

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA MÔI TRƯỜNG

ĐỒ ÁN TỐT NGHIỆP – CAPSTONE PROJECT

NGÀNH: MÔI TRƯỜNG
CHUYÊN NGÀNH: QUẢN LÝ TÀI NGUYÊN MÔI TRƯỜNG

ĐỀ TÀI:

**XÁC ĐỊNH RỦI RO SINH THÁI CỦA MỘT SỐ
KIM LOẠI NẶNG TRONG TRẦM TÍCH MẶT TÀI
SÔNG BÔNG MIÊU, TỈNH QUẢNG NAM VÀ DỰ
BÁO HIỆN TRẠNG SỨC KHỎE MÔI TRƯỜNG
BẰNG CÁC MÔ HÌNH HỌC MÁY**

Người hướng dẫn: PGS.TS LÊ PHƯỚC CƯỜNG

ThS. NGUYỄN THÀNH TRUNG

Sinh viên thực hiện: TRẦN MINH HÙNG

Số thẻ sinh viên: 117200039

Lớp: 20QLMT

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA MÔI TRƯỜNG

CAPSTONE PROJECT

NGÀNH: MÔI TRƯỜNG
CHUYÊN NGÀNH: QUẢN LÝ TÀI NGUYÊN MÔI TRƯỜNG

ĐỀ TÀI:

**XÁC ĐỊNH RỦI RO SINH THÁI CỦA MỘT SỐ KIM
LOẠI NẶNG TRONG TRẦM TÍCH MẶT TẠI SÔNG
BÔNG MIÊU, TỈNH QUẢNG NAM VÀ DỰ BÁO
HIỆN TRẠNG SỨC KHỎE MÔI TRƯỜNG BẰNG
CÁC MÔ HÌNH HỌC MÁY**

Người hướng dẫn: PGS.TS LÊ PHƯỚC CƯỜNG

ThS. NGUYỄN THÀNH TRUNG

Sinh viên thực hiện: TRẦN MINH HÙNG

Số thẻ sinh viên: 117200039

Lớp: 20QLMT

Đà Nẵng – 5/2025

NHẬN XÉT CỦA NGƯỜI HƯỚNG DẪN

TÓM TẮT

Tên đề tài: Xác định rủi ro sinh thái của một số kim loại nặng trong trầm tích mặt tại sông Bông Miêu, tỉnh Quảng Nam và dự báo hiện trạng sức khỏe môi trường bằng các mô hình học máy.

Sinh viên thực hiện: Trần Minh Hùng

Số thẻ SV: 117200039

Lớp: 20QLMT

GV hướng dẫn:

PGS.TS Lê Phước Cường

ThS. Nguyễn Thành Trung

Tóm tắt đề tài: Đề tài tiến hành xác định rủi ro sinh thái của một số kim loại nặng trong trầm tích mặt tại sông Bông Miêu, tỉnh Quảng Nam, khu vực đang chịu ảnh hưởng trực tiếp từ hoạt động khai thác vàng, bao gồm cả khai thác trái phép. Nghiên cứu sử dụng chỉ số rủi ro sinh thái tiềm năng (RI) do Lars Hakanson đề xuất để đánh giá mức độ nguy cơ tác động của các kim loại nặng đến hệ sinh thái thủy sinh trong khu vực. Song song đó, đề tài tiến hành khảo sát hiện trạng sức khỏe cộng đồng dân cư sống xung quanh sông thông qua phỏng vấn, thu thập số liệu từ cơ sở y tế địa phương, kết hợp phân tích mẫu trầm tích, nước mặt để đánh giá toàn diện các thành phần môi trường. Ngoài các phương pháp đánh giá truyền thống, đề tài ứng dụng các mô hình học máy (Machine Learning) nhằm dự báo xu hướng biến động chất lượng môi trường và sức khỏe cộng đồng trong tương lai. Kết quả nghiên cứu là cơ sở khoa học quan trọng cho việc đề xuất các biện pháp kiểm soát, giảm thiểu tác động tiêu cực của ô nhiễm kim loại nặng, góp phần bảo vệ sức khỏe cộng đồng và phục hồi môi trường sông Bông Miêu.

LỜI NÓI ĐẦU

Trong thời gian thực hiện đồ án tốt nghiệp, em đã nhận được nhiều sự giúp đỡ, đóng góp ý kiến của thầy cô, gia đình và bạn bè.

Em xin gửi lời cảm ơn chân thành đến PGS.TS. Lê Phước Cường - giảng viên Bộ môn Quản lý Môi Trường, trường Đại học Bách khoa, Đại học Đà Nẵng và ThS. Nguyễn Thành Trung – Phân viện phó, Phân viện Khoa học An toàn vệ sinh Lao động và Bảo vệ môi trường Miền trung, người đã tận tình hướng dẫn, chỉ bảo em trong suốt quá trình thực hiện đề tài thực tế: ***“Xác định rủi ro sinh thái của một số kim loại nặng trong trầm tích mặt tại sông Bông Miêu, tỉnh Quảng Nam và dự báo hiện trạng sức khỏe môi trường bằng các mô hình học máy”***.

Đề tài này là bước đầu đi vào thực tế, kiến thức của em còn hạn chế và còn nhiều bỡ ngỡ. Do vậy, không tránh khỏi những thiếu sót nên em rất mong nhận được những ý kiến đóng góp quý báu từ quý Thầy, Cô và các bạn học cùng lớp để kiến thức của em trong lĩnh vực này được hoàn thiện hơn.

Sau cùng, em xin kính chúc quý Thầy, Cô trong Khoa Môi Trường thật dồi dào sức khỏe để tiếp tục thực hiện sứ mệnh cao đẹp của mình là truyền đạt kiến thức cho thế hệ mai sau.

Trân trọng.

Đà Nẵng, ngày 20 tháng 05 năm 2025

Sinh viên thực hiện

Trần Minh Hùng

CAM ĐOAN

Em xin cam đoan đây là đề tài đồ án tốt nghiệp độc lập của riêng em. Các số liệu sử dụng phân tích trong đồ án tốt nghiệp có nguồn gốc rõ ràng, đã công bố theo đúng quy định. Các kết quả nghiên cứu trong đồ án tốt nghiệp phân tích một cách trung thực, khách quan và phù hợp với thực tiễn. Các kết quả chưa từng công bố trong bất kỳ đồ án tốt nghiệp nào khác.

Sinh viên thực hiện

Trần Minh Hùng

MỤC LỤC

| | |
|--|-------------|
| TÓM TẮT | i |
| LỜI NÓI ĐẦU | ii |
| CAM ĐOAN | iii |
| MỤC LỤC | iv |
| DANH MỤC BẢNG | vi |
| DANH MỤC HÌNH | vii |
| DANH MỤC CÁC TỪ VIẾT TẮT | viii |
| MỞ ĐẦU | 1 |
| CHƯƠNG 1. TỔNG QUAN | 3 |
| 1.1. Đặt vấn đề | 3 |
| 1.2. Tổng quan về khu vực nghiên cứu | 4 |
| 1.2.1. Điều kiện tự nhiên | 4 |
| 1.2.2. Đặc điểm kinh tế - xã hội | 6 |
| 1.3. Hiện trạng chất lượng môi trường trầm tích, nước mặt và sức khỏe người dân tại Sông Bồng Miêu | 7 |
| 1.3.1. Hiện trạng môi trường trầm tích, nước mặt tại KVNC | 7 |
| 1.3.2. Các vấn đề nổi cộm đối với môi trường nước mặt | 8 |
| 1.3.3. Hiện trạng chung về sức khỏe người dân tại tỉnh Quảng Nam và khu vực Sông Bồng Miêu | 9 |
| 1.4. Tổng quan về học máy và ứng dụng các mô hình học máy trong quản lý môi trường | 10 |
| 1.4.1. Tổng quan về học máy | 10 |
| 1.4.2. Ứng dụng các mô hình học máy trong quản lý môi trường | 12 |
| 1.5. Tổng quan về kim loại nặng | 17 |
| 1.5.1. Khái niệm về kim loại nặng | 17 |
| 1.5.2. Nguồn gốc ô nhiễm và độc tính của một số kim loại nặng | 17 |
| CHƯƠNG 2: ĐÁNH GIÁ HIỆN TRẠNG CHẤT LƯỢNG MÔI TRƯỜNG NƯỚC VÀ TRẦM TÍCH MẶT, XÁC ĐỊNH RỦI RO SINH THÁI CỦA MỘT SỐ KIM | |

