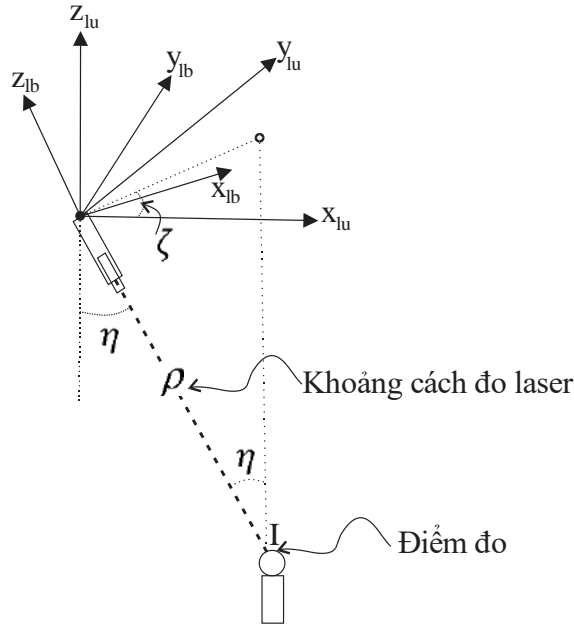
là ma trận xoay giữa hệ tọa độ IMU và hệ tọa độ bộ phát laser (ma trận Boresight của góc ω, ϕ, κ);

$$R_{lb}^{lu}(t) = \begin{bmatrix} \cos(\zeta)\cos(\eta) & -\sin(\zeta) & \cos(\zeta)\sin(\eta) \\ \sin(\zeta)\cos(\eta) & \cos(\zeta) & \sin(\zeta)\sin(\eta) \\ -\sin(\zeta) & 0 & \cos(\eta) \end{bmatrix}_{(t)} \quad (3)$$

là ma trận quay giữa hệ tọa độ laser và hệ tọa độ tia laser của các góc quét gương (ζ, η) đến điểm đo I tại thời điểm t (Hình 2);

$$r_I^{lb}(t) = \begin{bmatrix} 0 \\ 0 \\ -\rho \end{bmatrix}_{(t)} \quad (4)$$

là tọa độ của điểm đo I trong hệ tọa độ tia laser tại thời điểm t (Hình 2);



Hình 2. Tính chuyển tọa độ tia laser sang tọa độ bộ phát laser (Vẽ bởi: Trần Trung Anh)

Từ Hình 2 và công thức (3,4), tia laser của điểm đo I trong hệ bộ phát laser sẽ được tính như sau:

$$r_I^{lu} = R_{lb}^{lu}(t) * r_I^{lb}(t) = \begin{bmatrix} -\rho \cdot \cos(\zeta)\cos(\eta) \\ -\rho \cdot \sin(\zeta)\sin(\eta) \\ -\rho \cdot \cos(\eta) \end{bmatrix} \quad (5)$$

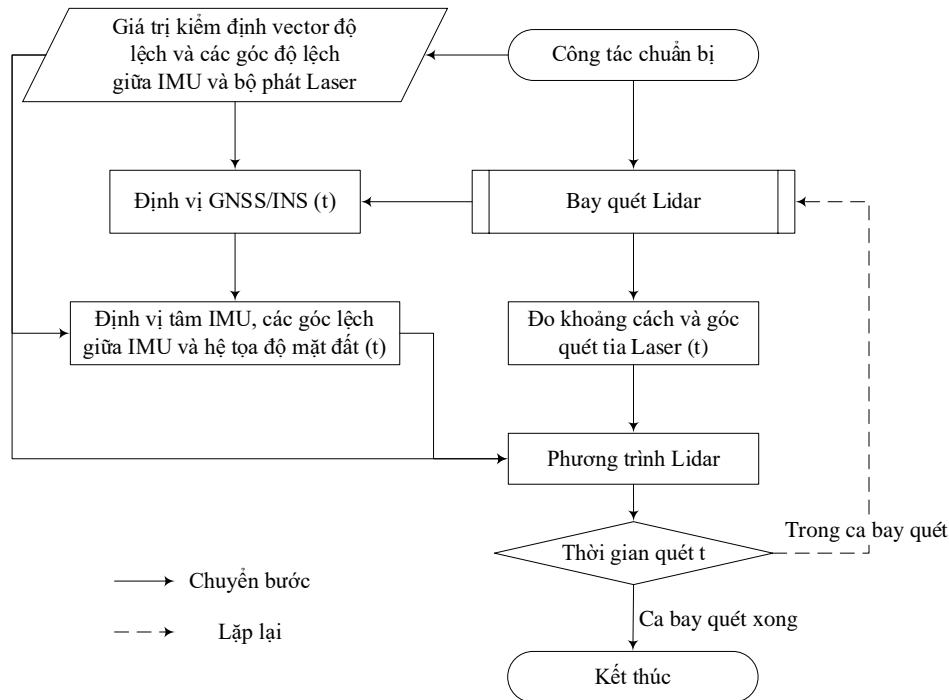
Từ Hình 1 và hình 2, phương trình LiDAR tính tọa độ điểm đo là tổng vector theo công thức (6):

$$r_I^G = r_b^G(t) + R_b^G(t) * r_{lu}^b + R_b^G(t) * R_{lu}^b * r_I^{lu} \quad (6)$$

Từ phương trình LiDAR theo (6) nhóm tác giả triển khai thuật toán để tính tọa độ điểm đo trong công nghệ bay quét LiDAR hàng không. Trong phương trình này yêu cầu cần có thông số kiểm định về độ lệch tọa độ, các góc xoay giữa tâm IMU và bộ quét LiDAR (thông số của nhà sản xuất), đây là thông số cố định chung khi sử dụng thiết bị bay quét; Khi triển khai ngoài thực địa, trong một lần quét ở thời điểm t các thông số dùng chung cho chùm quét là tọa độ tâm IMU, các góc xoay của hệ tọa độ IMU so với hệ tọa độ mặt đất (được định vị chính xác bởi công nghệ GNSS/INS); Tia laser quét cần trả về giá trị gốc của gương xoay để phục vụ tính chuyển về hệ tọa độ bộ phát laser, trong cùng một tia quét laser nếu trả về các lớp tín hiệu tương ứng với lớp điểm đo khác nhau, thì dùng chung giá trị gốc này còn giá trị khoảng cách sẽ khác nhau theo lời giải theo xung hoặc pha (Wang và nnk., 2024).

2.2. Phương pháp nghiên cứu

Sau khi xây dựng phương trình LiDAR, cần thiết phải xây dựng quy trình tính toán phương trình LiDAR. Trên cơ sở phương trình LiDAR (6), các thông số cần cung cấp, thông số đo ở thời điểm đo và các công thức liên quan (1-6), quy trình được trình bày trong Hình 3.



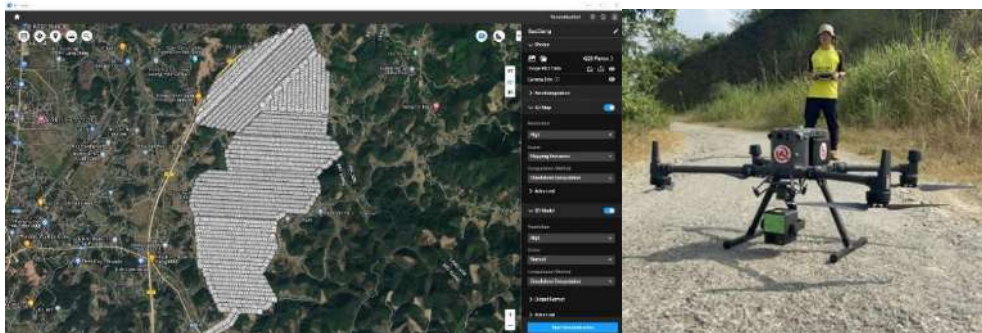
Hình 3. Quy trình của phương trình LiDAR (Trần Trung Anh, Trần Hồng Hạnh, Lưu Hải Âu)

Các bước xử lý lưới được trình bày trong quy trình ở Hình 3, trong đó công tác chuẩn bị gồm các công việc rất quan trọng như chuẩn bị phương tiện bay, pin, các thông số kỹ thuật của thiết bị bay, khả năng hoạt động bình thường,... bao gồm cả khả năng vận hành thiết bị của tác nghiệp viên, giấy cấp phép bay cho khu vực khảo sát và thiết kế bay quét LiDAR. Quy trình phương trình LiDAR rất cần rõ thông số kiểm định vector độ lệch, góc lệch giữa IMU và bộ phát laser vì liên quan đến tính chuyển tọa độ do IMU và bộ phát laser không trùng tâm, thông số này được nhà sản xuất cung cấp (thường đóng kín và ẩn dưới quá trình xử lý). Khi tiến hành bay quét LiDAR tại thực địa, kết nối và định vị GNSS/INS chính xác ở từng thời điểm quét (thông thường đặt ở tần số thu 5 Hz hoặc cao hơn, để đảm bảo tốc độ bay của phương tiện hàng không ảnh hưởng ít nhất đến độ trượt định vị). Từ trị đo tích hợp GNSS/INS cùng với bộ cảm biến IMU về góc, tính tọa độ và góc xoay chính xác của IMU so với hệ tọa độ mặt đất ở thời điểm phóng tia laser (t). Bộ quét laser phóng chùm tia laser ở thời điểm (t), từng tia laser được xác định các góc xoay, tính toán giá trị khoảng cách nhờ tín hiệu phản hồi của các lớp phản xạ của bề mặt ở thời điểm (t). Các thông số này phục vụ tính phương trình LiDAR (6).

Trong quá trình phân tích dữ liệu, nhóm tác giả đã sử dụng các công thức (1-6) để xây dựng thuật toán. Vì các thông số được đóng kín, khó tiếp cận dữ liệu tường minh gốc ở dạng text nên nhóm tác giả giả định giá trị của các thông số để áp dụng phương trình LiDAR tựa trên một vùng bay quét thực nghiệm cụ thể.

3. Thực nghiệm

Khu vực thực nghiệm là phường Tăng Tiến, thị xã Việt Yên, tỉnh Bắc Giang, cách trung tâm thị xã Việt Yên 5 km về phía Đông Nam, cách thành phố Bắc Giang khoảng 7 km theo hướng Đông Bắc theo Quốc lộ 1A mới; phía Tây giáp phường Nénh, phường Hồng Thái; phía Nam giáp huyện Yên Dũng; có vị trí địa lý 21°15'38" vĩ độ Bắc và 106°09'09" kinh độ Đông, diện tích khoảng 2000 ha.



Hình 4. Vị trí thực nghiệm bay quét LiDAR (Ảnh: Quách Mạnh Tuấn)

Nhóm tác giả tiến hành bay quét LiDAR bằng máy bay M350 và bộ quét Zenmuse L2 của hãng DJI, trong quá trình quét có định vị tâm chụp chính xác bằng phương pháp GNSS-RTK (Hình 4). Xử lý dữ liệu trên phần mềm DJI Terra nhận được sản phẩm đám mây điểm (Hình 5), đây là cơ sở tựa trên thực nghiệm cho kiểm chứng phương trình LiDAR.



Hình 5. Đám mây điểm LiDAR (Ảnh: Quách Mạnh Tuấn)

Vì các tệp dữ liệu đo của LiDAR được mã hóa, đóng kín, cùng thông số kiểm định các giá trị độ lệch giữa IMU và bộ phát Laser nên nhóm tác giả cung cấp số đo để kiểm chứng phương trình LiDAR.

Thông số kiểm định độ lệch giữa tâm IMU và bộ phát Laser là vector $r_{lu}^b = \begin{bmatrix} 0,147 \\ -0,247 \\ -0,476 \end{bmatrix} (m)$

Cùng với các góc xoay của 2 hệ tọa độ IMU và bộ phát Laser là: $\omega = 2,00^\circ$; $\phi = 1,00^\circ$, $\kappa = -2,00^\circ$.

Theo công thức (2) tính được ma trận xoay Boresight

$$R_{lu}^b = \begin{bmatrix} 0,9992386 & 0,0348942 & 0,0174524 \\ -0,0342695 & 0,9988033 & -0,0348942 \\ -0,0186491 & 0,0342695 & 0,9992386 \end{bmatrix}$$

Các giá trị r_{lu}^b và R_{lu}^b cố định từ kiểm định thiết bị bay quét LiDAR, phục vụ tính phương trình LiDAR.

Bảng 1. Tính toán theo phương trình LiDAR

t	Lớp phản xạ	Tia laser			Tính toán			Hệ tọa độ mặt đất (x+410000, y+2350000)	
		$\rho(m)$	ζ, η ($^\circ$)	r_l^{lu} (m)	α, β, γ ($^\circ$)	$R_b^G(t) * R_{lu}^b$ $* r_l^{lu}$ (m)	$R_b^G(t)$ $* r_{lu}^b$ (m)	$r_b^G(t)$ (m)	r_l^G (m)
1	1	102,000	5,03 1,08	$\begin{bmatrix} -1,915 \\ -0,169 \\ -101,982 \end{bmatrix}$	-2,00 1,00 2,00	$\begin{bmatrix} -5,226 \\ 7,196 \\ -101,612 \end{bmatrix}$	$\begin{bmatrix} 0,130 \\ -0,235 \\ -0,487 \end{bmatrix}$	$\begin{bmatrix} 941,863 \\ 1392,321 \\ 127,751 \end{bmatrix}$	$\begin{bmatrix} 936,768 \\ 1399,282 \\ 25,653 \end{bmatrix}$
	2	105,150		$\begin{bmatrix} -1,974 \\ -0,174 \\ -105,131 \end{bmatrix}$		$\begin{bmatrix} -5,387 \\ 7,418 \\ -104,750 \end{bmatrix}$			$\begin{bmatrix} 936,606 \\ 1399,504 \\ 22,515 \end{bmatrix}$
	3	118,815		$\begin{bmatrix} -2,231 \\ -0,196 \\ -118,794 \end{bmatrix}$		$\begin{bmatrix} -6,087 \\ 8,382 \\ -118,363 \end{bmatrix}$			$\begin{bmatrix} 935,906 \\ 1400,468 \\ 8,902 \end{bmatrix}$
...
i	1	106,665	10,01 6,68	$\begin{bmatrix} -12,219 \\ -2,157 \\ -105,941 \end{bmatrix}$	1,18 3,05 3,16	$\begin{bmatrix} -19,880 \\ 7,378 \\ -104,536 \end{bmatrix}$	$\begin{bmatrix} 0,125 \\ -0,218 \\ -0,496 \end{bmatrix}$	$\begin{bmatrix} 968,688 \\ 1396,236 \\ 127,123 \end{bmatrix}$	$\begin{bmatrix} 948,933 \\ 1403,396 \\ 22,091 \end{bmatrix}$
	2	114,600		$\begin{bmatrix} -13,128 \\ -2,317 \\ -113,822 \end{bmatrix}$		$\begin{bmatrix} -21,359 \\ 7,926 \\ -112,313 \end{bmatrix}$			$\begin{bmatrix} 947,455 \\ 1403,944 \\ 14,314 \end{bmatrix}$

t	Lớp phản xạ	Tia laser			Tính toán			Hệ tọa độ mặt đất (x+410000, y+2350000)	
		$\rho(m)$	ζ, η ($^{\circ}$)	r_l^{lu} (m)	α, β, γ ($^{\circ}$)	$R_b^G(t) * R_{lu}^b$ $* r_l^{lu}$ (m)	$R_b^G(t)$ $* r_{lu}^b$ (m)	$r_b^G(t)$ (m)	r_l^G (m)
	3	118,920		$\begin{bmatrix} -13,623 \\ -2,404 \\ -118,113 \end{bmatrix}$		$\begin{bmatrix} -22,164 \\ 8,225 \\ -116,546 \end{bmatrix}$			$\begin{bmatrix} 946,649 \\ 1404,243 \\ 10,081 \end{bmatrix}$
...

Từ dữ liệu ở Bảng 1 nhận thấy rằng: trong phương trình LiDAR có những thông số cố định tham gia tính toán toàn bộ dữ liệu như thông số kiểm định các giá trị độ lệch (vector tọa độ r_{lu}^b , góc xoay: ω, ϕ, κ để tính R_{lu}^b) giữa IMU và bộ phát laser nên thông số này cần có độ chính xác cao; có thông số dùng chung cho thời điểm phát chùm tia laser như: vector tọa độ $r_b^G(t)$ và góc xoay (α, β, γ) để tính $R_b^G(t)$ của IMU do vậy cũng cần xác định chính xác vì ảnh hưởng đến chùm quét laser; thông số riêng của từng tia quét laser là góc xoay (ζ, η) để đồng bộ tọa độ về hệ bộ phát laser, còn khoảng cách thì một tia laser phát ra có thể thu được nhiều lớp phản xạ khác nhau theo tầng dữ liệu mặt đất (có thể từ 1 đến 3 hoặc nhiều hơn tùy theo khả năng kỹ thuật của thiết bị quét và đối tượng mặt đất) tính được các khoảng cách ρ , từ đó có thể tính được tọa độ điểm đo ở nhiều lớp khác nhau.

4. Kết luận

Nghiên cứu này đã áp dụng phương trình LiDAR tính tọa độ điểm đo mặt đất trong công nghệ bay quét laser hàng không. Các kết quả thu được từ cơ sở toán học, hình học để xây dựng phương trình LiDAR (6), phân tích tính toán dữ liệu bay quét laser hàng không theo số liệu tựa trên vùng thực nghiệm. Từ đó nhận thấy rằng: trong phương trình LiDAR có những thông số cố định tham gia tính toán toàn bộ dữ liệu như thông số kiểm định các giá trị độ lệch (tọa độ, góc xoay) giữa IMU và bộ phát Laser nên thông số này cần có độ chính xác cao; có thông số dùng chung cho thời điểm phát chùm tia laser như: vector tọa độ và góc xoay của IMU do vậy cũng cần xác định chính xác vì ảnh hưởng đến chùm quét laser; thông số riêng của từng tia quét laser là góc xoay để đồng bộ tọa độ về hệ bộ phát laser, còn khoảng cách thì một tia laser phát ra có thể thu được nhiều lớp phản xạ khác nhau theo tầng dữ liệu mặt đất (có thể từ 1 đến 3 hoặc nhiều hơn tùy theo khả năng kỹ thuật của thiết bị quét và đối tượng mặt đất) dẫn đến có nhiều khoảng cách, từ đó có thể tính được tọa độ điểm đo ở nhiều lớp khác nhau. Quá trình xử lý theo phương trình LiDAR đã được chứng minh là hợp lý về hình học, chính xác về toán học. Điều này góp phần vào việc triển khai chủ động về thuật toán, lập trình xử lý mô phỏng. Kết quả nghiên cứu có ý nghĩa thực tiễn quan trọng trong việc hướng dẫn các kỹ sư trắc địa áp dụng phương trình LiDAR khi triển khai thực địa cần đảm bảo chính xác các yếu tố liên quan được dùng trong phương trình LiDAR. Quy trình của phương trình LiDAR đảm bảo sự thống nhất và tránh sai sót. Dựa trên kết quả của nghiên cứu này, nhóm tác giả đề xuất tiếp tục nghiên cứu để triển khai viết các ứng dụng xử lý dữ liệu LiDAR.

Tài liệu tham khảo

- Ayman Habib (2008). “Accuracy, Quality Assurance, and Quality Control of LiDAR Data”, Chapter 9 in Topographic Laser Ranging and Scanning: Principles and Processing.
- Trần Trung Anh, Trần Hồng Hạnh và Quách Mạnh Tuấn, 2022. Kết hợp lọc hình thái và lọc phân bố mật độ đám mây điểm UAV-LiDAR để xây dựng mô hình số địa hình, Tạp chí Khoa học kỹ thuật Mỏ - Địa chất, số 63, kỳ 5, tr. 1-10.
- Trần Trung Anh, Trần Hồng Hạnh, Quách Mạnh Tuấn và Trần Trường Sinh (2022). Công tác thiết kế kỹ thuật quét LiDAR tích hợp chụp ảnh số trên thiết bị bay không người lái phục vụ thành lập bản đồ địa hình tỷ lệ lớn, Hội nghị Khoa học Quốc gia về Công nghệ Địa không gian trong Khoa học Trái đất và Môi trường, 991-997.
- Dji, 2023. Matrice 350 RTK User Manual ver 1.0, <https://www.dji.com/> (accessed in 7/2024)
- Dji, 2022. Zenmuse L2 User Manual ver 1.0, <https://www.dji.com/> (accessed in 7/2024)
- El-Sheimy, N.; Valeo, C.; Habib, A. (2005). Digital Terrain Modeling: Acquisition, Manipulation and Applications, 1st ed.; Artech House Remote Sensing Library: Boston, MA, USA, pp. 1-256.
- Gonçalves, Gil & Jalobeanu, André (2011). LiDAR boresight calibration: a comparative study, Conference: AGU Fall Meeting AbstractsAt: San Francisco, CA.
- Habib, Ayman, Ki In Bang, Ana Paula Kersting, and Jacky Chow (2010). “Alternative Methodologies for LiDAR System Calibration”, Remote Sensing 2, No. 3: 874-907. <https://doi.org/10.3390/rs2030874>.
- Bùi Thị Thanh Lan và Đồng Thị Linh (2016). Nghiên cứu sự thay đổi tỷ số Lidar theo khoảng cách, Tạp chí Khoa học kỹ thuật Mỏ - Địa chất, số 53.
- Li, Zhen, Junxiang Tan, and Hua Liu (2019). “Rigorous Boresight Self-Calibration of Mobile and UAV LiDAR Scanning Systems by Strip Adjustment”. Remote Sensing 11, No. 4: 442.

Mostafa, M.M. (2001). Boresight calibration of integrated inertial/camera systems, https://www.researchgate.net/publication/266295376_Boresight_Calibration_of_Integrated_InertialCamera_Systems#fullTextFileContent (accessed in 7/2024)

Nguyễn Thị Hữu Phương (2017). Thuật toán K-means và k-NN trong phân loại đám mây điểm LiDAR, Tạp chí Khoa học kỹ thuật Mỏ - Địa chất, số 58, kỳ 5.

Nguyễn Thị Hữu Phương, Đặng Văn Đức, Nguyễn Trường Xuân, Phạm Hữu Lợi và Nguyễn Minh Thắng (2022). Phương pháp phân lớp đám mây điểm LiDAR sử dụng độ cao và cường độ phản xạ của điểm, Tạp chí Khoa học kỹ thuật Mỏ - Địa chất, số 63, kỳ 5, tr. 48-57.

Bùi Ngọc Quý, Lê Đình Hiền, Phạm Văn Hiệp, Vũ Sơn Tùng, Dương Anh Quân và Trần Thị Thu Trang (2023). Xây dựng chương trình tự động phân loại dữ liệu đám mây điểm, Tạp chí Khoa học kỹ thuật Mỏ - Địa chất, số 64, kỳ 4, tr. 1-11.

Nguyễn Trường Xuân, Nguyễn Thị Mai Dung, Trần Trung Chuyên, Trần Mai Hương, Trần Thị Hải Vân và Trần Phương Ly (2017). Hệ thống quét laser di động và ứng dụng, Tạp chí Khoa học kỹ thuật Mỏ - Địa chất, số 58, kỳ 5.

Schenk, T. (2001). Modeling and Analyzing Systematic Errors in Airborne Laser Scanners; Technical Report in Photogrammetry No. 19, Ohio State University: Columbus, OH, USA, pp. 1-42.

Scouarnec, Romain & Touzé, Thomas & Lacambre, J. & Seube, Nicolas (2014). A New Reliable Boresight Calibration Method for Mobile Laser Scanning Applications. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XL-3/W1. 10.5194/isprsarchives-XL-3-W1-67-2014.

Vaughn, C.R.; Bufton, J.L.; Krabill, W.B.; Rabine, D.L. (1996). Georeferencing of airborne laser altimeter measurements. Int. J. Remote Sens., 17, 2185-2200.

Wang, C., Yang, X., Xi, X., Nie, S., & Dong, P. (2024). Introduction to LiDAR Remote Sensing (1st ed.). CRC Press. <https://doi.org/10.1201/9781032671512>.

ABSTRACT

The LiDAR equation calculates object points in airborne laser scanning technology

Tran Trung Anh^{1,2*}, Tran Hong Hanh^{1,2}, Luu Hai Au³, Quach Manh Tuan⁴

¹*Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology*

²*Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology*

³*The Vietnam Institute of Surveying and Mapping*

⁴*Northern QT Construction and Trade Joint Stock Company*

This study applied the LiDAR equation to calculate ground point coordinates in airborne laser scanning technology. The results were obtained through mathematical and geometric foundations, constructing the LiDAR equation, and analyzing airborne laser scanning data based on field measurements. From this, several observations were made: LiDAR boresight calibration: The LiDAR equation involves fixed parameters for computing the entire dataset. These include accuracy-critical values such as the deviation (coordinates and rotation angles) between the Inertial Measurement Unit (IMU) and the Laser Transmitter. Ensuring high precision for these parameters is essential. Shared Parameters: Parameters related to laser beam emission time, such as IMU coordinates and rotation angles, also need accurate determination since they impact laser scanning. Additionally, each laser scan beam has its specific rotation angle for synchronizing coordinates within the laser transmitter system. Considering distance, a single laser beam can capture multiple layers of ground reflection (typically 1 to 3 or more, depending on the scanning device's technical capabilities and the ground object). This allows computation of point coordinates across different layers. LiDAR Process Assurance: The LiDAR processing workflow ensures consistency and minimizes errors. Experimental validation in the Tang Tien ward, Viet Yen town, Bac Giang province, involved airborne laser scanning using the DJI M350 aircraft and the Zenmuse L2 scanner. The validated LiDAR equations based on this field experiment demonstrate clarity and accuracy.

Keywords: LiDAR equation, airborne laser, object point, GNSS, IMU.

Lập bản đồ biến động hệ thực vật vùng đất ven biển khu vực Khánh Hòa bằng ảnh Landsat đa thời gian trên nền tảng Google Earth Engine

Trần Thanh Hà^{1,2*}

¹Trường Đại học Mở - Địa chất

²Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

TÓM TẮT

Vùng ven biển Khánh Hòa là một trong những khu vực có tính đa dạng sinh học cao, có hệ sinh thái động thực vật phong phú. Tuy nhiên, trong nhiều năm qua do sự phát triển nhanh chóng của dân số, đô thị hóa và các hoạt động phát triển du lịch đang làm thay đổi hiện trạng hệ thực vật vùng đất ven biển. Trong nội dung của nghiên cứu này, dữ liệu ảnh Landsat đa thời gian được phân loại theo phương pháp hướng đối tượng bằng thuật toán Random Forest, trên nền tảng Google Earth Engine để thành lập bản đồ biến động hệ thực vật vùng đất ven biển, điều này giúp cho việc quản lý cũng như hỗ trợ lập kế hoạch quản lý tài nguyên ven biển Khánh Hòa. Dữ liệu nhóm tác giả sử dụng là ảnh Landsat-8 OLI chụp năm 2019 và 2024 để thành lập bản đồ biến động rừng ngập mặn và hệ thực vật nông nghiệp giai đoạn 2019-2024. Kết quả cho thấy rừng ngập mặn tại khu vực ven biển tỉnh Khánh Hòa tập trung chủ yếu ở đầm Nha Phu, đầm Thủy triều, vịnh Cam Ranh và vịnh Vân Phong với diện tích tăng lên 61,3 ha và diện tích hệ thực vật nông nghiệp giảm 2,734 ha trong giai đoạn 2019-2024.

Từ khóa: Hệ thực vật vùng đất ven biển, ảnh Landsat-8, Google Earth Engine, Khánh Hòa, phương pháp phân loại hướng đối tượng.

1. Đặt vấn đề

Hệ thực vật vùng đất ven biển bao gồm rừng ngập mặn và hệ thực vật nông nghiệp tạo nên hệ sinh thái biển cực kỳ quan trọng, đóng vai trò trong phát triển kinh tế và duy trì hệ sinh thái sinh học của con người như: cung cấp thức ăn, là nơi trú ẩn cho cá và sinh vật biển, bảo vệ hệ sinh thái, ổn định đáy biển và duy trì chất lượng và giảm ô nhiễm nước (Le Cozannet, G và cộng sự, 2019; Pajak, M.J., 2002). Tuy nhiên, do quá trình đô thị hóa, sự gia tăng dân số và biến đổi khí hậu đã gây ra tác động lớn nhất đến lớp phủ thực vật, làm thay đổi bờ biển, suy thoái địa hình, mất đa dạng sinh học, xâm nhập nước biển và ô nhiễm nước ngầm ngày một tăng, làm suy thoái đất và không khí dọc theo các vùng ven biển (Ivan Valiela và cộng sự, 2001).

Hiện nay có nhiều phương pháp giám sát biến động hệ thực vật vùng đất ven biển khác nhau như: phương pháp dựa vào số liệu thống kê hàng năm, số liệu kiểm kê, hay số liệu điều tra ngoài thực địa. Các phương pháp này thường tốn nhiều thời gian, kinh phí do sự thay đổi nhanh chóng của hệ thực vật vùng đất ven biển (Rawat và cộng sự, 2013). Trong những năm gần đây, các kỹ thuật phân tích dữ liệu viễn thám sử dụng trong các nghiên cứu môi trường ven biển được chứng minh là rất hiệu quả (Lymburner, L và cộng sự, 2020). Dữ liệu ảnh viễn thám có nhiều loại khác nhau từ ảnh viễn thám có độ phân giải thấp, trung bình đến cao, siêu cao kết hợp với các thuật toán và phần mềm đã chứng minh được hiệu quả trong xác định chính xác về sự thay đổi và biến đổi của lớp phủ bề mặt đất tại các vùng nghiên cứu (Mohammady và cộng sự, 2015). Việc xử lý, phân loại ảnh viễn thám nói chung và phân loại lớp phủ hệ thực vật mặt đất nói riêng về cơ bản thường sử dụng hai công nghệ phổ biến hiện nay đó là phân loại theo điểm ảnh hay pixel (Pixel - based) và phân loại hướng đối tượng (Object - Oriented hay Object - based) (Pham, T.D. và cộng sự, 2019). Trong đó phương pháp phân loại theo hướng đối tượng được sử dụng phổ biến nhất để thành lập bản đồ hiện trạng và biến động hệ thực vật ven biển, đặc biệt là rừng ngập mặn (Pham, T.D và cộng sự, 2019). Ví dụ, như Giri và các cộng sự, đã sử dụng ảnh Landsat đa thời gian để nghiên cứu sự phân bố diện tích rừng ngập mặn trên quy mô toàn cầu bằng phương pháp phân loại có giám sát (Giri, C và cộng sự, 2011) với độ chính xác đạt được dao động từ 79% đến 86%. Tại Việt Nam, Nguyễn Hải Hòa (Nguyễn Hải Hoà, 2016) đã ứng dụng ảnh viễn thám Landsat đa thời gian để đánh giá biến động diện tích rừng ngập mặn huyện Tiên Yên, Quảng Ninh, kết quả độ chính xác đạt được lên đến 83%. Hay Nguyễn Thanh Sơn (Nguyen-Thanh, S. và cộng sự, 2015) đã sử dụng Landsat đa thời gian dựa trên phương pháp phân tích hướng đối tượng để theo dõi diễn biến diện tích rừng ngập mặn ở bán đảo Cà Mau, kết quả cho thấy độ chính xác đạt được lên đến 82%.

Gần đây, với sự phát triển mạnh mẽ của phần mềm mã nguồn mở trên nền tảng Google Earth Engine (GEE) cùng tư liệu viễn thám, công tác theo dõi biến động diện tích rừng ngập mặn đã được ứng dụng rất rộng rãi và

* Tác giả liên hệ

Email: tranthanhha@humg.edu.vn

cho độ chính xác cao. Có thể kể đến như nghiên cứu của Nguyễn Thanh Sơn (Yancho, J.M.M. và cộng sự, 2020) đã sử dụng ảnh Landsat Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), Operational Land Imager (OLI) và Sentinel-2 trên nền tảng GEE để thành lập bản đồ biến động rừng ngập mặn, kết quả đạt được tương đối cao. Phạm Văn Duẩn (Phạm Văn Duẩn và cộng sự, 2019) đã sử dụng ảnh Landsat và Sentinel đa thời gian để nghiên cứu thiết lập bản đồ phân bố từng ngập mặn tại Việt Nam trên nền tảng Google Earth Engine với kết quả độ chính xác chung đạt được lên đến 93%. Như vậy có thể thấy, nhu cầu về ứng dụng viễn thám trong nghiên cứu giám sát hiện trạng cũng như biến động hệ thực vật vùng ven biển để phục vụ công tác quản lý nhà nước trong lĩnh vực tài nguyên và môi trường nói chung và môi trường ven biển nói riêng là rất lớn.

Vùng ven biển Khánh Hòa có độ dài khoảng 385 km với nhiều cửa lạch, đầm, vịnh, cùng khoảng 200 đảo lớn, nhỏ. Với quần thể phong phú như hiện nay vùng biển Khánh Hòa được coi là một trong những khu vực có tính đa dạng sinh học cao (nơi lưu giữ nguồn lợi sinh vật biển vô cùng phong phú và là nơi sinh sống, sinh sản nhiều loài hải sản có giá trị kinh tế cao như tôm hùm, mực, cá, cua ghẹ và rất nhiều loài hải đặc sản khác), có hệ sinh thái động thực vật phong phú và đa dạng (Võ Xuân Mai và cộng sự, 2010). Tuy nhiên, nhiều năm qua, phong trào chặt bỏ cây rừng để mở rộng vùng nuôi trồng thủy sản đã ảnh hưởng rất lớn đến hệ thực vật ven biển, việc chặt bỏ cây rừng khiến cấu trúc rừng ngập mặn và hệ thực vật nông nghiệp bị suy thoái nghiêm trọng. Tình trạng khai thác và sử dụng tài nguyên chưa hợp lý đã dẫn đến diện tích rừng ngập mặn và hệ thực vật nông nghiệp trên địa bàn Khánh Hòa ngày càng bị thu hẹp, gây ảnh hưởng xấu đến môi trường sinh thái, hủy diệt các loại động, thực vật ven biển, làm cạn kiệt nguồn nước ngọt, tăng khả năng xâm nhập mặn. Do đó, việc thu thập thông tin, dữ liệu về biến động hệ thực vật vùng đất ven biển là rất cần thiết để thành lập bản đồ phục vụ cho quản lý bền vững vùng ven biển cũng như hỗ trợ lập kế hoạch quản lý tài nguyên ven biển tỉnh Khánh Hòa.

2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1. Giới thiệu về khu vực nghiên cứu

Khu vực nghiên cứu bao gồm khu vực ven biển tỉnh Khánh Hòa với đường bờ biển tỉnh Khánh Hòa có độ dài khoảng 385 km tính theo mép nước từ xã Đại Lãnh tới cuối vịnh Cam Ranh. Vùng bờ ven biển tỉnh Khánh Hòa có hình thái đa dạng và phức tạp với hệ thống vũng vịnh, bán đảo, đảo, đầm phá, cửa sông và thềm lục địa. Đặc điểm vùng bờ là núi đồi thường kéo sát bờ biển, tạo địa mạo đặc thù và nhiều mũi nhô bởi các bờ đá gốc bị mòn. Hiện nay, Khánh Hòa là một trong những khu vực có nhiều đảo ven bờ, chỉ sau Quảng Ninh, Kiên Giang và Hải Phòng.



Hình 1. Vị trí khu vực nghiên cứu

2.2. Dữ liệu sử dụng

Trong bài báo sử dụng dữ liệu ảnh bề mặt phản xạ Landsat đa thời gian được thu thập trên nền tảng GEE được sử dụng để thành lập bản đồ hiện trạng hiện trạng thực vật vùng ven biển. Trong nghiên cứu sử dụng 02 cảnh ảnh Landsat với độ phủ mây nhỏ hơn 10% đã được hiệu chỉnh ảnh hưởng khí quyển và được thu nhận vào buổi sáng để tránh ảnh hưởng của thủy triều tới mực nước biển.

Bảng 7. Dữ liệu ảnh Landsat-8 OLI/TIRs khu vực Khánh Hòa

STT	Tên ảnh	Ngày chụp	Độ phân giải	Kênh phổ sử dụng
1	Landsat-8	27/4/2019	30 m	Kênh 1 (Aerosol), kênh 2 (xanh lam), kênh 3 (xanh lục), kênh 4 (đỏ), kênh 5 (NIR), kênh 6 và kênh 7 (SWIR)
2	Landsat-8	24/4/2024	30 m	

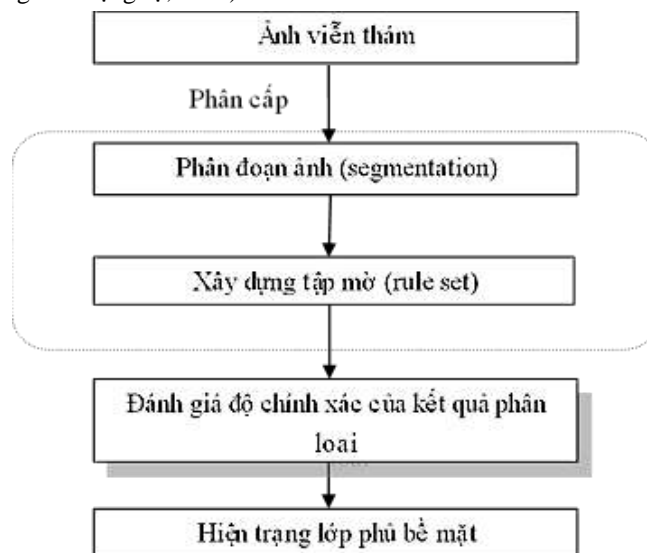


Hình 2. Ảnh được cắt của khu vực nghiên cứu

2.3. Phương pháp nghiên cứu

Trong nghiên cứu sử dụng phương pháp phân loại hướng đối tượng để chiết xuất thông tin về hệ thực vật vùng đất ven biển. Đây là phương pháp phân loại dựa trên đối tượng hoặc hướng đối tượng, phương pháp này sử dụng giá trị phổ và những đặc trưng không gian của đối tượng ảnh như hình dạng, kích thước, kết cấu... để phân loại đối tượng. Phân loại dựa trên đối tượng là một quá trình gồm hai bước, đầu tiên ảnh được phân đoạn hoặc chia thành các đối tượng riêng biệt với và sau đó từng đối tượng được phân loại. Có hai phương pháp cơ bản trong phân đoạn ảnh là: các phương pháp theo vùng và các phương pháp theo đường nét. Các phương pháp theo vùng là việc nhóm các pixel có đặc điểm tương tự nhau và quan hệ không gian. Các phương pháp phân đoạn theo đường nét là việc xác định các ranh giới. Trong đó phương pháp phân đoạn theo vùng thường hay được sử dụng và có thể giải thích như sau: Quá trình phân đoạn ảnh là chia ảnh thành các đối tượng ảnh (coi mỗi pixel là một đối tượng riêng biệt), sau đó các đối tượng liền kề nhau được nhóm lại và đồng nhất theo một tiêu chí đặt ra. Các tiêu chí, có thể là hình dạng, giá trị quang phổ, màu sắc, cấu trúc, độ mịn,... để tạo thành các phân đoạn lớn hơn. Quá trình này, làm cho độ bất đồng nhất tăng lên ở mức nhỏ nhất trong giới hạn định trước. Nếu giới hạn này vượt qua ngưỡng định sẵn thì quá trình gộp sẽ dừng lại. Với cách làm như vậy phân vùng đa phân giải là một thủ tục tối ưu hóa mang tính cục bộ (Sun Xiaoxia và cộng sự, 2005).

Trong nội dung nghiên cứu này, phương pháp phân loại định hướng đối tượng với thuật toán Random Forest đã được sử dụng. Quy trình của phương pháp chiết xuất thông tin theo định hướng đối tượng được trình bày như hình sau đây (Mariana Belgiu và cộng sự, 2014):



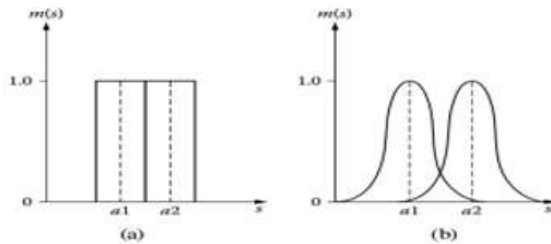
Hình 3. Quy trình chiết xuất thông tin lớp phủ bề mặt theo hướng tiếp cận đối tượng

Xây dựng tập mờ

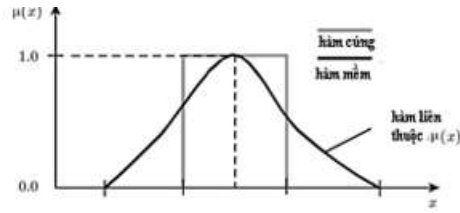
Theo (Bardossy A và cộng sự, 1995), chuỗi mờ được đưa ra bởi L. A Zadeh năm 1965 và đã được ứng dụng trong một vài lĩnh vực như xây dựng quyết định và điều khiển. Có nhiều tác giả định nghĩa khái niệm về tập mờ và nhìn tổng quan lại thì tập mờ được định nghĩa như sau:

X là một miền. A được gọi là tập con của X nếu A là một chuỗi trong cặp trường hợp: $A = \{(x, \mu_A(x)); x \in X, \mu_A(x) \in [0,1]\}$, nơi mà $\mu_A(x)$ là mức giá trị thành viên của đối tượng x thuộc A . Trong đó: hàm $\mu_A(x)$ được gọi là hàm liên thuộc.