| | |
|--|-----------|
| LOẠI NẶNG VÀ DỰ BÁO SỨC KHỎE MÔI TRƯỜNG BẰNG CÁC MÔ HÌNH HỌC MÁY..... | 24 |
| 2.1. Khảo sát chất lượng môi trường, hiện trạng sức khỏe môi trường và người dân tại khu vực sông Bồng Miêu, xã Tam Lãnh, huyện Phú Ninh, tỉnh Quảng Nam | 24 |
| 2.1.2. Thời gian lấy phiếu khảo sát | 25 |
| 2.1.3. Phân tích kết quả khảo sát..... | 25 |
| 2.2. Khảo sát hiện trạng chất lượng môi trường nước và trầm tích mặt tại Sông Bồng Miêu | 29 |
| 2.3. Xác định chỉ số rủi ro sinh thái RI theo Lars Hakanson (1980). | 40 |
| 2.3.1. Tổng quan về chỉ số rủi ro sinh thái RI (Lars Hakanson - 1980)..... | 40 |
| 2.3.2. Xác định chỉ số rủi ro sinh thái (RI) tại khu vực nghiên cứu (sông Bồng Miêu, tỉnh Quảng Nam) | 41 |
| 2.4. Xác định mô hình học máy phù hợp để dự báo chất lượng môi trường trầm tích mặt tại khu vực Sông Bồng Miêu, Tỉnh Quảng Nam | 44 |
| 2.4.1. Sơ đồ thực hiện mô hình học máy..... | 44 |
| 2.2.2. Thu thập số liệu..... | 46 |
| 2.2.3. Giới thiệu về ngôn ngữ R..... | 46 |
| 2.2.4. Tải R – Studio và cài đặt..... | 47 |
| 2.2.5. Các bước chạy mô hình học máy | 48 |
| CHƯƠNG 3: DỰ BÁO DIỄN BIẾN CHẤT LƯỢNG MÔI TRƯỜNG THEO THỜI GIAN VÀ CÁC GIẢI PHÁP QUẢN LÝ - KỸ THUẬT | 62 |
| 3.1. Dự báo hiện trạng môi trường diễn biến theo thời gian dựa trên mô hình tối ưu đã lựa chọn..... | 62 |
| 3.1.1. Dự báo chỉ số rủi ro sinh thái tại 10 vị trí nghiên cứu đến tháng 1/2026..... | 62 |
| 3.1.2. Dự báo nồng độ của từng kim loại nặng tại 10 vị trí nghiên cứu đến tháng.. | 63 |
| 3.2. Các biện pháp quản lý và giải pháp kỹ thuật | 67 |
| 3.2.1 Biện pháp quản lý..... | 68 |
| 3.2.2 Biện pháp kỹ thuật | 69 |
| KẾT LUẬN | 71 |
| KIẾN NGHỊ | 72 |
| TÀI LIỆU THAM KHẢO..... | 74 |

DANH MỤC BẢNG

| | |
|--|----|
| Bảng 1.1. Đặc điểm kinh tế - xã hội khu vực nghiên cứu | 6 |
| Bảng 2.1. Thông tin về tình trạng sức khỏe của người dân tại khu vực nghiên cứu do Bệnh viện Đa Khoa tỉnh Quảng Nam cung cấp | 28 |
| Bảng 2.2. Vị trí lấy mẫu nước sông Bông Miêu..... | 29 |
| Bảng 2.3. Phương pháp và thiết bị phân tích..... | 32 |
| Bảng 2.4. Kết quả phân tích mẫu nước sông Bông Miêu (22/4/2025- 29/4/2025) | 36 |
| Bảng 2.5. Giá trị nền (C_{n i}) và hệ số độc tính (Tr_i) của 7 kim loại nặng [15] | 41 |
| Bảng 2.6. Nồng độ kim loại nặng trong trầm tích và chỉ số rủi ro sinh thái RI (10 vị trí được thực hiện lấy mẫu và phân tích vào 5/2025) | 41 |
| Bảng 2.7. Kết quả chạy mô hình từ dữ liệu Training | 58 |
| Bảng 2.8. Kết quả chạy mô hình từ dữ liệu Validation | 59 |
| Bảng 2.9. So sánh hiệu suất của từng mô hình giữa tập Training và Validation | 60 |
| Bảng 2.10. Xếp hạng các mô hình tối ưu hóa dữ liệu nhất | 61 |
| Bảng 3.1. Tổng hợp dữ liệu sau dự báo (1/2026)..... | 67 |

DANH MỤC HÌNH

| | |
|---|----|
| Hình 1.1. Bản đồ vị trí khu vực nghiên cứu – Sông Bồng Miêu, Quảng Nam | 10 |
| Hình 2.1. Vị trí lấy phiếu khảo sát hộ dân xung quanh sông Bồng Miêu | 25 |
| Hình 2.2. Khảo sát, điều tra thực tế tại Bệnh viện Đa Khoa tỉnh Quảng Nam | 28 |
| Hình 2.3. Vị trí lấy mẫu nước và trầm tích mặt tại sông Bồng Miêu..... | 31 |
| Hình 2.4. Hình ảnh thực tế lấy mẫu nước và trầm tích mặt | 34 |
| Hình 2.5. Bảo quản mẫu tại Phòng thí nghiệm | 35 |
| Hình 2.6. Phân tích các chỉ tiêu nước sông Bồng Miêu tại PTN | 35 |
| Hình 2.7. Biểu tượng của phần mềm R trên desktop | 47 |
| Hình 2.8. Màn hình cửa sổ của R- Studio | 47 |
| Hình 2.9. Tải và nạp thành công các gói dữ liệu (Packages) | 50 |
| Hình 2.10. Thể hiện đường dẫn và đọc dữ liệu kim loại nặng vào học máy..... | 50 |
| Hình 2.11. Kết quả sau khi kiểm tra bộ dữ liệu..... | 51 |
| Hình 2.12. Kết quả sau khi tóm tắt dữ liệu đầu vào | 51 |
| Hình 2.13. Biểu đồ tương quan giữa các biến trong bộ dữ liệu | 52 |
| Hình 2.14. Môi tương quan ma trận giữa các biến..... | 53 |
| Hình 2.15. Các dòng dữ liệu sau khi chia ngẫu nhiên bộ dữ liệu..... | 53 |
| Hình 2.16. Kết quả sau khi thực hiện 5 mô hình học máy từ tập dữ liệu Training..... | 55 |
| Hình 2.17. Kết quả chạy mô hình từ tập dữ liệu Validation | 57 |
| Hình 3.1. Kết quả dự báo chỉ số rủi ro sinh thái (RI) cho tháng 1/2026 | 63 |
| Hình 3.2. Kết quả dự báo nồng độ các kim loại nặng tháng 1/2026 | 66 |
| Hình 3.3. Hiện trạng khai thác vàng và môi trường tại sông Bồng Miêu | 68 |

DANH MỤC CÁC TỪ VIẾT TẮT

| Từ viết tắt | Giải thích |
|--------------------|---|
| BOD | : Biochemical Oxygen Demand (Nhu cầu oxy sinh hóa) |
| BTNMT | : Bộ tài nguyên môi trường |
| COD | : Chemical Oxygen Demand (Nhu cầu oxy hóa học) |
| CRAN | : Comprehensive R Archive Network |
| GWQI | : Groundwater Quality Index (Chỉ số chất lượng nước dưới đất) |
| KNN | : K-Nearest Neighbors (K- Điểm dữ liệu gần nhất) |
| MAE | : Mean Squared Error (Sai số tuyệt đối trung bình) |
| ML | : Machine Learning (Học máy) |
| QCVN | : Quy chuẩn Việt Nam |
| R ² | : R- Square (Độ tin cậy) |
| RF | : Random Forest (Rừng ngẫu nhiên) |
| RMSE | : Root Mean Squared Error (Sai số toàn phương trung bình) |
| SVM | : Support vector machine (Máy hỗ trợ vector) |
| KVNC | : Khu vực nghiên cứu |
| RI | : Risk Index (Chỉ số rủi ro sinh thái) |
| BVMT | : Bảo vệ môi trường |
| TNHH | : Trách nhiệm hữu hạn |

MỞ ĐẦU

1. Mục tiêu đề tài

❖ Mục tiêu tổng quát:

- Xác định mức độ rủi ro sinh thái của một số kim loại nặng trong trầm tích mặt tại sông Bồng Miêu, tỉnh Quảng Nam.
- Ứng dụng mô hình học máy để dự báo hiện trạng sức khỏe môi trường tại khu vực chịu ảnh hưởng của hoạt động khai thác vàng.

❖ Mục tiêu cụ thể

- Xác định hàm lượng một số kim loại nặng (As, Pb, Cd, Hg, Cu, Cr, Zn) trong trầm tích mặt sông Bồng Miêu.
 - Tính toán chỉ số rủi ro sinh thái tiềm năng (RI) cho từng kim loại và tổng thể hệ sinh thái theo phương pháp của Lars Hakanson (1980).
 - Ứng dụng các mô hình học máy (Machine Learning) như Random Forest, SVM... để dự báo hiện trạng sức khỏe môi trường khu vực dựa trên dữ liệu quan trắc kim loại nặng và các thông số môi trường liên quan.
 - Đề xuất các giải pháp quản lý – kỹ thuật để kiểm soát và giảm thiểu ô nhiễm kim loại nặng tại khu vực nghiên cứu.

2. Đối tượng, phạm vi nghiên cứu

Đối tượng:

- Trầm tích mặt và nước mặt khu vực sông Bồng Miêu, tỉnh Quảng Nam;
- Sức khỏe môi trường xung quanh khu vực sông Bồng Miêu, tỉnh Quảng Nam (thông qua phiếu khảo sát người dân và cơ quan quản lý).

Phạm vi:

- Khu vực xung quanh sông Bồng Miêu, tỉnh Quảng Nam
- Đề xuất mô hình dự báo chất lượng trầm tích mặt bằng mô hình học máy (Machine learning) tại khu vực đánh giá, khảo sát, từ đó dự báo hiện trạng môi trường diễn biến theo thời gian tại khu vực.
- Đề xuất biện pháp quản lý và giải pháp kỹ thuật bảo vệ môi trường.

3. Phương pháp

3.1. Khảo cứu tài liệu

- Thu thập các dữ liệu về điều kiện tự nhiên: địa chất, địa hình, khí tượng thủy văn khu vực xã Tam Lãnh, huyện Phú Ninh, tỉnh Quảng Nam
- Tài liệu về mô hình học máy (Machine learning)
- Thu thập dữ liệu về kim loại nặng tại khu vực nghiên cứu của các năm trước đây

3.2. Phương pháp

- Phương pháp lập phiếu khảo sát điều tra: khảo sát ý kiến người dân về sự thay đổi chất lượng môi trường sống và sức khỏe của người dân sống xung quanh mỏ vàng Bông Miêu
- Phương pháp lấy mẫu, phân tích: sử dụng các thiết bị đo nhanh để lấy mẫu và đo đạc các thông số liên quan, lấy mẫu phân tích tại phòng thí nghiệm để xác định nồng độ các chất ô nhiễm trong môi trường nước mặt. Trên cơ sở đó đánh giá được hiện trạng môi trường khu vực;
- Phương pháp dự báo chất lượng môi trường trầm tích mặt: ứng dụng mô hình học máy (Machine learning) trong dự báo chất lượng trầm tích mặt thông qua ứng dụng R/R-Studio

4. Cấu trúc đề án tốt nghiệp

4.1. Thuyết minh

- Chương 1: Tổng quan.
- Chương 2: Đánh giá hiện trạng chất lượng môi trường nước và trầm tích, xác định rủi ro sinh thái của một số kim loại nặng và dự báo sức khỏe môi trường bằng các mô hình học máy tại khu vực nghiên cứu
- Chương 3: Dự báo diễn biến chất lượng môi trường theo thời gian và đề xuất các giải pháp quản lý – kỹ thuật

4.2. Bản vẽ

BV1: Kết quả khảo sát ý kiến người dân về hiện trạng môi trường khu vực sông Bông Miêu, tỉnh Quảng Nam

BV2: Bản đồ vị trí lấy mẫu quan trắc khu vực sông Bông Miêu, tỉnh Quảng Nam

BV3: Đánh giá hiện trạng ô nhiễm môi trường nước mặt và ô nhiễm kim loại nặng trong trầm tích mặt tại sông Bông Miêu, tỉnh Quảng Nam

BV4: Sơ đồ thực hiện mô hình học máy

BV5: Kết quả mô hình học máy

CHƯƠNG 1. TỔNG QUAN

1.1. Đặt vấn đề

Kể từ sau công cuộc đổi mới năm 1986 cho đến nay, kinh tế Việt Nam có xu hướng phát triển và ngày càng có nhiều thành tựu. Trong quá trình tăng trưởng này, quy mô sản xuất trong các lĩnh vực công nghiệp và nông nghiệp không ngừng được mở rộng. Song song với sự phát triển đó, vấn đề ô nhiễm môi trường – đặc biệt là ô nhiễm kim loại nặng trong trầm tích – ngày càng trở nên đáng lo ngại. Hiện tượng này là hệ quả của quá trình lắng đọng và tích tụ lâu dài các kim loại nặng từ nguồn nước vào nền trầm tích.

Nước đóng vai trò quan trọng trong hệ sinh thái, là nhân tố then chốt quyết định chất lượng môi trường sống và là tài nguyên thiết yếu, tái tạo được, phục vụ mọi nhu cầu của sự sống và hoạt động kinh tế xã hội của con người.

Kim loại nặng là thuật ngữ chỉ những kim loại có tỷ trọng lớn hơn 5mg/cm^3 (Crom, kẽm, đồng, thủy ngân, asen, chì, cadimi,...).[1] Các kim loại nặng này có tính tồn lưu, tích lũy sinh học, độc tính cao đối với môi trường thủy sinh[2]. Kim loại nặng phát sinh từ nước thải công nghiệp, hoạt động nông nghiệp, đặc biệt từ các hoạt động khai thác khoáng sản,... Khi các chất thải chứa kim loại bị thải ra môi trường, một phần nhỏ sẽ hòa tan trong nước, còn phần lớn sẽ kết hợp với các hạt lơ lửng và lắng đọng trong trầm tích. Do đó, việc xác định mức độ tích lũy kim loại nặng trong trầm tích là cơ sở để đánh giá sức khỏe môi trường nước.

Xác định rủi ro sinh thái là công cụ giúp nhận diện, đánh giá và dự đoán các tác động tiềm ẩn từ các chất ô nhiễm đối với hệ sinh thái, thông qua chỉ số rủi ro sinh thái tiềm năng (RI) do Lars Hakanson đề xuất năm 1980. Chỉ số này cho phép đánh giá tác động tổng thể đến hệ sinh thái và đã được ứng dụng rộng rãi trong nhiều nghiên cứu trên thế giới.

Tại tỉnh Quảng Nam, với mạng lưới sông ngòi phong phú, sông Bồng Miêu là một nhánh thuộc hệ thống Vu Gia - Thu Bồn, có vai trò quan trọng về môi trường và kinh tế. Tuy nhiên, chất lượng nước và trầm tích khu vực này đang chịu ảnh hưởng nghiêm trọng từ các hoạt động khai thác vàng, đặc biệt là khai thác trái phép đang dần trở lại mạnh mẽ trong thời gian gần đây.

Xuất phát từ những vấn đề trên, việc xác định rủi ro sinh thái của kim loại nặng trong trầm tích mặt và dự báo hiện trạng sức khỏe môi trường khu vực sông Bồng Miêu là hết sức cần thiết. Đây là cơ sở quan trọng để đề xuất các giải pháp kiểm soát và giảm

thiếu ô nhiễm hiệu quả. Vì vậy, đề tài: “**Xác định rủi ro sinh thái của một số kim loại nặng trong trầm tích mặt tại sông Bồng Miêu, tỉnh Quảng Nam và dự báo hiện trạng sức khỏe môi trường bằng các mô hình học máy**” được thực hiện với mục tiêu đánh giá mức độ ô nhiễm môi trường trầm tích mặt, nước mặt, từ đó dự báo và cảnh báo sức khỏe môi trường thông qua ứng dụng các mô hình học máy hiện đại.