Khi $\mu_A(x)$ càng gần 1, thì x càng thuộc A và ngược lại, $\mu_A(x)$ càng gần 0 thì x càng ít thuộc A . Nếu $[0,1]$ được thay thế bởi chuỗi hai yếu tố (Le Cozannet, G và cộng sự, 2019) thì A có thể được xác định là tập con của X .



Hình 4. (a) tập cứng, mức độ thành viên của nhóm a_1 hoặc a_2 chỉ là 0 hoặc 1. (b) tập mờ, cho phép sự chồng phủ giữa 2 lớp



Hình 5. Hàm liên thuộc

Theo Hình 4 và Hình 5 cho ta thấy rằng sự khác biệt giữa tập tường minh (cứng) và tập mờ (mềm) là dạng membership function (hàm liên thuộc). Hàm liên thuộc trong một tập tường minh với kết quả chỉ có thể chọn là {yes, no} hoặc {0,1} và mỗi đối tượng chỉ thuộc duy nhất một nhóm. Hay nói cách khác, quan hệ của đối tượng trong tập này là quan hệ 1 - 1. Tập mờ thì mềm dẻo hơn, giá trị liên thuộc của đối tượng trong tập mờ là khoảng từ 0 tới 1. Đối tượng đó có thể thuộc các nhóm khác nhau và kết quả sẽ thuộc nhóm nào nếu trọng số thuộc của đối tượng đó với nhóm có trọng số cao nhất. Trong phân loại mờ, độ chính xác phụ thuộc vào loại hàm liên thuộc được sử dụng và sự sắp xếp các hàm này trong hệ thống suy luận mờ. Như vậy, khi áp dụng hai tập này vào việc chiết xuất thông tin trên ảnh viễn thám, tập mờ sẽ linh hoạt và mềm dẻo hơn tập tường minh.

Khi phân loại các đối tượng theo tập mờ ta có thể sử dụng ba thuật toán khác nhau: thuật toán giá trị trung bình mờ (fuzzy c-means algorithm); phân loại xác suất lớn nhất mờ (fuzzy maximum likelihood classification); Phân loại dựa vào quy tắc mờ (fuzzy rule base). Trong nghiên cứu này quy tắc mờ được lựa chọn sử dụng trong phân loại ảnh.

Đánh giá độ chính xác của kết quả phân loại

Bản đồ hiện trạng hệ thực vật vùng đất ven biển năm 2019 và 2024 được thành lập dựa trên tư liệu ảnh Landsat-8 OLI, quá trình lấy mẫu và phân loại được thực hiện trên nền tảng GEE. Trong nghiên cứu này các mẫu được chọn ngẫu nhiên trên ảnh có 80% trên tổng số mẫu đã chọn được sử dụng cho phân loại ảnh và 20% còn lại được sử dụng cho đánh giá độ chính xác của phân loại theo công thức (Nguyễn Ngọc Thạch, 2011):

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

$$OA = \frac{\sum_{i=1}^r x_{ii}}{N} \quad (2)$$

trong đó:

K: Hệ số Kappa;

OA (Overall Accuracy): Độ chính xác tổng thể;

r: Số lượng cột trong ma trận ảnh;

N: Tổng pixel quan sát được trong ma trận ảnh;

X_{ii} : Số lượng pixel quan sát được tại hàng i và cột i (trên đường chéo chính);

X_{i+} : Tổng pixel quan sát tại hàng i;

X_{+i} : Tổng pixel quan sát tại cột i;

Hệ số Kappa nằm trong khoảng từ 0 đến 1 là đạt yêu cầu (Komatsu, T và cộng sự, 2012).

Nền tảng Google Earth Engine

Google Earth Engine (GEE) là nền tảng điện toán đám mây để phân tích dữ liệu viễn thám, thông số môi trường, khí tượng cấp độ từ khu vực nhỏ tới quy mô toàn cầu, cho phép người dùng chạy các phân tích không gian địa lý trên cơ sở nền tảng của Google. GEE ra đời xuất phát từ những ý tưởng kết hợp tri thức khoa học với nguồn dữ liệu khổng lồ và các nguồn lực công nghệ mới nhất của Google. Sự kết hợp này đem đến những hiệu quả rất lớn như tốc độ xử lý và khả năng tùy biến phát triển ứng dụng. Tốc độ tính toán, xử lý trên GEE nhanh chưa từng có (Pande CB và cộng sự, 2023). Trên GEE có tích hợp sẵn một kho lưu trữ hàng chục petabyte dữ liệu ảnh viễn thám miễn phí từ Cơ quan Hàng không Vũ trụ Mỹ (NASA), Cục Địa chất liên bang Hoa Kỳ (USGS), Cơ quan Hàng không vũ trụ châu Âu (ESA) và các dữ liệu khác. Hạ tầng điện toán đám mây của GEE được tối ưu hoá để xử lý dữ liệu không gian, kể cả xử lý chuỗi dữ liệu viễn thám trong khoảng thời gian dài với dung lượng rất lớn (Dong J. và cộng sự, 2016). Trong nghiên cứu này, tác giả sử dụng giao diện của Google Earth Engine - API (Google Earth Engine Application Program Interface) hay còn gọi là Code Editor. Giao diện này linh động hơn nhiều so với Explorer và còn có khả năng thực hiện các quy trình phân tích phức tạp theo yêu cầu người dùng. Trong Code Editor, tác giả sử dụng ngôn ngữ JavaScript để truy nhập và phân tích dữ liệu cũng như tải kết quả.

3. Kết quả và thảo luận

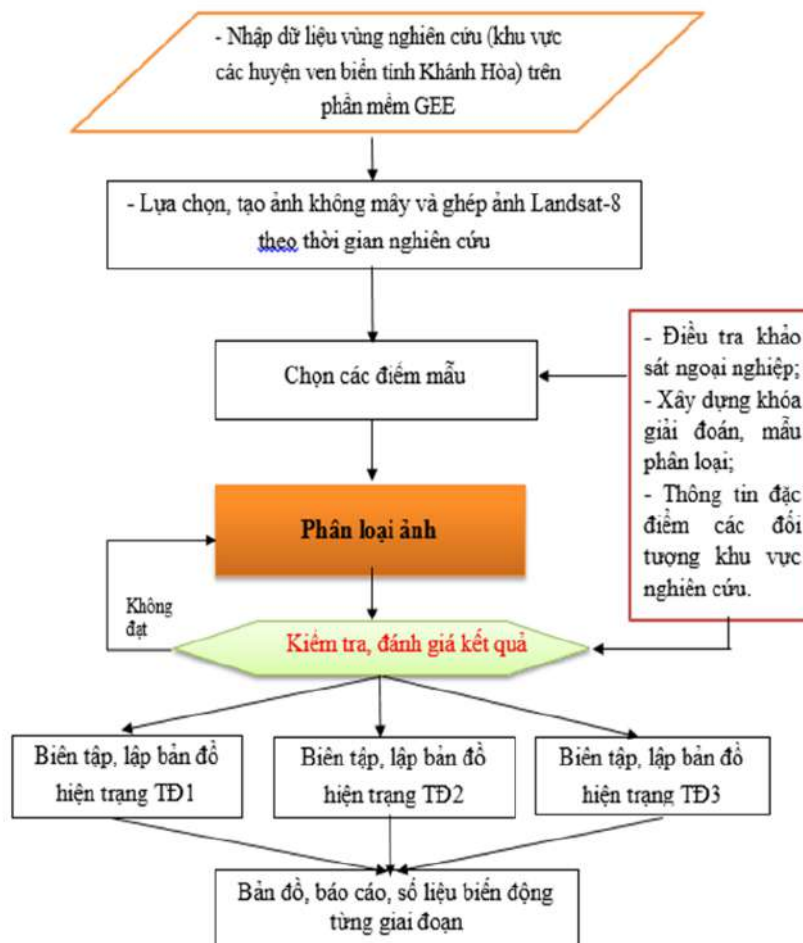
3.1. Quy trình thành lập bản đồ biến động hệ thực vật vùng đất ven biển

Đối với việc giám sát biến động hệ thực vật vùng đất ven biển (rừng ngập mặn và hệ thực vật nông nghiệp), cũng như các đối tượng xung quanh có liên quan, về cơ bản cũng tương tự như các phương pháp giám sát các đối tượng tự nhiên hoặc nhân tạo khác bằng công nghệ viễn thám. Trong đó, dữ liệu ảnh Landsat-8 OLI có độ phân giải không gian và thời gian phù hợp cho phép phát hiện và giải đoán thực vật vùng đất ven biển và các đối tượng cần quan tâm trong khu vực nghiên cứu. Có nhiều phương pháp xác định và đánh giá biến động lớp phủ bề mặt từ ảnh viễn thám đã được nghiên cứu và ứng dụng rộng rãi. Một số phương pháp tiêu biểu có thể kể ra như phương pháp phân tích sau phân loại, phương pháp phân loại trực tiếp từ ảnh đa thời gian, phương pháp nhận biết thay đổi phổ, phương pháp biến đổi số học,...

Trên cơ sở phân tích các đặc điểm, ưu thế và hạn chế của từng phương pháp đánh giá biến động sử dụng ảnh viễn thám, nhóm tác giả đã chọn phương pháp so sánh sau phân loại để giám sát sự biến động hệ thực vật vùng đất ven biển (rừng ngập mặn và hệ thực vật nông nghiệp) vì phương pháp này cho phép tiến hành phân loại độc lập tại từng thời điểm ảnh, thể hiện được sự biến động từ loại lớp phủ này sang loại lớp phủ khác và thuận lợi khi tiến hành chồng ghép so sánh trên các phần mềm ArcGIS.

Việc giải đoán các hệ thực vật vùng ven biển cần tiến hành trên ảnh tổ hợp màu tạo ra từ các kênh ảnh đa phổ của ảnh Landsat-8. Kênh ảnh chỉ số khác biệt thực vật chuẩn (NDVI) được tạo ra từ kênh cận hồng ngoại (NIR) và kênh đỏ, được sử dụng để hỗ trợ đánh giá, phân biệt các lớp thực phủ. Ngoài ra các kênh tổ hợp nhận biết vùng nước và vùng thực vật, giao thông, xây dựng trên ảnh Landsat-8 cũng được xử dụng để nhận dạng và xác định các điểm mẫu phân loại trên ảnh Landsat-8 một cách chính xác hơn.

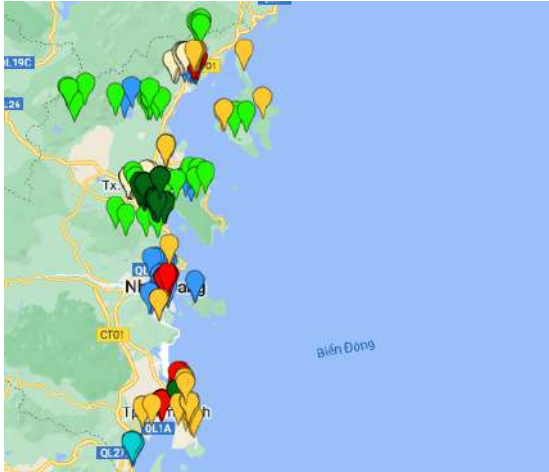
Quy trình xác định biến động lớp phủ hệ thực vật vùng đất ven biển trên nền tảng GEE sử dụng ảnh viễn thám đa thời gian sẽ được trình bày trong sơ đồ Hình 6.



Hình 6. Sơ đồ quy trình công nghệ giám sát biến động hệ thực vật vùng đất ven biển trên nền tảng GEE sử dụng dữ liệu ảnh viễn thám đa thời gian

3.2. Chọn mẫu phân loại

Trong nghiên cứu mẫu được lấy theo phương pháp Training point (lấy mẫu theo tập hợp điểm từ ảnh vệ tinh Landsat. Tổng số có 409 điểm mẫu được chọn ngẫu nhiên, trong đó 327 mẫu (80%) được dùng cho huấn luyện mẫu và 82 mẫu (20%) được dùng cho đánh giá độ chính xác.



Hình 7. Tập hợp các điểm mẫu phân loại



Hình 8. Kết quả phân loại lớp phủ thực vật khu vực nghiên cứu

3.3. Phân loại ảnh viễn thám

Sau khi lấy và khai báo mẫu xong, tạo hàm phân loại các đối tượng lớp phủ bề mặt và tiến hành phân loại bằng thuật toán Randomforest (RF)

```
//var classifier = ee.Classifier.smileCart().train({
//var classifier = ee.Classifier.libsvm().train({
//var classifier = ee.Classifier.smileRandomForest({numberOfTrees:150}).train({
// features: trainingPartition,
// classProperty: 'landcover',
// inputProperties: bands
//});
```

Sau khi phân loại, ta hiển thị kết quả (Hình 8).

3.4. Đánh giá độ chính xác

Kết quả phân loại được thể hiện trong Bảng 2 cho thấy độ chính xác tổng thể nhìn chung đều lớn hơn 92% và chỉ số Kappa đều lớn hơn 0,852. Do đó, các kết quả được hoàn toàn đáp ứng thành lập bản đồ hiện trạng và biến động cho giai đoạn 2019-2024.

Kết quả độ chính xác tổng thể năm 2019



Kết quả độ chính xác tổng thể năm 2024



3.5. Thành lập bản đồ biến động lớp phủ thực vật vùng đất ven biển (rừng ngập mặn và hệ thực vật nông nghiệp)

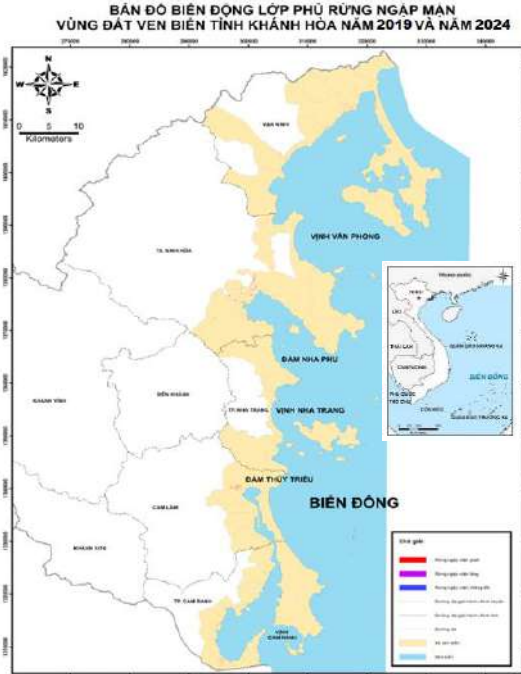
Sau khi có được ba bản đồ Hiện trạng Lớp phủ thực vật vùng đất ven biển (rừng ngập mặn và hệ thực vật nông nghiệp), sử dụng module Overlay trong ArcToolbox của phần mềm ArcGIS tiến hành chồng xếp lần lượt hai bản đồ hiện trạng của từng cặp thời gian kế tiếp nhau ta thu được các bản đồ biến động lớp phủ thực vật vùng đất ven biển (rừng ngập mặn và hệ thực vật nông nghiệp) và các số liệu thống kê về biến động cho giai đoạn 2019-2024.

a. Thành lập bản đồ biến động lớp phủ thực vật vùng đất ven biển rừng ngập mặn

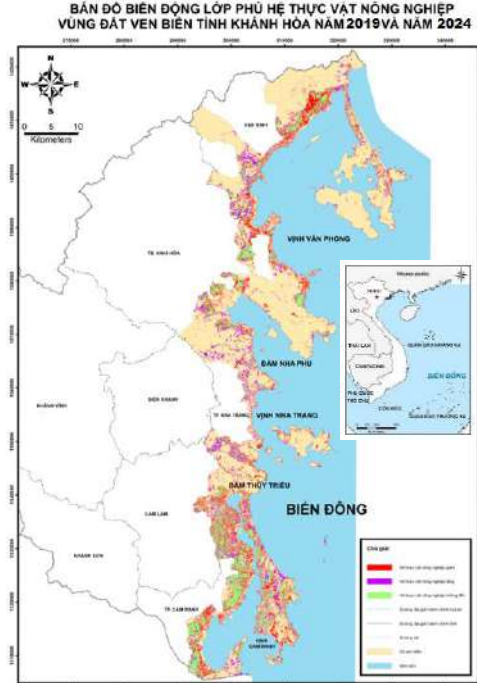
Bản đồ biến động lớp phủ thực vật vùng đất ven biển rừng ngập mặn giai đoạn 2019-2024 được thể hiện trong Hình 9 Số liệu thống kê về biến động các thời kỳ được thể hiện ở các Bảng 2.

b. Thành lập bản đồ biến động lớp phủ thực vật vùng đất ven biển hệ thực vật nông nghiệp

Bản đồ biến động lớp phủ thực vật vùng đất ven biển hệ thực vật nông nghiệp giai đoạn 2019-2024 được thể hiện trong hình 10. Số liệu thống kê về biến động các thời kỳ được thể hiện ở các Bảng 2.



Hình 9. Bản đồ biến động lớp phủ thực vật vùng đất ven biển rừng ngập mặn giai đoạn 2019-2024



Hình 10. Bản đồ biến động lớp phủ thực vật vùng đất ven biển hệ thực vật nông nghiệp giai đoạn 2019-2024

Bảng 2. Số liệu biến động hệ thực vật vùng đất ven biển (rừng ngập mặn và hệ thực vật nông nghiệp) từ giai đoạn 2019-2024

Diện tích rừng ngập mặn năm 2019 (ha)	Diện tích rừng ngập mặn năm 2024 (ha)	Số liệu biến động rừng ngập mặn (ha)	Số liệu hệ thực vật nông nghiệp năm 2019 (ha)	Số liệu hệ thực vật nông nghiệp năm 2024 (ha)	Số liệu biến động hệ thực vật nông nghiệp (ha)
85,1	146,4	+ 61,3	27.354	24.620	-2.734

Tỉnh Khánh Hòa có đường bờ biển dài gần 400 km, nhiều bãi triều rộng lớn ven vịnh Vân Phong, Nha Phu, Nha Trang, Cam Ranh,... là nơi sinh sống của các loại cây ngập mặn như bần, đước,... cô thụ. Nhiều năm qua, tình trạng khai thác và sử dụng tài nguyên chưa hợp lý đã dẫn đến diện tích rừng ngập mặn (RNM) trên địa bàn Khánh Hòa ngày càng bị thu hẹp. Trước năm 1990, toàn tỉnh có khoảng gần 2.500 ha RNM thì trong 2 năm 1990-2000, nhiều khu RNM đã bị người dân tàn phá nặng nề để lấy đất xây dựng ao, địa nuôi trồng thủy sản. Theo thống kê, thời điểm năm 2000, toàn tỉnh chỉ còn 25 ha RNM tập trung và RNM đang đứng trước nguy cơ bị xóa sổ (Nguyễn Xuân Hoà, 2009). Những năm 1990, khi nghề nuôi tôm đi vào thoái trào, bệnh trên tôm nuôi bùng phát, người dân mới thấy được tầm quan trọng của hệ sinh thái rừng đước nên bắt đầu tái tạo rừng đước, rừng ngập mặn. Qua kết quả thực nghiệm cho thấy rừng ngập mặn tại khu vực ven biển tỉnh Khánh Hòa tập trung chủ yếu ở đầm Nha Phu, đầm Thủy triều, vịnh Cam Ranh và vịnh Vân Phong. Kết quả tính toán diện tích rừng ngập mặn năm 2019 của khu vực là 85,1 ha đến năm 2024 là 146,4 ha, biến động diện tích rừng ngập mặn trong giai đoạn năm 2019-2024 đã tăng lên + 61,3 ha. Các kết quả tính toán là khách quan và phù hợp với việc tỉnh Khánh Hòa đã triển khai một số chương trình phục hồi trồng rừng ngập mặn trong những năm vừa qua.

Đối với hệ thực vật đất nông nghiệp trong giai đoạn 2019-2024 có xu thế giảm đều trong giai đoạn. Sự biến động này một phần là do yếu tố mùa vụ và sự biến động chuyển đổi cây trồng từ đất nông nghiệp sang đất dân cư của các dự án của toàn tỉnh Khánh Hòa. Về phân bố phù hợp với niên giám thống kê của tỉnh Khánh Hòa.

4. Kết luận

Ứng dụng chương trình Google Earth Engine (GEE) trong giám sát hệ thực vật vùng đất ven biển (rừng ngập mặn và hệ thực vật nông nghiệp) ở khu vực vùng rộng lớn đã tiết kiệm được nhiều thời gian công sức, cũng như

tiết kiệm kinh phí. Google Earth Engine (GEE) là một công cụ hỗ trợ đắc lực giúp cho các nhà nghiên cứu có thể dễ dàng truy cập và sử dụng các tài nguyên máy tính sẵn có và hạ tầng công nghệ thông tin của Google trong nghiên cứu, khai thác và sử dụng ảnh vệ tinh để quản lý và giám sát tài nguyên thiên nhiên và môi trường. GEE được xây dựng trên nền điện toán đám mây, giúp cho việc truy cập trở nên dễ dàng với tốc độ truy suất cao, cùng nguồn tài nguyên vô cùng rộng lớn. Hơn nữa, GEE lưu trữ và bao gồm rất nhiều dữ liệu về không gian địa lý được thu thập từ các các nguồn ảnh vệ tinh, với số lượng ảnh thường xuyên được cập nhật hằng ngày, nhằm phục vụ tốt hơn cho các nghiên cứu khác nhau như: dự báo khí hậu, thời tiết, tính toán độ che phủ rừng, địa hình và kinh tế- xã hội.

Qua kết quả thực nghiệm cho thấy rừng ngập mặn tại khu vực ven biển tỉnh Khánh Hòa tập trung chủ yếu ở đầm Nha Phu, đầm Thủy triều, vịnh Cam Ranh và vịnh Vân Phong. Kết quả tính toán diện tích rừng ngập mặn năm. Kết quả tính toán diện tích rừng ngập mặn năm 2019 của khu vực là 85,1 ha đến năm 2024 là 146,4 ha, biến động diện tích rừng ngập mặn trong giai đoạn năm 2019-2024 đã tăng lên + 61,3 ha. Các kết quả tính toán là khách quan và phù hợp với việc tỉnh Khánh Hòa, nhân dân ven biển và các tổ chức quốc tế đã triển khai một số chương trình phục hồi trồng rừng ngập mặn trong những năm vừa qua tại tỉnh Khánh Hòa.

Việc thực hiện phân tích, giải đoán các thông tin về lớp phủ bề mặt trên ảnh viễn thám đa thời gian cho phép theo dõi, giám sát biến động lớp phủ hệ thực vật ven bờ tại các khu vực nghiên cứu một cách khách quan và chính xác. Sử dụng ảnh Landsat-8 đa thời gian và tham khảo các thông tin bổ sung khác có thể cung cấp các số liệu kịp thời về hiện trạng, phạm vi, diện tích cũng như tình hình biến động hệ thực vật dưới nước ven biển (rong biển, cỏ biển) và hệ thực vật vùng đất ven biển (rừng ngập mặn và hệ thực vật nông nghiệp) theo chuỗi thời gian một cách nhanh chóng, hiệu quả tiết kiệm kinh phí.

Tài liệu tham khảo

- Le Cozannet, G.; Bulteau, T.; Castelle, B.; Ranasinghe, R.; Wöppelmann, G.; Rohmer, J.; Bernon, N.; Idier, D.; Louisor, J.; Salasy-Mélia, D. (2019). Quantifying uncertainties of sandy shoreline change projections as sea level rises. *Sci. Rep.*, 9, 42.
- Pajak, M.J. (2002). Leatherman, S. The high-water line as shoreline indicator. *J. Coast. Res.*, 18, 329-337.
- Ivan Valiela, Jennifer L, Bowem and Joanna K York., (2001). Mangrove Forests: One of the World's Threatened Major Tropical Environments. *BioScience Sciences*, October 2001, Vol. 51, No. 10.
- Rawat et al. (2013). J.S. Rawat, V. Biswas, M. Kumar Changes in land use/cover using geospatial techniques: a case study of Ramnagar town area, district Nainital, Uttarakhand, India Egypt. *J. Rem. Sens. Space Sci.*, 16 (2013), pp. 111-117.
- Lymburner, L.; Bunting, P.; Lucas, R.; Scarth, P.; Alam, I.; Phillips, C.; Ticehurst, C.; Held, A. (2020). Mapping the multi-decadal mangrove dynamics of the Australian coastline. *Remote Sens. Environ.*, 238, 111185.
- Mohammady et al. (2015). M. Mohammady, H.R. Moradi, H. Zeinivand, A. Temme A comparison of supervised, unsupervised and synthetic land use classification methods in the north of Iran *Int. J. Environ. Sci. Technol.*, 12(5), pp. 1515-1526.
- Pham, T.D.; Yokoya, N.; Bui, D.T.; Yoshino, K.; Friess, D.A. (2019). Remote Sensing Approaches for Monitoring Mangrove Species, Structure, and Biomass: Opportunities and Challenges. *Remote Sens*, 11, 230.
- Pham, T.D.; Xia, J.; Ha, N.T.; Bui, D.T.; Le, N.N.; Takeuchi, W. (2019). A Review of Remote Sensing Approaches for Monitoring Blue Carbon Ecosystems: Mangroves, Seagrasses and Salt Marshes during 2010-2018. *Sensors*, 19, 1933.
- Giri, C.; Ochieng, E.; Tieszen, L.L.; Zhu, Z.; Shingh, A.; Loveland, T. (2011). Status and distribution of mangrove forests of the world using earth observation satellite data. *Glob. Ecol. Biogeogr Sciences*, 20, 154-159.
- Nguyễn Hải Hòa (2016). Ứng dụng viễn thám Landsat đa thời gian và GIS đánh giá biến động diện tích rừng ngập mặn ven biển huyện Tiên Yên, tỉnh Quảng Ninh giai đoạn 1994-2015. *Tạp chí Khoa học Lâm nghiệp*, 1/2016, 4208-4217.
- Nguyen-Thanh, S.; Chen, C.-F.; Chang, N.-B.; Chen, C.-R.; Chang, L.-Y.; Bui, Xuan Thanh (2015). Mangrove Mapping and Change Detection in Ca Mau Peninsula, Vietnam, Using Landsat Data and Object-Based Image Analysis. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 8, 503-510.
- Yancho, J.M.M.; Jones, T.G.; Gandhi, S.R.; Ferster, C.; Lin, A.; Glass, L. (2020). The Google Earth Engine Mangrove Mapping Methodology (GEEMM). *Remote Sens.*, 12, 3758.
- Phạm Văn Duân, Lê Sỹ Doanh, Vũ Thị Thìn, Hoàng Văn Khiên, Phạm Thị Quỳnh (2019). Nghiên cứu thiết lập bản đồ phân bố rừng ngập mặn tại Việt Nam sử dụng ảnh Landsat-8 OLI và Sentinel 1 đa thời gian trên nền tảng điện toán đám mây của Google Earth Engine. *Tạp chí Khoa học Lâm nghiệp*, số 1, 100-110.
- Võ Xuân Mai, Hoàng Công Tín, Lê Như Hậu, Bùi Minh Lý, Trần Quang Thái, Võ Thành Trung, Ngô Thanh Trúc & Vũ Thị Mơ (2010). Ứng dụng công nghệ GIS và viễn thám trong xây dựng bản đồ vùng phân bố rong biển ở đảo Lý Sơn, tỉnh Quảng Ngãi. *Kỷ yếu hội nghị khoa học kỷ niệm 35 năm Viện Hàn lâm Khoa học và Công nghệ Việt Nam, Hà Nội*. Tr. 248-253.

Sun Xiaoxia Zhang Jixian Liu Zhengjun (2005). A Comparison of object-oriented and pixel-based classification approaches using quickbird imagery. International Society for Photogrammetry and Remote Sensing.

Mariana Belgiua, Lucian Dragut (2014). Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery, ISPRS J. Photogramm. Remote Sens., 96, 67-75.

Bardossy, A., L. Duckstein and I. Bogardi (1995). Fuzzy rule-based classification of atmospheric circulation patterns. Int. J. Climatol., 15: 1087-1097.

Nguyễn Ngọc Thạch (2011). Những nguyên lý cơ bản về viễn thám, hệ thống thông tin địa lý và hệ thống định vị toàn cầu, Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội.

Komatsu, T., Sagawa, T., Sawayama, S., Tanoue, H., Mohri, A., Sakanishi, Y. (2012). Mapping is a key for sustainable development of coastal waters. In: Sustainable Development, In Tech Publishing Co. Book 2 (Free Online Book), Croatia, pp. 145 - 160.

Pande CB, Moharir KN, Varade AM, Abdo HM, Mulla S (2023). Intertwined impacts of urbanization and land cover change on urban climate and agriculture in Aurangabad city (MS), India using google earth engine platform. *J Clean Prod* 422:138541.

Dong J, Xiao X, Menarguez MA, Zhang G, Qin Y, Thau D, Biradar C, Moore B (2016) Mapping paddy rice planting area in northeastern Asia with Landsat-8 images, phenology-based algorithm and Google Earth Engine. *Remote Sens Environ* 185:142-154.

Nguyễn Xuân Hòa (2009). Báo cáo Đề án môi trường tỉnh Khánh Hòa “Điều tra, thống kê diện tích, thành phần loài, đánh giá hiện trạng phân bố hệ sinh thái rừng ngập mặn, thảm cỏ biển và vai trò của chúng đối với kinh tế- xã hội, môi trường ở vùng biển ven bờ tỉnh Khánh Hòa - Đề xuất giải pháp quản lý và sử dụng bền vững”, Sở Tài nguyên và Môi trường tỉnh Khánh Hòa, Khánh Hòa.

ABSTRACT

Mapping coastal vegetation changes using Landsat time series based Google Earth Engine, case study of Khanh Hoa province

Tran Thanh Ha^{1,2*}

¹Hanoi University of Mining and Geology

²Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology

The coastal area of Khanh Hoa is one of the areas with high biodiversity and rich flora and fauna ecosystems. In recent years, due to rapid rise of population, urbanization and tourism development activities, the current status of coastal flora has been changing. In this study, multi-temporal Landsat image data is classified by the object-oriented method using the Random Forest algorithm, on the Google Earth Engine platform to establish a map of coastal flora changes, which helps in the management as well as supporting the planning of coastal resource management in Khanh Hoa. The data were Landsat-8 OLI images taken in 2019 and 2024 to establish a map of mangrove and agricultural flora changes in the period 2019-2024. The results showed that mangrove forests in the coastal area of Khanh Hoa province were mainly concentrated in Nha Phu Lagoon, Thuy Trieu Lagoon, Cam Ranh Bay and Van Phong Bay with an area increasing by 62.3 ha and the area of agricultural flora decreasing by 2,734 ha in the period 2019-2024.

Keywords: Coastal vegetation, Landsat-8 image, Google Earth Engine, Khanh Hoa, Object oriented classification.

Quan trắc trượt lở đất đá theo thời gian thực ở tầng khai thác mỏ lộ thiên dựa trên công nghệ GNSS/CORS

Phạm Công Khải^{1*}, Nguyễn Văn Hải²

¹Trường Đại học Mở - Địa chất

²Trường Đại học Thủy Lợi (Cơ sở 2), Thành phố Hồ Chí Minh

TÓM TẮT

Quan trắc theo thời gian thực để cảnh báo sớm trượt lở đất đá ở tầng khai thác mỏ lộ thiên là vấn đề rất quan trọng để đảm bảo cho quá trình khai thác mỏ được hiệu quả, kinh tế và an toàn. Ngày nay, có nhiều phương pháp để quan trắc trượt lở đất đá dựa trên những công nghệ và thiết bị khác nhau. Tuy nhiên, việc quan trắc trượt lở đất đá dựa trên công nghệ GNSS/CORS có nhiều ưu điểm như quan trắc được trong mọi điều kiện thời tiết, liên tục và theo thời gian thực trong không gian ba chiều. Mục đích của nghiên cứu này là thiết kế, xây dựng được một hệ thống quan trắc theo thời gian thực dựa trên công nghệ GNSS/CORS để cảnh báo sớm trượt lở đất đá ở tầng khai thác mỏ lộ thiên. Trong nghiên cứu này một hệ thống quan trắc trượt lở đất đá theo thời gian thực được thiết kế, xây dựng, bao gồm hệ thống trạm CORS và hệ thống trạm quan trắc. Hệ thống trạm CORS được thiết kế, xây dựng đặt tại thành phố Cẩm Phả, Quảng Ninh với ăng ten Zephyr 2 Geodetic của Trimble và bộ thu GNSS của hãng Stonex. Hệ thống trạm quan trắc được nhóm tác giả tự thiết kế phát triển cả về phần cứng và phần mềm, đảm bảo cho quá trình thu nhận, truyền dẫn, xử lý dữ liệu và cảnh báo diễn ra theo thời gian thực. Dữ liệu quan trắc thu được theo định dạng tiêu chuẩn NMEA0183 được xử lý bằng phương pháp và phần mềm tự thiết kế phát triển đã nâng cao được độ chính xác kết quả quan trắc đến milimet. Một thực nghiệm mô phỏng quan trắc trượt lở đất đá đã được thực hiện ở khu vực thành phố Cẩm Phả, Quảng Ninh và thực nghiệm thực tế ở trên bề mặt tầng khai thác ở mỏ than Cọc Sáu đã khẳng định được tính khả thi của giải pháp trong quan trắc trượt lở đất đá theo thời gian thực.

Từ khóa: Công nghệ GNSS/CORS, trượt lở đất đá, thời gian thực, mỏ than Cọc Sáu.

1. Đặt vấn đề

Từ trước tới nay, việc quan trắc trượt lở, chuyển dịch đất đá ở các mỏ lộ thiên Việt Nam thường được thực hiện bằng các phương pháp truyền thống với việc sử dụng các thiết bị đo đạc như máy thủy chuẩn, máy kinh vĩ, máy toàn đạc điện tử. Từ khi có hệ thống định vị toàn cầu GPS, nó đã được ứng dụng để quan trắc trượt lở đất đá và những tai biến địa chất khác (Z. Abidin H và cộng sự, 2004; M.-B. Su và cộng sự, 2009; J.A. Gili, 2000). Tuy nhiên công nghệ GPS hầu như cũng chỉ được ứng dụng để quan trắc chuyển dịch đất đá theo chu kỳ. Trong những năm gần đây với sự ra đời của nhiều hệ thống vệ tinh như GPS (Mỹ), Glonass (Nga), Galileo (Cộng đồng châu Âu) và Beidou (Trung Quốc) đã hình thành nên hệ thống vệ tinh dẫn đường toàn cầu (GNSS). Với công nghệ GNSS, các quốc gia trên thế giới đều đã xây mạng lưới các trạm tham chiếu hoạt động liên tục (CORS) làm cơ sở hạ tầng không gian quốc gia, đáp ứng cho các mục đích khác nhau, trong đó có mục đích quan trắc trượt lở đất đá. Do mạng lưới trạm CORS hoạt động liên tục trong mọi điều kiện thời tiết, cung cấp số liệu cải chính trong không gian 3 chiều, vì vậy nó rất phù hợp để ứng dụng trong quan trắc liên tục trượt lở đất đá theo thời gian thực. Hiện nay, việc ứng dụng công nghệ GNSS vào quan trắc trượt lở đất đá chủ yếu áp dụng giải pháp với trạm CORS đơn (Cina, M. Piras, 2015; T. Bellone và cộng sự, 2016; W.Y. Zhao và cộng sự, 2021). Ưu điểm của giải pháp này là khoảng cách từ trạm CORS đến các trạm quan trắc nhỏ nên có thể loại bỏ tốt hơn ảnh hưởng của các sai số khí quyển như độ trễ của tầng điện ly và tầng đối lưu đến kết quả quan trắc (N. Shen và cộng sự, 2021). Tuy nhiên, giải pháp này cũng có nhược điểm là bị giới hạn về khoảng cách giữa trạm CORS và các trạm quan trắc nên số lượng trạm CORS sẽ tăng lên, dẫn đến chi phí quan trắc cao, các trạm CORS cũng có xu hướng bị dịch chuyển cùng với các khối địa chất không ổn định (Y. Du và cộng sự, 2020).

Ngoài việc sử dụng giải pháp trạm CORS đơn để quan trắc liên tục trượt lở đất đá theo thời gian thực, thì giải pháp ứng dụng công nghệ trạm CORS mạng với kỹ thuật trạm tham chiếu ảo VRS cũng đã được sử dụng (Bao Shu và cộng sự, 2023; Pengxu Wang và cộng sự, 2022). Ưu điểm của công nghệ CORS mạng là khoảng cách từ trạm CORS đến các trạm quan trắc lớn nên không cần nhiều trạm CORS, vì vậy chi phí quan trắc sẽ giảm. Nhược điểm của giải pháp trạm tham chiếu ảo là sai số mô hình tầng điện ly và tầng đối lưu ảnh hưởng lớn đến kết quả quan trắc do khoảng cách giữa các trạm CORS quá dài (J. Paziewski, 2016; J. Cui và cộng sự, 2018). Kết quả là độ chính xác và độ ổn định của giải pháp CORS mạng kém hơn so với giải pháp CORS đơn (R. Odolinski và

* Tác giả liên hệ

Email: phamcongkhai@humg.edu.vn

cộng sự, 2015). Vì vậy, trong quan trắc trượt lở đất đá theo thời gian thực ứng dụng giải pháp trạm CORS đơn và máy thu GNSS tự phát triển là hiệu quả nhất, với chi phí thấp (Daniel Janos và cộng sự, 2022; Pengxu Wang và cộng sự, 2022).

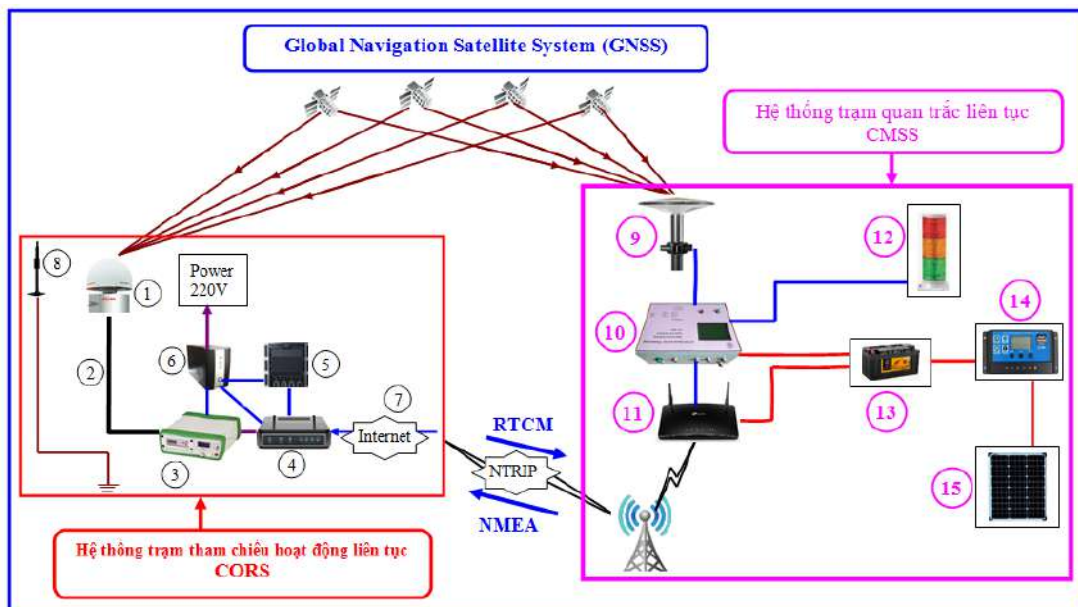
Hiện nay, một số mỏ than lộ thiên lớn của Việt Nam tập trung ở vùng Cẩm Phả tỉnh Quảng Ninh, độ sâu khai thác ngày càng lớn, điều kiện địa chất rất phức tạp, dẫn đến nguy cơ xảy ra trượt lở, dịch chuyển đất đá ở các tầng khai thác là rất cao, đe dọa đến tính mạng con người và thiết bị khai thác làm việc ở trong lòng mỏ. Để có thể cảnh báo sớm hiện tượng trượt lở, dịch chuyển đất đá ở các tầng khai thác ở mỏ lộ thiên, cần thiết phải có một hệ thống quan trắc liên tục cung cấp số liệu theo thời gian thực. Vì vậy, trong nghiên cứu này nhóm tác giả đề xuất giải pháp quan trắc trượt lở đất đá theo thời gian thực ở tầng khai thác mỏ lộ thiên Cọc Sáu dựa trên công nghệ trạm CORS đơn. Hệ thống trạm CORS được nhóm tác giả thiết lập ở khu vực thành phố Cẩm Phả, Quảng Ninh vào tháng 6 năm 2023 và cách khu vực khai thác khoảng 5 km. Hệ thống trạm quan trắc sử dụng máy thu GNSS do nhóm tác giả tự thiết kế, phát triển cho phép quan trắc. Việc quan trắc trượt lở đất đá theo thời gian thực bờ mỏ than Cọc Sáu ở thành phố Cẩm Phả, tỉnh Quảng Ninh, Việt Nam được thực hiện từ tháng 5 năm 2023. Dữ liệu quan trắc được truyền về máy tính chủ đặt tại Trường Đại học Mỏ - Địa chất ở Hà Nội và được xử lý tức thời bằng phần mềm chuyên dụng.