1.2. Tổng quan về khu vực nghiên cứu

1.2.1. Điều kiện tự nhiên

a. Vị trí địa lý [3]

Khu vực nghiên cứu sông Bồng Miêu thuộc ba huyện miền núi, huyện Phú Ninh, huyện Tiên Phước, huyện Bắc Trà My của tỉnh Quảng Nam. Bắt nguồn từ xã Trà Kót huyện Bắc Trà My sau đó chảy qua địa phận xã Tam Lãnh, huyện Phú Ninh và xã Tiên Lập, huyện Tiên Phước, sau đó đổ vào sông Tiên của huyện Tiên Phước.

b. Địa hình [3]

Địa hình khá phức tạp, địa hình có xu hướng thấp dần từ Tây Nam sang Đông Bắc, vùng núi cao chiếm trên 56%, vùng gò đồi chiếm 30%. Độ dốc lớn và chia cắt bởi sông suối. Độ cao trung bình 30 - 300m. Phía Tây có nhiều đồi núi cao như: núi Đá Đen (392,6m), núi Long Cám (306,6m), núi Đá Ngựa (283m).

Điều kiện địa chất: Có nhiều loại đất tại khu vực, điển hình là lớp đất loại sét pha lẫn sỏi sạn. Lớp đất hữu cơ lẫn rễ cây cỏ. Lớp đất cát pha màu xám vàng, trạng thái dẻo. Lớp đất sét màu xám xanh, trạng thái dẻo mềm. Lớp cát hạt thô lẫn sỏi sạn, màu xám vàng, kết cấu chặt vừa. Lớp sét pha lẫn sỏi sạn, màu nâu vàng, trạng thái dẻo cứng. Lớp sét pha lẫn nhiều sỏi sạn, màu nâu vàng, trạng thái nửa cứng.

c. Chế độ thủy văn [3]

Sông Bồng Miêu có hướng dòng chảy chủ yếu về phía Tây Bắc. Sông Bồng Miêu dài hơn chục kilomet (khoảng 11,7km), có đồng bằng ven sông rộng trung bình 1.500m (khu vực hạ lưu) và hẹp dần về phía thượng lưu.

Sông Bồng Miêu chảy theo hướng Tây rồi đổ về sông Tiên ở xã Tiên Lập. Sông Tiên nhập vào sông Thu Bồn ở Tân An, tiếp theo đó sông Thu Bồn chảy theo hướng Bắc và hướng Đông sau đó đổ vào cửa biển Hội An.

Lưu vực sông Bồng Miêu và các sông kế tiếp ở hạ lưu sông Tiên có tổng diện tích là 101km², trong đó lưu vực thượng nguồn ở Bồng Miêu là 35km². Các suối đầu nguồn sông Bồng Miêu ở phía Đông Nam, cao so với mực nước biển là 590m đến 658m.

d. Khí hậu [3]

Khí hậu tại khu vực nghiên cứu mang những tính chất và đặc điểm của kiểu khí hậu nhiệt đới gió mùa, hằng năm có 2 mùa rõ rệt:

- Mùa mưa bắt đầu từ tháng 9 đến hết tháng 12 hằng năm. Lượng mưa lớn tập trung ở các tháng 9, 10 và 11 thường có mưa lớn kèm gió bão và lũ lụt.

- Mùa khô bắt đầu từ tháng 2 đến hết tháng 8, vào mùa này khí hậu thường nóng bức, khô nắng.

** Chế độ nắng*

Khu vực và vùng phụ cận có số giờ nắng hằng năm khoảng 1.736 đến 2.259 giờ.

- Tháng 5 đến tháng 8 là các tháng có số giờ nắng cao nhất trong năm, dao động từ 137 đến 264 giờ.

- Tháng 9 đến tháng 1 là các tháng có số giờ nắng thấp nhất trong năm, dao động từ 19 đến 219 giờ.

** Chế độ mưa*

Lượng mưa hằng năm trên địa bàn tỉnh Quảng Nam biến động từ 2.000 mm đến 4.000 mm. Chế độ mưa thay đổi theo mùa, chia thành hai mùa rõ rệt trong năm: Mùa mưa và mùa khô.

- Mùa mưa bắt đầu vào tháng 9 và kết thúc vào tháng 1 năm sau: Lượng mưa trung bình tháng dao động từ 38 đến 238,5mm, chiếm khoảng 70 đến 80% tổng lượng mưa cả năm, tập trung chủ yếu từ tháng 9 đến tháng 11.

- Mùa khô bắt đầu từ tháng 2 đến tháng 8: Lượng mưa trung bình tháng dao động từ (10 – 160) mm lượng mưa thấp nhất từ tháng 2 đến tháng 5.

** Chế độ gió*

Do cơ chế hoàn lưu và địa hình nên chế độ gió tại khu vực nghiên cứu và vùng lân cận phân thành 2 mùa gió chính, đó là gió mùa đông và gió mùa hè.

- Từ tháng 10 đến tháng 3 năm sau, gió chủ yếu là Bắc, Đông Bắc và Đông, đặc biệt tháng 11 và 12 gió Đông Bắc.

- Tháng 4 gió Đông hoạt động nhiều nhất, tháng 6 và tháng 7 có gió mùa tây nam gây khô và nóng bức.

** Các hiện tượng thời tiết đặc biệt [4]*

Theo số liệu thống kê nhiều năm, trung bình nước ta chịu ảnh hưởng trực tiếp từ 5 - 6 cơn bão và áp thấp nhiệt đới (ATNĐ), tập trung từ tháng 7 đến tháng 11, nhiều nhất là tháng 9 và tháng 10.

Tỉnh Quảng Nam trung bình khoảng 1 năm có 1 đến 2 cơn bão hoặc ATNĐ ảnh

hường trực tiếp. Mùa bão tại Quảng Nam thường bắt đầu từ tháng 9 đến tháng 12 hàng năm, tập trung nhiều nhất vào tháng 10 và 11, cá biệt có những năm bão và ATNĐ xuất hiện sớm hơn (tháng 5, tháng 6 đã có bão và ATNĐ).

Khi bão hay ATNĐ hình thành ngoài biển đông và đổ bộ vào đất liền Quảng Nam thường kèm theo gió giật mạnh có tính chất xoáy, phá hỏng các công trình xây dựng trên đường của bão đi qua. Thống kê trong 30 năm trở lại đây, vận tốc gió trong bão trung bình đạt 17m/s (cấp 7), mạnh nhất đạt đến 33m/s (cấp 11). Ngoài ra, bão và ATNĐ còn kèm theo mưa lớn làm mực nước dâng cao gây ra lũ lụt và sạt lở đất.

1.2.2. Đặc điểm kinh tế - xã hội

Đặc điểm kinh tế - xã hội của khu vực nghiên cứu được thể hiện ở Bảng 1.1

Bảng 1.1. Đặc điểm kinh tế - xã hội khu vực nghiên cứu

| TT | Huyện | Tình hình phát triển kinh tế | Đặc điểm văn hóa - xã hội |
|----|----------------|--|--|
| 1 | Phú Ninh [1] | <ul style="list-style-type: none"> - Kinh tế duy trì tốc độ tăng trưởng khá, cơ cấu kinh tế chuyển dịch đúng định hướng: Nông nghiệp giảm còn 19%, Công nghiệp - Xây dựng tăng lên 45%, Thương mại - Dịch vụ chiếm 36%. - Thu nhập thực tế bình quân đầu người: trên 32,5 triệu/người/năm. - Tổng giá trị sản xuất đạt gần 5.253 tỷ đồng (giá hiện hành). | <ul style="list-style-type: none"> - Diện tích tự nhiên là 25.565ha, - Dân số trung bình năm 2021: 78.303 người. Mật độ dân số: 306,29 người/km² [4]. - Văn hóa - thông tin: Trên địa bàn huyện có 02 công trình văn hóa - thể dục thể thao cấp huyện, 103 công trình bưu điện, nhà văn hóa cấp thôn - xã. - Giáo dục: Toàn huyện có 36 trường học gồm: 11 trường mầm non; 13 trường tiểu học; 10 trường THCS và 2 trường THPT. Trong đó đã có 35 trường đạt chuẩn quốc gia.. |
| 2 | Bắc Trà My [2] | <ul style="list-style-type: none"> - Cơ cấu kinh tế đến năm 2020: Nông, lâm nghiệp 37,26%; Công nghiệp, tiểu thủ công nghiệp 23,97%; Thương mại, dịch vụ 38,77%. - Tổng diện tích gieo trồng cây hằng năm đạt khoảng | <ul style="list-style-type: none"> - Diện tích tự nhiên là 84.699ha, - Dân số trung bình năm 2021: 41.944 người, 10.396 hộ. Mật độ dân số: 49,52 người/km² [4]. - Văn hóa: là một huyện miền núi cao, với 20 thành phần dân tộc anh em cùng sinh sống. Trong đó, dân |

| | | | |
|---|----------------|--|---|
| | | <p>3.884ha; hình thức chăn nuôi nhỏ lẻ tại các hộ gia đình, chưa có các trang trại chăn nuôi tập trung quy mô lớn</p> <p>- Định hướng cơ cấu kinh tế giai đoạn 2025 và 2030: Tăng dần tỷ trọng công nghiệp, tiểu thủ công nghiệp và thương mại, dịch vụ, giảm dần tỷ trọng của nông lâm nghiệp</p> | <p>tộc Cor, Xê đăng, Mơ nông và nhóm người Cadong là người bản địa. Với đa thành phần dân tộc như trên đã tạo nên bức tranh văn hóa sống động, đa sắc màu góp phần làm giàu bản sắc văn hóa trên địa bàn huyện</p> <p>- Giáo dục: Toàn huyện có 44 trường học gồm: 15 trường mầm non; 14 trường tiểu học; 13 trường THCS và 2 trường THPT. Trong đó đã có 2 trường đạt chuẩn quốc gia.</p> |
| 3 | Tiên Phước [3] | <p>- Tổng diện tích gieo trồng cây hằng năm 3.137 ha</p> <p>- Tổng đàn gia súc 26.929; tổng đàn gia cầm 371.447</p> <p>- Giá trị sản xuất Công nghiệp - Tiểu thủ công nghiệp ước đạt 950 tỷ đồng (theo giá hiện hành), tăng 18,58% so với cùng kỳ năm 2021; giá trị Thương mại - dịch vụ ước đạt 2.074 tỷ đồng (theo giá hiện hành), tăng 15,67% so với cùng kỳ năm 2021</p> | <p>- Diện tích tự nhiên là 45.455ha,</p> <p>- Dân số trung bình năm 2021: 67.188 người. Mật độ dân số: 147,81 người/km² [4].</p> <p>- Văn hóa – thông tin: thực hiện tuyên truyền và phát thanh theo kế hoạch vào các ngày lễ lớn và sự kiện tại địa phương.</p> <p>- Giáo dục: Toàn huyện có 48 trường học gồm: 15 trường mầm non; 16 trường tiểu học; 15 trường THCS và 2 trường THPT và 1 trung tâm dạy nghề.</p> |

Nguồn: [1] *Thuyết minh quy hoạch vùng huyện Phú Ninh giai đoạn đến năm 2030*

[2] *Thuyết minh quy hoạch vùng huyện Bắc Trà My giai đoạn đến năm 2030*

[3] *Báo cáo tình hình Kinh tế - Xã hội, Quốc Phòng - An ninh huyện Tiên Phước*

năm 2022

1.3. Hiện trạng chất lượng môi trường trầm tích, nước mặt và sức khỏe người dân tại Sông Bồng Miêu

1.3.1. Hiện trạng môi trường trầm tích, nước mặt tại KVNC [5]

Trong các sông thuộc khu vực thượng nguồn hệ thống sông Thu Bồn, sông Bồng Miêu có chất lượng nước thấp nhất với nhiều thông số vượt cột A2 của QCVN 08-

MT:2015/BTNMT, thậm chí vượt cột B1 như: TSS, Fe, Pb, Coliform. Hàm lượng TSS và Fe có xu hướng tăng từ năm 2016-2018 với biên độ lớn, vượt giới hạn cho phép trong nhiều đợt quan trắc; đột biến vào tháng 3/2018 hàm lượng Fe vượt cột B1 đến 59,6 lần, TSS vượt 5,76 lần, nhưng đến năm 2019 lại giảm trở lại. Hàm lượng Pb quan trắc được trên sông Bồng Miêu cũng thường xuyên ở mức cao, vượt cột A2 với tần suất 100% từ năm 2015-2017, đột biến vào tháng 3/2018 vượt cột B1 đến 49,8 lần dẫn đến hiện tượng cá chết nhiều trên sông, đến năm 2019 hàm lượng Pb giảm dần còn vượt cột A2 vào một số thời điểm. Nguyên nhân của sự gia tăng hàm lượng Fe, Pb trên sông Bồng Miêu là do từ khi Công ty vàng Bồng Miêu ngừng hoạt động (năm 2016), lợi dụng địa bàn sẵn có nên tình trạng khai thác trái phép diễn ra phức tạp làm cho lượng đất đá và quặng thải chứa Fe, Pb cuốn trôi theo nước mưa đổ ra sông Bồng Miêu gây ô nhiễm; vào năm 2018 đã xảy ra sự cố vỡ đập hồ chứa quặng sau tuyển tại mỏ vàng Bồng Miêu (do Công ty CP Tập đoàn Khoáng sản công nghiệp 6666 làm chủ đầu tư) khiến 10 m³ bùn đất và nước thải chảy tràn ra suối Trang và sông Bồng Miêu làm gia tăng ô nhiễm đột biến vào thời điểm này.

1.3.2. Các vấn đề nổi cộm đối với môi trường nước mặt [5]

Hoạt động khai thác khoáng sản, nhất là khai thác trái phép, tiếp tục gây ra những ảnh hưởng tiêu cực đến chất lượng môi trường nước mặt phía thượng lưu hệ thống sông Vu Gia - Thu Bồn. Với đặc điểm phân bố phân tán tại các khu vực có địa hình hiểm trở nên công tác kiểm soát hoạt động khai thác khoáng sản gặp rất nhiều khó khăn, nhất là các hoạt động khai thác trái phép mang tính chất thủ công, tận thu không chú ý đến vấn đề BVMT, gây tiềm ẩn nguy cơ cao xảy ra sự cố môi trường và ô nhiễm môi trường nghiêm trọng.