2. Phương pháp nghiên cứu

2.1. Thiết kế, xây dựng hệ thống quan trắc trượt lở đất đá theo thời gian thực

Hệ thống quan trắc trượt lở, chuyển dịch đất đá theo thời gian thực dựa trên công nghệ GNSS/CORS được thiết kế phải đảm bảo hoạt động ổn định, liên tục từ thu nhận, truyền tải, xử lý dữ liệu và cảnh báo diễn ra theo thời gian thực. Sơ đồ của hệ thống quan trắc trượt lở đất đá theo thời gian thực được thiết kế, xây dựng gồm có hai phần chính (Hình 1). Phần thứ nhất là hệ thống trạm tham chiếu hoạt động liên tục (Continuously Operating Reference Station - CORS) và phần thứ hai là hệ thống trạm quan trắc liên tục (Continuously Monitoring Station System - CMSS).

Chức năng của hệ thống trạm CORS là cung cấp thông tin số liệu cải chính vị trí cho trạm quan trắc, xử lý số liệu của mạng lưới đo động xử lý tức thời RTK, hiệu chỉnh số nguyên đa trị của toàn mạng lưới trạm CORS, thiết lập mô hình cải chính sai số tầng đối lưu, tầng điện ly và quỹ đạo vệ tinh. Ngoài ra máy tính chủ của trạm CORS là nơi lưu trữ và xử lý số liệu của trạm quan trắc.



Hình 1. Sơ đồ hệ thống quan trắc trượt lở đất đá theo thời gian thực

Hệ thống trạm quan trắc gồm có các thành phần chính là ăng ten GNSS (9), bộ thu GNSS (10) và bộ định tuyến wifi 4G (11), module cảnh báo bằng âm thanh và ánh sáng (12), ắc quy (13), bộ chuyển đổi nguồn điện (14), tấm pin năng lượng mặt trời (15). Trong các thành phần của hệ thống trạm quan trắc, bộ thu GNSS đã được nhóm tác giả tự thiết kế, phát triển. Bộ thu này sử dụng bo mạch thu nhận và xử lý tín hiệu vệ tinh BD970 của Trimble và được tích hợp vào bo mạch kết nối để tạo thành module thu tín hiệu vệ tinh GNSS BD970 [20]. Module GNSS BD970 thu được 220 kênh của các hệ thống vệ tinh GPS (L1C/A, L2E, L2C, L5), GLONASS (L1C/A, L1P, L2C/A, L2P), SBAS (L1C/A, L5), Galileo (L1BOC, E5A, E5B, E5AltBOC1), BeiDou (B1, B2),

QZSS(L1C/A, L1 SAIF, L2C, L5). Độ chính xác định vị RTK là 8 mm+1ppm về mặt bằng và 15 mm+1 ppm về độ cao với khoảng cách định vị không quá 30 km (<https://oemgnss.trimble.com/>). Module này có một cổng mạng LAN Ethernet qua đó có thể kết nối với modem wifi để truyền dữ liệu quan trắc về máy tính chủ theo địa chỉ IP và công nhận dữ liệu đã được mở trên modem ở trạm CORS. Module GNSS BD970 kết hợp với một số module khác được kết nối với nhau tạo thành bộ thu GNSS (Hình 3a) cho hệ thống trạm quan trắc (Hình 3b).



Bảng 1. Cấu trúc của tin nhắn trị đo GGA và các thông tin chi tiết của nó

GGA-Thông tin về hệ thống định vị, thời gian, vị trí mặt bằng và độ cao máy thu nhận được												
1	2	3	4	5	6	7	8	9	10	11	12	13
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
\$GNGGA,hhmmss.ss,llll.ll,a,yyyyy.yy,a,x,xx,x.x,x.x,M,x.x,M*hh												
↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
1	2	3	4	5	6	7	8	9	10	11	12	13

Tin nhắn trị đo GST chứa các thông tin về thời gian hiệu chỉnh vị trí (1), phần dư của khoảng cách giả (2), giá trị bán trục lớn của elip sai số (3), giá trị bán trục nhỏ của elip sai số (4), góc định hướng của bán trục lớn của elip sai số (5), giá trị sai số theo trục Y (6), giá trị sai số theo trục X (7), sai số độ cao (8), mã kiểm tra chuỗi tin nhắn gửi về (9). Cấu trúc của tin nhắn trị đo GST và các thông tin chi tiết của nó được thể hiện như ở (Bảng 2).

Bảng 2. Cấu trúc của tin nhắn trị đo GST và các thông tin chi tiết của nó

GST – Thống kê về sai số định vị								
1	2	3	4	5	6	7	8	9
↓	↓	↓	↓	↓	↓	↓	↓	↓
\$GNGST,hhmmss.ss,a,a,b,b,c,c,d,d,e,e,f,f,g,g*hh								
↑	↑	↑	↑	↑	↑	↑	↑	↑
1	2	3	4	5	6	7	8	9

Chỉ số chất lượng định vị RTK theo công nghệ CORS có 6 mức độ khác nhau từ mức 0 đến mức 5. Nếu chỉ số này ở mức 4 thì có nghĩa là trị đo này đã được hiệu chỉnh từ trạm CORS và người đo sẽ lấy giá trị tọa độ này trong chuỗi trị đo GGA. Kết hợp với sai số định vị thể hiện trong chuỗi tin nhắn GST, sẽ chọn được những tọa độ định vị tốt nhất.

2.2.2. Phương pháp xử lý dữ liệu quan trắc trượt lở đất đá theo thời gian thực

Dữ liệu quan trắc trượt lở đất đá dựa trên công nghệ GNSS/CORS nhận được theo định dạng tiêu chuẩn NMEA0183 cần được xử lý để nâng cao độ chính xác của kết quả quan trắc. Trong dữ liệu quan trắc trượt lở đất đá có nhiều loại tin nhắn khác nhau nhưng có hai loại tin nhắn là GGA và GST có chứa các thông tin về tọa độ và độ cao cũng như sai số của nó. Dựa trên các thông tin này sẽ lọc được những trị đo tốt nhất.

Quá trình xử lý dữ liệu quan trắc trượt lở đất đá theo thời gian thực được tiến hành qua các bước sau:

1) Kiểm tra tính toàn vẹn của các thông tin trong chuỗi tin nhắn trị đo

Các dòng tin nhắn trị đo theo định dạng NMEA0183 được truyền từ trạm quan trắc về máy tính chủ qua giao thức NTRIP có thể bị lỗi dẫn tới thông tin không đầy đủ. Để loại bỏ những dòng tin nhắn này cần phải kiểm tra tính toàn vẹn của nó. Nếu trong các tin nhắn này không đầy đủ các thông tin thì sẽ loại bỏ các dòng tin nhắn trị đo này. Để kiểm tra tính toàn vẹn của các thông tin trong chuỗi tin nhắn trị đo cần phân tích tất cả các ký tự trong chuỗi từ ký tự "\$" đến ký tự "*", sau đó áp dụng thuật toán thao tác bit bắt đầu từ ký tự đầu tiên đến ký tự tiếp theo cho đến khi kết thúc chuỗi để nhận được mã phân tích chuỗi mới. Mã phân tích chuỗi này đem so sánh với mã kiểm tra (checksum) gửi kèm trong tin nhắn. Nếu hai mã này giống nhau, thì dòng tin nhắn trị đo này sẽ được chấp nhận.

2) Lọc những tin nhắn có tọa độ và độ cao đã được hiệu chỉnh từ trạm CORS.

Trị số tọa độ và độ cao nhận được trong chuỗi tin nhắn GGA có chỉ số báo chất lượng phép đo GNSS khác nhau và được đánh số từ 0 đến 5. Khi chỉ số báo chất lượng ở mức 4 thì trị đo này đã được hiệu chỉnh vị trí từ trạm CORS và cho chất lượng tốt nhất. Khi đó các giá trị tọa độ và độ cao trong dòng tin nhắn này được tách ra để sử dụng cho việc tính độ trượt lở đất đá.

3) Lọc những tin nhắn trị đo có sai số vị trí nhỏ nhất.

Để đảm bảo yêu cầu độ chính xác trong quan trắc trượt lở đất đá thì chỉ những tin nhắn có tọa độ và độ cao đã được hiệu chỉnh nhưng từ trạm CORS và có sai số nhỏ nhất mới được giữ lại. Để lọc ra những tin nhắn có tọa độ và độ cao có sai số nhỏ nhất cần tiến hành phân tích chuỗi tin nhắn GST. Chuỗi tin nhắn này có chứa các thông

tin về elip sai số và sai số vị trí điểm. Trong nghiên cứu này những tin nhắn có sai số mặt bằng nhỏ hơn 5 mm và sai số độ cao nhỏ hơn 8 mm sẽ được giữ lại.

4) Tính chuyển đổi hệ tọa độ WGS- 84 sang VN-2000

Tọa độ của các trạm quan trắc trượt lở đất đá được xác định trong hệ tọa độ WGS84, để thuận tiện trong việc xác định độ trượt lở đất đá cần được tính chuyển sang hệ tọa độ VN-2000 theo công thức (1) (Đăng Nam Chính và Đỗ Ngọc Đường, 2012).

$$\left. \begin{aligned} X &= \Delta X_0 + k(X' + \varepsilon_0 Y' - \psi_0 Z') \\ Y &= \Delta Y_0 + k(-\varepsilon_0 X' + Y' + \varpi_0 Z') \\ Z &= \Delta Z_0 + k(\psi_0 - \varpi_0 Y' + Z') \end{aligned} \right\} \quad (1)$$

trong đó:

X, Y, Z là tọa độ vuông góc không gian trong hệ tọa độ VN-2000, (m)

X', Y', Z' là tọa độ vuông góc không gian trong hệ tọa độ WGS-84, (m)

$\Delta X_0, \Delta Y_0, \Delta Z_0$ là các tham số dịch chuyển gốc tọa độ (m)

$\omega_0, \psi_0, \varepsilon_0$ là 3 góc xoay trục tọa độ tương ứng với các trục X, Y, Z, (radian)

k là hệ số tỷ lệ chiều dài giữa 2 hệ tọa độ.

Các tham số dịch chuyển tọa độ từ hệ tọa độ WGS84 sang hệ tọa độ VN-2000 do Bộ Tài nguyên và Môi trường công bố.

5) Đánh mức độ phân tán của chuỗi tọa độ và độ cao

Độ phân tán của chuỗi trị đo tọa độ và độ cao được xác định bằng đại lượng độ lệch chuẩn của một dãy trị đo được xác định theo các công thức (2) (Hoàng Ngọc Hà và Trương Quang Hiếu, 2003).

$$\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}, \quad \sigma_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2}, \quad \sigma_h = \sqrt{\frac{1}{n} \sum_{i=1}^n (H_i - \bar{H})^2} \quad (2)$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i, \quad \bar{H} = \frac{1}{n} \sum_{i=1}^n H_i$$

trong đó:

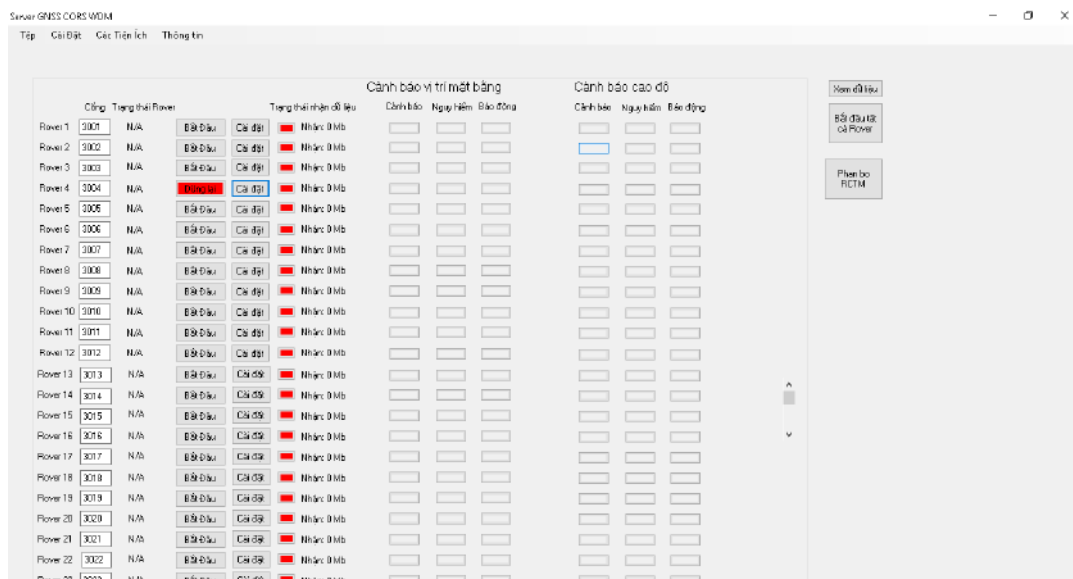
X_i, Y_i, H_i là tọa độ và độ cao xác định lần thứ i (m);

$\bar{X}, \bar{Y}, \bar{H}$ là tọa độ và độ cao trung bình trong chuỗi trị đo (m).

6) Xác định các đại lượng chuyển dịch, trượt lở đất đá theo các trục tọa độ xác định theo các công thức (3):

$$\left. \begin{aligned} \text{- Dịch chuyển theo trục X:} \quad Dx &= X_i(t+1) - X_i(t) \\ \text{- Dịch chuyển theo trục Y:} \quad Dy &= Y_i(t+1) - Y_i(t) \\ \text{- Dịch chuyển ngang:} \quad Dp &= \sqrt{Dx^2 + Dy^2} \\ \text{- Đại lượng dịch chuyển đứng:} \quad S &= H_i(t+1) - H_i(t) \end{aligned} \right\} \quad (3)$$

Quá trình xử lý dữ liệu quan trắc trượt lở đất đá theo thời gian thực được xử lý tức thời dựa trên phần mềm có tên là Server GNSS CORS WDM (Phạm Công Khải và nnk, 2023) (Hình 3)



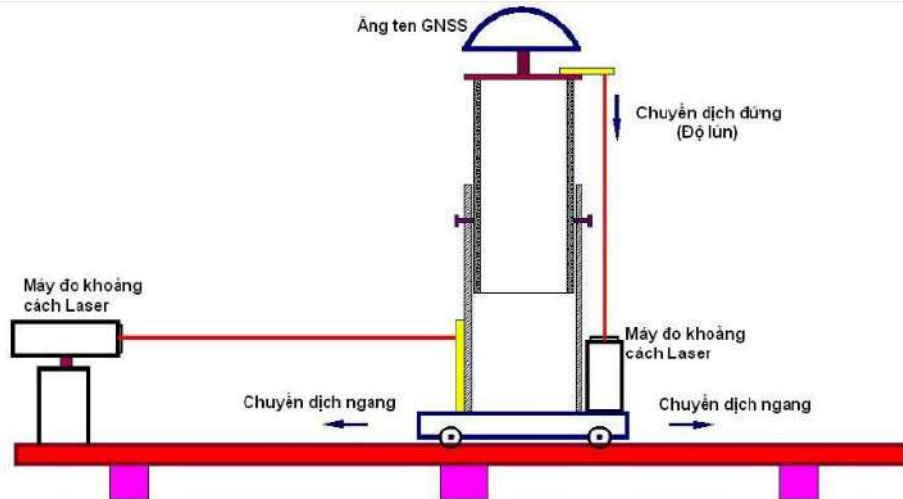
Hình 3. Giao diện phần mềm xử lý dữ liệu quan trắc trượt lở đất đá theo thời gian thực

3. Thực nghiệm và kết quả quan trắc trượt lở đất đá dựa trên công nghệ GNSS/CORS

3.1. Thực nghiệm mô phỏng quan trắc trượt lở đất đá

3.1.1. Nguyên lý thực nghiệm mô phỏng

Để kiểm chứng, đánh giá độ chính xác của hệ thống quan trắc trượt lở đất đá dựa trên công nghệ GNSS/CORS, một thực nghiệm mô phỏng đã được tiến hành ở khu vực thành phố Cẩm Phả, Quảng Ninh, ở đó nhóm nghiên cứu đã thiết lập một trạm CORS có số hiệu CORS-CAMPHA. Nguyên lý của thực nghiệm mô phỏng trượt lở đất đá là dựa trên sự thay đổi vị trí của ăng ten GNSS theo phương nằm ngang và phương thẳng đứng (Hình 4).



Hình 4. Nguyên lý quan trắc thực nghiệm mô phỏng trượt lở đất đá bằng công nghệ GNSS/CORS

Hệ thống thiết bị mô phỏng có thể đánh giá được cả chuyển dịch ngang và chuyển dịch đứng. Để tiến hành thực nghiệm mô phỏng một hệ thống thiết bị đã được thiết kế chế tạo gồm có hai ống thép hình trụ lồng vào nhau. Ống thép ngoài cố định và ống thép trong có thể di chuyển theo phương thẳng đứng bên trong lòng ống thép ngoài, phía trên ống thép trong có gắn một đĩa thép tròn, tâm của đĩa thép này có khoét một lỗ để lắp ăng ten thu tín hiệu vệ tinh GNSS. Toàn bộ trụ móc này được gắn cố định với 4 bánh xe. Các bánh xe này có thể di chuyển ngang trên một đường ray nằm ngang dùng để xác định chuyển dịch ngang. Sử dụng máy đo khoảng cách Laser cầm tay SD-120C của hãng SINCON để đo khoảng dịch chuyển theo phương ngang và theo phương thẳng đứng từ điểm tham chiếu cố định đến trụ móc có lắp ăng ten GNSS. Sau khi thu tín hiệu vệ tinh trong khoảng thời gian khoảng 30' thì thay đổi vị trí ăng ten theo phương ngang và phương thẳng đứng, rồi lại đo khoảng cách từ điểm tham chiếu đến trụ móc có lắp ăng ten và lại thu tín hiệu vệ tinh ở vị trí mới. Hiệu khoảng cách đo được giữa hai lần đo theo phương nằm ngang và phương thẳng đứng chính là độ chuyển dịch ngang và chuyển dịch đứng. Dữ liệu thu được ở trạm quan trắc sẽ được xử lý qua 6 bước như đã trình bày ở trên, từ đó xác định được độ chuyển dịch đất đá dựa trên công nghệ GNSS/CORS theo công thức (3). So sánh, đối chứng độ chuyển dịch xác định bằng hai phương pháp này sẽ đánh giá được độ tin cậy, độ chính xác của công nghệ GNSS/CORS.

3.1.2. Tiến hành thực nghiệm mô phỏng

Việc quan trắc thực nghiệm mô phỏng được tiến hành vào ngày 15 tháng 7 năm 2023 ở thành phố Cẩm Phả, Quảng Ninh với 6 lần thay đổi vị trí của ăng ten GNSS. Sau mỗi lần thay đổi vị trí ăng ten GNSS theo phương nằm ngang và thẳng đứng một cách chủ động để tạo ra độ chuyển dịch ngang và chuyển dịch đứng. Vị trí đặt trạm quan trắc thực nghiệm mô phỏng nằm bên lề đường bao biển Cẩm Phả - Hạ Long cách trạm CORS-CAMPHA khoảng 3,40 km (Hình 5).

Trạm quan trắc được kết nối với trạm CORS-CAMPHA thông qua địa chỉ IP của máy tính chủ là 118.70.171.179 và cổng nhận dữ liệu là 3004. Dữ liệu quan trắc theo định dạng tiêu chuẩn NMEA được gửi về máy tính chủ đặt ở Trường Đại học Mỏ - Địa chất, Hà Nội. Trong thực nghiệm mô phỏng này, nhóm nghiên cứu đã tiến hành thực hiện 6 lần chuyển dịch chủ động, mỗi lần lần đo với thời gian thu tín hiệu vệ tinh là 30 phút với tần suất thu tín hiệu là 1 giây. Dữ liệu của mỗi lần quan trắc được lưu vào một tệp riêng biệt với hai loại tin nhắn trị đo thu được là GGA và GST.



a) Trạm quan trắc thực nghiệm mô phỏng b) Trạm CORS- CAMPHA

Hình 5. Khu vực bố trí trạm quan trắc thực nghiệm mô phỏng

3.1.3. Dữ liệu và kết quả quan trắc thực nghiệm mô phỏng

Dữ liệu quan trắc nhận được theo định dạng tiêu chuẩn NMEA với hai loại tin nhắn là GGA và GST. Ở (Bảng 3) là một đoạn dữ liệu quan trắc thực nghiệm mô phỏng ở lần quan trắc thứ nhất.

Bảng 3. Một đoạn dữ liệu quan trắc mô phỏng ở lần quan trắc thứ nhất.

```
$GNGGA,045915.00,2059.77350765,N,10716.19035748,E,1,10,0.9,21.233,M,-23.050,M,*,*66
$GNGST,045917.00,8.083,2.096,1.647,1.6,2.096,1.648,4.100*49
$GNGGA,045922.00,2059.77299225,N,10716.18966106,E,1,11,0.8,15.804,M,-23.050,M,*,*6C
$GNGST,045922.00,8.005,1.893,1.497,177.9,1.893,1.498,3.565*4E
$GNGGA,045927.00,2059.77228694,N,10716.18956554,E,1,20,0.6,14.195,M,-23.050,M,*,*61
$GNGST,045927.00,9.950,1.409,1.262,155.7,1.385,1.289,2.801*4A
$GNGGA,045932.00,2059.77212546,N,10716.18967398,E,1,22,0.6,13.296,M,-23.050,M,*,*61
$GNGST,045932.00,9.242,1.373,1.171,165.9,1.362,1.184,2.646*4A
$GNGGA,045937.00,2059.76848666,N,10716.18868776,E,4,19,0.6,4.153,M,-23.050,M,1.0,0000*75
$GNGST,045937.00,0.031,0.004,0.003,163.7,0.004,0.003,0.010*45
$GNGGA,045927.00,2059.77228694,N,10716.18956554,E,1,20,0.6,14.195,M,-23.050,M,*,*61
$GNGST,045927.00,9.950,1.409,1.262,155.7,1.385,1.289,2.801*4A
$GNGGA,045932.00,2059.77212546,N,10716.18967398,E,1,22,0.6,13.296,M,-23.050,M,*,*61
$GNGST,045932.00,9.242,1.373,1.171,165.9,1.362,1.184,2.646*4A
$GNGGA,045937.00,2059.76848666,N,10716.18868776,E,4,19,0.6,4.153,M,-23.050,M,1.0,0000*75
.....
```

Dữ liệu quan trắc được xử lý bằng phần mềm Server GNSS CORS WDM theo 6 bước đã trình bày ở Mục 2.2.2 và xác định được tọa độ và độ cao trạm quan trắc cũng như độ phân tán của nó. Ở (Bảng 4) thể hiện kết quả xử lý dữ liệu quan trắc độ lún ở chu kỳ thứ nhất.

Dữ liệu quan trắc thực nghiệm mô phỏng được xử lý qua 6 bước như đã trình bày ở trên bằng phần mềm Server GNSS CORS WDM được kết quả như ở (Bảng 4).

Bảng 4. Kết quả xử lý dữ liệu ở lần quan trắc thứ nhất

TT	Tọa độ trắc địa			Tọa độ VN-2000			Độ lệch chuẩn (mm)		
	B (°, ', ")	L (°, ', ")	H (m)	X(m)	Y(m)	h(m)	Vx	Vy	Vh
1	20 59 46,10905	107 16 11,32120	4,176	2323503,981	735754,164	4,176	-0,5	-3,8	7,3
2	20 59 46,10905	107 16 11,32116	4,189	2323503,981	735754,163	4,189	-0,5	-2,6	-5,7
3	20 59 46,10909	107 16 11,32120	4,184	2323503,983	735754,164	4,184	-1,9	-3,7	-0,7
4	20 59 46,10903	107 16 11,32124	4,182	2323503,981	735754,165	4,182	0,0	-5,0	1,3
5	20 59 46,10913	107 16 11,32117	4,179	2323503,984	735754,163	4,179	-3,0	-2,8	4,3
6	20 59 46,10903	107 16 11,32122	4,187	2323503,981	735754,164	4,187	0,0	-4,3	-3,7
7	20 59 46,10908	107 16 11,32122	4,182	2323503,982	735754,164	4,182	-1,5	-4,4	1,3
8	20 59 46,10913	107 16 11,32120	4,189	2323503,984	735754,164	4,189	-2,9	-3,9	-5,7
9	20 59 46,10911	107 16 11,32120	4,179	2323503,983	735754,164	4,179	-2,3	-3,7	4,3
10	20 59 46,10893	107 16 11,32113	4,182	2323503,978	735754,162	4,182	3,2	-1,7	1,3
...
190	20 59 46,10892	107 16 11,32099	4,189	2323503,977	735754,158	4,189	3,5	2,1	-5,7
191	20 59 46,10895	107 16 11,32099	4,189	2323503,978	735754,158	4,189	2,5	2,3	-5,7
192	20 59 46,10890	107 16 11,32105	4,185	2323503,977	735754,160	4,185	4,0	0,4	-1,7
193	20 59 46,10889	107 16 11,32109	4,188	2323503,976	735754,161	4,188	4,4	-0,8	-4,7
194	20 59 46,10888	107 16 11,32110	4,184	2323503,976	735754,161	4,184	4,6	-0,9	-0,7
195	20 59 46,10890	107 16 11,32109	4,189	2323503,977	735754,161	4,189	4,2	-0,8	-5,7
196	20 59 46,10887	107 16 11,32108	4,188	2323503,976	735754,160	4,188	4,9	-0,3	-4,7
197	20 59 46,10894	107 16 11,32106	4,188	2323503,978	735754,160	4,188	2,9	0,1	-4,7
198	20 59 46,10890	107 16 11,32108	4,188	2323503,977	735754,160	4,188	4,1	-0,3	-4,7
199	20 59 46,10889	107 16 11,32105	4,184	2323503,977	735754,159	4,184	4,2	0,6	-0,7
TB	20 59 46,10903	107 16 11,32107	4,183	2323503,981	735754,160	4,183	0,0	0,2	-0,9

Độ lệch chuẩn của các thành phần tọa độ:

$$\sigma_x = 1,96 \text{ mm}$$

$$\sigma_y = 1,36 \text{ mm}$$

$$\sigma_h = 3,92 \text{ mm}$$

Dữ liệu quan trắc thực nghiệm mô phỏng ở các lần tiếp theo cũng được xử lý tương tự như ở lần thứ nhất. Từ kết quả xử lý dữ liệu quan trắc thực nghiệm mô phỏng của sáu lần đo, dựa vào tọa độ và độ cao của ăng ten xác định được, sẽ tính được độ chuyển dịch ngang, chuyển dịch đứng và chuyển dịch toàn phần. Các đại lượng chuyển dịch này được so sánh với giá trị chuyển dịch đo trực tiếp bằng thiết bị đo Laser. Các đại lượng chuyển dịch dựa và kết quả quan trắc và đại lượng chuyển dịch đo trực tiếp cũng như độ chênh lệch của nó được thể hiện như ở Bảng 5.

Bảng 5. Kết quả quan trắc thực nghiệm mô phỏng

TT	Chuyển dịch xác định bằng công nghệ GNSS/CORS				Chuyển dịch đo trực tiếp		Độ chênh lệch chuyển dịch	
	Q _x (m)	Q _y (m)	S (m)	Q _p (m)	Q _p (đo) (m)	S (đo) (m)	ΔQ _p (m)	ΔS (m)
1	0,032	-0,012	-0,068	0,034	0,030	-0,057	0,004	0,011
2	0,057	-0,007	-0,009	0,057	0,060	-0,015	0,003	0,006
3	0,064	0,006	-0,056	0,064	0,073	-0,040	0,009	0,016
4	0,105	0,003	-0,064	0,105	0,109	0,049	0,004	0,015
5	0,182	0,013	-0,074	0,182	0,188	-0,088	0,007	0,014

Từ kết quả quan trắc thực nghiệm mô phỏng như ở Bảng 3, nhận thấy rằng chênh lệch về chuyển dịch ngang lớn nhất là 9 mm và chuyển dịch đứng lớn nhất là 16 mm.

3.2. Thực nghiệm thực tế

Thực nghiệm được tiến hành trên một khu vực bờ mỏ phía Bắc của mỏ than Cọc Sáu thuộc địa phận thành phố Cẩm Phả, tỉnh Quảng Ninh (Hình 6). Thời gian bắt đầu quan trắc từ ngày 16 tháng 7 năm 2023 và quan trắc liên tục trong thời gian dài. Độ cao trung bình bờ mỏ đặt trạm quan trắc khoảng +250 m, độ cao đáy mỏ đến thời điểm hiện tại là -300 m. Như vậy độ chênh cao từ bề mặt bờ mỏ quan trắc đến đáy mỏ khoảng 550 m. Với độ chênh cao lớn như vậy, cộng thêm với các hoạt động khai thác làm cho bề mặt bờ mỏ chuyển dịch, trượt lở rất nhanh. Để tiến hành quan trắc thực nghiệm ngoài thực địa, có 5 mốc đã được thiết kế gia công bằng thép không gỉ dễ lắp được ăng ten GNSS theo hình thức định tâm bắt buộc.



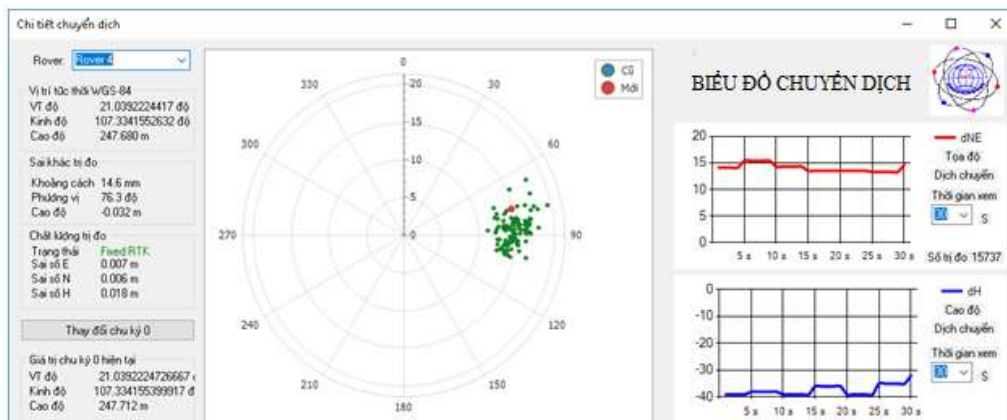
Hình 6. Vị trí khu vực quan trắc thực nghiệm (mỏ than Cốc Sáu, thành phố Cẩm Phả, tỉnh Quảng Ninh)

Các mốc được thi công và gắn lên bề mặt bờ mỏ bằng vữa xi măng ở những vị trí đã được thiết kế, lựa chọn và được đánh số từ P1 đến P5. Sau khi xi măng đông cứng, các mốc đã được gắn chắc chắn vào bề mặt bờ tầng thì tiến hành lắp đặt thiết bị cho trạm quan trắc. Ở (Hình 7) thể hiện một trạm quan trắc trượt lở đã được lắp đặt hoàn chỉnh trên bề mặt bờ mỏ Cốc Sáu.



Hình 7. Trạm quan trắc trượt lở bờ mỏ đã được lắp đặt

Các thao tác thực hiện tại trạm quan trắc trên bề mặt bờ mỏ cũng tương tự như thực nghiệm mô phỏng. Dữ liệu quan trắc theo định dạng tiêu chuẩn NMEA được tự động truyền về máy tính chủ của trạm CORS và được xử lý để xác định tọa độ và độ cao cho các mốc quan trắc. Quá trình xử lý dữ liệu quan trắc chuyển dịch bề mặt bờ mỏ cũng thực hiện tương tự như xử lý dữ liệu quan trắc mô phỏng, được thực hiện bằng phần mềm Server GNSS CORS WDM. Các thông tin về tọa độ, độ cao, cũng như sai số định vị và biểu đồ chuyển dịch được xử lý tức thời và hiển thị trên giao diện phần mềm được cài đặt trên máy tính chủ (Hình 8).



Hình 8. Thông tin chi tiết quan trắc dịch chuyển, trượt lở đất đá theo thời gian thực

Việc theo dõi quá trình chuyển dịch, trượt lở đất đá có thể được thực hiện bằng cách đăng nhập vào máy tính chủ theo địa chỉ IP và cổng kết nối đã được mở bằng điện thoại thông minh, do đó có thể theo dõi quá trình dịch chuyển, trượt lở đất đá theo thời gian thực ở mọi lúc, mọi nơi.

4. Kết luận

Nghiên cứu này đã ứng dụng thành công công nghệ GNSS/CORS trong quan trắc chuyển dịch, trượt lở bờ mỏ lộ thiên. Hệ thống quan trắc dựa trên công nghệ GNSS/CORS đã được thiết kế bao gồm hệ thống trạm CORS và hệ thống trạm quan trắc. Một đóng góp của nghiên cứu này là thiết kế phát triển được bộ thu GNSS cho hệ thống trạm quan trắc dựa trên nền tảng công nghệ của hãng Trimble. Bộ thu GNSS được phát triển thu được tín hiệu vệ tinh trong hệ thống GNSS, nhận được số liệu cải chính từ trạm CORS, giải mã tín hiệu vệ tinh GNSS để chuyển về dữ liệu theo định dạng tiêu chuẩn NMEA-0183. Dữ liệu quan trắc được phân tích và xử lý qua 6 bước và được thực hiện bằng phần mềm tự phát triển đã nâng cao được độ chính xác kết quả quan trắc đến milimet.

Để đánh giá độ chính xác của hệ thống quan trắc, một thực nghiệm quan trắc mô phỏng chuyển dịch, trượt lở đất đá được tiến hành ở khu vực thành phố Cẩm Phả, Quảng Ninh. Qua thực nghiệm mô phỏng cho thấy độ chênh lệch lớn nhất về chuyển dịch ngang là 9 mm và chuyển dịch đứng là 16 mm. Thực nghiệm với số liệu thực tế được thực hiện trên bờ mỏ của mỏ than Cọc Sáu với 5 trạm quan trắc thực hiện trong một thời gian dài. Dữ liệu quan trắc được xử lý và xác định được tọa độ mặt bằng và độ cao của các mốc quan trắc với độ lệch chuẩn nhỏ hơn 4 mm. Các thông tin chi tiết của trạm quan trắc cũng như biểu đồ chuyển dịch được hiển thị trực quan trên màn hình máy tính chủ theo thời gian thực.

Tài liệu tham khảo

- Z. Abidin H, Andreas H, Gamal M, Surono S, Hendrasto M (2004). Studying Landslide Displacements in Megamendung (Indonesia) Using GPS Survey Method. *Itbj.Eng.Sci* 36(2):109-123.
- M.-B. Su, I.-H. Chen, C.-H. Liao (2009). Using TDR Cables and GPS for landslide monitoring in high mountain area. *J. Geotech. Geoenviron. Eng.*, 135(8), pp. 1113-1121.
- J.A. Gili, J. Corominas, J. Rius. (2000). Using global positioning system techniques in landslide monitoring. *Eng. Geol.*, 55(3), pp. 167-192.
- Cina, M. Piras (2015). Performance of low-cost GNSS receiver for landslides monitoring: test and results. *Geomat. Nat. Haz. Risk*, 6(5-7), pp. 497-514.
- T. Bellone, P. Dabove, A.M. Manzano, C. Taglioretti (2016). Real-time monitoring for fast deformations using GNSS low-cost receivers. *Geomat. Nat. Haz. Risk*, 7(2), pp. 458-470.
- W.Y. Zhao, M.Z. Zhang, J. Ma, B. Han, S.Q. Ye, Z. Huang (2021). Application of CORS in landslide monitoring. *IOP Conf. Ser. Earth Environ. Sci.*, 861(4), Article 042049.
- N. Shen, L. Chen, L. Wang, H. Hu, X. Lu, C. Qian, J. Liu, S. Jin, R. Chen. (2021). Short-term landslide displacement detection based on GNSS real-time kinematic positioning *IEEE Trans. Instrum. Meas.*, 70, pp.1-14.
- Y. Du, G. Huang, Q. Zhang, Y. Gao, Y. Gao. (2020). Asynchronous RTK Method for detecting the stability of the reference station in GNSS deformation monitoring. *Sensors*, 20(5), p. 1320.
- Bao Shu, Yuanhao He, Li Wang, Qin Zhang, Xinrui Li, Xuanyu Qu, Guanwen Huang, Wei Qu. Real-time high-precision landslide displacement monitoring based on a GNSS CORS network. *Measurement*, Volume 217, August 2023, 113056.
- Pengxu Wang, Hui Liu, Guigen Nie, Zhixin Yang, Jiaji Wu, Chuang Qian, Bao Shu (2022). Performance evaluation of a real-time high-precision landslide displacement detection algorithm based on GNSS virtual reference station technology. *Measurement*, Volume 199, August 2022, 111457.
- J. Paziewski (2016). Study on desirable ionospheric corrections accuracy for network-RTK positioning and its impact on time-to-fix and probability of successful single-epoch ambiguity resolution, *Adv. Space Res.* 57(4), 1098-1111.
- J. Cui, W. Tang, L. Jin, C. Deng, X. Zou, S. Gu (2018). An improved ionosphere interpolation algorithm for network RTK in low-latitude regions, *GPS Solut.* 22(4), 109.
- R. Odolinski, P.J.G. Teunissen, D. Odijk. Combined GPS+ BDS for short to long baseline RTK positioning, *Meas. Sci. Technol.* 26(4) (2015), 045801.
- Daniel Janos, Przemysław Kuras, Łukasz Ortyl (2022). Evaluation of low-cost RTK GNSS receiver in motion under demanding conditions. *Measurement*, Vol. 201, 30 September, 111647.
- Đặng Nam Chính, Đỗ Ngọc Đường (2012). Giáo trình định vị vệ tinh. Nhà xuất bản Khoa học và Kỹ thuật, Hà Nội.
- Hoàng Ngọc Hà, Trương Quang Hiếu (2003). Cơ sở toán học xử lý số liệu trắc địa. Nhà xuất bản Giao thông vận tải, Hà Nội.
- Phạm Công Khải và nnk., (2023). Nghiên cứu thiết kế, chế tạo hệ thống quan trắc chuyển dịch công trình theo thời gian thực trên địa bàn thành phố Hà Nội. Báo cáo tổng hợp kết quả nghiên cứu đề tài cấp thành phố, mã số 01C-01/02-2020-3. Sở Khoa học và Công nghệ Hà Nội.