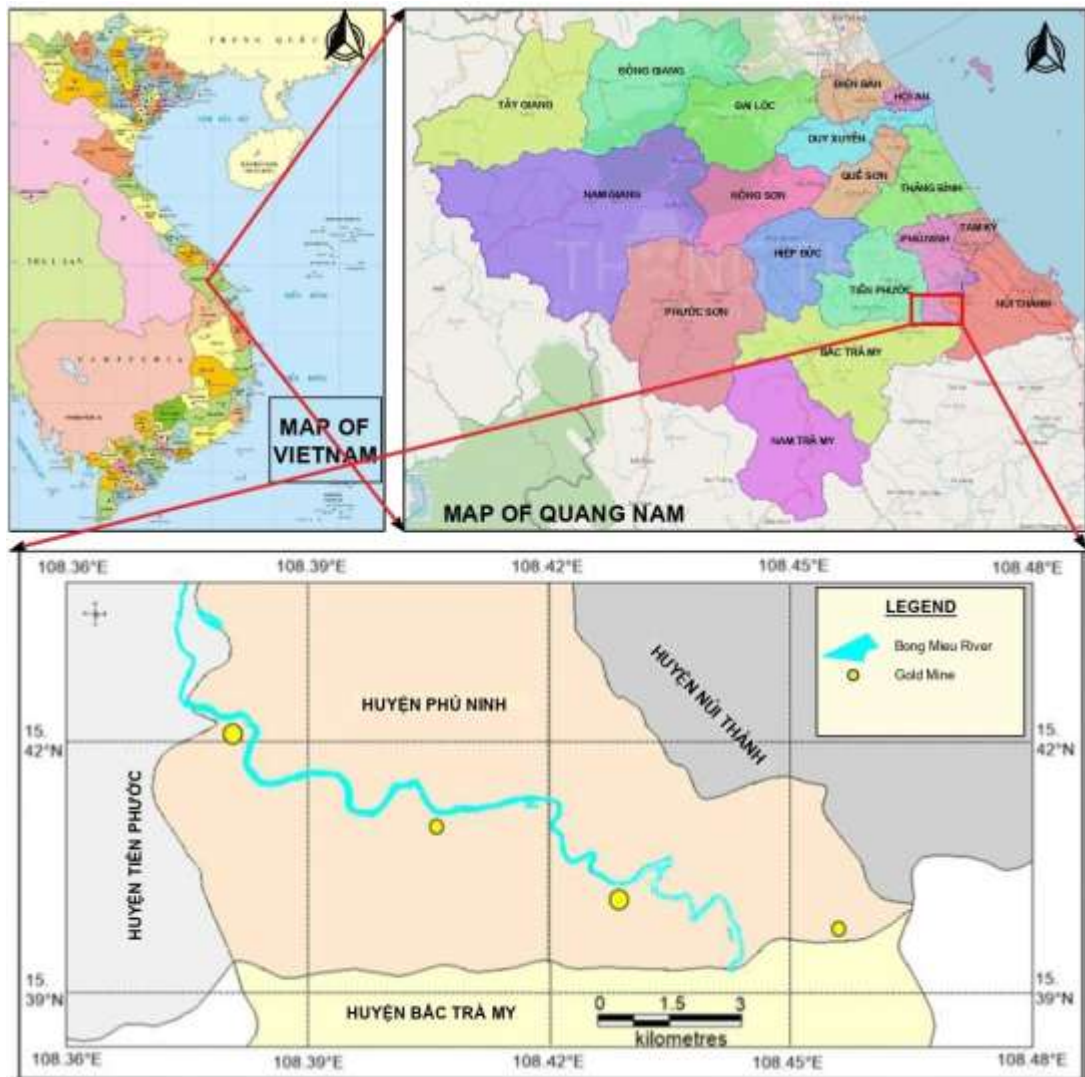
Nhiều trường hợp khai thác vàng trái phép sau khi Công ty TNHH vàng Bồng Miêu thuộc Tập đoàn Besra hết thời hạn cấp phép khai thác đã gây mất an ninh trật tự, an toàn xã hội trên địa bàn và ô nhiễm nghiêm trọng nước mặt sông Bồng Miêu. Vào năm 2018 sự cố vỡ đập hồ chứa quặng sau tuyển đã xảy ra tại mỏ vàng Bồng Miêu của Công ty Cổ phần Tập đoàn Khoáng sản công nghiệp 6666 làm cho 10 m³ bùn đất và nước thải chảy tràn ra suối Trang và sông Bồng Miêu dẫn đến nước đục, cá chết hàng loạt, quan trắc chất lượng nước sông Bồng Miêu tại thời điểm đó cho thấy hàm lượng TSS, Fe và Pb vượt nhiều lần giới hạn cho phép...

1.3.3. Hiện trạng chung về sức khỏe người dân tại tỉnh Quảng Nam và khu vực Sông Bồng Miêu [5]

Trong giai đoạn từ đầu năm 2016 -2019, cả tỉnh có hơn 38,3 nghìn ca mắc các bệnh liên quan đến ô nhiễm không khí như viêm phổi, viêm phế quản; hơn 45 nghìn ca mắc liên quan thần kinh và giác quan (tai, mắt, mũi, da...), hơn 90 nghìn ca mắc các bệnh liên quan đến ô nhiễm nguồn nước như sốt xuất huyết, tả, thương hàn, lỵ, tiêu chảy,... Các thiệt hại kinh tế do đầu tư khám chữa bệnh mắc phải, nghỉ việc vì ốm đau (đối với người lớn), chăm sóc trẻ em cũng như người mắc bệnh tạo nên gánh nặng cho toàn cộng đồng.

Nguồn nước sông ở một số địa phương hiện cung cấp nước sinh hoạt, sản xuất chính cho dân cư đô thị và vùng nông thôn. Ở một số lưu vực sông có những dấu hiệu ô nhiễm từ các hoạt động khai khoáng, nuôi trồng thủy sản, nguồn thải từ hoạt động các vùng dân cư, khu cụm công nghiệp. Các tác nhân chính trong nguồn nước bị ô nhiễm như chuỗi vi sinh, hóa chất, dầu mỡ, kim loại nặng... trong nguồn nước mặt sẽ tác động trực tiếp sẽ gây kích ứng da, ngứa rát, dị ứng mẩn đỏ, đau mắt... Các bệnh về da thường mắc phải như ghẻ lở, nấm da, viêm da dị ứng, mề đay... và đặc biệt phụ nữ ở nông thôn dễ mắc các bệnh viêm nhiễm phụ khoa. Dân cư nông thôn có nguy cơ mắc bệnh cao hơn do tiếp xúc thường xuyên nguồn nước sinh hoạt từ sông, ao hồ không đảm bảo tiêu chuẩn vệ sinh. Ngoài ra tác nhân do động vật và vật nuôi chết, các nguồn thải hữu cơ góp phần tạo mầm bệnh nguy hiểm gây nên dịch bệnh ở người và trên súc vật. Trong giai đoạn từ năm 2015 đến nay, Quảng Nam có phần lớn diện tích nuôi trồng thủy sản có nguồn nước từ ao nuôi bị ô nhiễm. Nguồn nước tù đọng cũng là môi trường sống chủ đạo của các loài côn trùng đốt làm truyền nhiễm các bệnh sốt xuất huyết, sốt rét, viêm não...

Bệnh lý liên quan đến ô nhiễm nguồn nước xảy ra hằng năm với tổng số trường hợp mắc bệnh lên đến hàng chục nghìn ca. Riêng về lĩnh vực an toàn vệ sinh thực phẩm trong 5 năm qua (từ năm 2016 đến tháng 6/2020) chỉ tiêu số ca mắc/100.000 người dân là <7, trung bình trong cả giai đoạn số ca mắc/100.000 người dân không chế ở mức 1,4 (đạt chỉ tiêu theo yêu cầu chung). Nguyên nhân chủ yếu do tính ngộ độc thực phẩm không phải do yếu tố vi sinh, hóa chất... mà chủ yếu do ngộ độc tự nhiên là chính, một số trường hợp không xác định được nguyên do. Một số trường hợp ngộ độc thực phẩm xảy ra trong thời gian qua cũng có liên quan đến việc sử dụng nguồn nước không hợp vệ sinh (do không xử lý trước khi đưa vào sử dụng).



Hình 1.1. Bản đồ vị trí khu vực nghiên cứu – Sông Bồng Miêu, Quảng Nam

1.4. Tổng quan về học máy và ứng dụng các mô hình học máy trong quản lý môi trường

1.4.1. Tổng quan về học máy

1.4.1.1. Lịch sử ra đời

Học máy (Machine learning) ra đời vào cuối những năm 80 và đầu những năm 90. Là tập hợp con của trí tuệ nhân tạo (AI - *Artificial Intelligence*) nó cho phép máy tính hành động và đưa ra quyết định dựa trên dữ liệu để thực hiện một nhiệm vụ nhất định. Các chương trình này là các thuật toán được thiết kế theo cách mà chúng có thể học hỏi và cải thiện theo thời gian khi tiếp xúc với dữ liệu mới.

1.4.1.2. Khái niệm

- Học máy (Machine learning) là một thuật ngữ để chỉ hành động máy tính tự học để cải thiện một nhiệm vụ mà nó đang thực hiện;

- Học máy là một nhánh của trí thông minh nhân tạo, cho phép các hệ thống máy tính để học trực tiếp từ các ví dụ, dữ liệu và kinh nghiệm;

- Hiệu suất của máy tính khi thực hiện một nhiệm vụ sẽ trở nên tốt hơn khi thực hiện nhiệm vụ đó nhiều lần. Hay nói cách khác khả năng cơ bản nhất của Machine learning là sử dụng các thuật toán để phân tích các thông tin có sẵn học hỏi từ nó rồi đưa ra quyết định hoặc dự đoán về một thứ gì đó có liên quan thay vì tạo ra một phần mềm mới với những hành động hướng dẫn chi tiết để thực hiện một nhiệm vụ cụ thể;

- Máy tính được huấn luyện bằng cách sử dụng một lượng lớn dữ liệu và các thuật toán, nhờ đó học cách thực hiện nhiệm vụ một cách tự động. Nếu không có học máy, trí tuệ nhân tạo sẽ bị giới hạn vì học máy cung cấp cho máy tính khả năng tự khám phá mà không cần lập trình cụ thể cho từng tác vụ.

1.4.1.3. Phân loại học máy

Học máy được phân làm 4 nhóm [6]:

- ❖ **Học có giám sát (Supervised learning):** là thuật toán dự đoán đầu ra của một dữ liệu mới dựa trên các cặp *input, outcome* đã biết từ trước. Cặp dữ liệu này còn được gọi là dữ liệu, nhãn. Supervised learning là nhóm phổ biến nhất trong các thuật toán Machine Learning.
- ❖ **Học không giám sát (Unsupervised learning):** không biết được kết quả hay nhãn mà chỉ có dữ liệu đầu vào. Thuật toán Unsupervised learning sẽ dựa vào cấu trúc của dữ liệu để thực hiện một công việc nào đó, ví dụ như phân nhóm (clustering) hoặc giảm số chiều của dữ liệu (dimension reduction) để thuận tiện trong việc lưu trữ và tính toán.
- ❖ **Học bán giám sát (Semi-supervised learning):** là các bài toán khi chúng ta có một lượng lớn dữ liệu nhưng chỉ một phần trong chúng được gán nhãn. Những bài toán thuộc nhóm này nằm giữa hai nhóm được nêu bên trên. Một ví dụ điển hình của nhóm này là chỉ có một phần ảnh hoặc văn bản được gán nhãn (ví dụ bức ảnh về người, động vật hoặc các văn bản khoa học, chính trị) và phần lớn các bức ảnh/văn bản khác chưa được gán nhãn được thu thập từ internet. Thực tế cho thấy rất nhiều các bài toán Machine Learning thuộc vào nhóm này vì việc thu thập dữ liệu có nhãn tốn rất nhiều thời gian và có chi phí cao. Rất nhiều loại dữ

liệu thậm chí cần phải có chuyên gia mới gán nhãn được (ảnh y học chẳng hạn). Ngược lại, dữ liệu chưa có nhãn có thể được thu thập với chi phí thấp từ internet.

❖ **Học tăng cường (Reinforcement learning):** là các bài toán giúp cho một hệ thống tự động xác định hành vi dựa trên hoàn cảnh để đạt được lợi ích cao nhất (maximizing the performance). Hiện tại, Reinforcement learning chủ yếu được áp dụng vào Lý Thuyết Trò Chơi (Game Theory), các thuật toán cần xác định nước đi tiếp theo để đạt được điểm số cao nhất.

1.4.2. Ứng dụng các mô hình học máy trong quản lý môi trường

Trong những năm gần đây, các mô hình học máy (machine learning) đã được ứng dụng ngày càng rộng rãi trong lĩnh vực môi trường trên toàn thế giới nhằm phục vụ công tác giám sát, phân tích, dự báo và ra quyết định quản lý. Trên thế giới, các mô hình như Random Forest, Support Vector Machine (SVM), Artificial Neural Network (ANN), và Gradient Boosting đã được sử dụng thành công trong các bài toán dự báo chất lượng không khí, mô phỏng ô nhiễm nước, xác định vùng đất bị ô nhiễm kim loại nặng, và giám sát suy giảm rừng thông qua dữ liệu ảnh vệ tinh. Chẳng hạn, mô hình LSTM đã được ứng dụng để dự báo chỉ số chất lượng không khí (AQI) tại các đô thị lớn như Bắc Kinh, New Delhi với độ chính xác cao, góp phần hỗ trợ chính sách kiểm soát ô nhiễm.

Tại Việt Nam, các nghiên cứu ứng dụng học máy trong môi trường tuy còn mới nhưng đang có xu hướng phát triển mạnh. Một số đề tài đã áp dụng các mô hình như Random Forest và SVM để phân tích chất lượng nước mặt, dự báo nồng độ kim loại nặng trong nước ngầm, hoặc đánh giá nguy cơ sức khỏe môi trường tại các khu vực khai thác khoáng sản như Thái Nguyên, Quảng Nam. Những mô hình này đã cho thấy tiềm năng lớn trong việc hỗ trợ cơ quan chức năng và nhà khoa học đưa ra các quyết định nhanh chóng và chính xác hơn trong quản lý và bảo vệ môi trường.

Nhìn chung, việc ứng dụng học máy trong lĩnh vực môi trường đang trở thành xu thế tất yếu, đặc biệt trong bối cảnh ô nhiễm ngày càng gia tăng và nhu cầu xử lý dữ liệu lớn (big data) từ các thiết bị cảm biến, vệ tinh, và hệ thống quan trắc tự động ngày càng cao.

1.4.2.1. RF – Random Forest (Rừng ngẫu nhiên) [7]

RF là 1 sự kết hợp của cây dự đoán, như vậy là mỗi cây phụ thuộc vào giá trị của vector ngẫu nhiên được lấy mẫu độc lập (independently) và với sự phân bố như nhau cho tất cả các cây có trong rừng. RF là một loại của thuật toán tổng hợp (ensemble) được gọi là tổng hợp (aggregation) bootstrap và là một trong những phương pháp ML phổ

biến nhất. RF thực hiện (perform) một dự đoán dựa vào biến đầu vào (x) xảy ra ở bước tiếp theo:

(1) Lấy mẫu bootstrap và chọn ngẫu nhiên một tập hợp con (subset) của tập dữ liệu học có giám sát;

(2) Điều chỉnh một cây bằng cách sử dụng phân vùng đệ quy nhị phân (binary recursive partitioning);

(3) Bắt đầu với tất cả quan sát trong một nút duy nhất và lặp lại đệ quy các bước cho mỗi nút ngắt (unsplit) cho đến khi tiêu chí dừng được áp dụng. Trong bước này, 2 tham số turning, mtry (số lượng các mẫu biến ngẫu nhiên như các ứng cử viên ở mỗi phần) và ntree (số lượng cây) được sử dụng. Cuối cùng, đưa ra một dự đoán của một điểm mới x. Đối với hồi quy, giá trị trung bình của một số cây hồi quy là kết quả cuối cùng.

1.4.2.2. SVM – Support vector machines (Máy hỗ trợ vector) [7]

SVM là một trong những thuật toán phân lớp phổ biến và hiệu quả. SVM là một khái niệm trong thống kê và khoa học máy tính cho một tập hợp các phương pháp học có giám sát liên quan đến nhau để phân loại và phân tích hồi quy. SVM dạng chuẩn nhận dữ liệu vào và phân loại chúng vào hai lớp khác nhau. Do đó, SVM là một thuật toán phân loại nhị phân. Với một bộ các ví dụ luyện tập thuộc hai thể loại cho trước, thuật toán luyện tập SVM xây dựng một mô hình SVM để phân loại các ví dụ khác vào hai thể loại đó. Một mô hình SVM là một cách biểu diễn các điểm trong không gian và lựa chọn ranh giới giữa hai thể loại sao cho khoảng cách từ các ví dụ luyện tập tới ranh giới là xa nhất có thể. Các ví dụ mới cũng được biểu diễn trong cùng một không gian và được thuật toán dự đoán thuộc một trong hai thể loại tùy vào ví dụ đó nằm ở phía nào của ranh giới.

1.4.2.3. KNN – K nearest neighbors [K – dữ liệu gần nhất] [7]

KNN (K-Nearest Neighbors) là một thuật toán học máy thuộc nhóm học có giám sát (supervised learning), được sử dụng phổ biến cho các bài toán phân loại (classification) và hồi quy (regression). Về bản chất, KNN là một thuật toán dựa trên trường hợp (instance-based) hoặc dựa trên bộ nhớ (memory-based learning), do không xây dựng mô hình huấn luyện cụ thể mà chỉ lưu trữ toàn bộ tập dữ liệu huấn luyện. Do đó, KNN được xem là một thuật toán học lười (lazy learning), bởi tất cả các tính toán chỉ được thực hiện trong giai đoạn dự đoán khi có dữ liệu mới cần phân loại hoặc dự đoán giá trị.