ABSTRACT

Real-time landslide monitoring at open-pit mining slope based on GNSS/CORS technology

Pham Cong Khai^{1*}, Nguyen Van Hai²

¹*Hanoi University of Mining and Geology*

²*Thuy Loi University (Second campus), Ho Chi Minh City*

Real-time monitoring for early warning of landslides at open-pit mining slope is a very important issue to ensure the mining process is effective, economical and safe. Nowadays, there are many methods for landslide monitoring based on different technologies and equipment. However, landslide monitoring based on GNSS/CORS technology has many advantages such as monitoring in all weather conditions, continuously and in real time in three-dimensional space. The purpose of this study is to design and build a real-time monitoring system based on GNSS/CORS technology for early warning of landslides at open-pit mining levels. In this study, a real-time landslide monitoring system is designed and built, including a CORS station system and a monitoring station system. The CORS station system was designed and built in Cam Pha city, Quang Ninh with Trimble's Zephyr 2 Geodetic antenna and Stonex's GNSS receiver. The monitoring station system was self-designed and developed in both hardware and software, ensuring the process of data collection, transmission, processing and warning in real time. The monitoring data collected in standard NMEA0183 format is processed by self-designed and developed methods and software, improving the accuracy of monitoring results to mm. A simulated landslide monitoring experiment was conducted in Cam Pha city, Quang Ninh and a real experiment at Coc Sau coal mine slope confirmed the feasibility of the solution in real-time landslide monitoring.

Keywords: GNSS/CORS Technology, Landslide, Real-time, Coc Sau coal mine.

Dự báo thay đổi lớp phủ bề mặt bằng thuật toán học máy

Trần Thị Hòa^{1*}, Trần Thị Ngọc¹, Trần Đình Trí²

¹*Trường Đại học Mở - Địa chất*

²*Hội Trắc địa - Bản đồ - Viễn thám Việt Nam*

TÓM TẮT

Lớp phủ mặt đất là bề mặt hình thái vật lý của lớp vỏ Trái đất, bao gồm thực vật, nước, bề mặt tự nhiên và những đối tượng nhân tạo. Những hoạt động diễn ra trên bề mặt Trái đất từ các hoạt động tự nhiên cho đến hoạt động phát triển kinh tế - xã hội, đô thị hóa đều dẫn đến sự thay đổi đáng kể của lớp phủ bề mặt. Tài nguyên đất là một trong những nguồn tài nguyên quan trọng và có ảnh hưởng sâu sắc đến môi trường nói chung và mục tiêu phát triển bền vững nói riêng, vì vậy, việc theo dõi và giám sát sự thay đổi của lớp phủ bề mặt, cũng như phân tích dự báo xu hướng thay đổi là một nhiệm vụ quan trọng. Các thuật toán học máy hiện nay đã có những đóng góp đáng kể trong việc tận dụng khả năng của máy tính và đảm bảo tính khách quan trong phân tích các vấn đề liên quan đến dự đoán. Bài báo xem xét khả năng ứng dụng của một số thuật toán học máy (Machine Learning) khác nhau trong dự đoán biến động lớp phủ bề mặt, cụ thể: cây quyết định (Decision Trees), rừng ngẫu nhiên (Random Forest) và vector máy hỗ trợ (SVM).

Từ khóa: Biến động lớp phủ, dự đoán, học máy, Random Forest, SVM.

1. Đặt vấn đề

Lớp phủ bề mặt là một thuật ngữ chung dùng để chỉ các vật liệu hoặc các yếu tố tự nhiên và nhân tạo bao phủ bề mặt Trái đất, bao gồm: thực vật (rừng, đồng cỏ, hay cây bụi); nước (sông suối, ao hồ hay biển); đất canh tác (đất trồng lúa, đất trồng màu hay đất trồng cây ăn quả - cây lâu năm); đất trống (cát, hoang mạc, đất không có cây cỏ); đất đô thị và công nghiệp (đường xá, khu vực xây dựng, khu vực có công trình hạ tầng) và khu vực bị băng tuyết bao phủ quanh năm (Giri, 2020). Những yếu tố này phân bố xen kẽ nhau tạo nên sự đa dạng và hình thái đặc trưng của mỗi vùng địa lý trên bề mặt Trái đất. Sự phân bố của các lớp phủ đóng vai trò quan trọng trong hệ sinh thái và môi trường.

Sự thay đổi của lớp phủ bề mặt có thể diễn ra dưới nhiều hình thức khác nhau, ví dụ sự thay mặt phân bố không gian hay diện tích hay sự biến đổi về cấu trúc - thành phần của một hay nhiều loại lớp phủ theo thời gian. Ví dụ như: cùng là lớp phủ thực vật nhưng từ rừng chuyển sang đồng cỏ hay sự thay đổi từ lớp phủ thực vật sang lớp phủ đất đô thị và công nghiệp. Nguyên nhân dẫn đến sự thay đổi lớp phủ bề mặt chia thành hai nhóm - yếu tố tự nhiên khách quan và yếu tố chủ quan - con người. Các yếu tố tự nhiên có thể gây ra sự thay đổi về lớp phủ một cách từ từ như sự thay đổi về khí hậu (nhiệt độ và lượng mưa) hoặc đột ngột như sự xuất hiện của các hiện tượng thiên tai như lũ lụt, hạn hán, cháy rừng hay động đất (Gregorio & Jansen, 2005) (Giri, 2020). Bên cạnh đó, các yếu tố về con người có thể kể đến như quá trình khai thác tài nguyên thiên nhiên, công nghiệp hóa, đô thị hóa, canh tác nông nghiệp hay các chính sách quản lý tài nguyên. Nhìn chung, những thay đổi về lớp phủ đều có những ảnh hưởng nhất định đến môi trường và xã hội, trong đó bao gồm giảm đa dạng sinh học, giảm chất lượng nước và không khí, biến đổi khí hậu và sự phát triển kinh tế vùng. Vì thế, công tác giám sát và quản lý lớp phủ bề mặt là công tác quan trọng đảm bảo sự phát triển bền vững và bảo vệ môi trường.

Những nghiên cứu về biến động lớp phủ thường chủ yếu liên quan đến sự thay đổi về mặt phân bố không gian (diện tích) ứng dụng công nghệ hiện đại như viễn thám, viễn thám đa thời gian, kết hợp với hệ thống tin địa lý (GIS) (Giri, 2020). Một số nghiên cứu nổi bật có thể kể đến như nghiên cứu về sự biến động rừng nhiệt đới amazon sử dụng dữ liệu vệ tinh được cung cấp bởi Cơ quan Vũ trụ hàng không quốc gia Hoa Kỳ (NASA) và Cơ quan Vũ trụ liên minh châu Âu (ESA) (Hansen et al., 2013). Nghiên cứu đã chỉ tốc độ phá rừng nhanh trong những năm gần đây thông qua sự mất mát phần lớn diện tích rừng và sự thay thế bởi đất trồng, đất nông nghiệp và chăn nuôi. Sự thay đổi này làm phá vỡ cấu trúc bền vững của hệ sinh thái rừng tự nhiên, gây ảnh hưởng tiêu cực đến đa dạng sinh học và thậm chí làm ảnh hưởng đến khí hậu toàn cầu. Hay nghiên cứu về ảnh hưởng của hiện tượng hoang mạc hóa ở châu Phi (Herrmann & Hutchinson, 2005) dẫn đến sự sụt giảm đáng kể của đất canh tác và lan rộng của sa mạc. Liên quan đến đô thị hóa, các rất nhiều nghiên cứu từ ảnh vệ tinh chỉ ra rằng tốc độ mở rộng đô thị nhanh ở những thành phố lớn đã làm gia tăng nhanh chóng các khu vực xây dựng, giảm thiểu diện tích đất nông nghiệp và các vùng xanh (Seto, Güneralp & Hutyrá, 2012) (Phạm & Nguyễn, 2023) (Nguyễn & Đặng, 2021) (Trần T. T., 2020). Những thay đổi trên có tác động lớn đến suy thoái môi trường, ảnh hưởng đến sinh kế, hiện tượng ngập lụt cục bộ và ô nhiễm không khí. Vì vậy, việc dự đoán sự thay đổi về lớp phủ bề mặt là

* Tác giả liên hệ

Email: hoatran106@gmail.com

điều cần thiết và có ý nghĩa quan trọng trong nhiều khía cạnh khác nhau, đặc biệt liên quan đến quản lý tài nguyên thiên nhiên và quy hoạch - phát triển bền vững.

Trong những năm gần đây, việc sử dụng trí tuệ nhân tạo (AI) mang lại nhiều vượt trội so với mô hình GIS truyền thống (Kumar & Sharma, 2023). Ví dụ như mô hình AI có khả năng xử lý và phân tích lượng lớn dữ liệu từ nhiều nguồn trong thời gian ngắn và phát hiện nhanh chóng thay đổi với thời gian thực. Bên cạnh đó, các thuật toán học sâu cho phép khả năng nhận diện và phân loại lớp phủ bề mặt với độ chính xác cao hơn (Zhang, Li, & Wang, 2023) (Johnson & Lee, 2023) (García & Rodríguez, 2022). Các mô hình AI có khả năng tự động cải thiện độ chính xác qua thời gian thông qua việc học từ dữ liệu thực tế. Ngoài ra, mô hình AI rất linh hoạt, có thể áp dụng được trong nhiều ngữ cảnh khác nhau, trong khi mô hình truyền thống thường bị giới hạn (Zhou & Liu, 2022). Ví dụ như không thể dùng mô hình dự báo sa mạc hóa truyền thống vào giám sát đô thị được, yếu tố can thiệp của con người cũng là một trong những nguyên nhân gây sai sót trong mô hình truyền thống mà AI có thể giảm thiểu được thông qua khả năng ứng biến và tự động hóa (Patel & Desai, 2021).

Nội dung của bài báo đi sâu vào tìm hiểu ưu nhược điểm và khả năng ứng dụng của một số thuật toán học sâu trong mô hình trí tuệ nhân tạo trong dự báo sự biến động của lớp phủ trong tương lai gần. Các thuật toán này khá thông dụng trong phân loại lớp phủ tự động như cây quyết định (Decision Trees), rừng ngẫu nhiên (Random Forest) hay máy hỗ trợ vector (SVM). Qua đó, đề xuất quy trình ứng dụng các thuật toán này trong dự báo biến động lớp phủ với việc xác định các tham số đầu vào vào nguồn tư liệu khai thác.

2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1. Lớp phủ bề mặt và sự thay đổi về lớp phủ bề mặt

Bề mặt tự nhiên của Trái đất là một tập hợp đa dạng các dạng địa hình và hệ sinh thái, bao gồm rừng, núi, sa mạc, đại dương và sông ngòi. Sự phân bố xen kẽ về mặt không gian và mối tác động tương quan với khí hậu và con người đã tạo ra những hình thái đặc trưng khác nhau về địa hình và sinh thái ở các khu vực địa lý khác nhau mà thể hiện bên ngoài của nó là lớp phủ bề mặt. Lớp phủ bề mặt Trái đất bao gồm hình thái của các thành phần tự nhiên và nhân tạo như đất, cây cối, đá, băng và các công trình xây dựng. Việc nghiên cứu lớp phủ bề mặt Trái đất là một phần quan trọng trong các lĩnh vực khoa học như địa lý, sinh thái học, nông nghiệp và môi trường.

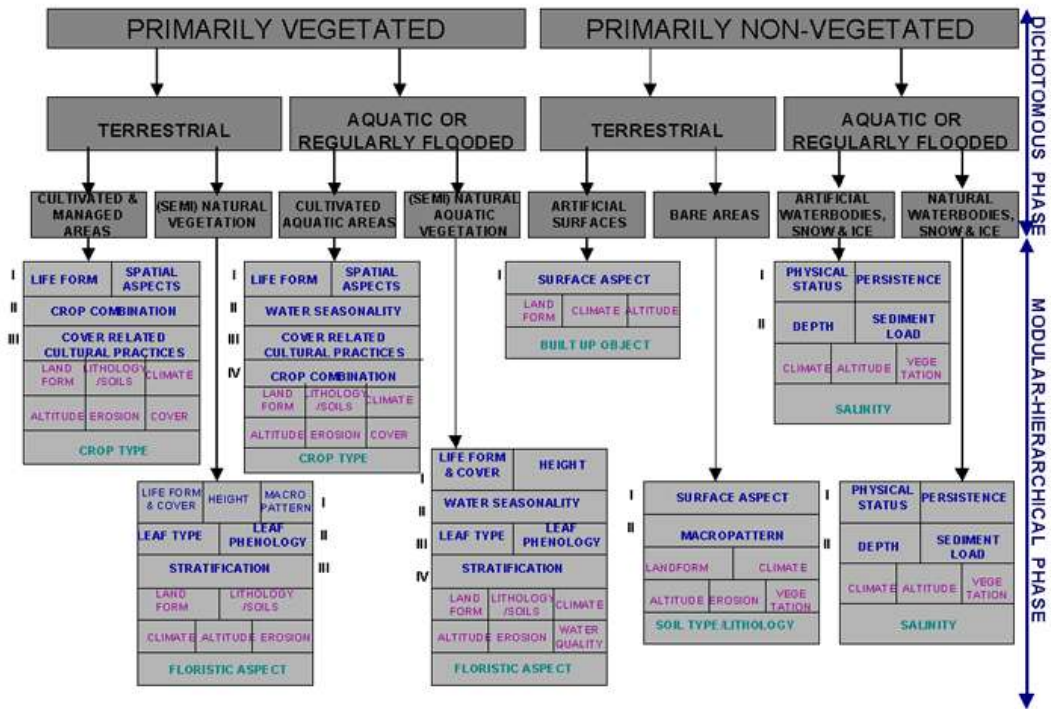
Mỗi một lĩnh vực nghiên cứu đều có cách thức phân loại lớp phủ bề mặt Trái đất với mức độ chi tiết khác nhau. Ví dụ như phân loại theo hệ sinh thái, lớp phủ được chia thành 05 nhóm: rừng, đồng cỏ, sa mạc, băng và nước, Phương pháp phân loại theo mục đích sử dụng, lớp phủ bề mặt được chia thành 04 nhóm lớn, bao gồm đất nông nghiệp, đô thị, rừng và đất không sử dụng. Còn đối với đặc điểm địa hình, lớp phủ chia thành 03 nhóm: địa hình cao, địa hình thấp và địa hình đặc biệt như hang động hay hẻm núi. Tổ chức Nông Lương thế giới (FAO) có cách phân chia theo ý nghĩa sử dụng của đất (Thackway & Atyeo, 2006); USGS - Cục khảo sát địa chất Hoa Kỳ có phương pháp riêng để phân loại lớp phủ bề mặt dựa trên đặc điểm đối tượng tự nhiên xuất hiện trên bề mặt lớp phủ. Tương tự, Cơ quan Vũ trụ châu Âu (ESA) cũng lựa chọn phương pháp phân chia lớp phủ của tổ chức liên hợp quốc dựa trên loại thực vật có trên bề mặt (the UN LCCS) (Gregorio & Jansen, 2005). Bên cạnh đó, để phù hợp với đặc điểm địa lý của từng khu vực cũng như mục đích sử dụng, mỗi quốc gia đều xây dựng hệ phân loại lớp phủ - lớp sử dụng đất khác nhau.

Về cơ bản, việc phân loại lớp phủ bề mặt nhằm mục đích để theo dõi, giám sát những tác động của tự nhiên và con người được thuận tiện và dễ dàng hơn. Qua đó, công tác bảo vệ, quản lý và quy hoạch sẽ đạt được.

Bảng 1 và Hình 1 thể hiện hệ thống phân loại mà FAO hiện đang sử dụng hiện nay trong phân loại lớp phủ bề mặt và lớp hiện trạng sử dụng đất đối với các tài liệu và bản đồ công bố bởi tổ chức này. Đây cũng là hệ thống phân loại khá phổ biến được áp dụng trong các nghiên cứu về biến động hay hiện trạng lớp phủ hiện nay. Giai đoạn phân đôi (dichotomous phase) xác định ba cấp phân loại dựa theo thực vật (Bảng 1), còn giai đoạn phân cấp (modular - hierarchical phase) (Gregorio & Jansen, 2005) thì mỗi lớp ở cấp 3 của giai đoạn 1 sẽ chia thành nhiều loại hơn nữa phụ thuộc vào tính chất của môi trường (như khí hậu, loại đất, khả năng xói mòn) và tính chất kỹ thuật, ví dụ như trong đất nông nghiệp có đất trồng lúa hay đất trồng màu. Đặc điểm này cho phép những nghiên cứu về sao không bị giới hạn về số lượng lớp phủ mà tùy thuộc vào khu vực, các nhà nghiên cứu có thể bổ sung thêm định nghĩa của lớp phủ bề mặt mới.

Bảng 1. Hệ phân loại phân đôi UN LCCS (Dichotomous Phase) (Thackway & Atyeo, 2006)

Cấp 1	Cấp 2	Cấp 3
Đất có thực vật	Thực vật trên cạn	Thực vật trên cạn có quản lý
		Thực vật tự nhiên hoặc bán tự nhiên trên cạn
	Thực vật ngập nước hoặc thường xuyên ngập nước	Thực vật vùng trồng ngập nước Thực vật vùng tự nhiên và bán tự nhiên
Đất gần như không có thực vật	Trên cạn	Bề mặt nhân tạo Đất trống
		Hồ nhân tạo, băng và tuyết
	Vùng ngập nước hoặc thường xuyên ngập nước	Vùng nước tự nhiên, băng và tuyết phủ theo mùa



Hình 1. Hệ thống phân loại theo giai đoạn: giai đoạn phân đôi và giai đoạn phân cấp
(Gregorio & Jansen, 2005)

Trong nghiên cứu này, nhóm tác giả chủ yếu đi sâu vào mô hình phân loại của ESA, có kế thừa phương pháp phân loại của UN LCCS. Cụ thể gồm 11 lớp, trình bày ở Bảng 2.

Bảng 2. Hệ thống phân loại của ESA (ESA)

Giá trị	Màu sắc	Tên tiếng Anh	Tên tiếng Việt
10	#006400	Tree cover	Thực phủ
20	#ffbb22	Shrubland	Cây bụi
30	#ffff4c	Grassland	Đồng cỏ
40	#f096ff	Cropland	Đất nông nghiệp
50	#fa0000	Built-up	Đất xây dựng
60	#b4b4b4	Bare/sparse vegetation	Đất trống/ít thực vật
70	#f0f0f0	Snow and ice	Tuyết và băng
80	#0064c8	Permanent water bodies	Nước mặt
90	#0096a0	Herbaceous wetland	Thực vật thân thảo ngập nước
95	#00cf75	Mangroves	Rừng ngập mặn
100	#fae6a0	Moss and lichen	Đất ẩm có rêu và địa y

2.2. Một số thuật toán học máy trong dự đoán thay đổi lớp phủ bề mặt

Học máy và học sâu là hai lĩnh vực con của trí tuệ nhân tạo AI. Các thuật toán học sâu sử dụng các mạng nơ-ron nhân tạo để biểu diễn dữ liệu phức tạp và trừu tượng. Những mạng nơ-ron này thường nhiều tầng và lớp tự động tìm ra các đặc trưng dữ liệu và hạn chế xử lý dữ liệu thủ công. Tuy nhiên, đối với học sâu, dữ liệu cần lớn về cả dung lượng và thời gian để đạt hiệu suất tốt, đồng thời yêu cầu tài nguyên tính toán mạnh mẽ. Với lĩnh vực học máy, các thuật toán thường nhắm đến việc máy tính có thể tự học từ các đặc trưng lựa chọn thủ công mà không cần lập trình rõ ràng từng nhiệm vụ cụ thể. Các thuật toán học máy thường đơn giản và ít tham số hơn nhưng cần kiến thức chuyên gia trong lựa chọn đặc trưng. Chính vì vậy, học máy phù hợp với yêu cầu ít dữ liệu huấn luyện, không cần máy tính cấu hình cao, dễ dàng triển khai và hiệu quả trong các ứng dụng với dữ liệu thực tế. Cụ thể đối với dự đoán thay đổi lớp phủ bề mặt, dữ liệu đầu vào là ảnh vệ tinh đa thời gian nhưng không liên tục, cần có đánh giá của các chuyên gia trong trích xuất đặc trưng từng lớp phủ. Vì vậy, sử dụng học máy là một lựa chọn hợp lý. Dưới đây, các tác giả trình bày ba thuật toán học máy được ứng dụng phổ biến nhất trong phân loại lớp phủ bề mặt, về đề xuất quy trình ứng dụng với đặc điểm cụ thể của khu vực nghiên cứu.

2.2.1. Thuật toán cây quyết định (Decision Tree)

Thuật toán cây quyết định (DT) là một thuật toán học máy mạnh mẽ được sử dụng hiệu quả trong phân loại và dự đoán thay đổi lớp phủ từ dữ liệu ảnh vệ tinh (Pal & Mather, 2003). DT hoạt động bằng cách chia dữ liệu thành các nhánh dựa trên các tiêu chí quyết định, giúp phân loại và dự đoán các loại lớp phủ bề mặt khác nhau. Trong đó mỗi nút trong cây quyết định đại diện cho một thuộc tính của dữ liệu và mỗi nhánh là kết quả của một thuộc tính. Các quyết định này được tiến hành thực hiện cho đến khi đạt được sự thống nhất một dự đoán cụ thể ở lá cây (Breiman, Friedman, Stone & Olshen, 1984).

Công thức của thuật toán này được mô tả như sau (Quinlan, 1986):

- Entropy (S): đo lường mức độ hỗn loạn của tập dữ liệu:

$$H(S) = -\sum_{i=1}^c p_i \log_2 p_i \quad (1)$$

trong đó: p_i là tỷ lệ phần tử thuộc lớp thứ i trong tập dữ liệu S.

- Information Gain (IG): đo lường mức độ giảm hỗn loạn sau khi chia tách tập dữ liệu theo một thuộc tính A.

$$IG(S, A) = H(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} H(S_v) \quad (2)$$

trong đó: S_v là tập con của A với thuộc tính A có giá trị v .

Một trong những ưu điểm nổi bật của DT là tính diễn giải và trực quan giúp cho các nhà nghiên cứu phân định các yếu tố quyết định đến sự thay đổi của lớp phủ bề mặt thông qua xác định yếu tố hồi quy để ra quyết định phân loại.

2.2.2. Thuật toán rừng ngẫu nhiên (Random Forest)

Rừng ngẫu nhiên (RF) là một tập hợp của nhiều cây quyết định được huấn luyện độc lập trên các mẫu dữ liệu khác nhau. Kết quả cuối cùng được quyết định bằng cách lấy trung bình hoặc phiếu đa số của các cây quyết định. Điều này giúp giảm thiểu hiện tượng quá khớp dữ liệu (overfitting) so với việc sử dụng một cây đơn lẻ (Breiman, 2001).

Công thức của thuật toán này được mô tả như sau (Breiman, 2001):

- Bootstrap sampling: mỗi cây trong rừng được huấn luyện trên một mẫu dữ liệu bootstrap (chọn mẫu mới thay thế từ tập dữ liệu huấn luyện).

- Random feature selection: tại mỗi nút chia tách, một tập con ngẫu nhiên của các thuộc tính được chọn để tìm kiếm thuộc tính tốt nhất để chia tách.

- Ensemble Prediction: dự đoán tập hợp:

Đối với bài toán phân loại, kết quả được xác định bằng cách lấy đa phiếu từ các cây:

$$\hat{y} = \text{mode}(\{T_i(x) \mid i = 1, 2, \dots, N\}) \quad (3)$$

trong đó: $T_i(x)$ là dự đoán của cây thứ i cho mẫu x .

Đối với bài toán hồi quy, kết quả được xác định bằng cách lấy trị trung bình của các dự đoán:

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N T_i(x) \quad (4)$$

trong đó, $T_i(x)$ là dự đoán của cây thứ i cho mẫu x .

So với thuật toán cây quyết định, RF linh hoạt hơn và có độ chính xác cao do có tính chất tổng hợp của nhiều cây quyết định. Bên cạnh đó, RF có thể xử lý hiệu quả các bộ dữ liệu lớn và phức tạp, bao gồm cả dữ liệu số và dữ liệu phân loại. Ngoài ra, việc sử dụng nhiều cây quyết định cũng giúp giảm thiểu hiện tượng overfitting, mang lại kết quả dự đoán ổn định hơn đặc biệt với dữ liệu vệ tinh khi mà các pixel nhiễu khó phân loại tương đối phổ biến (Belgiu & Drăguț, 2016) (Rodriguez-Galiano, Ghimire, Chica-Olmo, & Rigol-Sanchez, 2012). Đặc biệt, RF có thể cung cấp thông tin về tầm quan trọng của các biến đầu vào, giúp hiểu rõ hơn về các yếu tố ảnh hưởng đến dự đoán.

2.2.1. Thuật toán vector máy hỗ trợ (Support Vector Machine)

Vector máy hỗ trợ (SVM) là một thuật toán học máy có giám sát, chủ yếu được sử dụng cho các bài toán phân loại và hồi quy. SVM hoạt động bằng cách tìm kiếm một siêu phẳng (hyperplane) tối ưu để phân chia các lớp dữ liệu trong không gian nhiều chiều sao cho khoảng cách từ siêu phẳng đó đến các điểm dữ liệu gần nhất của mỗi lớp là lớn nhất, hay còn gọi là biên an toàn (margin) (Vapnik, 1998).

Về công thức của thuật toán SVM bao gồm các thành phần:

- Siêu phẳng phân tách có dạng:

$$w \cdot x + b = 0 \quad (5)$$

trong đó: w là vector trọng số và b là bias.

- Margin: khoảng cách từ siêu phẳng đến điểm dữ liệu gần nhất là: $\frac{1}{|w|}$

- Hàm mục tiêu:

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (6)$$

Với các ràng buộc:

$$y_i(w \cdot x_i + b) \geq 1, \quad \forall i \quad (7)$$

- Hàm mục tiêu với soft margin (trong trường hợp SVM có lỗi cho phép):

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \quad (8)$$

Với các ràng buộc:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad \forall i \quad (9)$$

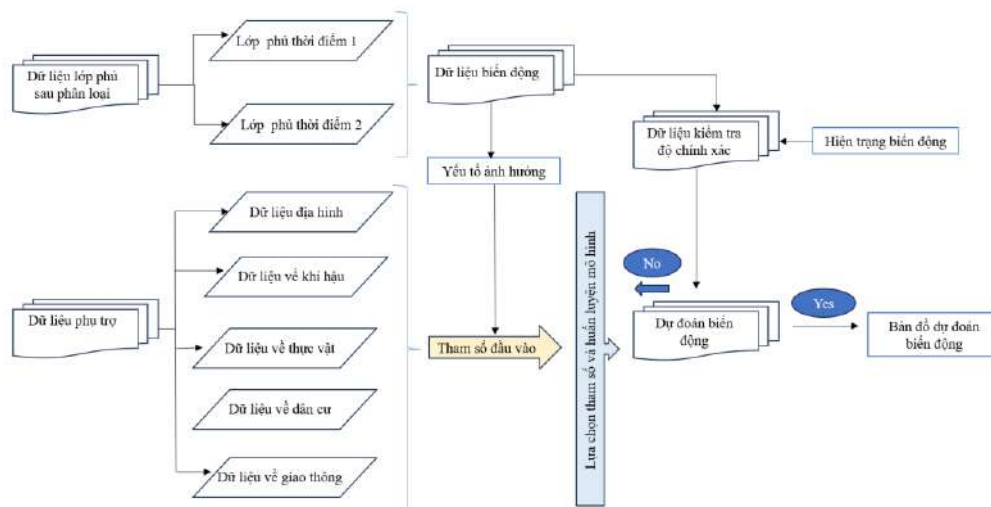
trong đó: ξ_i là biến lỗi cho phép và CCC là tham số điều chỉnh mức độ cho phép lỗi.

- Kernel Trick: SVM có thể sử dụng hàm kernel để ánh xạ dữ liệu vào không gian cao hơn, để dễ dàng phân tách. Một số kernel thường được dùng: Kernel tuyến tính; Kernel đa thức; RBF kernel.

SVM có hiệu quả khi làm việc với dữ liệu có nhiều thuộc tính hay đặc trưng và có thể tìm ra các mô hình phân loại trong không gian nhiều chiều (ví dụ thêm về chiều thời gian). Bên cạnh đó, SVM có khả năng tổng quát hóa tốt, đặc biệt với các bộ dữ liệu nhỏ đến vừa phải và có thể giảm thiểu hiện tượng overfitting (Melgani & Bruzzone, 2004) (Mountrakis, Im, & Ogole, 2011) (Huang, Davis & Townshend, 2002). Ngoài ra, bằng cách sử dụng các hàm kernel ánh xạ dữ liệu vào không gian cao hơn, SVM có thể phân tách dữ liệu dễ dàng hơn.

2.3. Đề xuất quy trình dự đoán biến động lớp phủ bề mặt

Việc dự đoán biến động lớp phủ bề mặt tương đối phức tạp do quá trình này có thể bị tác động bởi những yếu tố tự nhiên và nhân tạo, đôi khi là sự tương tác qua lại giữa các yếu tố này nên để có thể triển khai được mô hình học máy, cần nhiều dữ liệu tham số đầu vào trong đó vai trò chính vẫn là hiện trạng lớp phủ ở các thời điểm gián đoạn thu nhận hoặc cập nhật dữ liệu. Quy trình đề xuất được mô hình hóa ở Hình 2.



Hình 2. Quy trình đề xuất ứng dụng thuật toán học máy trong dự đoán biến động

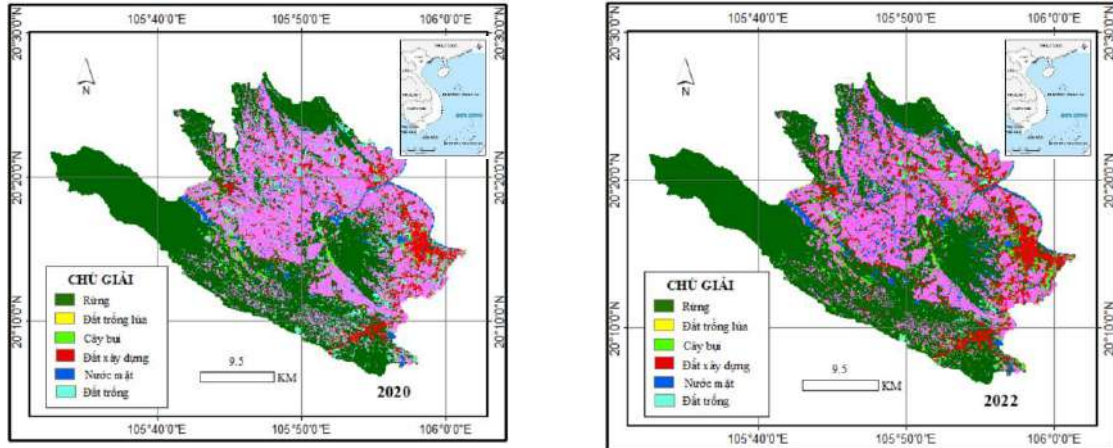
Trong đó, dữ liệu đầu vào bao gồm: dữ liệu ảnh vệ tinh để thu thập thông tin về hiện trạng lớp phủ; dữ liệu địa lý chung với thông tin về địa hình, khí hậu và các yếu tố tự nhiên khác (loại đất) và dữ liệu kinh tế - xã hội (dân cư, tốc độ đô thị hóa). Các dữ liệu này được chuẩn hóa về mặt hình học và độ phân giải, phân loại và phân để trích xuất các đặc trưng.

Các thông tin về phân loại và đặc trưng sau đó được lọc để xử lý giá trị nhiễu và tạo bộ dữ liệu mẫu chuẩn. Bộ dữ liệu mẫu được chia thành các tập huấn luyện và kiểm tra để đánh giá độ chính xác của mô hình.

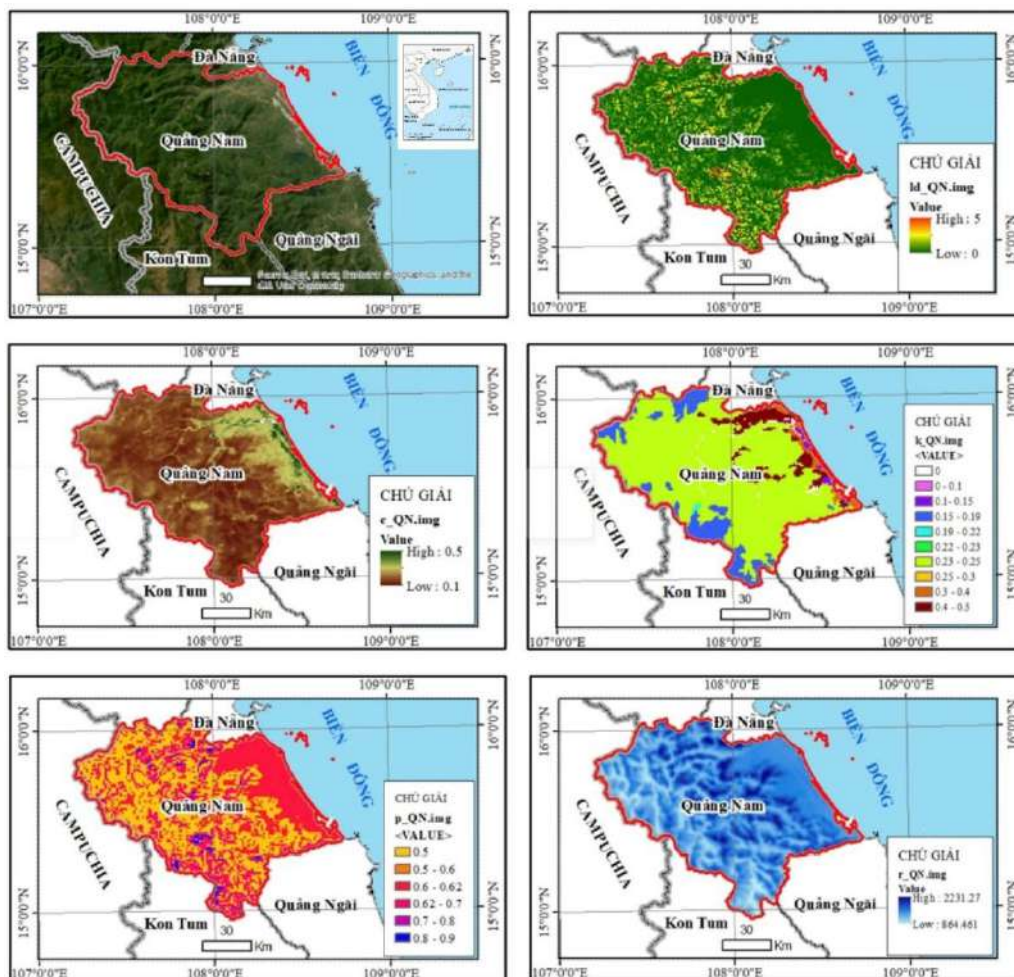
Trong quá trình huấn luyện, các tham số sẽ được tinh chỉnh để tối ưu hóa hiệu suất của mô hình. Việc đánh giá hiệu suất của mô hình được tiến hành thông qua đối sánh với tập dữ liệu kiểm tra thông qua các chỉ số như độ chính xác, độ đặc hiệu và F1 score. Ngoài ra, để đảm bảo mô hình không bị overfitting có khả năng tổng quát hóa tốt, các tác giả đề xuất sử dụng phương pháp k-fold. Khi mô hình đạt đến hiệu suất mong muốn, tiến hành triển khai mô hình vào hệ thống và dự đoán thay đổi đối với bộ dữ liệu mới, tiếp tục đối sánh để nâng cao chất lượng của mô hình. Bên cạnh việc giám sát hoạt động của mô hình, việc cập nhật dữ liệu mới định kỳ cũng đóng vai trò quan trọng trong duy trì độ chính xác và hiệu suất của dự đoán.

3. Kết quả và thảo luận

Hiện tại, nghiên cứu đang trong quá trình lựa chọn khu vực nghiên cứu và xây dựng bộ dữ liệu mẫu. Trong đó, theo dự kiến khu vực thực nghiệm ở hai khu vực: khu vực có tốc độ đô thị hóa cao và khu vực thường xảy ra các tai biến tự nhiên như xói lở. Hình 3 là bản đồ hiện trạng lớp phủ một trong hai khu vực dự kiến thực nghiệm. Hình 4 là khu vực tỉnh Quảng Ngãi với các tham số hiệu chỉnh phục vụ xói mòn đất.



Hình 3. Bản đồ hiện trạng lớp phủ 2020 và 2022 khu vực tỉnh Ninh Bình



Hình 4. Khu vực thực nghiệm tỉnh Quảng Ngãi với tham số hiệu chỉnh chuẩn hóa cho mô hình biến động do xói mòn đất với các chỉ số ls và P liên quan đến địa hình, c liên quan đến lớp phủ và lớp hiện trạng sử dụng đất, k liên quan đến loại đất, và r liên quan đến khí hậu (Trần, Trần & Đoàn, 2023)

Do điều kiện thu thập và khai thác dữ liệu ở mỗi khu vực có những hạn chế khác nhau, điều này cũng phản ánh thực tế trong những khó khăn của việc thu thập và xử lý dữ liệu ở Việt Nam nói chung. Tuy nhiên, với những dữ liệu hiện có, các tác giả sẽ tiếp tục tập trung về mặt làm hài hòa dữ liệu và tiến hành phát triển các thuật toán trong điều kiện hạn chế với các thông số đầu vào.

4. Kết luận

Nghiên cứu đã đề cập tới lý thuyết chung về lớp phủ bề mặt và ý nghĩa của việc theo dõi giám sát của lớp phủ bề mặt. Thêm vào đó, nghiên cứu cũng tiến hành phân tích ưu và nhược điểm của các thuật toán học máy phổ biến trong phân loại và dự đoán biến động lớp phủ bề mặt. Các thuật toán cây quyết định, rừng ngẫu nhiên và vector máy hỗ trợ đều là những thuật toán có tính tổng quát và trực quan có khả năng áp dụng đối với nghiên cứu về lớp phủ và biến động lớp phủ. Những thuật toán này dựa vào việc chiết xuất đặc trưng và chuyên gia phân tích để tạo ra các thông tin đặc trưng cho bộ dữ liệu mẫu, giúp máy tính học tốt hơn và phù hợp với khả năng tính toán trên máy tính cá nhân. Bên cạnh đó, với bộ dữ liệu không liên tục và không đủ lớn, các thuật toán này là lựa chọn phù hợp so với các thuật toán về mạng nơron nhân tạo. Nghiên cứu đã xác định khu vực thực nghiệm và đầu vào dữ liệu dự kiến. Trong thời gian tới, nghiên cứu sẽ tiến hành thiết lập bộ dữ liệu mẫu và huấn luyện mô hình. Kết quả về hiệu suất và độ chính xác của các thuật toán sẽ được đem ra so sánh để đề xuất thuật toán phù hợp có khả năng áp dụng ở những khu vực khác ở Việt Nam.

Lời cảm ơn

Bài báo là kết quả tìm hiểu và nghiên cứu của các tác giả khi thực hiện đề tài cấp cơ sở trường Đại học Mở Địa Chất, mã số: T24-02.

Một số hình ảnh trong bài viết như Hình 1 và Hình 4 được sử dụng tuân thủ theo quy định Copyright.

Tài liệu tham khảo

- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*(114), 24-31.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and Regression Trees*. . Wadsworth International Group.
- ESA. (n.d.). Worldwide land cover mapping. Retrieved from <https://esa-worldcover.org/en>
- García, L., & Rodríguez, A. (2022). Predicting land cover changes with support vector machines: A case study in South America. *Journal of Environmental Management*, 308, 114528.
- Giri, C. P. (2020). *Remote Sensing of Land Use and Land Cover*. CRC Press.
- Gregorio, A. D., & Jansen, L. J. (2005). *Land Cover Classification System - Classification concepts and user manual*. Rome: FAO. Retrieved from [fao.org: https://www.fao.org/4/x0596e/x0596e01f.htm#:~:text=The%20Land%20Cover%20Classification%20System,and%20Semi%2DNatural%20Terrestrial%20Vegetation](https://www.fao.org/4/x0596e/x0596e01f.htm#:~:text=The%20Land%20Cover%20Classification%20System,and%20Semi%2DNatural%20Terrestrial%20Vegetation)
- Hansen, M. C., Potapov, P. V., Hancher, M., Turubanova, S. A., Tyukavina, A. ..., & Townshend, J. R. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850-853.
- Herrmann, S. M., & Hutchinson, C. F. (2005). The changing contexts of the desertification debate. *Journal of Arid Environments*, 63(3), 538-555.
- Huang, C., Davis, L. S., & Townshend, J. R. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23(4), 725-749.
- Johnson, R., & Lee, S. (2023). Random forest-based models for forecasting land use and land cover changes. *Environmental Research Letters*, 18(3).
- Kumar, P., & Sharma, A. (2023). Machine learning approaches for land cover change detection: A review. *Environmental Monitoring and Assessment*, 195(2), 345.
- Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), 1779-1790.
- Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247-259.
- Nguyễn, V. H., & Đăng, T. T. (2021). Đánh giá biến động lớp phủ mặt đất tại thành phố Đà Nẵng bằng ảnh viễn thám Landsat. *Tạp chí Môi trường*.
- Pal, M., & Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86(4), 554-565.
- Patel, D., & Desai, M. (2021). Machine learning models for predicting agricultural land cover changes: A comparative study. *Computers and Electronics in Agriculture*, 192.
- Phạm, M. T., & Nguyễn, T. L. (2023). Ứng dụng viễn thám trong đánh giá biến động lớp phủ đất tại vùng đồng bằng sông Hồng. *Tạp chí Khoa học và Công nghệ*.

- Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning*, 1(1), 81-106.
- Rodriguez-Galiano, V. F., Ghimire, B. R., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*(67), 93-104.
- Seto, K. C., Güneralp, B., & Hutya, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, 109(4), 16083-16088.
- Thackway, R., & Atyeo, C. (2006). *Classifying Australia LandCover*.
- Trần, T. H., Trần, T. H., & Đoàn, C. L. (2023). A novel approach of neural networks and USLE in smart soil erosion modeling, case study in Southern coastal of Vietnam. *International conference GIS IDEAS 2023* (pp. 369-379). Ha Noi: Publishing house for Science and Technology.
- Trần, T. T. (2020). Phân tích biến động lớp phủ đô thị bằng dữ liệu ảnh vệ tinh tại thành phố Hồ Chí Minh. *Tạp chí Khoa học và Phát triển*.
- Vapnik, V. (1998). *Statistical Learning Theory*. Wiley.
- Zhang, H., Li, Y., & Wang, X. (2023). Application of deep learning in predicting land cover changes using satellite imagery. *Remote Sensing*, 15(1), 1-15.
- Zhou, F., & Liu, Y. (2022). Predicting land cover changes in arid regions using machine learning algorithms. *Journal of Arid Environments*, 199.

ABSTRACT

Predicting land cover changes using Machine learning algorithms

Tran Thi Hoa^{1*}, Tran Thi Ngoc¹, Tran Dinh Tri²

¹*Hanoi University of Mining and Geology*

²*Vietnam Association of Geography, Cartography and Remotesensing*

Land cover refers to the physical surface morphology of the Earth's crust, encompassing vegetation, water, natural surfaces, and artificial objects. Activities occurring on the Earth's surface, ranging from natural processes to socio-economic development and urbanization, lead to significant changes in land cover. Soil resources are among the critical resources with profound impacts on the environment in general and sustainable development goals in particular. Therefore, monitoring and observing changes in land cover, as well as analyzing and forecasting trends, are crucial tasks. Current machine learning algorithms have made significant contributions by leveraging computational power and ensuring objectivity in analyzing related predictive issues. This paper examines the potential applications of various machine learning algorithms in predicting land cover changes, specifically: Decision Trees, Random Forests, and Support Vector Machines (SVM).

Keywords: Land cover changes, predicting, Machine Learning, Random Forest, SVM.

Xây dựng cơ sở dữ liệu GIS phục vụ công tác tìm kiếm và cứu nạn khu vực Vịnh Bắc Bộ

Nguyễn Quang Minh^{1*}, Nguyễn Văn Trung^{2,3}, Bùi Tiến Diệu⁴

¹Cục Bản đồ - Bộ Tổng tham mưu

²Trường Đại học Mở - Địa chất

³Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất

⁴Đại học Đông Nam Na Uy

TÓM TẮT

Phương pháp ứng dụng công nghệ hệ thống tin địa lý (GIS) được áp dụng rộng rãi để giải quyết nhiều vấn đề trong thực tiễn. Tuy nhiên, khả năng ứng dụng GIS trong các kế hoạch tìm kiếm và cứu nạn trên biển vẫn là một thách thức ở nước ta. Việc đưa công cụ GIS áp dụng vào xây dựng cơ sở dữ liệu (CSDL) phục vụ giải quyết bài toán tối ưu về thời gian di chuyển từ trạm cứu hộ đến các vị trí tìm kiếm và cứu hộ là cần thiết ở các vùng biển ở nước ta. Bài báo này tập trung vào các công tác xử lý, chuẩn hóa các dữ liệu để xây dựng các CSDL thành phần bao gồm dữ liệu lịch sử tìm kiếm và cứu hộ, dữ liệu về đội tàu và các trạm cứu hộ và dữ liệu về khí tượng hải văn. Việc xây dựng các CSDL thành phần này cần đảm bảo tiêu chuẩn để sử dụng làm dữ liệu đầu vào cho mô hình tính toán đường đi tối ưu về thời gian di chuyển trên biển. Bên cạnh đó, các CSDL thành phần xây dựng được cũng được đề cập để cung cấp thông tin về khả năng đáp ứng đối với các kế hoạch tìm kiếm và cứu nạn ở vịnh Bắc Bộ.

Từ khóa: Vịnh Bắc Bộ, cứu hộ, GIS, khí tượng, độ cao sóng biển.

1. Đặt vấn đề

So với tìm kiếm, cứu hộ trên đất liền, hoạt động tìm kiếm, cứu hộ trên biển gặp nhiều khó khăn, thách thức lớn hơn do đặc thù khu vực vùng biển. Những khó khăn, thách thức chính trong tìm kiếm, cứu hộ trên biển bao gồm: diện tích vùng biển rộng, điều kiện thời tiết trên biển phức tạp, nguy hiểm và hạn chế về trang thiết bị tìm kiếm, cứu hộ.

Các hệ thống tìm kiếm, cứu hộ trên biển đã được phát triển nhiều trên thế giới bao gồm: hệ thống an toàn và cứu nạn hàng hải toàn cầu (GMDSS) là một hệ thống dùng để liên lạc tín hiệu khẩn cấp tự động cho tàu trên biển do Tổ chức Hàng hải quốc tế (IMO) phát triển năm 1988; Hệ thống quản lý thông tin chương trình tìm kiếm và cứu nạn (SISAR) của cảnh sát biển Canada chứa cơ sở dữ liệu về các sự cố tìm kiếm và cứu nạn hàng hải (SAR); hệ thống nhận dạng tự động (AIS) do Đại học Rostock (Liên bang Đức) đã thực hiện dự án nghiên cứu liên ngành AgaPaS nhằm phát triển hệ thống báo động tự động trong tìm kiếm, cứu hộ để rút ngắn đáng kể thời gian cứu nạn trên biển; hệ thống SARGIS chứa phần mềm dựa trên GIS được Guoxing và cộng sự (2010) phát triển để cung cấp thông tin hỗ trợ cho dịch vụ tìm kiếm và cứu nạn hàng hải và hệ thống giám sát dựa trên dữ liệu viễn thám vệ tinh (RS-MSR) phục vụ tìm kiếm cứu nạn hàng hải phát triển bởi Peng và Yang (2011).

Sử dụng cơ sở dữ liệu (CSDL) GIS để xây dựng các mô hình tìm kiếm cứu hộ, cứu nạn trong hệ thống SISAR như Shahrabi (2003) đã phát triển một kỹ thuật phân tích không gian và thời gian về các sự cố đánh bắt cá và giao thông hàng hải ngoài khơi bờ biển Nova Scotia, trong đó dự đoán vị trí và thời gian xảy ra sự cố trên biển. Bằng cách sử dụng hệ thống thông tin địa lý (GIS) và nhiều phương pháp thống kê không gian, chẳng hạn như ước tính mật độ (kernel density estimation) và phương pháp phân cụm theo cấp bậc (hierarchical clustering method), Shahrabi có thể xác định các khu vực có nguy cơ sự cố cao hơn đối với ngư dân và tàu đánh bắt cá ngoài khơi khu vực nghiên cứu (Shahrabi, 2003).

Nghiên cứu của Venalainen (2014) đã sử dụng công nghệ GIS trong xác định thời gian phản hồi trong công tác tìm kiếm, cứu hộ với các sự cố do hoạt động chèo thuyền ở vịnh Phần Lan (Venalainen, 2014). Phương pháp mô hình hóa được áp dụng để tính thời gian phản hồi cho toàn bộ khu vực nghiên cứu bằng cách sử dụng các thông tin như dữ liệu sự cố, đặc điểm tàu tìm kiếm, cứu hộ, kích bản về chiều cao sóng biển, gió biển. Thống kê từ Cơ quan Bảo vệ bờ biển Phần Lan cho thấy, chỉ có 40,76% sự cố xảy ra mà có các tàu tìm kiếm, cứu hộ đang tuần tra. Như vậy, có tới gần 60% các sự cố trên biển xảy ra mà không có các tàu SAR đang hoạt động, do vậy cần sự điều động từ các trạm tìm kiếm, cứu hộ.

Tác giả Zhou và cộng sự (2020) đã phát triển một mô hình xác định thời gian phản hồi của tàu cứu hộ nhằm đánh giá khả năng tìm kiếm, cứu hộ hàng hải ở khu vực Biển Đông, tập trung vào 4 quốc gia gồm Trung Quốc,

* Tác giả liên hệ

Email: nguyenvanquangminh.sar@gmail.com

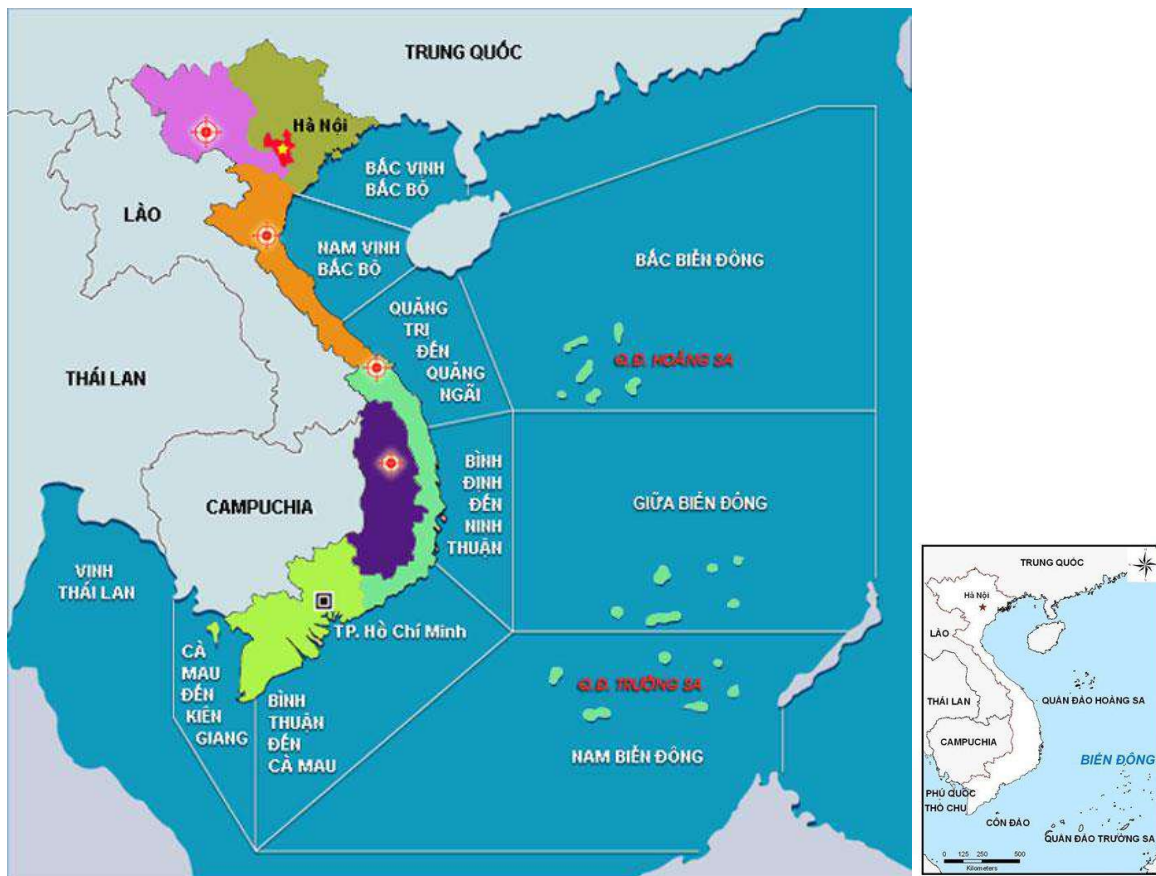
Việt Nam, Malaysia và Philippin. Mô hình này gồm 3 bước: (1) Đầu tiên, mô hình thời gian phản hồi dựa trên GIS được phát triển để ước tính thời gian phản hồi của tàu cứu hộ, (2) Thứ hai, nhu cầu về dịch vụ SAR hàng hải được tính toán bằng cách kết hợp dữ liệu vị trí tàu và dữ liệu sự cố hàng hải; (3) Thứ ba, ba tiêu chí đánh giá gồm phạm vi phủ sóng chính, phạm vi phủ sóng có trọng số và thời gian truy cập trung bình, được sử dụng để đánh giá định lượng khả năng phản hồi của hệ thống SAR hàng hải (Zhou và cộng sự, 2020).

Ashrafi và cộng sự (2024) đề xuất một khung mô hình hóa dựa trên tác nhân (ABM - agent based modelling) để đánh giá hiệu suất của các hoạt động tìm kiếm, cứu hộ (SAR) nhằm giải quyết những thách thức liên quan đến điều kiện thời tiết ở Bắc Cực. Phương pháp mô phỏng Monte Carlo được sử dụng để mô hình hóa động lực và sự không chắc chắn của điều kiện thời tiết và biển bằng cách sử dụng dữ liệu lịch sử. Sau đó, quy trình đánh giá của chuyên gia sẽ được sử dụng để định lượng tác động của các mức độ nghiêm trọng đó đối với hoạt động của lực lượng cứu hộ cũng như tổng thời gian cứu hộ. Tại khu vực thử nghiệm ở Biển Barents (Na Uy) cho thấy, tổng thời gian giải cứu dài nhất là từ tháng 12 đến tháng 2 năm sau và thời gian giải cứu ngắn nhất là từ tháng 5 đến tháng 8 (Ashrafi và cộng sự, 2024).

Có thể nhận thấy, cho đến nay trên thế giới đã có nhiều nghiên cứu sử dụng phương pháp mô hình hóa trên nền tảng GIS nhằm xác định khoảng cách, thời gian tối ưu trong tiếp cận tìm kiếm, cứu hộ trên biển. Mặc dù vậy, các nghiên cứu này đều đối mặt thách thức do sự hạn chế về dữ liệu đầu vào, nhất là thông tin về điều kiện khí tượng, hải văn khu vực nghiên cứu.

2. Dữ liệu sử dụng và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu



Hình 1. Vị trí địa lý các vùng biển Việt Nam

Vịnh Bắc Bộ có tọa độ địa lý ở vĩ tuyến 18°31'19" Bắc, kinh tuyến 108°41'17" Đông, qua đảo Cồn Cỏ của Việt Nam đến một điểm trên bờ biển của Việt Nam có tọa độ địa lý là vĩ tuyến 16°57'40" Bắc và kinh tuyến 107°08'42" Đông (Hình 1). Vịnh Bắc Bộ có tài nguyên thiên nhiên phong phú với nhiều ngư trường lớn, có tiềm năng về dầu mỏ và khí đốt. Với đặc điểm trên, vịnh Bắc Bộ có tầm quan trọng đặc biệt đối với sự nghiệp phát triển kinh tế, thương mại quốc tế cũng như quốc phòng, an ninh của nước ta. Ngày 25/12/2000, Việt Nam và Trung Quốc đã ký Hiệp định Phân vịnh Bắc Bộ để xác định ranh giới lãnh hải, vùng đặc quyền kinh tế và thềm lục địa của mỗi nước trong vịnh Bắc Bộ. Hiệp định có hiệu lực thi hành từ ngày 30/6/2004.

Về lượng mưa hàng năm, khu vực các tỉnh ven biển Quảng Ninh - Hải Phòng và Nghệ An - Quảng Bình có lượng mưa cao nhất với gần 1400 mm/năm. Vào mùa đông, lượng mưa giảm mạnh, dao động trong khoảng 150 - 600 mm và giảm dần theo hướng Tây Nam - Đông Bắc. Bão là hiện tượng thiên tai phổ biến tại vịnh Bắc Bộ. Theo số liệu thống kê của Trung tâm Dự báo Khí tượng Thủy văn Quốc gia, trung bình mỗi năm có khoảng 5 - 6 cơn bão ảnh hưởng trực tiếp đến vịnh Bắc Bộ. Bên cạnh đó, bão thường bắt đầu vào tháng 5 (muộn vào tháng 6) và kết thúc vào tháng 10 hoặc tháng 11. Sức gió mạnh nhất trong các cơn bão có thể đạt tới 50 m/s. Trong những năm gần đây, do ảnh hưởng của biến đổi khí hậu, tần suất và cường độ của bão tại vịnh Bắc Bộ có xu hướng gia tăng. Cơn bão đầu tiên trong năm 2024 (cơn bão số 1, tháng 5/2024) ảnh hưởng đến vịnh Bắc Bộ có sức gió lên đến cấp 11. Thủy triều và triều lưu trong vịnh Bắc Bộ mang tính nhật triều là chủ yếu. Thủy triều tại khu vực phía Bắc dao động với biên độ lớn với biên độ khoảng 3,8 m đến 4,2 m. Khu vực cửa vịnh thì biên độ nhỏ hơn dao động từ 1 m đến 1,5 m (Nguyễn Đình Hải và cộng sự, 2023). Bên cạnh đó, độ lớn thủy triều giảm từ Bắc xuống Nam.

Vịnh Bắc Bộ có diện tích rộng, là một trong những tuyến đường hàng hải quan trọng trên thế giới với giao thông hàng hải nhộn nhịp, các tai nạn hàng hải, sự cố môi trường diễn ra rất phức tạp, nhất là trong điều kiện ảnh hưởng của biến đổi khí hậu. Vùng biển vịnh Bắc Bộ và các khu vực ven biển thường xuyên chịu ảnh hưởng của các yếu tố thiên tai bão, lũ, triều cường, mưa lớn, thời tiết nguy hiểm trên biển và các hệ quả thiên tai như sóng, nước dâng do bão, ngập lụt, bồi lắng, xói lở, xâm nhập mặn,... gây ảnh hưởng đến sự ổn định và phát triển kinh tế - xã hội. Hàng năm, khu vực chịu ảnh hưởng của nhiều cơn bão và áp thấp nhiệt đới, kéo theo đó các hệ thống sông đổ ra biển ở vịnh Bắc Bộ cũng chịu ảnh hưởng của lũ thượng nguồn, mưa lớn, triều cường, sóng lớn và nước dâng do bão.

2.2. Dữ liệu sử dụng

2.2.1. Dữ liệu sự cố

Dữ liệu về sự cố và đối tượng cần tìm kiếm, cứu nạn là yêu cầu đầu vào rất quan trọng trong các mô hình tìm kiếm, cứu hộ trên biển. Dữ liệu sự cố trên biển là các thông tin được thu thập và lưu trữ về các sự cố, tai nạn và rủi ro xảy ra trên biển. Những dữ liệu này bao gồm:

- Thông tin về sự cố: Thời gian xảy ra sự cố, vị trí, địa điểm, số lượng và thông tin tàu thuyền có liên quan, tình trạng thời tiết, mô tả sự cố, thiệt hại về người và tài sản trong quá trình xảy ra sự cố.
- Nguyên nhân và ảnh hưởng của sự cố: Lỗi con người, lỗi kỹ thuật, điều kiện khí hậu thời tiết bất lợi; nguy cơ ô nhiễm môi trường, gián đoạn của hoạt động giao thông hàng hải.
- Quá trình ứng phó và khắc phục: Các biện pháp được áp dụng để ứng phó, cứu hộ và khôi phục hoạt động.

Việc xây dựng và phân tích bộ dữ liệu sự cố là rất quan trọng để nâng cao khả năng vận hành, quản lý rủi ro và cải thiện hiệu quả hoạt động của các tổ chức vì các lý do sau:

- Xác định mức độ nguy hiểm và nhu cầu cần thiết: Thông tin về tính chất, quy mô của sự cố giúp đánh giá mức độ nguy hiểm và số lượng, tình trạng của tàu thuyền, người cần tìm kiếm, cứu hộ. Điều này cho phép lên kế hoạch cung cấp các nguồn lực, thiết bị, phương tiện phù hợp để đáp ứng nhu cầu cứu hộ một cách hiệu quả.
- Định vị và tiếp cận nạn nhân: Dữ liệu về vị trí, đặc điểm của đối tượng cần tìm kiếm, cứu hộ giúp xác định chính xác nơi cần tìm kiếm và tiếp cận đối tượng, nhất là trong điều kiện tìm kiếm, cứu hộ trên biển.
- Lên kế hoạch cứu hộ phù hợp: Thông tin về tình trạng, vị trí của sự cố giúp xác định các biện pháp xử lý sự cố, phương tiện di chuyển cần thiết. Dữ liệu về quy mô, tính chất sự cố và số lượng người cần cứu hộ giúp điều phối hiệu quả các lực lượng, phương tiện tham gia.

Như vậy, có thể nhận thấy dữ liệu về sự cố và đối tượng là lớp thông tin đầu vào đặc biệt quan trọng của các mô hình tìm kiếm, cứu hộ, giúp xác định nhu cầu, lên kế hoạch và phối hợp các hoạt động tìm kiếm, cứu hộ một cách chính xác và hiệu quả nhất.

2.2.2. Dữ liệu khí tượng, thủy văn, hải văn

Thông tin về khí tượng, thủy văn, hải văn là một lớp dữ liệu đầu vào quan trọng của các mô hình tìm kiếm, cứu hộ trên biển. Việc kết hợp dữ liệu khí tượng, thủy văn, hải văn với các nguồn dữ liệu khác sẽ giúp nâng cao hiệu quả các hoạt động tìm kiếm, cứu nạn, cứu hộ trong nhiều tình huống khẩn cấp, hỗ trợ xây dựng kế hoạch cũng như ra quyết định phù hợp trong điều động lực lượng, phương tiện tìm kiếm, cứu hộ.

Nhiệt độ có ảnh hưởng đáng kể đến hoạt động tìm kiếm, cứu hộ trên biển. Trước hết, trong trường hợp nhiệt độ thấp, điều này sẽ ảnh hưởng đến sức khỏe và an toàn của thủy thủ đoàn và nhân viên cứu hộ, có thể dẫn đến tình trạng giảm nhiệt, ướt lạnh. Bên cạnh đó, nhiệt độ thấp có thể làm giảm hiệu quả hoạt động, khả năng di chuyển và sử dụng thiết bị. Thiết bị và phương tiện cứu hộ có thể bị hư hỏng do nhiệt độ lạnh. Ngược lại, nhiệt độ cao có thể gây ra nguy cơ về sức khỏe như say nắng, kiệt sức cho người tham gia cứu hộ, ảnh hưởng đến khả năng hoạt động, dẫn đến suy giảm năng suất và hiệu quả. Nhiệt độ nước biển cũng ảnh hưởng đến khả năng sống sót của người gặp nạn trên biển.

Bên cạnh đó, độ ẩm cũng là một yếu tố có vai trò quan trọng trong hoạt động tìm kiếm, cứu hộ trên biển. Độ ẩm cao làm giảm khả năng nhìn thấy các vật thể trên biển, ảnh hưởng đến việc phát hiện người, tàu thuyền gặp nạn. Đồng thời, độ ẩm cao có thể gây hư hỏng thiết bị như radar, máy ảnh, thiết bị điện tử sử dụng trong hoạt động tìm kiếm, cứu hộ, làm giảm hiệu quả hoạt động.

Ngoài ra, các yếu tố khác như mưa, sương mù, gió có thể ảnh hưởng đến khả năng tiếp cận, di chuyển tới vị trí xảy ra sự cố trên biển. Những hiện tượng thời tiết cực đoan như bão, sương mù có thể gây nguy hiểm cho hoạt động tìm kiếm, cứu nạn trên biển.

Các thông tin về thủy, hải văn như thủy triều, dòng chảy rất quan trọng cho các hoạt động tìm kiếm, cứu hộ trên biển, giúp lực lượng tìm kiếm, cứu nạn lập kế hoạch tiếp cận khu vực xảy ra sự cố một cách an toàn, khả thi. Hơn nữa, các dữ liệu về độ sâu, địa hình đáy biển có vai trò quan trọng trong xác định vị trí, tuyến đường di chuyển an toàn của các tàu tìm kiếm, cứu nạn.

2.3. Phương pháp xây dựng các lớp thông tin đầu vào

2.3.1. Lớp dữ liệu tàu tìm kiếm, cứu hộ trên biển

Lớp dữ liệu tàu tìm kiếm, cứu hộ trên biển là một yếu tố đầu vào có vai trò quan trọng trong xây dựng mô hình ước lượng khoảng cách hiệu dụng hỗ trợ công tác tìm kiếm, cứu hộ trên biển. Để mô hình hóa khoảng cách hiệu dụng, người lập kế hoạch cần biết có những loại đơn vị tìm kiếm, cứu nạn nào; vị trí của các đơn vị tìm kiếm, cứu nạn; các đặc điểm, thông số của các trang thiết bị tìm kiếm, cứu nạn trên biển. Thông số về các tàu tìm kiếm, cứu nạn cho phép xây dựng kế hoạch ứng phó khả thi trong từng điều kiện thời tiết cụ thể ở khu vực xảy ra sự cố. Hình 2 trình bày hình ảnh tàu tìm kiếm, cứu nạn CN-09 của bộ đội biên phòng thành phố Hải Phòng.



Hình 2. Hình ảnh tàu cứu hộ CN-09 của bộ đội biên phòng thành phố Hải Phòng

Bảng 1 trình bày đặc điểm và các thông số của các tàu tìm kiếm, cứu nạn trên biển được sử dụng trong nghiên cứu.

Bảng 1. Thông số lực lượng tàu tìm kiếm cứu nạn trên biển

Đơn vị	Số hiệu	Vận tốc (Hải lý/giờ)		Tầm hoạt động (Hải lý)
		V_{kt}	V_{max}	
Hải quân	TS: 01, 02, 04, 06, 08, 10, 12, 14, 16, 18, 20, 21, 22.	10 (TS 12: 7)	12 (TS 12: 10.7)	4000 (TS 12: 3500)
Biên phòng	CN-09	20	22	07 ngày ($V = 10HL/h$)
Cục Kiểm ngư	KN-579	22	25	13 ngày
	KN-568	15	15	13 ngày

2.3.2. Lớp dữ liệu đối tượng đất liền - mặt nước

Lớp dữ liệu đối tượng đất liền - mặt nước là một tập hợp các dữ liệu mô tả vị trí và đặc điểm của các khu vực đất liền và mặt nước trên biển. Trong đó, các đối tượng mặt nước có giá trị là 0, các đối tượng đất liền có giá trị > 0 (thông thường các đối tượng đất liền có giá trị giống nhau và được lấy bằng 1) (Siljander và cộng sự, 2015).

Lớp dữ liệu đối tượng đất liền - mặt nước được xây dựng nhằm hỗ trợ các mục đích sau:

+ Đầu vào cho quy trình tính toán ảnh lấy gió: Dữ liệu được sử dụng để xác định giới hạn những vị trí cần tính toán trong ảnh lấy gió và dữ liệu khoảng cách hiệu dụng ở bước tiếp theo.

+ Xác định giới hạn ranh giới lớp dữ liệu độ sâu: Lớp dữ liệu độ sâu được xác định tại các vị trí mặt nước. Chính vì vậy, việc xây dựng dữ liệu đối tượng đất liền - mặt nước làm giảm chi phí tính toán, tăng thời gian xử lý dữ liệu.

2.3.3. Lớp dữ liệu hướng gió

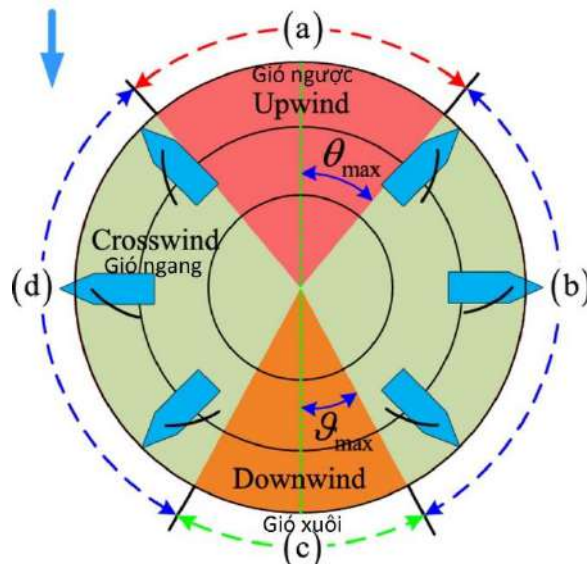
Hướng gió được xác định theo 8 hướng (Bắc, Đông Bắc, Đông, Đông Nam, Nam, Tây Nam, Tây, Tây Bắc). Bên cạnh đó, người dùng cũng có thể đo hướng gió theo đơn vị từ 0° đến 360° . Sự ảnh hưởng của hướng gió tác động tới các đơn vị tàu như sau (Hoffmann và cộng sự, 2021; Zhang và cộng sự, 2022):

+ Gió thuận (*gió xuôi* - *downwind*): Gió thổi từ phía sau tàu (góc 0° - 180° so với hướng di chuyển của tàu) có thể tạo ra lực đẩy bổ sung, giúp tàu tăng tốc độ khi di chuyển trên mặt biển.

+ Gió ngược (*upwind*): Gió thổi từ phía trước tàu (góc 180° - 360°) có thể tạo ra lực cản, khiến tàu giảm tốc độ và tiêu hao nhiều nhiên liệu hơn. Khi hướng gió đối diện hoặc gần đối diện với hướng di chuyển của tàu, nó tạo ra lực cản lớn, làm giảm tốc độ di chuyển của tàu và yêu cầu nhiều năng lượng hơn để duy trì tốc độ. Trong trường hợp này, tàu có thể cần điều chỉnh hướng hoặc thay đổi tốc độ để vượt qua lực cản.

+ Gió ngang (*crosswind*): Gió thổi vuông góc với hướng di chuyển của tàu có thể khiến tàu bị lệch hướng.

Mô tả ảnh hưởng của hướng gió, bao gồm gió thuận, gió ngược và gió ngang được thể hiện trên Hình 3 (Zhang và cộng sự, 2022).



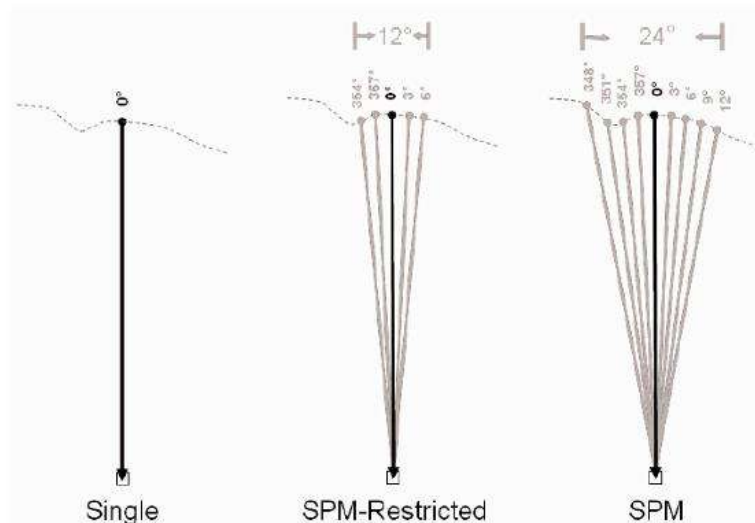
Hình 3. Mô tả ảnh hưởng của hướng gió đến sự di chuyển của tàu trên biển

2.3.4. Lớp dữ liệu tốc độ gió điều chỉnh

Dữ liệu về tốc độ gió có ảnh hưởng lớn đến hiệu quả và an toàn của các hoạt động tìm kiếm, cứu hộ trên biển. Trước hết, cường độ của gió ảnh hưởng đến các hoạt động của đơn vị tìm kiếm, cứu hộ cũng như các trang thiết bị hoạt động trên biển. Gió mạnh có thể tạo ra sóng lớn và biến động bề mặt biển, làm cho việc duy trì định vị và điều khiển tàu thuyền trở nên khó khăn. Ngoài ra, gió mạnh cũng tạo ra lực đẩy lớn, có thể làm tăng tốc độ di chuyển của tàu hoặc ngược lại, tùy thuộc vào hướng gió tương đối với hướng di chuyển. Bên cạnh đó, tốc độ gió có mối quan hệ chặt chẽ và ảnh hưởng rất lớn đến chiều cao sóng trên biển. Tốc độ gió càng mạnh thì chiều cao sóng càng lớn. Đây là mối quan hệ tỷ lệ thuận. Gió mạnh sẽ tạo ra những lượng năng lượng lớn được truyền từ gió xuống mặt biển, làm cho sóng cao và dữ dội hơn (Siljander và cộng sự, 2015).

2.3.5. Lớp dữ liệu ảnh lấy gió

Lấy gió (*wind fetch*) được định nghĩa là khoảng cách liên tục mà gió thổi trên mặt nước, không bị ảnh hưởng bởi địa hình hoặc các vật cản khác. Thông tin về dữ liệu lấy gió được tính bằng khoảng cách theo đường thẳng tính từ vị trí quan tâm đến vị trí xa nhất mà gió thổi liên tục trên mặt biển. Hay nói cách khác, lấy gió là khoảng cách không bị cản trở mà gió có thể di chuyển trên mặt nước theo một hướng không đổi. Khoảng cách lấy gió càng dài sẽ dẫn đến chiều cao sóng lớn hơn. Mô hình của Cục Khảo sát địa chất Hoa Kỳ (USGS) cung cấp ba phương pháp khác nhau để tính toán dữ liệu lấy gió, bao gồm đơn, SPM hạn chế và SPM (Rohweder và cộng sự, 2008; Venalainen, 2014) - Hình 4.



Hình 4. Các phương pháp tính toán thông tin lấy gió khác nhau

2.3.6. Lốp dữ liệu gia tốc trọng trường

Gia tốc trọng trường, đại diện bởi giá trị g , là gia tốc mà một vật thể trên bề mặt Trái đất trọng lực. Gia tốc trọng trường không ảnh hưởng trực tiếp đến độ cao sóng trong môi trường không khí (Cartro, 2014).

Độ cao sóng trong môi trường không khí được xác định bởi một số yếu tố như tốc độ gió, sự chênh lệch nhiệt độ, độ ẩm và áp suất không khí. Những yếu tố này ảnh hưởng đến sự truyền dẫn và giao thoa của sóng âm, sóng ánh sáng và các loại sóng khác.

Tuy nhiên, gia tốc trọng trường có thể có một tác động gián tiếp đến độ cao sóng thông qua ảnh hưởng của nó đến khí quyển. Gia tốc trọng trường cao hơn có thể dẫn đến mật độ khí quyển cao hơn ở độ cao thấp hơn. Điều này có thể làm cho âm thanh truyền qua không khí một cách nhanh chóng hơn và có thể ảnh hưởng đến tần số và độ dài sóng của âm thanh (Cartro, 2014).

2.3.7. Lốp dữ liệu độ sâu

Dữ liệu độ sâu và địa hình đáy biển có ảnh hưởng đáng kể đến quá trình hình thành và phát triển của sóng biển, đặc biệt là độ cao của sóng. Trước hết, khi sóng di chuyển vào vùng nước nông ven bờ, quá trình dịch chuyển của sóng bị ảnh hưởng bởi độ sâu. Đồng thời, nếu sóng bị biến dạng thì đỉnh sóng sẽ dốc hơn dẫn đến độ cao sóng tăng lên (Herterich và Dias, 2019; Salamed và cộng sự, 2019).

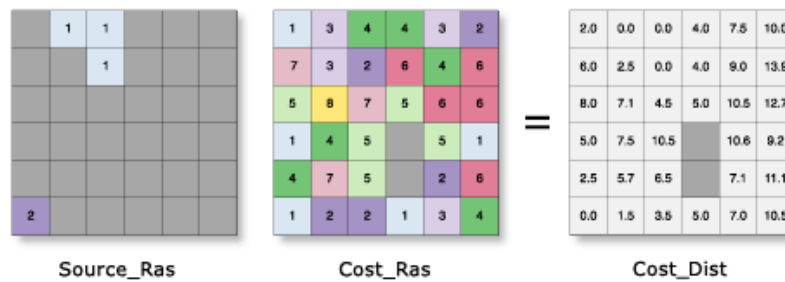
Bên cạnh đó, các đụn cát, vách đá, vật thể dưới đáy biển sẽ tạo ra các hiệu ứng phản xạ, giao thoa sóng. Điều này làm độ cao sóng tăng lên tại các vị trí nhất định. Ngoài ra, trong điều kiện, đường bờ không đều, sóng sẽ bị uốn cong, tập trung hoặc phân tán khi di chuyển dọc theo đường bờ không đều. Điều này dẫn đến các khu vực có độ cao sóng lớn hơn hoặc nhỏ hơn so với trung bình (Herterich và Dias, 2019; Salamed và cộng sự, 2019).

2.4. Phương pháp mô hình hóa khoảng cách hiệu dụng

2.4.1. Khái niệm về khoảng cách hiệu dụng

Khoảng cách hiệu dụng hay khoảng cách chi phí (Cost distance) là một phép đo khoảng cách giữa các vị trí, trong đó có tính đến chi phí hay sức cản khi di chuyển từ vị trí này đến vị trí khác. Khác với khoảng cách Euclidean là khoảng cách đường thẳng ngắn nhất giữa hai điểm, khoảng cách chi phí là khoảng cách ngắn nhất được xác định dựa trên một hoặc nhiều tuyến đường di chuyển tối ưu (Greenberg cộng sự, 2011). Đây là thông tin quan trọng phục vụ xây dựng các kế hoạch trong tìm kiếm, cứu nạn, nhất là tìm kiếm, cứu nạn trên biển, giúp xác định được tuyến đường tiếp cận tối ưu và nhanh chóng trong điều kiện cụ thể về khí hậu, hải văn.

Mô hình hoá khoảng cách hiệu dụng phục vụ tìm kiếm, cứu hộ trên biển được đánh giá thông qua thời gian di chuyển ngắn nhất từ vị trí trạm cứu nạn tới điểm cần cứu nạn (Hình 5). Mô hình được xây dựng với các hướng gió và độ cao sóng khác nhau. Đây là các yếu tố ảnh hưởng tới tốc độ di chuyển của đơn vị cứu hộ. Bên cạnh đó, thời gian phản hồi được mô hình hóa bằng cách sử dụng đơn vị cứu hộ nhanh nhất có thể đến địa điểm xảy ra sự cố và không được xem xét liệu đơn vị đó có phù hợp với nhiệm vụ hay không.



Hình 5. Ví dụ về xác định khoảng cách hiệu dụng

2.4.2. Sơ đồ quy trình

Trên cơ sở phân tích điều kiện tự nhiên, khí hậu, thủy-hải văn khu vực vịnh Bắc Bộ cũng như kinh nghiệm từ các nghiên cứu trên thế giới, trong đề tài đã xây dựng bộ dữ liệu đầu và phục vụ mô hình hóa khoảng cách hiệu dụng hỗ trợ công tác tìm kiếm, cứu nạn trên biển. Từ bộ dữ liệu này, trong đề tài đã xây dựng quy trình xác định thời gian phản hồi (khoảng cách hiệu dụng) trong tìm kiếm, cứu nạn trên biển, áp dụng cho khu vực vịnh Bắc Bộ như Hình 6.



Hình 6. Quy trình xử lý dữ liệu mô hình hóa khoảng cách hiệu dụng trong tìm kiếm, cứu nạn trên biển

Dữ liệu được xây dựng ban đầu bao gồm:

+ Hướng gió: Được xác định là hướng gió trung bình theo tháng hoặc khoảng thời gian quan tâm. Đầu vào của hướng gió được xác định bởi người dùng. Tuy nhiên, quá trình xây dựng hướng gió cần đảm bảo một số yêu cầu như sau:

- Hướng gió được xác định có thể theo đơn vị từ 0° đến 360° . Người xử lý dữ liệu có thể đưa ra bất kỳ giá trị nào trong khoảng đơn vị đó. Đồng thời, người xử lý cần đưa thêm tỉ lệ phần trăm xuất hiện hướng gió đó theo khoảng thời gian đánh giá. Bên cạnh đó, người dùng có thể chỉ đánh giá khu vực nghiên cứu theo một hướng gió duy nhất.

- Tổng giá trị phần trăm xuất hiện của các hướng gió phải là 100%. Nếu người xử lý dữ liệu chỉ muốn đánh giá khu vực quan tâm theo một hướng gió cố định, đầu vào của hướng gió sẽ bao gồm hai giá trị là hướng gió và 100 (không cần ghi thêm dấu %).

+ Ảnh đối tượng đất liền - nước mặt: Xây dựng dựa trên các vector khu vực nghiên cứu và biển. Vector được chuyển đổi về raster theo độ phân giải không gian phù hợp với hướng xử lý (đảm bảo quá trình tính toán không bị tràn bộ nhớ). Giá định trong nghiên cứu là $100\text{ m} \times 100\text{ m}$. Đồng thời, đối tượng nước mặt có giá trị 0, đối tượng đất liền có giá trị 1.

+ Tốc độ gió hiệu chỉnh: Xác định theo tốc độ gió chủ yếu. Trong nội dung này, người dùng cần đưa ra các tốc độ gió theo hướng gió đã được đưa ra ở phần trên và thời gian xảy ra. Đầu vào sẽ yêu cầu ba tham số:

- Hướng gió

- Tốc độ gió

- Thời gian xuất hiện tốc độ gió. Thời gian ở đây được hiểu là ngày xảy ra hướng gió và tốc độ gió đó.

Tham số thời gian được định nghĩa dưới dạng dd-mm-yy (ví dụ 061021, có nghĩa là ngày 06 tháng 10 năm 2021).

+ Gia tốc trọng trường đang được sử dụng mặc định.

- + Bản đồ độ sâu được xây dựng dựa trên dữ liệu điểm độ sâu kết hợp với thuật toán Spline.
- + Dữ liệu tàu, tốc độ tàu theo độ cao sóng. Nghiên cứu sử dụng các thông số lực lượng tàu tìm kiếm cứu nạn trên biển đã được mô tả chi tiết trong Mục 2.4.3 làm dữ liệu thử nghiệm.

2.4.3. Các bước thực hiện

Quá trình tính toán các lớp dữ liệu đầu vào phục vụ mô hình hóa khoảng cách hiệu dụng trong tìm kiếm, cứu nạn trên biển khu vực vịnh Bắc Bộ được thực hiện như sau:

- + Lấy gió (*wind fetch*) được tính theo phương pháp SPM. Sau khi tính toán được dữ liệu lấy gió, dữ liệu này được sử dụng cho tính toán chiều cao sóng hiệu chỉnh.
- + Mô hình tính toán chiều cao sóng hiệu chỉnh với các công thức dưới được lấy từ Sổ tay Kỹ thuật Bờ biển (USACE, 2002):

$$\begin{aligned} C_d &\approx 0.001 * (1.1 + (0.035 * U_A)) \\ U_* &= (C_d)^{1/2} * U_A \\ x^{\wedge} &= (g * x) / (U_*)^2 \\ H_{m0}^{\wedge} &= \lambda_1 * (x^{\wedge})^{m1} \\ H_{m0} &= H_{m0}^{\wedge} * ((U_*)^2 / g) \end{aligned} \quad (1)$$

trong đó:

- + U_A = tốc độ gió được điều chỉnh (m/s);
- + z = độ cao quan sát được của tốc độ gió (m);
- + t = số giây để đi được một dặm;
- + C_d = hệ số lực cản;
- + U^* = vận tốc ma sát;
- + $\lambda_1 = 0,0413$;
- + $m1 = 1/2$;
- + H_{m0}^{\wedge} = chiều cao sóng có ý nghĩa không thứ nguyên;
- + H_{m0} = chiều cao sóng đáng kể (m);
- + x^{\wedge} = lấy gió không thứ nguyên;
- + x = lấy gió (m);
- + g = gia tốc trọng trường (9,82 m/s²).

Thời gian đi hết một đơn vị không gian được tính cho tàu tìm kiếm, cứu nạn theo mô hình của USACE như sau (USACE, 2002):

Đầu tiên tính thời gian không hiệu chỉnh: $t = \frac{s}{v}$ với s là độ phân giải không gian, v là tốc độ của tàu. Sau đó, tiến hành chia các khoảng độ cao sóng với các vận tốc khác nhau, mỗi độ cao sóng tính toán các tham số:

$$\begin{aligned} a &= \frac{t_1 - t_2}{wh_1 - wh_2} \\ b &= \frac{wh_1 * t_2 - wh_2 * t_1}{wh_1 - wh_2} \end{aligned} \quad (2)$$

Khi đó, thời gian đi hết 1 đơn vị không gian được tính bằng: $t = wh * a + b$ với wh là độ cao sóng, a và b lần lượt là các tham số ứng với độ cao wh .

Cuối cùng, thời gian phản hồi tìm kiếm và cứu nạn được tính toán theo khoảng cách hiệu dụng với đầu vào là vị trí các trạm cứu nạn cứu hộ và dữ liệu thời gian đi hết một đơn vị không gian.

3. Kết quả xây dựng dữ liệu đầu vào

Trước hết, nghiên cứu tiến hành thu thập các số liệu khí tượng, thủy văn tại khu vực nghiên cứu. Từ đó, các tác giả tiến hành thực nghiệm với 3 kịch bản như sau:

Kịch bản 1: Sử dụng tàu TS-01 của đơn vị Hải quân tìm kiếm, cứu hộ tàu gặp nạn. Thời tiết trời nắng, gió thổi hướng Bắc (0°), tốc độ gió là 5 m/s.

Kịch bản 2: Sử dụng tàu CN-09 của đơn vị Biên phòng tìm kiếm, cứu hộ tàu gặp nạn. Thời tiết trời nắng, gió thổi hướng Đông Bắc (45°), tốc độ gió là 10 m/s.

Kịch bản 3: Sử dụng tàu KN-579 của đơn vị Cục kiểm ngư tìm kiếm, cứu hộ tàu gặp nạn. Thời tiết trời nắng, gió thổi hướng Nam (180°), tốc độ gió là 15 m/s.

3.1. Kết quả xây dựng lớp dữ liệu tàu tìm kiếm, cứu hộ trên biển

Bảng 2. Các tàu tìm kiếm, cứu hộ trên biển sử dụng trong thực nghiệm

Tàu	Đơn vị	Vận tốc tối đa (hải lý/giờ)
TS-01	Hải quân	12
CN-09	Biên phòng	22
KN-579	Cục Kiểm ngư	25

Độ cao sóng thu thập được tại khu vực nghiên cứu không vượt quá 4 m. Chính vì vậy, nghiên cứu giả sử vận tốc tàu thay đổi, giảm dần với các độ cao sóng khác nhau như trong Bảng 3.

Bảng 3. Vận tốc tàu thay đổi với các độ cao sóng khác nhau (hải lý/giờ)

Tàu \ Chiều cao sóng (m)	TS-01	CN-09	KN-579
0	12	22	25
0,5	12	22	25
1	10	20	22
2	9	18	20
3	8	16	18
4	7	14	16

Tiếp theo, dữ liệu thời gian được biểu diễn dưới dạng raster. Một đơn vị không gian sử dụng trong nghiên cứu là 100 m (có nghĩa là 1 điểm ảnh có độ phân giải không gian là 100 m x 100 m). Với số liệu vận tốc tàu thay đổi dựa trên các độ cao sóng khác nhau ở Bảng 3, thời gian để một đơn vị tìm kiếm cứu nạn trên biển đi hết một đơn vị không gian được tính toán trong Bảng 4.

Bảng 4. Thời gian tàu đi hết một đơn vị không gian với các chiều cao sóng khác nhau (s)

Tàu \ Chiều cao sóng (m)	TS-01	CN-09	KN-579
0	16,20	8,84	7,78
0,5	16,20	8,84	7,78
1	19,44	9,72	8,84
2	21,60	10,80	9,72
3	24,30	12,15	10,80
4	27,77	13,89	12,15

Sau đó, các tham số a và b trong phương trình $t = wh * a + b$ dựa trên công thức (2) được tính toán, phục vụ cho quá trình xác định thời gian phản hồi ở các bước tiếp theo. Các tham số a và b ứng với các đơn vị thực nghiệm đối với mỗi khoảng chiều cao sóng được tính toán trong Bảng 5.

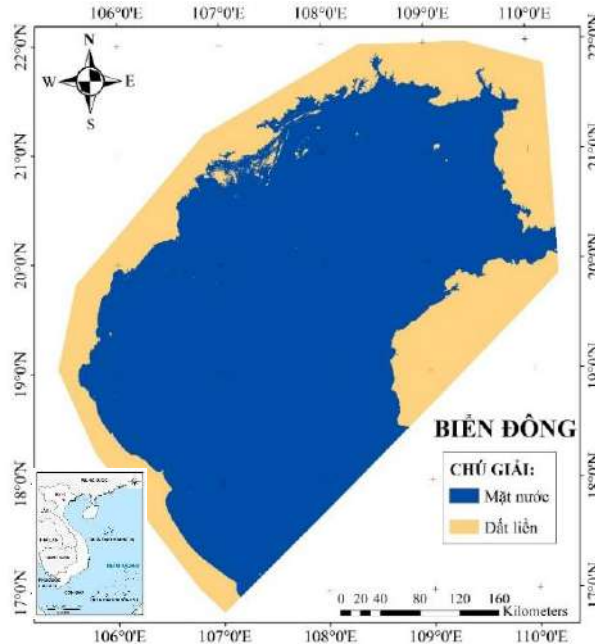
Bảng 5. Các tham số a và b ứng với các đơn vị thực nghiệm đối với mỗi khoảng chiều cao sóng khác nhau

Tàu \ Chiều cao sóng (m)	TS-01		CN-09		KN-579	
	a	b	a	b	a	b
0-0,5	0,00	16,20	0,00	8,84	0,00	7,78
0,5-1	6,48	12,96	1,77	7,95	2,12	6,72
1-2	2,16	17,28	1,08	8,64	0,88	7,95
2-3	2,70	16,20	1,35	8,10	1,08	7,56
3-4	3,47	13,89	1,74	6,94	1,35	6,75

3.2. Kết quả xây dựng lớp dữ liệu đối tượng đất liền - mặt nước

Hình 3.1 trình bày kết quả lớp dữ liệu đối tượng đất liền - mặt nước. Trong hình biểu thị hai màu xanh biển và

vàng, trong đó màu vàng tượng trưng cho các đối tượng đất liền, màu xanh biển tượng trưng cho các đối tượng mặt nước. Về mặt giá trị, các đối tượng đất liền có giá trị 1 và các đối tượng mặt nước có giá trị là 0.



Hình 7. Kết quả lớp dữ liệu đối tượng đất liền - mặt nước

3.3. Kết quả xây dựng lớp dữ liệu hướng gió

Bảng 6 cung cấp thông tin chi tiết về tần suất các hướng gió tại vịnh Bắc Bộ, cho thấy cái nhìn rõ ràng hơn về phân bố hướng gió trong khu vực này. Theo bảng, hướng gió phổ biến nhất là hướng Đông Bắc với góc 45°, chiếm 17,39% tổng số lần khảo sát. Theo sau đó là hướng Bắc với góc 0°, chiếm 15,65%, và hướng Nam với góc 180°, chiếm 13,91%. Cùng chiếm 13,91% là hướng Tây Bắc với góc 315°. Hướng Đông Nam với góc 135° và hướng Tây Nam với góc 225° lần lượt chiếm 12,17% và 10,43%. Hướng Đông với góc 90° chiếm 8,70%, trong khi hướng Tây với góc 270° là ít phổ biến nhất, chỉ chiếm 7,83%.

Dựa trên những con số này, có thể thấy rằng gió thổi từ phía Đông Bắc và Bắc chiếm tỉ lệ cao hơn so với các hướng khác. Sự phân bố này cho thấy một xu hướng gió rõ ràng tại vịnh Bắc Bộ, điều này có thể ảnh hưởng lớn đến các hoạt động hàng hải và các lĩnh vực khác trong khu vực. Đó cũng chính là lý do, các tác giả giả định hướng gió Đông Bắc, Bắc và Nam trong ba kịch bản thử nghiệm.

Bảng 6. Tần suất hướng gió thu thập tại Vịnh Bắc Bộ

Hướng gió	Góc	Tần suất	Phần trăm
Đông Bắc	45°	20	17,39%
Bắc	0°	18	15,65%
Nam	180°	16	13,91%
Tây Bắc	315°	16	13,91%
Đông Nam	135°	14	12,17%
Tây Nam	225°	12	10,43%
Đông	90°	10	8,70%
Tây	270°	9	7,83%

3.4. Kết quả xây dựng lớp dữ liệu tốc độ gió điều chỉnh

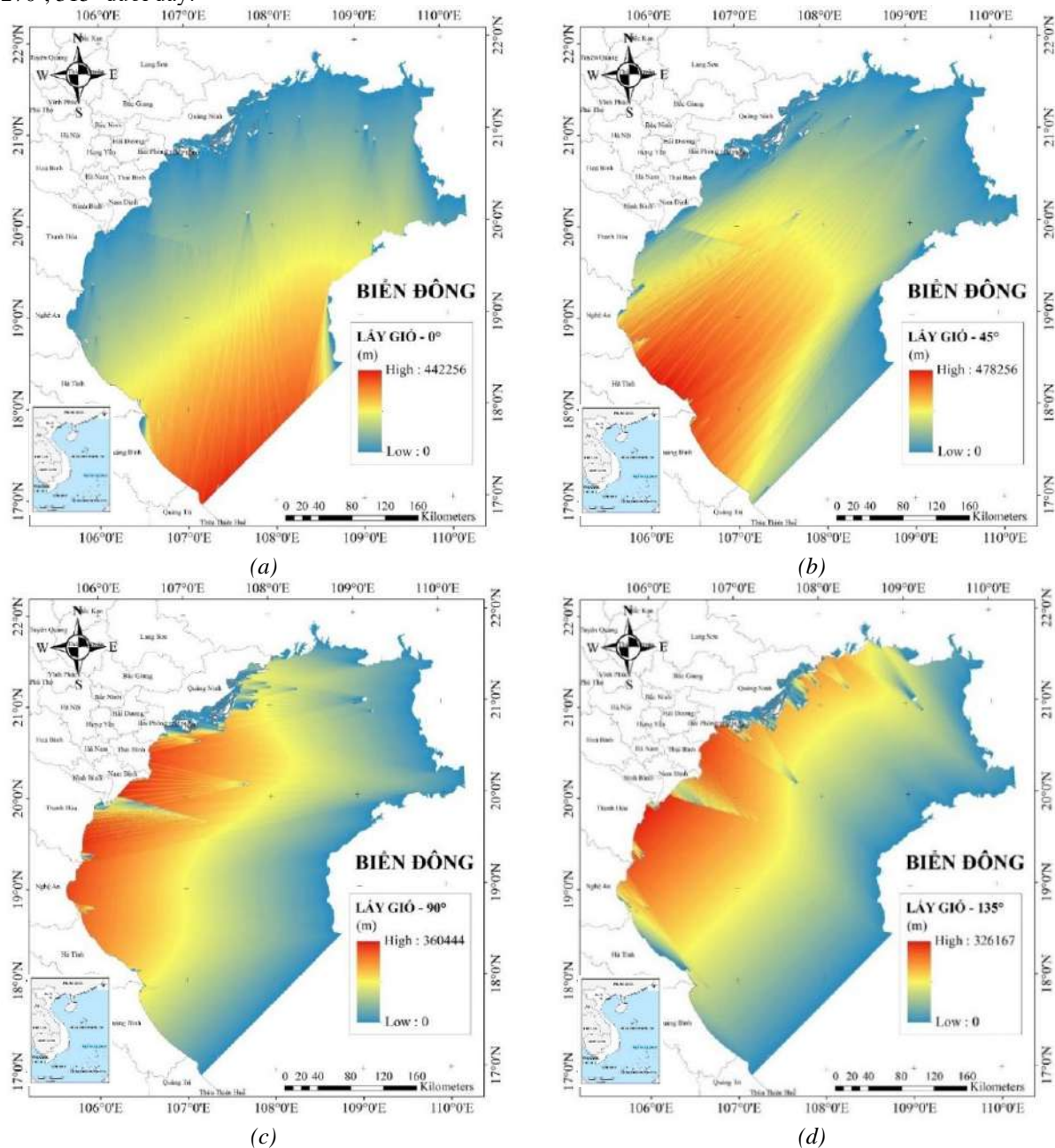
Bảng 7 biểu diễn tần suất xuất hiện của các cấp độ gió đo được tại vịnh Bắc Bộ. Có thể thấy rằng, cấp độ gió từ 1 m/s đến 5 m/s xuất hiện 20 lần, chiếm 17,39% số lần đo. Bên cạnh đó, cấp độ gió từ 6 m/s đến 10 m/s xuất hiện 38 lần, chiếm 33,04% số lần đo. Hơn thế nữa, cấp độ gió từ 10 m/s đến 15 m/s xuất hiện 48 lần, chiếm 41,74% số lần đo. Ngoài ra, cấp độ gió lớn hơn 15 m/s xuất hiện 9 lần, chiếm 7,83% số lần đo. Điều này cho thấy vịnh Bắc Bộ chủ yếu chịu ảnh hưởng của gió nhẹ và gió vừa, với tần suất xuất hiện cao nhất là cấp độ gió từ 10 m/s đến 15 m/s. Dựa trên kết quả thu thập tại Bảng 6, nghiên cứu giả định ba tốc độ gió là 5 m/s, 10 m/s và 15 m/s tương ứng với ba kịch bản thử nghiệm khác nhau.

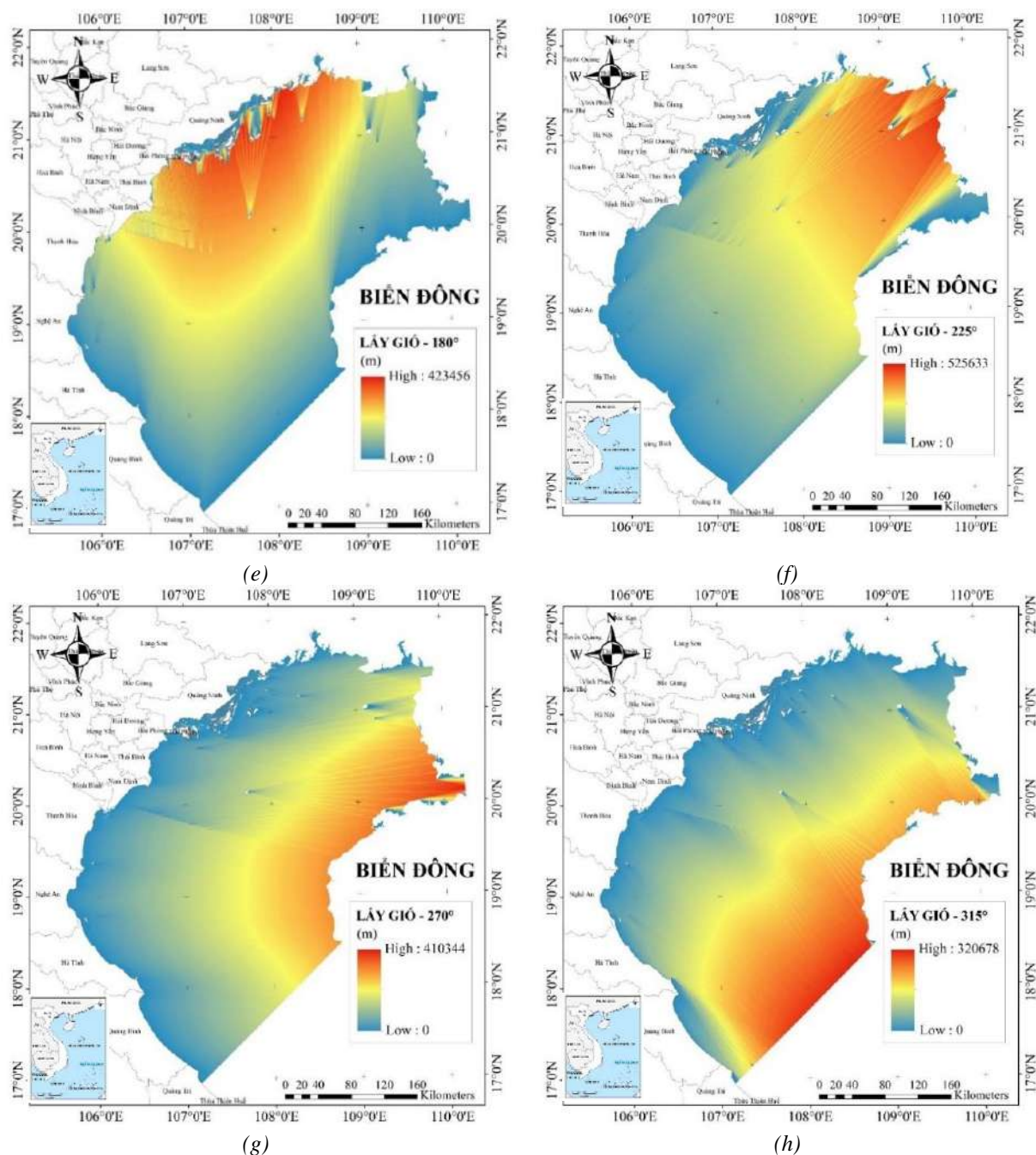
Bảng 7. Tần suất tốc độ gió thu thập tại vịnh Bắc Bộ

Tốc độ gió (m/s)	Tần suất (x 4696 lần)	Phần trăm (%)
1 - 5	20	17,39
6 - 10	38	33,04
10 - 15	48	41,74
> 15	9	7,83

3.5. Kết quả xây dựng lớp dữ liệu ảnh lấy gió

Hình 8 trình bày kết quả các lớp dữ liệu lấy gió (wind fetch) với hướng gió 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315° dưới đây:

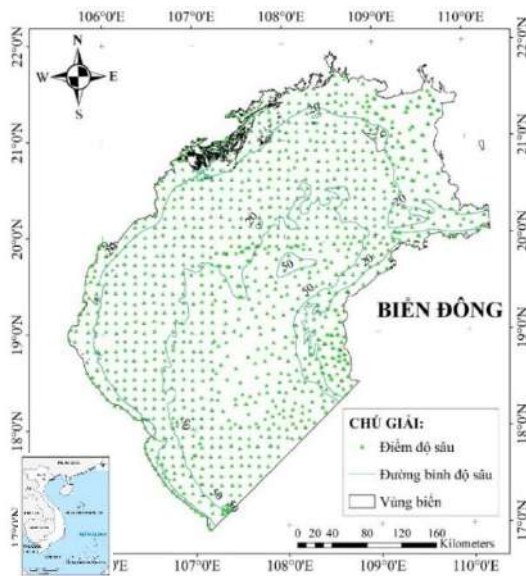




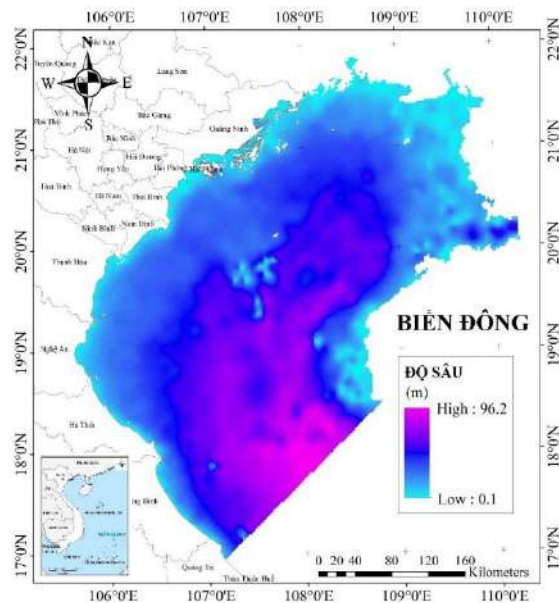
Hình 8. Kết quả các lớp dữ liệu lấy gió (wind fetch) với hướng gió: a) 0°, b) 45°, c) 90°, d) 135°, e) 180°, f) 225°, g) 270°, h) 315°

3.6. Kết quả xây dựng lớp dữ liệu độ sâu

Hình 9 biểu diễn dữ liệu điểm độ sâu thực tế thu thập tại khu vực vịnh Bắc Bộ. Dữ liệu bao gồm 1180 điểm với độ sâu từ 0,1 m đến 96 m. Tiếp theo, các tác giả sử dụng thuật toán Spline để tạo bản đồ độ sâu dạng raster như đã trình bày trong phần trên, xây dựng dữ liệu đầu vào cho mô hình tính toán chiều cao sóng. Dữ liệu độ sâu dạng raster được hiển thị trong Hình 10. Các giá trị độ sâu được tính toán có giá trị từ 0,1 m đến 96,2 m.



Hình 9. Dữ liệu điểm độ sâu thực tế tại khu vực vịnh Bắc Bộ



Hình 10. Dữ liệu bản đồ độ sâu xác định bằng thuật toán Spline

4. Kết luận

Bài báo đã trình bày cơ sở khoa học phương pháp mô hình hoá khoảng cách hiệu dụng phục vụ tìm kiếm và cứu hộ khu vực biển vịnh Bắc Bộ. Trước hết, các yếu tố liên quan tới đặc điểm vịnh Bắc Bộ bao gồm địa hình, khí hậu, thủy - hải văn, môi trường và thông tin về sự cố được thu thập và xây dựng. Bên cạnh đó, nghiên cứu đưa ra các cơ sở khoa học khởi tạo các lớp dữ liệu đầu vào, phục vụ cho xây dựng quy trình mô hình hóa khoảng cách hiệu dụng phục vụ tìm kiếm, cứu hộ trên biển.

Ngoài ra, bài báo cũng đề xuất quy trình mô hình hóa khoảng cách hiệu dụng với thời gian phản hồi tìm kiếm, cứu hộ làm đầu ra cho quá trình xử lý dữ liệu. Bên cạnh đó, các CSDL đầu vào đã được xây dựng gồm: lớp dữ liệu tàu tìm kiếm, cứu hộ trên biển, lớp dữ liệu đối tượng đất liền - mặt nước, lớp dữ liệu hướng gió, lớp dữ liệu tốc độ gió điều chỉnh, lớp dữ liệu ảnh lấy gió và lớp dữ liệu độ sâu.

Tài liệu tham khảo

- Ashrafi B., Kim G., Naseri M., Barabady J., Dhar S., Heo G., Baek S. (2024). An agent-based modelling framework for performance assessment of search and rescue operations in the Barents Sea, *Safety in Extreme Environments*. <https://doi.org/10.1007/s42797-024-00101-2>.
- Castro R. (2014). Gravitational acceleration equation with wavelength and speed of light without using the universal gravitational constant of Newton, *ffhal-00947254v5*, 12 pp.
- Greenberg, J. A., Rueda, C., Hestir, E. L., Santos, M. J., & Ustin, S. L. (2011). Least cost distance analysis for spatial interpolation. *Computers & Geosciences*, 37(2), 272-276.
- Guoxing L., Maofeng L. (2010). SARGIS: A GIS-Based Decision-making Support System for Maritime Search and Rescue, 2010 International Conference on E-Business and E-Government, Doi: 10.1109/ICEE.2010.398.
- Herterich J.G., Dias F. (2019). Extreme long waves over a varying bathymetry, *Journal of Fluid Mechanics*, 878, 481-501.
- Hoffmann M., Roy S., Berger A., Bergmann W., Chan K., Shubbak M., Langhorst J., Schnauder T., Strub O., Buskens C. (2021). Wind affected maneuverability of Tunboat-controlled ships, *IFAC-PapersOnLine* 54(16):70-75. DOI:10.1016/j.ifacol.2021.10.075.
- Nguyễn Đình Hải, Nguyễn Gia Trọng, Phạm Văn Tuấn, Bùi Văn Tòng, Nguyễn Tiến Thành (2023). Nghiên cứu phương pháp quy chuyển độ cao giữa hải đồ và bản đồ địa hình đáy biển dựa trên dữ liệu thủy triều tại khu vực Vịnh Bắc Bộ, *Tạp chí Khí tượng Thủy văn*, 752, 51-61; Doi:10.36335/VNJHM.2023(752), 51-61.
- Peng J., Yang D. (2011). Remote sensing monitoring system for maritime search and rescue, 2011 International Conference of Soft Computing and Pattern Recognition (SoCPaR), Doi 10.1109/SoCPaR.2011.6089122.

Rohweder, J., J.T. Rogala, B.L. Johnson, D. Anderson, S. Clark, F. Chamberlin, D. Potter & K. Runyon (2012). Application of wind fetch and wave models for habitat rehabilitation and enhancement projects, *Contract report prepared for U.S. Army Corps of Engineers' Upper Mississippi River Restoration - Environmental Management Program*. 52 p.

Salameh E., Frappart F., Almar R., Baptista P., Heygster G., Lubac B.,... & Laignel B. (2019). Monitoring beach topography and nearshore bathymetry using spaceborne remote sensing: A review. *Remote Sensing*, 11(19), 2212.

Siljander, M., Venäläinen, E., Goerlandt, F., & Pellikka, P. (2015). GIS-based cost distance modelling to support strategic maritime search and rescue planning: A feasibility study. *Applied Geography*, 57, 54-70.

Shahrabi J. (2003). Spatial and temporal analysis of maritime fishing and shipping traffic incidents, *PhD dissertation*, Dalhousie University.

USACE (2002). *Coastal Engineering Manual 10-2-1100*. Parts 1-6, Washington, DC.

Venäläinen E. (2014). Evaluating voluntary emergency response in the Gulf of Finland, Master's thesis of Geography and Geoinformatics, 118 pp.

Zhang L., Li J., Liu C., Zhang W. (2022). A robust fuzzy speed regulator for unmanned sailboat robot via the composite ILOS guidance, *Nonlinear Dynamics* 110(3):1-16. Doi:10.1007/s11071-022-07763-2.

Zhou X., Cheng L., Min K., Zuo X., Yan Z., Ruan X., Chu S., Li M. (2020). A framework for assessing the capability of maritime search and rescue in the East Sea, *International Journal of Disaster Risk Reduction* 47, 101568, <https://doi.org/10.1016/j.ijdrr.2020.101568>.

ABSTRACT

Establishment of GIS database to serve search and rescue operations in the Gulf of Tonkin

Nguyen Quang Minh^{1*}, Nguyen Van Trung^{2,3}, Bui Tien Dieu⁴

¹Department of Cartography, The General Staff

²Hanoi University of Mining and Geology

³Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology

⁴University of South-Eastern Norway

The Geographic Information System (GIS) technology is widely applied to solve many practical problems. However, the ability to apply GIS in search and rescue plans at sea is still a challenge in our country. Applying GIS tools to build databases to serve the problem of optimizing travel time from rescue stations to search and rescue locations is necessary in our country's sea areas. This article focuses on processing and standardizing data to build component databases including historical search and rescue data, data on fleets and rescue stations, and data on hydrometeorology. The construction of these component databases must ensure standards for use as input data for the model of calculating optimal routes in terms of travel time at sea. In addition, the constructed component databases are also mentioned to provide information on the response capacity for search and rescue plans in the Gulf of Tonkin.

Keywords: Gulf of Tonkin, rescue, GIS, meteorology, wave height.

Phân tích lượng mưa tích lũy và nguy cơ trượt lở đất: Nghiên cứu và mô phỏng số tại Km10+950, Quốc lộ 34, tỉnh Hà Giang

Đỗ Văn Đăng^{1*}, Nguyễn Châu Lân², Hà Thị Hằng³, Bùi Duy Quỳnh³,
Dương Công Hiếu⁴, Lưu Thị Diệu Chinh¹

¹Khoa Công trình thủy, Trường Đại học Xây dựng Hà Nội

²Khoa Công Trình, Trường Đại học Giao thông Vận tải

³Khoa Cầu đường, Trường Đại học Xây dựng Hà Nội

⁴Viện Công nghệ Trắc địa xây dựng, Trường Đại học Xây dựng Hà Nội

TÓM TẮT

Sạt lở đất là mối nguy hiểm đáng kể ở các vùng miền núi, đặc biệt ở các khu vực có địa hình phức tạp như tỉnh Hà Giang, Việt Nam. Nghiên cứu này tập trung phân tích lượng mưa tích lũy với cơ chế hình thành trượt lở đất tại điểm Km10+950 trên Quốc lộ 34 trên cơ sở sử dụng dữ liệu từ 35 trạm đo mưa tự động và ảnh vệ tinh PlanetScope. Bên cạnh đó, nghiên cứu này còn mô phỏng quá trình trượt lở đất tại đây bằng mô hình độ số độ cao DEM được xây dựng từ dữ liệu UAV. Kết quả cho thấy lượng mưa lớn liên tục, kéo dài, đặc biệt từ ngày 02/7 đến ngày 12/7/2024, với tổng cộng là 491,4 mm, đã làm bão hòa đất, từ đó làm tăng đáng kể nguy cơ trượt lở đất. Phần mềm RAMMS được sử dụng để mô phỏng dòng chảy bùn đá, cung cấp thông tin chi tiết về động lực của vụ trượt lở đất này. Các phát hiện này nêu bật nhu cầu thiết yếu về việc xác định ngưỡng mưa tích lũy chính xác trong việc kích hoạt trượt lở đất và thực hiện các biện pháp giảm thiểu hiệu quả. Nghiên cứu này không chỉ nâng cao hiểu biết về cơ chế trượt lở ở các địa hình phức tạp mà còn đóng góp các giải pháp thiết thực nhằm tăng cường an toàn cơ sở hạ tầng và bảo vệ cộng đồng địa phương.

Từ khóa: Trượt lở đất, lượng mưa tích lũy, mô phỏng sạt trượt, RAMMS, Hà Giang.

1. Đặt vấn đề

Lở đất là một trong những loại hình thiên tai phổ biến và gây thiệt hại nghiêm trọng ở các khu vực đồi núi trên toàn thế giới (Froude, Melanie J. và Petley, David N., 2018). Tại Việt Nam, tỉnh Hà Giang được biết đến là một trong những vùng có địa hình phức tạp và thường xuyên xảy ra hiện tượng lở đất, đặc biệt là trong mùa mưa. Trượt lở đất không chỉ gây thiệt hại về tài sản và cơ sở hạ tầng, mà còn đe dọa nghiêm trọng đến tính mạng của người dân. Hệ thống cảnh báo sớm về các mối nguy hiểm tự nhiên là công cụ quan trọng để giảm thiểu rủi ro thiên tai (Guzzetti, Fausto và các cộng sự, 2020). Do đó, việc nghiên cứu và dự báo nguy cơ trượt lở đất là vô cùng quan trọng để có thể đưa ra các biện pháp phòng chống và giảm thiểu thiệt hại.

Nghiên cứu về lượng mưa tích lũy và nguy cơ trượt lở đất đã thu hút được sự quan tâm rộng rãi từ các nhà khoa học trên toàn thế giới. Các nghiên cứu trước đây đã chứng minh rằng lượng mưa tích lũy đóng vai trò quan trọng trong việc kích hoạt các hiện tượng trượt lở đất, đặc biệt ở các khu vực đồi núi có địa hình phức tạp. Iverson, Richard M. (2000) đã phân tích mối quan hệ giữa lượng mưa tích lũy và trượt lở đất, cho thấy rằng lượng mưa liên tục và dài hạn là yếu tố chính dẫn đến sự bão hòa của đất và gây ra trượt lở đất. Guzzetti, Fausto và các cộng sự (2022) đã xây dựng mô hình dự báo trượt lở đất dựa trên dữ liệu lượng mưa và địa chất, cung cấp các bản đồ nguy cơ trượt lở đất có độ chính xác cao. Baum, Rex L. và Jonathan W. Godt (Baum, Rex L. và Godt, Jonathan W., 2009) đã sử dụng mô hình số để mô phỏng quá trình trượt lở đất, kết hợp dữ liệu lượng mưa và điều kiện địa chất để dự báo nguy cơ trượt lở đất trong thời gian thực. Nguyen Duc Ha và các cộng sự (2018) đã xây dựng công cụ xác định ngưỡng mưa và nguy cơ sạt trượt cho khu vực tỉnh Bắc Kạn. Hầu hết các nghiên cứu trước đây đều đã thiết lập một nền tảng vững chắc cho việc hiểu rõ hơn về mối quan hệ giữa lượng mưa tích lũy và nguy cơ trượt lở đất. Tuy nhiên, vẫn còn nhiều thách thức trong việc áp dụng các mô hình số để dự báo trượt lở đất với độ chính xác cao, đặc biệt là tại các khu vực có địa hình, địa chất phức tạp và điều kiện khí hậu biến đổi.

Mục tiêu của nghiên cứu này là tiếp tục phát triển và ứng dụng các mô hình số tiên tiến trong phân tích lượng mưa tích lũy với nguy cơ trượt lở đất, mô phỏng quá trình trượt lở tại vị trí Km10+950, Quốc lộ 34, tỉnh Hà Giang. Bằng cách sử dụng dữ liệu quan trắc thực tế và phương pháp mô hình số, nghiên cứu này nhằm cung cấp các cơ sở khoa học vững chắc trong nhận định nguy cơ hình thành trượt lở, làm cơ sở cho việc đề xuất các biện pháp phòng chống hiệu quả, góp phần giảm thiểu thiệt hại do trượt lở đất gây ra.

* Tác giả liên hệ

Email: dang1511464@huce.edu.vn

2. Khu vực nghiên cứu

Khu vực nghiên cứu được xác định tại Km10+950 trên Quốc lộ 34, thuộc thôn Tả Mò, xã Yên Định, huyện Bắc Mê, tỉnh Hà Giang, đây là một trong những tuyến đường giao thông quan trọng nối tỉnh Hà Giang với các tỉnh miền núi lân cận. Địa hình tại đây chủ yếu là đồi núi dốc, với độ dốc lớn, đất đá không ổn định và dễ bị tác động bởi các yếu tố thời tiết. Cấu trúc địa chất của khu vực này chủ yếu gồm các loại đất đá bờ rời, dễ bị phong hóa và xói mòn dưới tác động của nước mưa. Đặc biệt, lớp đất bề mặt thường bị nứt nẻ và thấm nước, làm giảm độ bền cơ học của đất đá (Hình 1).



Hình 1. Km10+950 trên Quốc lộ 34, thuộc thôn Tả Mò, xã Yên Định, huyện Bắc Mê, tỉnh Hà Giang
(Ảnh chụp từ UAV)

3. Dữ liệu và phương pháp nghiên cứu

3.1. Số liệu sử dụng

a. Dữ liệu lượng mưa

Dữ liệu lượng mưa được thu thập từ thu thập từ 35 trạm đo tự động, gồm 26 trạm của hệ thống đo mưa tự động Vrain và 9 trạm của hệ thống cơ sở ngành Thủy lợi, quan trắc liên tục trong 12 ngày, từ 01-12/7 tại khu vực tỉnh Hà Giang. Các trạm đo mưa tự động này ghi nhận lượng mưa theo giờ, cung cấp thông tin chi tiết về sự biến đổi của lượng mưa trong từng khoảng thời gian cụ thể.

b. Dữ liệu ảnh vệ tinh

Dữ liệu ảnh vệ tinh cung cấp thông tin về sự thay đổi bề mặt đất trước và sau khi xảy ra trượt lở đất, giúp đánh giá mức độ và phạm vi ảnh hưởng. Nghiên cứu này sử dụng ảnh vệ tinh PlanetScope từ cổng thông tin ESA Earth Online với thời gian chụp từ ngày 01/7/2024 đến ngày 16/7/2024, bao gồm các ảnh chụp trước và sau sự kiện trượt lở. Ảnh vệ tinh PlanetScope là dữ liệu ảnh có độ phân giải cao (3 m) nên cho phép quan sát chi tiết các thay đổi trên bề mặt đất và hiện tượng trượt lở.

c. Dữ liệu bay chụp UAV (Unmanned Aerial Vehicle)

Dữ liệu bay chụp UAV được sử dụng để thu thập hình ảnh chi tiết và xây dựng mô hình số độ cao (DEM) với độ phân giải không gian cao tại khu vực nghiên cứu vào ngày 02/8/2024. Dữ liệu DEM được xây dựng này sẽ cung cấp thông tin chi tiết về độ dốc, hướng dòng chảy nước mưa và các yếu tố khác liên quan đến nguy cơ sạt lở.

d. Dữ liệu khảo sát thực địa

Quan sát hiện trường để ghi nhận các đặc điểm địa hình, cấu trúc đất đá và các dấu hiệu trượt lở trên thực địa, từ đó xác định nguyên nhân, đặc điểm và hậu quả của vụ trượt lở đất. Bên cạnh đó, cũng tiến hành thu thập các báo cáo và số liệu thống kê từ Sở Giao thông vận tải và các cơ quan liên quan.

3.2. Phương pháp nghiên cứu

a. Phân tích lượng mưa tích lũy

Tổng hợp dữ liệu lượng mưa từ các trạm đo để tính toán lượng mưa tích lũy trong khoảng thời gian nghiên cứu. Trên cơ sở đó, lập biểu đồ lượng mưa tích lũy và so sánh với các sự kiện sạt lở đất trước đó để xác định ngưỡng kích hoạt sạt lở.

b. Phân tích dữ liệu ảnh vệ tinh

Sử dụng kỹ thuật phát hiện thay đổi (change detection) trong phần mềm ArcGIS Pro để xác định sự thay đổi bề mặt đất trước và sau sự kiện sạt lở trên ảnh vệ tinh PlanetScope. Bên cạnh đó, sử dụng phần mềm Agisoft Metashape để xử lý dữ liệu bay chụp từ UAV và xây dựng mô hình số độ cao DEM của khu vực nghiên cứu. Dữ liệu DEM này sẽ được sử dụng làm nguồn dữ liệu đầu vào trong mô phỏng số quá trình trượt lở.

c. Phương pháp mô phỏng số quá trình trượt lở

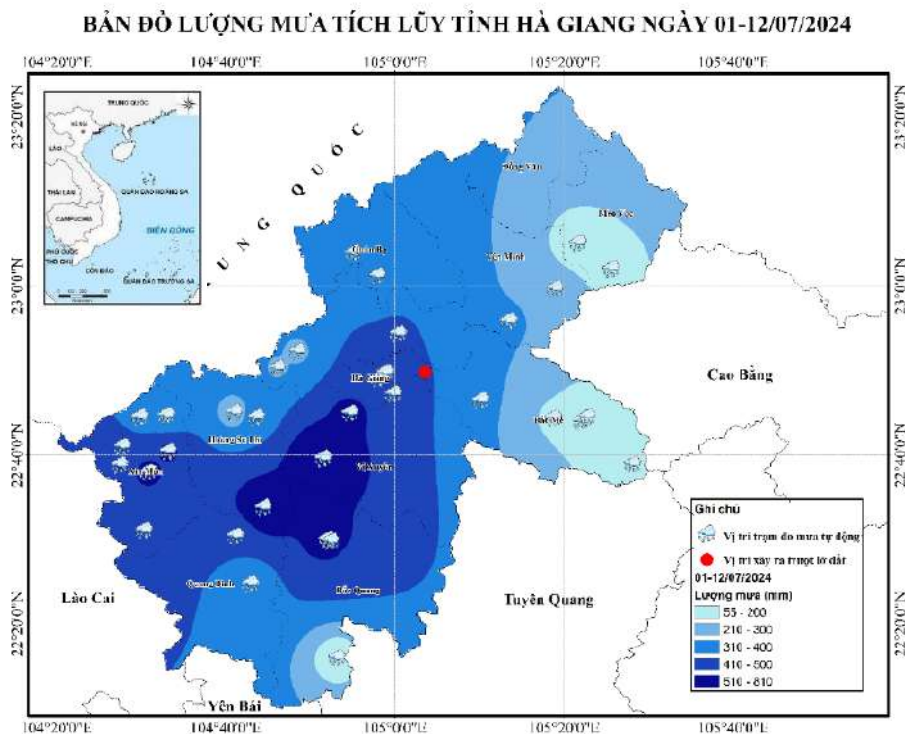
Phần mềm RAMMS (Rapid Mass Movement Simulation) là một phần mềm được phát triển bởi Viện Nghiên cứu Tuyết và Lở tuyết Thụy Sĩ (WSL Institute for Snow and Avalanche Research SLF) nhằm mô phỏng và dự báo các hiện tượng di chuyển khối lượng nhanh như lở tuyết, lở đất, lở đá, và dòng bùn. Phần mềm này sử dụng các thuật toán tiên tiến và mô hình toán học để mô phỏng các hiện tượng di chuyển khối lượng. Nó tính toán dựa trên mô hình địa hình số (DEM) và các thông số vật lý của vật liệu. RAMMS bao gồm nhiều module chuyên biệt để mô phỏng các loại hiện tượng khác nhau, như Avalanche để mô phỏng lở tuyết; Debris flow để mô phỏng chuyển vị của bùn và đất đá; Rockfall để mô phỏng lở đá.

Nghiên cứu này sử dụng module Debris flow trong phần mềm RAMMS để mô phỏng sự chuyển vị của khối đất đá từ trên sườn núi xuống các điểm thấp hơn dựa trên các tính chất cơ lý của vật liệu và địa hình của khu vực xảy ra sạt lở. Các dữ liệu đầu vào cho mô hình bao gồm: DEM độ phân giải cao thu được từ dữ liệu UAV, dữ liệu về kích thước và loại vật liệu của khối trượt.

4. Kết quả và thảo luận

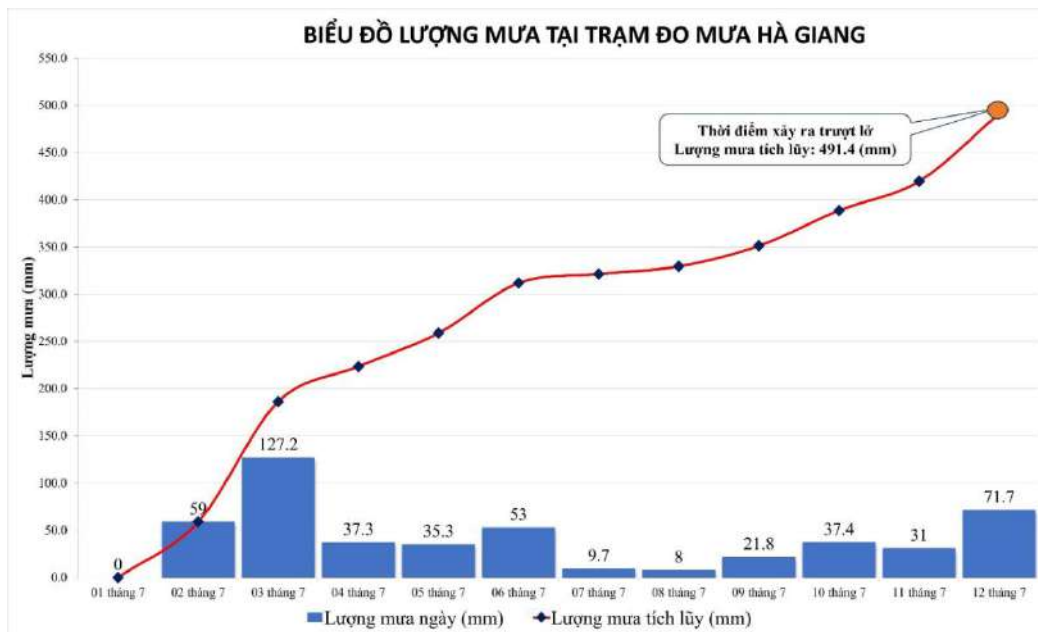
4.1. Phân tích lượng mưa tích lũy với cơ chế hình thành trượt lở

Bản đồ phân vùng lượng mưa tích lũy cho toàn bộ khu vực tỉnh Hà Giang trong khoảng thời gian từ ngày 01/7/2024 đến 12/7/2024 được xây dựng từ dữ liệu lượng mưa thu thập được trong phần mềm ArcGIS Pro (Hình 2). Đối với vị trí xảy ra sạt lở đất vào rạng sáng ngày 13/7, cụ thể tại Km 10+900 trên Quốc lộ 34 thuộc địa phận thôn Tả Mò, xã Yên Định, huyện Bắc Mê, tỉnh Hà Giang, qua phân tích dữ liệu lượng mưa thu thập tại các trạm đo mưa trên địa bàn tỉnh Hà Giang gần vị trí xảy ra sạt lở cho thấy, ngày 01/7 chưa xuất hiện mưa, từ ngày 02/7, mưa xuất hiện liên tục và kéo dài nhiều ngày, trong đó, lượng mưa lớn nhất ghi nhận được là trong ngày 03/7 và ngày 12/7 với tổng lượng mưa tích lũy là 491,4 mm (Hình 3).



Hình 2. Bản đồ lượng mưa tích lũy tại tỉnh Hà Giang từ ngày 01-12/7/2024

Để làm rõ hơn mối quan hệ tương quan giữa lượng mưa với cơ chế hình thành sạt lở đất tại vị trí xảy ra sạt lở, tiến hành xem xét dữ liệu mưa tích lũy trên biểu đồ lượng mưa của sau 8 ngày là 350 mm, sau 11 ngày (từ 02/7 đến 12/7), là 491,4 mm (Hình 3). Lượng mưa tích lũy tăng dần cho đến ngày xảy ra sự cố trượt lở đất, vào rạng sáng ngày 13/7. Thực tế này cho thấy, sạt lở đất có mối quan hệ mật thiết với lượng mưa trong khu vực, khi lượng mưa lớn kéo dài liên tục trong nhiều ngày, làm tăng độ bão hòa nước trong đất, điều này làm cho lực dính giữa các hạt đất giảm đi, đất trở nên yếu hơn và dễ bị trượt lở. Lượng mưa lớn kéo dài nhiều ngày còn làm tăng dòng chảy bề mặt, tạo ra lực tác động mạnh lên trên bề mặt đất và có thể gây ra hiện tượng xói mòn và sạt lở đất. Ngoài ra, với đặc thù của một tuyến quốc lộ vùng núi, tại Km10+900 trên Quốc lộ 34, đoạn thuộc địa phận thôn Tả Mò, xã Yên Định, huyện Bắc Mê, tỉnh Hà Giang, vốn là khúc cua gấp với nhiều đoạn quanh co, độ dốc địa hình lớn, địa hình dốc theo khe tụ thủy, góp phần khiến cho tuyến đường này có nguy cơ sạt lở đất cao khi có yếu tố tác động từ bên ngoài, cụ thể ở đây là lượng mưa. Đồ thị lượng mưa tích lũy và so sánh với các sự kiện sạt lở trước đó cho thấy ngưỡng lượng mưa để kích hoạt sạt lở tại khu vực nghiên cứu nằm trong khoảng 400-500 mm.



Hình 3. Biểu đồ lượng mưa tích lũy tại trạm đo mưa Hà Giang từ ngày 1-12/7/2024 (trạm đo gần điểm sạt lở Km10+950 Quốc lộ 34)

4.2. Phân tích hiện trạng sạt lở đất qua nguồn ảnh vệ tinh độ phân giải siêu cao

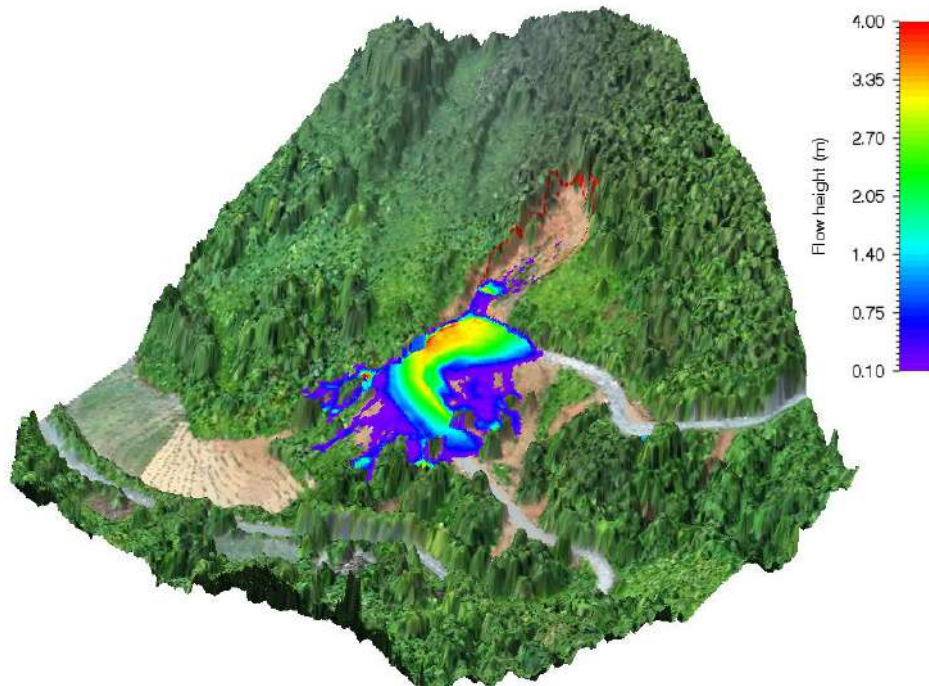
Qua phân tích hình ảnh trên dữ liệu ảnh vệ tinh PlanetScope trước và sau khi xảy ra sự cố sạt trượt tại khu vực Km10+900 trên Quốc lộ 34, có thể nhận thấy rõ những thay đổi trên bề mặt đất, cụ thể là sự thay đổi bề rộng mặt đường, sự thay đổi bề mặt lớp phủ thực vật, sự biến đổi địa hình, sự hư hỏng cơ sở hạ tầng (Hình 4). Kết quả phân tích này cho thấy, sạt lở đất tại Km10+900 trên Quốc lộ 34 có đặc điểm là mảng trượt lớn, rãnh trượt sâu đã gây thiệt hại nghiêm trọng về người và tài sản, khối lượng đất bị xói mòn và trượt lở rộng lớn.



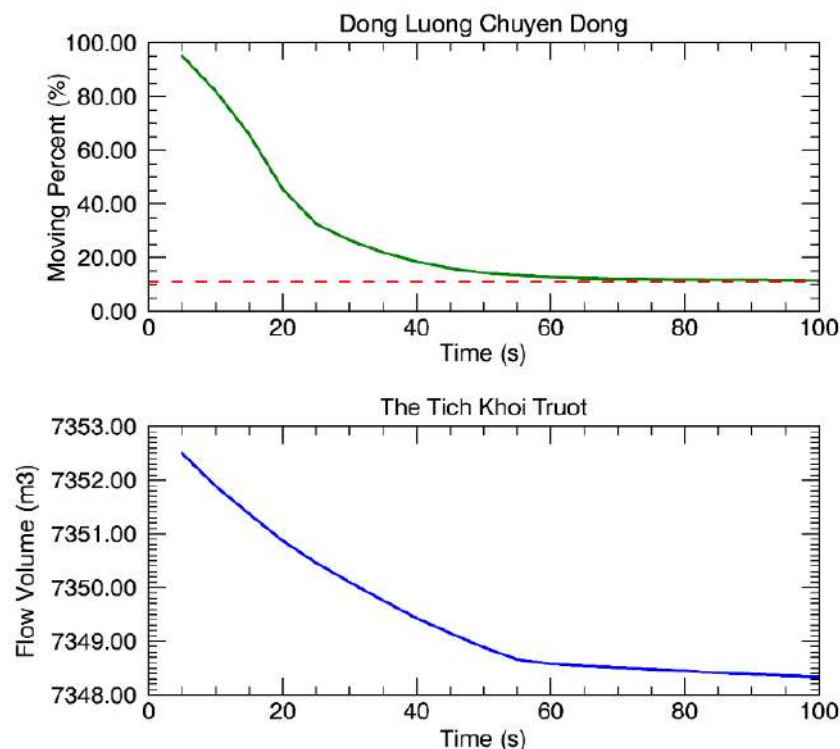
Hình 4. Khu vực Km10+900 trên Quốc lộ 34 thuộc địa phận thôn Tả Mò, xã Yên Định, huyện Bắc Mê tỉnh Hà Giang trước và sau sự cố sạt lở ngày 13/7/2024.

4.3. Kết quả mô phỏng số quá trình trượt lở

Hình 5 mô phỏng lại quá trình sạt lở đất tại Km10+950 trên Quốc lộ 34 qua tỉnh Hà Giang vào rạng sáng ngày 13/7/2024. Mô phỏng cho thấy sự di chuyển của dòng bùn đá từ khu vực sạt lở (khu vực màu đỏ) xuống sườn đồi và tràn ra đường bên dưới. Mô phỏng sử dụng màu sắc để biểu thị chiều sâu của dòng bùn đá (Flow height), với màu đỏ tương ứng với chiều sâu lớn nhất, giảm dần qua các màu sắc khác như cam, vàng, xanh lá và xanh dương. Hình ảnh cung cấp cái nhìn trực quan về mức độ nguy hiểm và phạm vi ảnh hưởng của sự cố sạt lở, giúp đánh giá rủi ro, có thể kết hợp tính toán được áp lực đất đá phục vụ thiết kế các giải pháp gia cố phù hợp.



Hình 5. Mô hình mô phỏng sạt lở tại đoạn Km10+950 trên Quốc lộ 34, tỉnh Hà Giang trong phần mềm RAMMS



Hình 6. Biểu đồ thể hiện sự chuyển động và thể tích của khối trượt theo thời gian

Hình 6 thể hiện kết quả mô phỏng của khối trượt theo thời gian trên hai biểu đồ.

- Đối với biểu đồ phía trên, trục tung thể hiện phần trăm khối lượng đang chuyển động (%), trục hoành thể hiện thời gian tính bằng giây (s). Đường màu xanh lá cây thể hiện sự giảm dần của phần trăm khối lượng đang chuyển động theo thời gian. Ban đầu, phần trăm khối lượng chuyển động là 100%, nhưng giảm nhanh chóng và ổn định dần khi thời gian tiến triển. Điều này cho thấy khối trượt ban đầu có sự chuyển động mạnh mẽ, nhưng sau đó dần ổn định và gần như ngừng lại sau khoảng 80 giây.

- Đối với biểu đồ phía dưới, trục tung thể hiện thể tích dòng chảy tính bằng mét khối (m^3), trục hoành thể hiện thời gian tính bằng giây (s). Đường màu xanh lam cho thấy sự giảm dần của thể tích dòng chảy theo thời gian. Ban đầu, thể tích nằm ở mức 7353 m^3 và giảm dần theo thời gian, điều này cho thấy khối lượng vật chất trong khối trượt dần giảm do sự di chuyển và phân tán của nó.

Tóm lại, kết quả mô phỏng này cho thấy khối trượt ban đầu có sự chuyển động và phân tán mạnh mẽ, nhưng sau một khoảng thời gian, quá trình này dần ổn định và gần như ngừng lại, với sự giảm dần của cả phần trăm khối lượng đang chuyển động và thể tích dòng chảy.

5. Kết luận

Nghiên cứu này đã thực hiện phân tích lượng mưa tích lũy với cơ chế hình thành trượt lở đất và mô phỏng số quá trình trượt lở tại Km10+950 trên Quốc lộ 34, tỉnh Hà Giang, dựa trên dữ liệu lượng mưa quan trắc thực tế, dữ liệu ảnh vệ tinh và dữ liệu DEM từ UAV. Kết quả nghiên cứu đã chỉ ra rằng lượng mưa tích lũy có tác động lớn đến sự hình thành các sự cố trượt lở đất, đặc biệt trong điều kiện địa hình phức tạp của khu vực này. Cụ thể, ngưỡng lượng mưa để kích hoạt trượt lở đất tại khu vực nghiên cứu đã được xác định nằm trong khoảng 400-500 mm. Bên cạnh đó, việc mô phỏng quá trình trượt lở đất sử dụng module Debris flow trong phần mềm RAMMS cũng cho thấy phần trăm khối lượng chuyển động và thể tích dòng chảy di chuyển theo thời gian. Nhìn chung, với việc sử dụng và tổng hợp dữ liệu lượng mưa kết hợp với phân tích ảnh vệ tinh đã phần nào cho thấy tính hiệu quả của công nghệ mới trong đánh giá hiện trạng sạt lở đất, cùng với việc mô phỏng số quá trình trượt lở trong phần mềm RAMMS cũng góp phần dự báo quỹ đạo, vận tốc và sức mạnh của các dòng di chuyển trong khu vực. Những kết quả đạt được này có thể cung cấp những thông tin quan trọng trong việc định hướng, khoanh vùng những khu vực có nguy cơ sạt lở trong tương lai, như những cung đường có khúc cua hẹp, địa hình dốc cùng với lượng mưa tích lũy lớn trong nhiều ngày, đây sẽ là cơ sở phục vụ cho công tác cảnh báo sớm cũng như phòng tránh và giảm thiểu thiệt hại do thiên tai. Bên cạnh đó, các kết quả nghiên cứu này có thể áp dụng được trong việc thiết lập hệ thống cảnh báo sớm và quy hoạch các biện pháp phòng chống trượt lở đất hiệu quả, góp phần vào việc giảm thiểu rủi ro thiên tai tại các khu vực đồi núi phức tạp.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi dự án VLIR-UOS TEAM (VN2022TEA533A105), “Cơ sở hạ tầng dữ liệu GEOdata và khoa học công dân hỗ trợ sự phát triển bền vững của cộng đồng nông thôn tỉnh Quảng Nam (GEOSCIRE)”.

Tài liệu tham khảo

- Baum, Rex L., Godt & Jonathan W. (2009). "Early warning of rainfall-induced shallow landslides and debris flows in the USA", *Landslides*, 7(3), pp. 259-272.
- Froude, Melanie J., Petley & David N. (2018). "Global fatal landslide occurrence from 2004 to 2016", *Natural Hazards and Earth System Sciences*, 18(8), pp. 2161-2181.
- Guzzetti, Fausto et al. (2020). "Geographical landslide early warning systems", *Earth-Science Reviews*. 200.
- Guzzetti, Fausto et al. (2022). "Chapter 15 - Rainfall and landslide initiation", In: Morbidelli, Renato (ed.), *Rainfall*, Elsevier, pp. 427-450.
- Iverson, Richard M. (2000), "Landslide triggering by rain infiltration", *Water Resources Research*, 36(7), pp. 1897-1910.
- Nguyen Duc Ha et al. (2018). "TXT-tool 2.084-3.1 Rainfall Thresholds for Triggering Geohazards in Bac Kan province (Vietnam)", In: *Landslide Dynamics: ISDR-ICL Landslide Interactive Teaching Tools : Volume 1: Fundamentals, Mapping and Monitoring*, K. Sasa et al, Springer International Publishing, Cham, pp. 351-360, <http://link.springer.com/10.1007/978-3-319-57774-6>.

ABSTRACT

Analysis of accumulated rainfall and landslide risk and numerical model simulation: Case study at Km10+950, National Highway 34, Ha Giang province

Do Van Dang¹, Nguyen Chau Lan², Ha Thi Hang³, Bui Duy Quynh³, Duong Cong Hieu⁴, Luu Thi Dieu Chinh¹

¹*Faculty of Hydraulic Engineering, Hanoi University of Civil Engineering*

²*Faculty of Engineering, University of Transport and Communication*

³*Faculty of Bridges and Roads, Hanoi University of Civil Engineering*

⁴*Institute of Geodetic Engineering Technology, Hanoi University of Civil Engineering*

Landslides are a significant hazard in mountainous regions, particularly in areas with complex terrain like Ha Giang province, Vietnam. This study focuses on analyzing cumulative rainfall and its correlation with landslide risks, specifically at the Km10+950 mark on National Highway 34. Using data from 35 automatic rain gauge stations, satellite images, and UAV data, we have developed a digital elevation model (DEM) to simulate landslide events. The results indicate that continuous heavy rainfall, particularly from July 2 to July 12, 2024, totalling 491.4 mm, saturated the soil, thereby significantly increasing landslide risks. The RAMMS software was used to simulate debris flow, providing detailed insights into the dynamics of the landslide. The findings highlight the critical need for precise rainfall thresholds to predict landslides and implement effective mitigation measures. This research not only advances the understanding of landslide mechanisms in complex terrains but also contributes practical solutions to enhance infrastructure safety and protect local communities.

Keywords: Landslide, cumulative rainfall, landslide simulation, RAMMS, Ha Giang.

Nghiên cứu một số vấn đề chuyển đổi tọa độ từ WGS-84 về hệ tọa độ trắc địa sử dụng tại Việt Nam

Hoàng Ngọc Hà^{1*}

¹Trường Đại học Mở - Địa chất

TÓM TẮT

Hiện nay công nghệ GNSS được sử dụng rộng rãi trong công tác Trắc địa - Bản đồ - Viễn thám. Kết quả xử lý số liệu GNSS được xác định trong hệ tọa độ WGS-84. Tuy nhiên, từ khi công bố hệ tọa độ VN-2000, ở Việt Nam sử dụng Elipsoit với kích thước như WGS-84 nhưng được định vị lại. Trong bài báo đề cập vấn đề nghiên cứu chuyển đổi tọa độ trên cơ sở ứng dụng lý thuyết Molodenski. Xem xét một số bài toán cụ thể, trong đó có bài toán quan trắc biến dạng. Các kết quả này góp phần xây dựng cơ sở khoa học xây dựng cơ sở dữ liệu không gian ở nước ta.

Từ khóa: GNSS, chuyển đổi tọa độ, quan trắc biến dạng.

1. Đặt vấn đề

Công nghệ GNSS đã tạo ra bước tiến lớn trong lĩnh vực trắc địa- bản đồ nói chung và quan trắc biến dạng nói riêng. Công nghệ GNSS ở nước ta ứng dụng hệ tọa độ WGS-84. Tuy nhiên, các yêu cầu về pháp lý cần phải xác định trong hệ tọa độ VN-2000. Việc tính toán này đôi khi gặp khó khăn do chúng ta vẫn cần bộ tham số tính chuyển tin cậy trong các khu vực khác nhau. Trong một số trường hợp cụ thể có thể không nhất thiết chuyển đổi tọa độ mà vẫn đạt được mục tiêu đề ra một cách chặt chẽ. Trong bài báo đã nghiên cứu vấn đề này với bài toán phân tích biến dạng sử dụng công nghệ GNSS.

Trong đo đạc truyền thống, việc xác định biến dạng chuyển dịch lún thường phải thực hiện bằng việc xây dựng lưới thủy chuẩn và đo theo chu kỳ. Công việc này đòi hỏi thời gian đo dài và tốn nhiều công sức. Ứng dụng công nghệ GNSS cho phép xây dựng lưới với độ chính xác cao với thời gian ngắn, trong phạm vi rộng và cho phép xác định vector chuyển dịch vector chuyển dịch đứng của vỏ Trái đất. Công nghệ GNSS có ưu việt hơn công nghệ truyền thống là không đòi hỏi sự thông hướng giữa các điểm, đo đạc được tiến hành trong mọi điều kiện thời tiết. Một số công bố quốc tế đã tập trung xây dựng các mô hình ứng dụng phép lọc Kalman để ứng dụng trong lĩnh vực phân tích biến dạng như (Welsch, W. M. and Heunecke, O., 2001), phân tích chuyển dịch phục vụ điều tra sạt lở (Yalcinkaya, M. and Bayrak, T., 2003),...

Ở Việt Nam, công nghệ GNSS đã được ứng dụng trong một số nghiên cứu về chuyển dịch của đới đứt gãy Sông Đà và đới đứt gãy Sơn La - Bim Sơn (Vy Quốc Hải, Trần Đình Tô và Dương Chí Công, 2005) của Viện Địa chất, nghiên cứu chuyển dịch của vỏ Trái đất trên khu vực đứt gãy Lai Châu - Điện Biên (Hà Minh Hòa và cộng sự, 2004).

Trong bài báo ngoài vấn đề hệ tọa độ trong phân tích biến dạng còn tập trung nghiên cứu ứng dụng lý thuyết bình sai lưới GNSS kết hợp với phép lọc Kalman để giải quyết bài toán phân tích, đánh giá sự chuyển dịch đứng.

2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1. Nghiên cứu đánh giá sự dịch chuyển do biến dạng trong hệ tọa độ WGS-84 và VN-2000

Để tính toán chuyển đổi tọa độ trắc địa giữa 2 hệ tọa độ thường dùng công thức Molodenski chuẩn với 5 tham số chuyển đổi: a, df, X_0, Y_0, Z_0 . Dạng ma trận của công thức này là:

$$\begin{bmatrix} B \\ L \\ H \end{bmatrix}_2 = \begin{bmatrix} B \\ L \\ H \end{bmatrix}_1 + \begin{bmatrix} dB \\ dL \\ dH \end{bmatrix} \quad (1)$$

*Tác giả liên hệ:

Email: hoanghakhcn@gmail.com

$$\begin{bmatrix} \frac{M+H}{\rho} dB \\ \frac{N+H}{\rho} \cos B dL \\ dH \end{bmatrix} = \begin{bmatrix} \frac{1}{a} Ne^2 \sin B \cos B & (\frac{a}{b} M + \frac{b}{aN}) \sin B \cos B \\ 0 & 0 \\ \frac{a}{N} & \frac{b}{a} N \sin^2 B \end{bmatrix} \begin{bmatrix} da \\ df \end{bmatrix} + \begin{bmatrix} \sin B \cos L & -\sin B \sin L & \cos B \\ -\sin L & \cos B & 0 \\ \cos B \cos L & \cos B \sin L & \sin B \end{bmatrix} \begin{bmatrix} dX_o \\ dY_o \\ dZ_o \end{bmatrix} \quad (2)$$

Đối với 2 hệ tọa độ WGS-84 và VN-2000 có cùng kích thước Elipxoit nên $da=0$ và $df=0$. Khi đó công thức(1) sẽ có dạng sau:

$$\begin{bmatrix} dB \\ dL \\ dH \end{bmatrix} = S^{-1} G \begin{bmatrix} dX_o \\ dY_o \\ dZ_o \end{bmatrix} \quad (3)$$

Ở đây:

$$S = \begin{bmatrix} \left(\frac{M+H}{\rho}\right) & 0 & 0 \\ 0 & \left(\frac{N+H}{\rho}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$G = \begin{bmatrix} \sin B \cos L & -\sin B \sin L & \cos B \\ -\sin L & \cos B & 0 \\ \cos B \cos L & \cos B \sin L & \sin B \end{bmatrix} \quad (5)$$

Đối với bài toán quan trắc biến dạng, so sánh giá số tọa độ giữa 2 chu kỳ, có thể thấy:

$$\begin{bmatrix} \Delta B \\ \Delta L \\ \Delta H \end{bmatrix}_{WGS-84} - \begin{bmatrix} \Delta B \\ \Delta L \\ \Delta H \end{bmatrix}_{VN2000} = S^{-1} O \begin{bmatrix} dX_o \\ dY_o \\ dZ_o \end{bmatrix} \quad (6)$$

Ở đây ma trận O trong công thức (6) gồm các thành phần O_i gần bằng 0.

2.2. Xác định độ cao trắc địa từ kết quả bình sai lưới tự do với ẩn số là tọa độ B,L,H

Hệ phương trình các số hiệu chỉnh tương ứng với n trị đo ΔX_{ij} , ΔY_{ij} , ΔZ_{ij} bình sai lưới không gian GNSS trong hệ tọa độ địa tâm (X, Y, Z)

$$V = AD + L \quad (7)$$

Với A là ma trận hệ số với các hệ số +1 và -1 ứng với các thành phần tọa độ tương ứng trong trị đo (ΔX_{ij} , ΔY_{ij} , ΔZ_{ij}), Δx là vector ẩn số, V và L tương ứng là vector số hiệu chỉnh và số hạng tự do trong hệ phương trình. Ma trận trọng số P có dạng:

$$P = \begin{pmatrix} P_1 & 0 & 0 & 0 \\ 0 & P_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & P_n \end{pmatrix} \quad (8)$$

$$P_i = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} = Q_i^{-1} \quad (9)$$

Với Q_i là ma trận hiệp phương sai của các trị đo (ΔX_{ij} , ΔY_{ij} , ΔZ_{ij}).

Đối với mỗi điểm lưới GNSS, chúng ta có công thức sau:

$$\begin{pmatrix} dB_i \\ dL_i \\ dH_i \end{pmatrix} = \begin{pmatrix} (M_i + H_i)\sin B_i \cos L_i & (N_i + H_i)\cos B_i \sin L_i & -\cos B_i \cos L_i \\ (M_i + H_i)\sin B_i \sin L_i & (N_i + H_i)\cos B_i \cos L_i & -\cos B_i \sin L_i \\ (M_i + H_i)\cos B_i & 0 & -\sin B_i \end{pmatrix} \begin{pmatrix} dX_i \\ dY_i \\ dZ_i \end{pmatrix} \quad (10)$$

Như vậy, nếu ẩn số là B, L, H thì hệ phương trình các số hiệu chỉnh (1) có thể viết lại như sau:

$$V = A_z Dz + L \quad (11)$$

Ở đây:

Ma trận $A_z = P A$

$$P = \begin{pmatrix} P_1 & & \\ & P_2 & \\ & \dots & \\ & & P_n \end{pmatrix} \quad (12)$$

$$P_i = \begin{pmatrix} (M_i + H_i)\sin B_i \cos L_i & (N_i + H_i)\cos B_i \sin L_i & -\cos B_i \cos L_i \\ (M_i + H_i)\sin B_i \sin L_i & (N_i + H_i)\cos B_i \cos L_i & -\cos B_i \sin L_i \\ (M_i + H_i)\cos B_i & 0 & -\sin B_i \end{pmatrix} \quad (13)$$

Từ hệ phương trình số hiệu chỉnh, ta lập được hệ phương trình chuẩn có dạng:

$$R_z Dz + b_z = 0 \quad (14)$$

$$R_z = A_z^T P A_z, b_z = A_z^T P L \quad (15)$$

Ở đây ma trận R_z suy biến.

$$Dz = -\tilde{R}_z b_z \quad (16)$$

Trong đó, \tilde{R}_z là ma trận nghịch đảo tổng quát:

$$\tilde{R}_z = (R_z + CC^T)^{-1} - TT^T \quad (17)$$

Với:

$$T^T = B(C^T G)^{-1} \quad (18)$$

$$G^T = (G_1 \quad G_2 \quad \dots \quad G_n) \quad (19)$$

G_i là ma trận Hermet, khi các trị đo là baseline ($\Delta X, \Delta Y, \Delta Z$), ma trận G có dạng:

$$G_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (20)$$

$$C^T = (G_1 \quad \dots \quad G_k \quad 0 \quad \dots \quad 0) \quad (21)$$

Trong trường hợp $n = k$ (tức $C = G$), ma trận nghịch đảo tổng quát được tính theo công thức:

$$R_z^+ = (R_z + GG^T)^{-1} - TT^T \quad (22)$$

$$T^T = G(G^T G)^{-1} \quad (23)$$

2.3. Mô hình động và phép lọc Kalman

Triển khai công thức Taylor của hàm số $H_j^{(k+1)}$ theo biến số thời gian t , mô hình động theo thời gian với độ cao trắc địa, vận tốc và gia tốc được biểu diễn bằng công thức sau (Markuze Y U. I., Hoàng Ngọc Hà, 1991):

$$H_j^{(k+1)} = H_j^{(k)} + (t_{k+1} - t_k)v_{Hj} + 1/2 (t_{k+1} - t_k)^2 a_{Hj} \quad (24)$$

$H_j^{(k+1)}$: Độ cao trắc địa điểm j ở thời điểm chu kỳ (t_{k+1})

$H_j^{(k)}$: Độ cao trắc địa điểm j ở thời điểm chu kỳ (t_k)

v_{Hj} : Vận tốc chuyển dịch của H của điểm j

a_{Hj} : Gia tốc chuyển dịch của H của điểm j
 $k=1, 2, \dots, i$ (i: số thứ tự chu kỳ đo)
 $j=1, 2, \dots, n$ (n: số thứ tự của điểm trong lưới)
 Như vậy:

$$\left. \begin{aligned} H_j^{(k+1)} &= H_j^{(k)} + (t_{k+1}-t_k) + 0,5(t_{k+1}-t_k)^2 \\ v_{Hj}^{(k+1)} &= v_{Hj}^{(k)} + (t_{k+1}-t_k)a_{Hj}^{(k)} \\ a_{Hj}^{(k+1)} &= a_{Hj}^{(k)} \end{aligned} \right\} \quad (25)$$

Biểu thức (25) có thể viết lại dưới dạng ma trận như sau:

$$\begin{pmatrix} H_j^{(k+1)} \\ v_{Hj}^{(k+1)} \\ a_{Hj}^{(k+1)} \end{pmatrix} = \begin{pmatrix} 1 & (t_{k+1}-t_k) & 0,5(t_{k+1}-t_k)^2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} H_j^{(k)} \\ v_{Hj}^{(k)} \\ a_{Hj}^{(k)} \end{pmatrix} \quad (26)$$

Ký hiệu:

$$T_j^{(k+1)} = \begin{pmatrix} 1 & (t_{k+1}-t_k) & 0,5(t_{k+1}-t_k)^2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad (27)$$

$$y_j^{(k+1)} = \begin{pmatrix} H_j^{(k+1)} \\ v_{Hj}^{(k+1)} \\ a_{Hj}^{(k+1)} \end{pmatrix} \quad (28)$$

$$y_j^{(k)} = \begin{pmatrix} H_j^{(k)} \\ v_{Hj}^{(k)} \\ a_{Hj}^{(k)} \end{pmatrix} \quad (29)$$

$$y_j^{(k+1)} = T_j^{(k+1)} y_j^{(k)} \quad (30)$$

Phép lọc Kalman được sử dụng để dự đoán vector trạng thái hiện tại bằng cách sử dụng thông tin vector trạng thái của các thông số chuyển động đã biết ở chu kỳ t_k và các phép đo ở chu kỳ T_{k+1} . được thu thập ở kỳ T_{k+1} . Vector trạng thái của tham số chuyển động bao gồm các tham số vị trí, vận tốc chuyển động và gia tốc. Hệ phương trình dạng ma trận của mô hình chuyển động được sử dụng để dự báo các tham số chuyển động bằng kỹ thuật lọc Kalman trong lưới có thể được biểu diễn như sau:

$$\bar{Y}_{(k+1)} = T_{(k+1)} \bar{Y}_{(k+1)} \quad (31)$$

Ở đây:

$$\bar{Y}_{(k+1)} = \begin{pmatrix} y_1^{(k+1)} \\ y_2^{(k+1)} \\ \dots \\ y_n^{(k+1)} \end{pmatrix} \quad (32)$$

$$\bar{Y}_{(k)} = \begin{pmatrix} y_1^{(k)} \\ y_2^{(k)} \\ \dots \\ y_n^{(k)} \end{pmatrix} \quad (33)$$

$$T_{(k+1)} = \begin{pmatrix} T_1^{(k+1)} \\ \dots \\ T_n^{(k+1)} \end{pmatrix} \quad (34)$$

$\hat{Y}_{(k+1)}^T$ - vector giá trị dự báo của tọa độ, vận tốc và gia tốc

\hat{Y}_k^T - vector trạng thái tại thời điểm T_k .

T_{k+1} - ma trận chuyển dịch.

Phương trình (31) là phương trình dự báo trong phép lọc Kalman

Phương trình có yếu tố nhiễu sẽ như sau:

$$\hat{Y}_{(k+1)} = T_{(k+1)} \hat{Y}_k + S_{k+1,k} \quad (35)$$

$$C_{\hat{Y}(k+1)} = T_{k+1,k} C_Y T_{k+1,k}^T + S_{k+1,k} C_S S_{k+1,k}^T \quad (36)$$

Ở đây, ma trận C_Y là ma trận hiệp phương sai của vector trạng thái \hat{Y}_k đã được lọc ở thời điểm t_k .

Ma trận C_S là ma trận hiệp phương sai của vector T nhiễu hệ thống ở thời điểm T_k .

Tại thời điểm t_{k+1} chúng ta thực hiện đo đặc lưới GNSS, có thể thành lập hệ phương trình đo (phương trình lọc) như sau:

$$v_{l,k+1} = A_{k+1} \hat{Y}_k - l_{k+1} \quad (37)$$

Hay là:

$$l_{k+1} + v_{l,k+1} = A_{k+1} \hat{Y}_k \quad (38)$$

Từ các biểu thức (30) và (34) chúng ta có công thức sau:

$$V = \hat{A}_{k+1} \hat{Y}_{k+1} + L \quad (39)$$

Ký hiệu:

$$V = \begin{pmatrix} v_{k+1} \\ v_{l,k+1} \end{pmatrix} \quad (40)$$

$$\hat{A}_{k+1} = \begin{pmatrix} E \\ A_{k+1} \end{pmatrix} \quad (41)$$

$$L = - \begin{pmatrix} \hat{Y}_{k+1} \\ l_{k+1} \end{pmatrix}$$

Ma trận hiệp phương sai:

$$C_{k+1} = \begin{pmatrix} C_{\hat{Y}(k+1)} & 0 \\ 0 & C_{l,k+1} \end{pmatrix} \quad (42)$$

Ma trận A_{k+1} nếu lấy giá trị đo bằng chính giá trị bình sai ở chu kỳ $k+1$ sẽ là: $A_{k+1}^T = (100)$

Ma trận $E_{(k \times k)}$ - Ma trận đơn vị, k số ẩn số tọa độ.

3. Kết quả nghiên cứu và thảo luận

3.1. Kết quả nghiên cứu

3.1.1. Khảo sát đánh giá sự dịch chuyển do biến dạng trong hệ tọa độ WGS-84 và VN-2000

Bảng 1. Giá trị tọa độ trắc địa trong hệ WGS-84

Tên điểm	B			L			H
	WGS-84						
1	22	36	21,13080	104	05	09,91421	320,198
2	22	10	01,68093	104	43	49,64803	109,290
3	21	48	11,15885	105	03	33,23984	89,926
4	21	24	24,01892	105	28	17,80833	22,965
5	21	08	01,69835	105	47	29,73300	10,787
6	22	39	12,16087	104	54	53,06101	111,167
7	22	11	40,18437	105	12	24,38548	71,243
8	21	48	41,28093	105	26	11,20578	85,977
9	21	30	17,50599	105	48	52,27588	23,207
10	20	52	59,22781	105	47	01,81859	6,000
11	21	08	01,69765	105	47	29,73294	10,781

Bảng 2. Giá trị tọa độ trắc địa trong hệ VN-2000

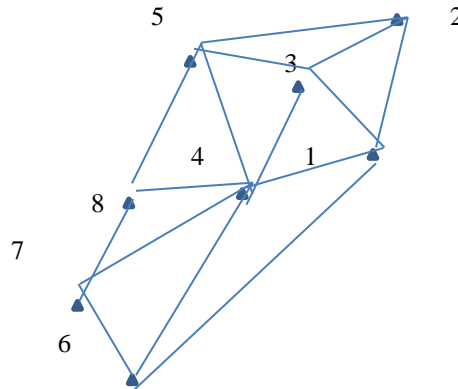
Tên điểm	B			L			H
	VN-2000						
1	22	36	17,68273	104	05	16,77838	285,098
2	22	09	58,19687	104	43	56,48721	77,046
3	21	48	07,65447	105	03	40,05940	59,406
4	21	24	20,49083	105	28	24,60631	-5,495
5	21	07	58,15295	105	47	36,51583	-16,137
6	22	39	08,67816	104	54	59,92280	78,595
7	22	11	36,68081	105	12	31,22273	40,453
8	21	48	37,76112	105	26	18,02291	56,639
9	21	30	13,96557	105	48	59,07557	-4,349
10	20	52	55,67882	105	47	08,59018	-20,471
11	21	07	58,15226	105	47	36,51577	-16,142

Bảng 3. Độ lệch giá trị biến dạng giữa 2 hệ WGS-84 và VN-2000

Tên điểm	$dB_{2000}-dB_{84}$	$dL_{2000}-dL_{84}$	$dH_{2000}-dH_{84}$
	Độ lệch giữa 2 hệ		
1	0,00000	0,00000	0,000
2	0,00001	0,00000	-0,001
3	0,00000	0,00000	0,001
4	0,00000	0,00000	0,000
5	0,00001	0,00000	0,001
6	0,00000	0,00001	0,001
7	0,00000	-0,00001	0,000
8	0,00000	0,00000	0,000
9	-0,00001	0,00000	0,000
10	0,00000	-0,00001	0,001
11	0,00000	0,00000	0,000

3.1.2. Phân tích biến dạng độ cao sử dụng độ cao trắc địa GNSS

Để tính toán, thử nghiệm nhóm tác giả sử dụng lưới đo GNSS tại khu vực miền Trung (lưới Thạch Bàn - Cát Tiên) gồm 8 điểm (Hình 1).



Hình 1. Sơ đồ lưới GNSS

Thực hiện đo lặp 4 chu kỳ 2015, 2016, 2017, 2018 với giãn cách thời gian giữa các chu kỳ là một năm. Kết quả bình sai cho giá trị độ cao trắc địa tại các điểm lưới và sai số trung phương tương ứng của các chu kỳ được trình bày ở các Bảng 1- 7.

Bảng 4. Giá trị độ cao trắc địa và sai số trung phương của các điểm trong chu kỳ 1 (2015)

	Điểm 1	Điểm 2	Điểm 3	Điểm 4	Điểm 5	Điểm 6	Điểm 7	Điểm 8
H (m)	233,660	422,073	46,164	462,312	478,985	433,701	132,895	1530,043
m _H (m)	0,0015	0,002	0,002	0,002	0,002	0,002	0,002	0,002

Bảng 5. Giá trị độ cao trắc địa và sai số trung phương của các điểm trong chu kỳ 2 (2016)

	Điểm 1	Điểm 2	Điểm 3	Điểm 4	Điểm 5	Điểm 6	Điểm 7	Điểm 8
H (m)	233,712	422,062	46,150	462,315	478,945	433,055	132,860	1530,039
m _H (m)	0,001	0,002	0,002	0,002	0,001	0,002	0,001	0,002

Bảng 6. Giá trị độ cao trắc địa và sai số trung phương của các điểm trong chu kỳ 3 (2017)

	Điểm 1	Điểm 2	Điểm 3	Điểm 4	Điểm 5	Điểm 6	Điểm 7	Điểm 8
H (m)	233,671	421,997	46,117	462,256	478,938	433,026	132,838	1530,029
m _H (m)	0,001	0,001	0,002	0,001	0,001	0,001	0,001	0,001

Bảng 7. Giá trị độ cao trắc địa và sai số trung phương của các điểm trong chu kỳ 4 (2018)

	Điểm 1	Điểm 2	Điểm 3	Điểm 4	Điểm 5	Điểm 6	Điểm 7	Điểm 8
H (m)	233,673	422,043	46,153	462,287	478,951	433,044	132,873	1530,031
m _H (m)	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002

Bảng 8. Các tham số chuyển dịch độ cao giữa năm 2015 - 2016 và các tham số tốc độ chuyển dịch

	Điểm 1	Điểm 2	Điểm 3	Điểm 4	Điểm 5	Điểm 6	Điểm 7	Điểm 8
dH (m)	0,052	-0,012	-0,015	0,003	-0,039	-0,009	-0,035	-2,1E-05
v _H (m/năm)	0,062	-0,014	-0,018	0,003	-0,047	-0,035	-0,042	-2,5E-05

Bảng 9. Các tham số chuyển dịch độ cao giữa năm 2015 - 2017 và các tham số tốc độ và gia tốc theo chu kỳ 1 năm

	Điểm 1	Điểm 2	Điểm 3	Điểm 4	Điểm 5	Điểm 6	Điểm 7	Điểm 8
dH (m)	0,052	-0,076	-0,047	-0,056	-0,049	-0,009	-0,057	-0,014
v_H (m/năm)	-0,086	-0,091	-0,040	-0,090	0,009	-0,036	-0,015	-0,011
a_H (m/năm r^2)	-0,092	-0,053	-0,017	-0,062	0,032	-0,022	0,0131	-0,003

Bảng 10. Các tham số chuyển dịch độ cao giữa năm 2015 - 2018 và các tham số tốc độ và gia tốc theo chu kỳ 1 năm

	Điểm 1	Điểm 2	Điểm 3	Điểm 4	Điểm 5	Điểm 6	Điểm 7	Điểm 8
dH (m)	0,012	-0,076	-0,011	-0,025	-0,034	-0,024	-0,008	-0,012
v_H (m/năm)	-0,092	0,032	-0,040	0,076	0,024	0,047	0,064	0,004
a_H (m/năm r^2)	-0,072	0,065	-0,01687	0,09	0,021	0,050	0,057	0,005

3.2. Thảo luận

Qua 2 thực nghiệm có thể thấy một số kết quả sau:

Đối với bài toán nghiên cứu biến dạng dịch chuyển, có thể sử dụng trực tiếp kết quả bình sai lưới không gian GNSS mà không cần chuyển đổi về hệ VN-2000 vì cho kết quả tương đương.

Giải pháp ứng dụng công nghệ GNSS trong vấn đề độ cao và chuyển dịch độ cao trong bài toán quan trắc biến dạng đã đạt mục tiêu đã đề ra. Giải pháp này giúp giải bài toán quan trắc biến dạng trong các phạm vi khác nhau mà không cần xây dựng mạng lưới thủy chuẩn. Các tham số chuyển dịch ngoài biển đổi độ cao như trong mô hình tĩnh, mà còn xác định các tham số vận tốc và gia tốc chuyển dịch.

Kết hợp công nghệ GNSS với các phương pháp xử lý số liệu trắc địa hiện đại như bình sai lưới tự do không gian GNSS và lọc Kalman cho phép hình thành công nghệ phân tích quan trắc chuyển dịch biến dạng đứng không nhất thiết phải xây dựng lưới thủy chuẩn truyền thống.

Để có kết luận đầy đủ cần mở rộng thử nghiệm ở nhiều khu vực và với số chu kỳ nhiều hơn để đánh giá về khả năng dự báo một cách tin cậy.

4. Kết luận

Trong bài báo đã xây dựng cơ sở khoa học và tính toán thực nghiệm chứng minh cho vấn đề chuyển đổi tọa độ giữa 2 hệ WGS-84 và VN-2000 trên cơ sở ứng dụng lý thuyết Molodenski đối với bài toán phân tích, đánh giá chuyển dịch biến dạng đứng sử dụng công nghệ GNSS. Tính ưu việt của giải pháp này được thể hiện ở sự phát triển của công nghệ GNSS với độ chính xác cao. Ngoài ra với số chu kỳ đo lớn có thể có khả năng dự báo về chuyển dịch biến dạng trôi, lún để có các phương án ứng phó phù hợp.

Hiện nay, với tình hình sụt, lún ở các tỉnh Đồng bằng sông Cửu Long, có thể xây dựng các trạm quan trắc, dựa trên ứng dụng GNSS và áp dụng giải pháp đã trình bày ở trên để đánh giá, phân tích và xây dựng cơ sở dữ liệu về chuyển dịch biến dạng lún của vùng.

Tài liệu tham khảo

Hoàng Ngọc Hà, Trương Quang Hiếu (2000). Cơ sở toán học xử lý số liệu trắc địa. Nhà xuất bản Giao thông vận tải.

Hoàng Ngọc Hà (2020). Bình sai tính toán lưới Trắc địa và GPS/GNSS. Nhà xuất bản Khoa học và Kỹ thuật, Hà Nội.

Hoàng Ngọc Hà, Phạm Thanh Thảo (2016). Ứng dụng lý thuyết bình sai tự do lưới không gian trong phân tích độ chuyển dịch đứt gãy sông Sài Gòn. Tài nguyên và môi trường, số 23(253), Kỳ 1,12/2016.

Hoang, N. H. (2020). Modernization of Height System in Vietnam Using GNSS and Geoid Model. In: Tien Bui D., Tran H.T., Bui XN. (eds). Proceedings of the International Conference on Innovations for Sustainable and Responsible Mining. Lecture Notes in Civil Engineering, Vol. 108. Springer, 10/2020.

Hoang, N. H. Monitoring the displacement of the Ba River fault zone of Vietnam using GNSS technology. FIG Conference, Orlando, Florida, USA 28 May - 1 June 2023.

Hofmann-Wellenhof, Herbert Lichtenegger, Elmar Wasle (2008). GNSS. Springer Wien New York.

Vy Quốc Hải, Trần Đình Tô và Dương Chí Công, 2005. Xác định chuyển dịch hiện đại đứt gãy Sông Đà và đứt gãy Sơn La - Bim Sơn bằng số liệu GPS. Địa chất và Khoáng sản, số 9, trang 257-265.

Hà Minh Hòa, Nguyễn Ngọc Lâu, Dương Chí Công. Nghiên cứu ứng dụng công nghệ GPS để xác định chuyển dịch vỏ Trái đất trên khu vực đứt gãy Lai châu - Điện Biên. Đề tài KHCN cấp Bộ Tài nguyên và Môi trường giai đoạn 2002 - 2004. Viện Khoa học Đo đạc và Bản đồ.

R.E. Kalman, A new approach to linear filtering and prediction problems, J Basic Eng 82D (1960), 82(1): 3545.

Markuze Y U. I., Hoàng Ngọc Hà (1991). Bình sai các mạng lưới không gian mặt đất và vệ tinh, Nhà xuất bản Nedra Matxcova. Sách chuyên khảo (tiếng Nga).

Markuze Y U. I, Gólubev V.V. (2010). Lý thuyết xử lý toán học lưới trắc địa. Alma Mater (Tiếng Nga).

Welsch, W. M. and Heunecke, O. (2001). Models and terminology for the analysis of geodetic monitoring observations, in: Proceedings of the 10th International Symposium on Deformation Measurements, Orange, California, USA, 390-412.

Yalcinkaya, M. and Bayrak, T. (2003). Dynamic model for monitoring landslides with emphasis on underground water in Trabzon province, Northeastern Turkey, J. Surv. Eng.-ASCE, 129(3), 115-124.

ABSTRACT

Research on some issues of converting coordinates from WGS-84 to the geodetic coordinate system used in Vietnam

Hoang Ngoc Ha^{1*}

¹Hanoi University of Mining and Geology

Currently, GNSS technology is widely used in Geodesy - Mapping - Remote sensing work. GNSS data processing results are determined in the WGS-84 coordinate system. However, since the announcement of the VN-2000 coordinate system in Vietnam, it uses the Ellipse but has been repositioned. In the article, the issue of studying coordinate conversion based on the application of Molodenski theory is mentioned. Consider some specific problems, including the problem of deformation monitoring. These results contribute to building the scientific basis for building spatial databases in our country.

Keywords: GNSS, coordinate systems, adjustment, molodenski theory, deformation monitoring.

NHIỆM VỤ CHÍNH

Xây dựng các văn bản quy phạm pháp luật, tiêu chuẩn, quy chuẩn kỹ thuật quốc gia, định mức kinh tế - kỹ thuật;

Thực hiện các đề án, dự án, nhiệm vụ chuyên môn về đo đạc cơ bản trên đất liền và trên biển;

Xây dựng cơ sở dữ liệu nền địa lý quốc gia, thành lập bản đồ địa hình quốc gia, bản đồ địa hình đáy biển;

Khảo sát, đo đạc, thành lập hệ thống bản đồ địa hình đáy biển, hải đảo và các loại bản đồ khác phục vụ công tác quản lý nhà nước;

Quản lý, vận hành và khai thác các trạm định vị vệ tinh (DGPS/DGNSS) cung cấp tín hiệu cải chính phân sai phục vụ công tác định vị dẫn đường trên biển thuộc phạm vi quản lý.



ĐỒNG HÀNH CÙNG CÁC BẠN TRÊN BIỂN

CÁC LOẠI HÌNH DỊCH VỤ

Chuyên cung cấp dịch vụ trắc địa, bản đồ phục vụ các công tác khảo sát khả thi, tiền lắp đặt, lắp đặt, hoàn công và sửa chữa, bảo dưỡng các công trình trên biển và đất liền, giao biển bằng nhiều công nghệ hiện đại cụ thể:



Đo sâu hồi âm bằng công nghệ đơn tia, đa tia, thành lập bản đồ địa hình đáy biển các tỷ lệ; tư vấn lập hồ sơ giao khu vực biển, hải đảo; lân biển, nhận chìm ở biển.



Quét ảnh phát hiện địa vật dưới nước, quét khảo sát hiện trạng công trình ngầm nước bằng công nghệ quét sườn, quét 360° 2D, 3D độ phân giải cao.



Định vị RTK, DGNSS bề mặt, định vị ngầm USBL, định hướng, xác định độ nghiêng độ chính xác cao phục vụ các công tác khảo sát địa chất, khoan thăm dò dầu khí, xây dựng công trình, bay chụp ảnh hàng không.



Đo đạc, quan trắc lún, biến dạng công trình; cung cấp các dịch vụ về khảo sát, thành lập bản đồ, địa hình, địa mạo, địa vật lý, khí tượng thủy văn và các loại bản đồ chuyên ngành khác; đánh giá địa hình, địa mạo các vùng cửa biển, cửa sông, bãi bồi và các dạng tai biến thiên nhiên, biến động địa hình đáy biển, xói lở, bồi tụ bờ biển.

GIỚI THIỆU

Trung tâm Trắc địa và Bản đồ biển (Seamap) là đơn vị Nhà nước, được thành lập vào năm 1998.

Hiện nay Trung tâm trực thuộc Cục Đo đạc, Bản đồ và Thông tin địa lý Việt Nam - Bộ Tài nguyên và Môi trường.

Với bề dày hơn 25 năm xây dựng và phát triển, Trung tâm có đội ngũ nhân lực gồm 160 người có trình độ chuyên môn cao, giàu kinh nghiệm và trình độ ngoại ngữ tốt. Trong số đó có 18 Thạc sĩ, 96 Đại học và Cao đẳng, hàng chục chuyên gia được đào tạo chuyên sâu về ngành đo đạc và bản đồ tại các trường của các nước có nền khoa học, kỹ thuật phát triển như Liên Xô cũ, Na Uy, Singapore.

Trung tâm có năng lực về kỹ thuật, thiết bị và công nghệ đo đạc, định vị và thành lập bản đồ biển hàng đầu tại Việt Nam.

LIÊN HỆ



(+84) +24 3 7547194



(+84) +24 3 7547365



contact@seamap.com.vn

CÔNG TY CỔ PHẦN LONG MÃ



Công ty Cổ phần Long Mã, được thành lập từ năm 2010, tự hào là một trong những đơn vị tiên phong trong lĩnh vực khảo sát và thiết kế công trình Nông nghiệp & phát triển nông thôn, công trình đê điều tại Việt Nam. Với đội ngũ chuyên gia giàu kinh nghiệm và trang thiết bị hiện đại, chúng tôi cam kết cung cấp các dịch vụ chất lượng cao, đáp ứng mọi yêu cầu khắt khe của khách hàng.

Khảo sát địa hình và địa chất công trình

Công ty Cổ phần Long Mã có chứng chỉ năng lực hoạt động khảo sát Hạng 1, chuyên thực hiện khảo sát địa hình và địa chất các công trình xây dựng, giao thông, thủy lợi. Chúng tôi cung cấp các dịch vụ khảo sát chính xác và chi tiết, từ việc xác định các yếu tố địa chất cơ bản đến phân tích sâu rộng về điều kiện địa chất công trình, đảm bảo sự an toàn và ổn định cho các công trình xây dựng.

Thiết kế và thẩm tra công trình nông nghiệp và phát triển nông thôn

Chúng tôi có năng lực thiết kế và thẩm tra các công trình nông nghiệp và phát triển nông thôn đến Cấp 1. Công ty Cổ phần Long Mã cam kết mang đến các giải pháp thiết kế tối ưu, phù hợp với các tiêu chuẩn kỹ thuật và yêu cầu thực tiễn của ngành nông nghiệp và phát triển nông thôn.

Thiết kế công trình đê điều

Công ty Cổ phần Long Mã chuyên thiết kế công trình đê điều đến Cấp 1. Với kiến thức chuyên sâu và kinh nghiệm dày dặn trong lĩnh vực này, chúng tôi đảm bảo thiết kế các công trình đê điều vững chắc và hiệu quả, góp phần bảo vệ an toàn cho các khu vực ven biển và sông ngòi.

Thiết kế công trình giao thông

Chúng tôi cũng có năng lực thiết kế công trình giao thông đến Cấp 3. Dịch vụ của chúng tôi bao gồm thiết kế các loại công trình giao thông như cầu, đường, và hệ thống giao thông khác, đảm bảo tính khả thi và hiệu quả cho các dự án hạ tầng giao thông.

Ứng dụng công nghệ BIM trong thiết kế

Các kỹ sư và chuyên gia của Công ty Cổ phần Long Mã có khả năng ứng dụng công nghệ BIM (Building Information Modeling) vào công việc tính toán và thiết kế công trình. Công nghệ BIM giúp cải thiện tính chính xác và hiệu quả của các dự án thông qua mô hình hóa thông tin 3D, cho phép dự đoán và giải quyết các vấn đề tiềm ẩn từ giai đoạn thiết kế đến thi công. Điều này đảm bảo sự đồng bộ và tối ưu hóa trong quản lý dự án và giảm thiểu rủi ro trong quá trình thực hiện.

Với cam kết chất lượng và sự tận tâm trong công việc, Công ty Cổ phần Long Mã là đối tác tin cậy cho mọi dự án xây dựng và phát triển. Chúng tôi sẵn sàng hợp tác và mang đến các giải pháp tối ưu nhất cho khách hàng.

Trụ sở chính: Số 9 khu Niềm, P. Kinh Bắc,

TP. Bắc Ninh;

VP đại diện: Biệt thự B01-L20 Khu đô thị

Dương Nội, Q. Hà Đông, Hà Nội;



0989050566; 0983083572; 0971444169



ctylongma@gmail.com



TRUNG TÂM NGHIÊN CỨU ỨNG DỤNG CÔNG NGHỆ MỚI TRẮC ĐỊA - BẢN ĐỒ



Giám đốc

PGS.TS Lê Đức Tình

Trung tâm Nghiên cứu ứng dụng công nghệ mới Trắc địa - Bản đồ được thành lập theo Quyết định số 206/QĐ.TCCB ngày 25 tháng 9 năm 1995 của Trường Đại học Mở - Địa chất.

- Giấy chứng nhận đăng ký hoạt động khoa học và công nghệ số: 63/ĐK-KHCNMT ngày 31/12/1995, cấp lại 11/7/2024 của Sở Khoa học và Công nghệ thành phố Hà Nội.
- Giấy phép hoạt động đo đạc và bản đồ số 00580 do Cục Đo đạc, Bản đồ và Thông tin địa lý Việt Nam - Bộ Tài Nguyên & Môi trường cấp ngày 8/10/2020.
- Chứng chỉ năng lực hoạt động xây dựng số: BXD-00057386 do Cục QLHĐXD (Bộ Xây dựng) cấp ngày 13/6/2024 chức năng khảo sát xây dựng: Địa hình Hạng I.



**Chức
năng
chính**

- Nghiên cứu thực nghiệm, triển khai áp dụng các công nghệ mới, các kỹ thuật mới, các phương pháp mới và biện pháp tổ chức mới về Trắc địa, Bản đồ, Địa kỹ thuật.
- Triển khai kết quả nghiên cứu ứng dụng công nghệ mới vào sản xuất, tạo nên các sản phẩm mới với trình độ công nghệ cao.
- Tham gia, phối hợp, thành lập các quy trình, quy phạm, các chỉ tiêu kỹ thuật Trắc địa, Bản đồ trên cơ sở ứng dụng công nghệ mới.
- Thực hiện các dịch vụ khoa học công nghệ, thông tin, tư vấn nghiệp vụ Trắc địa, Bản đồ, Địa kỹ thuật.
- Tham gia, phối hợp tổ chức đào tạo, bồi dưỡng cán bộ có trình độ Đại học và trên Đại học, nâng cao trình độ chuyên môn, nghiệp vụ.
- Hợp tác quốc tế về đào tạo khoa học công nghệ thuộc lĩnh vực Trắc địa, Bản đồ, Địa kỹ thuật.
- ...

Để thực hiện các công tác chuyên môn Trung tâm có các loại máy móc, thiết bị hiện đại; các thiết bị phần cứng và phần mềm xử lý nội nghiệp,...

Cán bộ trong biên chế kiêm nhiệm, hợp đồng dài hạn của Trung tâm là 35 người trong đó có: 5 PGS, 8 TS, 10 ThS, 12 KS. Ngoài ra, trung tâm còn hợp tác hoạt động khoa học công nghệ với hàng chục cộng tác viên là các nhà khoa học có uy tín ở trong và ngoài Trường Đại học Mở - Địa chất.

1 1 1N1 1/ 1 1 1 11 1 1 1N1 1 1 1 1 1 1 1 1 1 1 1 1

1 1LeNQebN1L1L eddeaebdN 1L QVTcaQTdN

CÔNG TY CỔ PHẦN THIẾT BỊ KHOA HỌC CÔNG NGHỆ TƯỜNG ANH



Địa chỉ ĐKKD: 234/4A Bùi Đình Túy, Phường 12, Quận Bình Thạnh, TP. HCM
VPGD HCM: 251/9 Nguyễn Văn Trỗi, Phường 10, Quận Phú Nhuận, TP. HCM
VPGD HN: 25 Ngõ 12 Đường Văn Phú, Phường Phú La, Quận Hà Đông, TP. Hà Nội
Tel : (+84)942646644; Website: www.tuonganh.vn
Email: sales.tuonganh@gmail.com

Công ty Cổ phần Thiết bị Khoa học Công nghệ Tường Anh thành lập năm 2017 tại TP. HCM, với thành viên sáng lập là các Kỹ sư Trắc địa dày dạn kinh nghiệm tốt nghiệp Khoa Trắc địa – Bản đồ và Quản lý đất đai thuộc Trường Đại học Mỏ - Địa chất. Với nền tảng kiến thức khoa học và công nghệ vững chắc, Công ty Tường Anh đã nhanh chóng vươn lên trở thành nhà cung cấp thiết bị và dịch vụ uy tín hàng đầu về giải pháp định vị vệ tinh GNSS. Chỉ sau một năm thành lập, Công ty Tường Anh đã chính thức trở thành Đại diện phân phối độc quyền tại Việt Nam cung cấp các sản phẩm của Hãng Stonex Srl đến từ Italia và tiếp tục đảm nhiệm thành công vai trò này cho đến nay, góp phần đưa thương hiệu Stonex trở thành một trong những thương hiệu mạnh tại Việt Nam. Công ty Tường Anh đang là đối tác cung cấp thiết bị và giải pháp công nghệ cho hàng chục văn phòng đăng ký đất đai các tỉnh, các công ty đo đạc bản đồ, khảo sát địa chính – địa hình trực thuộc Nhà nước cũng như tư nhân trên toàn quốc.

Hiện nay, chúng tôi đang tư vấn và cung cấp các giải pháp công nghệ trọn gói như sau:

- Giải pháp trạm định vị vệ tinh GNSS tham chiếu thường trực.
- Thiết bị định vị vệ tinh GNSS (Static, RTK, PPK, PPP).
- Thiết bị khảo sát truyền thống: Máy toàn đạc điện tử, Toàn đạc Robotic, thủy chuẩn.
- Giải pháp quan trắc tự động đa sensor (GNSS, Toàn đạc Robotic, sensor địa kỹ thuật,...).
- Giải pháp GIS (thiết bị và phần mềm thu thập dữ liệu, phần mềm CSDL GIS).
- Giải pháp 3D (Thiết bị quét laser 3D, SLAM, XVS vSLAM và phần mềm xử lý, mô hình hóa).
- Giải pháp thiết bị trắc địa điều khiển tự động trong xây dựng thông minh (Machine Control).



SURVEYING
GNSS | TS | GIS



SOFTWARE
FIELD | GIS | 3D



3D SCANNING
HANDHELD | TERRESTRIAL



MACHINE CONTROL
SOLAR | MINING
AGRICULTURE | MARINE



CÔNG TY CỔ PHẦN TƯ VẤN PHÚ THÁI



Các lĩnh vực hoạt động chính

- Khảo sát xây dựng công trình (địa hình, địa chất).

Thiết kế xây dựng công trình, giao thông, hạ tầng kỹ thuật, nông nghiệp và phát triển nông thôn,...

- Tư vấn thu hồi đất, đền bù GPMB.
- Tư vấn thẩm định công trình xây dựng.
- Tư vấn đấu thầu.

Thông tin liên hệ:

Địa chỉ: Số 40, Ngõ 2 phố Nguyễn Ngọc Nại, phường Khương Mai, quận Thanh Xuân, TP. Hà Nội.

Điện thoại: 024. 3565.7695

Website: tuvanphuthai.com

Email: tuvanphuthai@gmail.com

Giới thiệu

- Với hơn 20 năm hoạt động, chúng tôi đã hợp tác với nhiều đối tác lớn trong và ngoài nước, mang lại những giải pháp thiết kế tối ưu và bền vững.
- Đội ngũ chuyên gia hàng đầu: Đội ngũ kỹ sư và chuyên gia của chúng tôi có trình độ chuyên môn cao, luôn cập nhật các công nghệ và phương pháp mới nhất trong lĩnh vực giao thông.
- Cam kết chất lượng: Chúng tôi cam kết mang đến những sản phẩm và dịch vụ chất lượng, đáp ứng mọi yêu cầu kỹ thuật và thời gian của khách hàng.

TRUNG TÂM HỖ TRỢ PHÁT TRIỂN KHOA HỌC KỸ THUẬT

Trung tâm Hỗ trợ phát triển khoa học kỹ thuật được thành lập theo Quyết định số 190/TCCB ngày 21/01/1991 và chuyển đổi theo Quyết định số 1893/QĐ-BGDĐT ngày 30/5/2014 của Bộ Giáo dục và Đào tạo.

- Giấy chứng nhận đăng ký hoạt động khoa học và công nghệ số A-944 do Bộ khoa học và Công nghệ cấp ngày 04/4/2016.
- Giấy phép hoạt động đo đạc và bản đồ số 00248 do Cục Đo đạc, Bản đồ và Thông tin địa lý Việt Nam cấp ngày 08/12/2019, cấp bổ sung ngày 20/12/2022.
- Chứng chỉ năng lực hoạt động xây dựng số: HAN-00044087 do Sở Xây dựng thành phố Hà Nội cấp ngày 05/8/2020.

Lĩnh vực hoạt động khoa học và công nghệ

Nghiên cứu và sản xuất thử nghiệm các sản phẩm mới theo công nghệ mới trong lĩnh vực Mỏ, Địa chất, Dầu khí và Trắc địa; Nghiên cứu thiết kế, áp dụng các biện pháp thi công tiên tiến và thẩm định nước dân dụng và công nghiệp; Áp dụng công nghệ tiên tiến trong khai thác mỏ và dầu khí.

Lập dự án, thiết kế kỹ thuật - dự toán nhiệm vụ đo đạc và bản đồ cơ bản, bản đồ chuyên ngành; Kiểm tra chất lượng sản phẩm đo đạc và bản đồ; Xây dựng mạng lưới tọa độ, độ cao; Xây dựng cơ sở dữ liệu nền địa lý quốc gia tỷ lệ 1:2.000, 1:5.000; Đo đạc thành lập bản đồ địa hình tỷ lệ 1:200, 1:500, 1:1.000, 1:2.000, 1:5.000 và 1:10.000; Đo đạc thành lập bản đồ địa hình đáy biển khu vực cửa sông, cảng biển; Đo đạc thành lập bản đồ địa giới hành chính; Đo đạc thành lập bản đồ địa chính; Thành lập bản đồ hành chính cấp tỉnh, cấp huyện; Đo đạc thành lập bản công trình.

Tư vấn đào tạo trong lĩnh vực Mỏ, Địa chất, Dầu khí, Trắc địa, Bản đồ, Viễn thám, Hệ thông tin địa lý, Công nghệ thông tin và Môi trường; Tư vấn lập các dự án đầu tư trong lĩnh vực Mỏ, Địa chất, Dầu khí, Trắc địa, Bản đồ, Viễn thám, Hệ thông tin địa lý, Công nghệ thông tin và Môi trường, khảo sát đánh giá tác động môi trường và các công trình hạ tầng; Tư vấn phát triển hợp tác đầu tư trong lĩnh vực thăm dò, khai thác khoáng sản, đánh giá trữ lượng, chất lượng tài nguyên và đo đạc các loại bản đồ; Lập trình phần mềm ứng dụng trong các lĩnh vực liên quan; Chế tạo gia công cơ khí; Khoan thăm dò, khảo sát địa chất, địa chất công trình và địa kỹ thuật; Quy hoạch; Kiểm định và sửa chữa thiết bị đo đạc và thực hiện các dịch vụ khoa học và công nghệ khác trong lĩnh vực dầu khí và công nghệ tin.

Khảo sát địa hình trong lĩnh vực xây dựng.



Tầng 1, Nhà A, Đại học Mỏ - Địa chất, 18 Phố Viên, Phường Đức Thắng, Quận Bắc Từ Liêm, TP. Hà Nội



0354806361 / 0983355935



tthotro.mdc@gmail.com

CÔNG TY CP TƯ VẤN KHẢO SÁT VÀ DỊCH VỤ TRẮC ĐỊA BẢN ĐỒ



Địa chỉ : P308, Nhà C5, KĐT Mỹ Đình I, Tổ 4, P. Cầu Diễn, Q. Nam Từ Liêm, Hà Nội
Tel : (+84)966151618
Website: www.tomas-tracdiabando.com

Công ty chúng tôi chuyên cung cấp các dịch vụ:

1. Khảo sát xây dựng công trình:

Hạng I (Dân dụng cấp I, Công nghiệp cấp I, Giao thông cấp I, Hạ tầng kỹ thuật cấp I, Thủy lợi cấp I).

- Khảo sát, lập thiết kế kỹ thuật - dự toán công trình đo đạc và bản đồ.
- Kiểm tra chất lượng sản phẩm đo đạc và bản đồ địa chính, địa hình.
- Đo vẽ, thành lập bản đồ địa hình tỷ lệ 1/500, 1/1000, 1/2000, 1/5000.
- Đo đạc, thành lập bản đồ địa chính.
- Đo đạc, thành lập bản đồ công trình.
- Bay chụp - Xử lý ảnh bay chụp - Thành lập bản đồ bằng công nghệ bay chụp UAV.

2. Thiết kế, thẩm tra thiết kế xây dựng công trình:

Hạng III (Công nghiệp cấp III, Giao thông cấp II, Hạ tầng kỹ thuật cấp II, Thủy lợi cấp III).

3. Giám sát thi công xây dựng công trình: Hạng III.

TRẮC ĐỊA BẢN ĐỒ KHẢO SÁT ĐỊA HÌNH

- ✓ Đội ngũ cán bộ kỹ thuật giàu kinh nghiệm
- ✓ Hợp tác với nhiều đối tác lớn trong ngành
- ✓ Là doanh nghiệp khảo sát hạng I

Chứng chỉ BXD-00006974 - HAN-00006974

KHẢO SÁT ĐỊA HÌNH - XÂY DỰNG MỐC THỦY CÔNG THIẾT KẾ CẦU ĐƯỜNG - CẮM MỐC GPMB

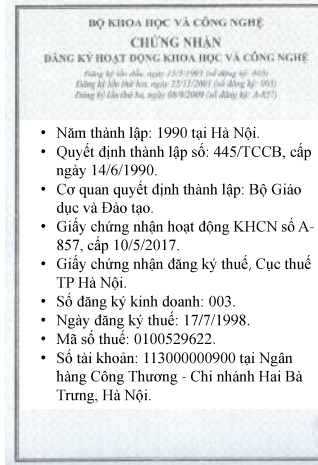
- ✓ Phục vụ mọi công trình dân dụng, giao thông, thủy điện...
- ✓ Ứng dụng các trang thiết bị máy móc hiện đại

TƯ VẤN TẬN TÌNH TIẾN ĐỘ NHANH CHÓNG CHI PHÍ PHẢI CHĂNG

MÁY TRẮC ĐỊA

Máy toàn đạc điện tử - Máy GPS - Máy bay không người lái

- ✓ Khảo sát, quan trắc địa hình công trình chính xác
- ✓ Nhập khẩu chính hãng từ nhiều thương hiệu lớn
- ✓ Thiết kế chắc chắn, hiện đại, tích hợp công nghệ cao



Tiếp nhận thiết bị phục vụ nghiên cứu công nghệ khai thác mỏ



Trao đổi nghiên cứu về công nghệ khai thác bauxit, Tây Nguyên



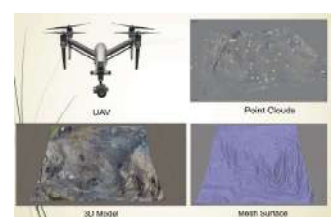
Tham gia các Hội nghị quốc tế Những tiến bộ trong Khai thác mỏ và Xây dựng ngầm - 2008



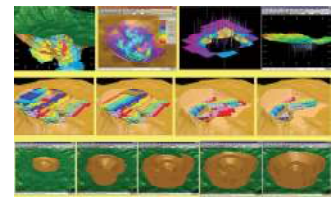
Khảo sát thực tế và hợp tác với chuyên gia nước ngoài mỏ đa kim Núi Pháo - Thái Nguyên và mỏ đồng Sin Quyền - Lào Cai, 2019



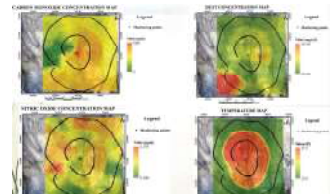
Tham gia hội thảo và hợp tác với chuyên gia tại Ba Lan 2019, 2020 và 2023



Xây dựng bản đồ 3D mô phỏng phục vụ giám sát an toàn và môi trường mỏ



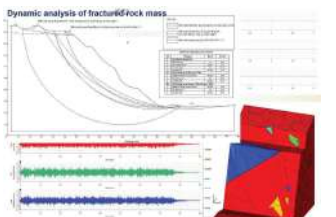
Nghiên cứu tối ưu hóa biên giới mỏ lộ thiên bằng phần mềm chuyên dụng



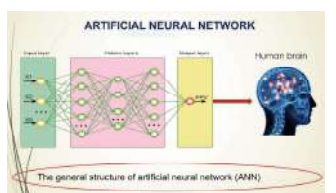
Xây dựng các mô hình dự báo chất lượng không khí trên mỏ lộ thiên



Xây dựng các mô hình cải tạo, phục hồi môi trường sau khai thác mỏ

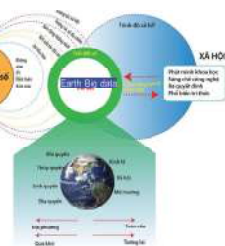


Xây dựng các mô hình ổn định bờ mỏ lộ thiên



Ứng dụng AI trong việc đánh giá, mô phỏng các hoạt động khai thác mỏ

Hướng phát triển trong thời gian tới



Bảng doanh thu 3 năm gần nhất

TT	Năm	Doanh thu, VND
1	2021	11.771.666.098
2	2022	15.805.155.490
3	2023	15.271.077.604

Số đề tài, dự án 3 năm gần đây

TT	Năm	Số lượng
1	2021	26
2	2022	54
3	2023	49

Phụ trách Trung tâm: **PGS.TS. Nguyễn Anh Tuấn**
Điện thoại: 0912507096

Email: nguyenanhtuan@humg.edu.vn

Tên đơn vị: **TRUNG TÂM NGHIÊN CỨU CƠ ĐIỆN MỎ**

Địa chỉ trụ sở chính: Phòng N.01 Nhà N, Khu B, Trường Đại học Mỏ - Địa chất, Tổ dân phố số 4, Phường Cổ Nhuế 2, Quận Bắc Từ Liêm, TP. Hà Nội

Điện thoại: 0243.8681888

Email: codienmo@humg.edu.vn

NHÀ XUẤT BẢN KHOA HỌC TỰ NHIÊN VÀ CÔNG NGHỆ

Nhà A16 - Số 18 Hoàng Quốc Việt, Cầu Giấy, Hà Nội

Điện thoại: Phòng Phát hành: **024.22149040**;

Phòng Biên tập: **024.37917148**;

Phòng Quản lý Tổng hợp: **024.22149041**;

Fax: **024.37910147**; Email: **nxb@vap.ac.vn**; Website: **www.vap.ac.vn**

**HỘI NGHỊ KHOA HỌC QUỐC GIA
CÔNG NGHỆ ĐỊA KHÔNG GIAN VÌ SỰ PHÁT TRIỂN BỀN VỮNG
NATIONAL CONFERENCE
ON GEOSPATIAL TECHNOLOGY FOR SUSTAINABLE DEVELOPMENT
(NCGSD - 2024)**

Chịu trách nhiệm xuất bản

Giám đốc, Tổng biên tập

PHẠM THỊ HIẾU

Biên tập:

Nguyễn Thị Chiên

Trình bày kỹ thuật:

PGS.TS Nguyễn Văn Trung

Trình bày bìa:

Nguyễn Minh Hải

Liên kết xuất bản:

GS.TS Trần Thanh Hải - Trường Đại học Mỏ - Địa chất

Địa chỉ: 18 Phố Viên, Đông Ngạc, Bắc Từ Liêm, Hà Nội

ISBN: 978-604-357-305-3

In 100 cuốn, khổ 21×29,7 cm, tại Công ty Cổ phần Khoa học và Công nghệ Hoàng Quốc Việt. Địa chỉ: Số 11 ngách 1, ngõ 1 Võ Chí Công, P. Nghĩa Đô, Q. Cầu Giấy, TP. Hà Nội. Số xác nhận đăng ký xuất bản: 3231-2024/CXBIPH/02-36/KHTNVCN. Số quyết định xuất bản: 50/QĐ-HKTNCN, tháng 10 năm 2024. In xong và nộp lưu chiểu quý IV năm 2024.

ISBN: 978-604-357-305-3



SÁCH KHÔNG BÁN