Nguyên lý hoạt động của KNN dựa trên việc tìm kiếm K điểm dữ liệu gần nhất trong tập huấn luyện đối với một mẫu mới cần phân loại, sau đó dựa vào các đặc điểm của các điểm lân cận này để quyết định đầu ra cho mẫu mới. Việc đo lường khoảng cách giữa các điểm dữ liệu thường sử dụng khoảng cách Euclidean, tức là khoảng cách giữa hai điểm trong không gian đặc trưng. Số lượng K là một tham số do người dùng xác định trước, và giá trị này ảnh hưởng trực tiếp đến hiệu quả của thuật toán.

Do KNN dựa hoàn toàn vào việc so sánh khoảng cách với các điểm lân cận gần nhất, nên có hai yếu tố quan trọng cần được chú ý: (1) cách thức xác định và lựa chọn hàm đo khoảng cách phù hợp, và (2) hiệu quả của việc tính toán khoảng cách trong các tập dữ liệu lớn hoặc có số chiều cao.

KNN có ưu điểm nổi bật là không yêu cầu giai đoạn huấn luyện nên độ phức tạp tính toán gần như bằng 0 trong giai đoạn này, đồng thời không đòi hỏi giả định về phân phối của các lớp dữ liệu. Tuy nhiên, nhược điểm lớn của KNN là rất nhạy cảm với nhiễu khi giá trị K nhỏ, đồng thời đòi hỏi tính toán khoảng cách với toàn bộ tập huấn luyện trong mỗi lần dự đoán, gây tiêu tốn nhiều thời gian và tài nguyên, đặc biệt trong các bài toán có tập dữ liệu lớn hoặc nhiều chiều. Việc lưu trữ toàn bộ dữ liệu trong bộ nhớ cũng có thể làm giảm hiệu năng của thuật toán.

Tóm lại, KNN là một thuật toán đơn giản nhưng hiệu quả trong nhiều trường hợp, với nguyên lý dựa vào thông tin của K điểm dữ liệu gần nhất để xác định đầu ra cho dữ liệu mới. Tuy nhiên, khi áp dụng, cần đặc biệt cân nhắc đến hiệu quả tính toán và lựa chọn tham số K phù hợp để đảm bảo độ chính xác và hiệu suất của mô hình.

1.4.2.4. Cubist (Lập thể) [8]

Lập thể (Cubist) là một thuật toán dựa trên các nguyên tắc (rules) được sử dụng để xây dựng các mô hình dự báo dựa trên việc phân tích dữ liệu đầu vào. Nó được phát triển dựa trên sự mở rộng của mô hình cây quyết định M5 của Quinlan, John R với khả năng xử lý lên tới hàng nghìn biến đầu vào. Sau đó, cây quyết định M5 đã được cải thiện bởi Wang, Yong và Ian H Witten với hiệu suất tốt hơn. Dựa trên thuật toán cây quyết định M5 và sự cải thiện của Wang, Yong và Ian H Witten, thuật toán Cubist đã được phát triển dựa trên các nguyên tắc của thuật toán M5.

Về mặt lý thuyết, thuật toán Cubist được thực hiện qua 4 bước:

- Bước 1: Phân nhánh/chia tách dữ liệu để phát triển một cây quyết định hoàn chỉnh;

- Bước 2: Phát triển mô hình hồi quy tại mỗi nút để dự báo dựa trên dữ liệu đã được phân nhánh/chia tách;
- Bước 3: “Cắt tỉa” cây để tránh vấn đề “thiếu ăn khớp” (under-fitting) hoặc “thừa ăn khớp” (over-fitting) của các giá trị dự báo so với các giá trị thực tế. Trong bước này, một số biện pháp cắt tỉa được áp dụng để chuẩn hóa dữ liệu và làm cho mô hình “hiểu” dữ liệu hơn trong quá trình học/huấn luyện/huấn luyện;
- Bước 4: “Làm mịn” cây để bù cho sự gián đoạn quá phân nhánh/chia tách dữ liệu.

Thuật toán Lập thể là một mô hình hồi quy hướng dự đoán được mô tả bởi Quinlan (Quinlan, 1992) với sửa chữa bổ sung dựa vào lân cận gần nhất (nearest neighbors) trong tập huấn luyện, như đã mô tả ở Quinlan (Quinlan, 1993). Mô hình Lập thể đã được kết hợp bằng cách sử dụng kết hợp tuyến tính như sau:

$$\hat{y}_{par} = a \cdot x \cdot \hat{y}_k + (1 - a) \cdot x \cdot \hat{y}_p$$

Trong đó: \hat{y}_k là sự dự đoán từ thuật toán hiện hành; \hat{y}_p là mô hình mẹ ở trên nó trong cây; a là hệ số làm mịn.

Lập thể là một mô hình cây, đồng thời (simultaneously) tạo thành (constitute) một tập hợp các quy tắc (a set of rules) và một số mô hình hồi quy tuyến tính. Nói cách khác, Lập thể sử dụng câu lệnh “Nếu” và “Sau đó”, “Nếu” là một điều kiện và “Sau đó” là một phương trình (equation) tuyến tính để tính toán ra kết quả. Hai tham số của Lập thể có thể mặc định (default) hoặc điều chỉnh được (tunable), bao gồm các ban-số lượng các hoạt động thúc đẩy (boosting) và lân cận-số trường hợp được sử dụng để sửa dự đoán dựa trên quy tắc.

1.4.2.5. LR – Linear Regression (Hồi quy tuyến tính) [6]

Trong thống kê, hồi quy tuyến tính là một cách tiếp cận tuyến tính để mô hình hóa mối quan hệ giữa một phản ứng vô hướng và một hoặc nhiều biến giải thích (còn được gọi là các biến phụ thuộc và độc lập). Trường hợp của một biến giải thích được gọi là hồi quy tuyến tính đơn giản; đối với nhiều hơn một, quá trình này được gọi là hồi quy nhiều tuyến tính. Thuật ngữ này khác với hồi quy tuyến tính đa biến, trong đó nhiều biến phụ thuộc tương quan được dự đoán, thay vì một biến vô hướng đơn lẻ.

Trong hồi quy tuyến tính, các mối quan hệ được mô hình hóa bằng cách sử dụng các hàm dự báo tuyến tính mà các tham số mô hình chưa biết được ước tính từ dữ liệu. Các mô hình như vậy được gọi là mô hình tuyến tính. Thông thường nhất, giá trị trung

biên có điều kiện của phản hồi cho các giá trị của các biến giải thích (hoặc các yếu tố dự đoán) được giả định là một hàm liên kết của các giá trị đó; ít phổ biến hơn, trung vị có điều kiện hoặc một số lượng tử khác được sử dụng. Giống như tất cả các hình thức phân tích hồi quy, hồi quy tuyến tính tập trung vào phân phối xác suất có điều kiện của phản hồi cho các giá trị của các yếu tố dự đoán, thay vì phân phối xác suất chung của tất cả các biến này, là lĩnh vực của phân tích đa biến.

Nghiệm cho bài toán Linear Regression:

Hàm mất mát * $\mathcal{L}(w)$ có đạo hàm tại mọi w . Việc tìm giá trị tối ưu của w có thể được thực hiện thông qua việc giải phương trình đạo hàm của $\mathcal{L}(w)$ theo w bằng không. Đạo hàm của hàm Linear Regression:

$$\frac{\nabla \mathcal{L}(w)}{\nabla(w)} = \frac{1}{N} X(X^T w - y) \quad (1)$$

Giải phương trình đạo hàm bằng không:

$$\frac{\nabla \mathcal{L}(w)}{\nabla(w)} = 0 \Leftrightarrow XX^T w = Xy \quad (2)$$

Nếu phương trình XX^T khả nghịch, phương trình (2) có nghiệm duy nhất $w = (XX^T)^{-1} Xy$

Nếu ma trận XX^T không khả nghịch, phương trình (2) sẽ vô nghiệm hoặc vô số nghiệm. Lúc này, một nghiệm đặc biệt của phương trình có thể được xác định dựa vào *giả nghịch đảo (pseudo inverse)*. Người ta chứng minh được rằng với mọi ma trận X , luôn tồn tại duy nhất một giá trị w có l_2 norm nhỏ nhất giúp tối thiểu $\|X^T w - y\|_F^2$. Cụ thể, $w = (XX^T)^+ Xy$. Trong đó $(XX^T)^+$ là giả nghịch đảo của XX^T . Giả nghịch đảo của một ma trận A luôn luôn tồn tại, thậm chí cả khi ma trận đó không vuông. Khi A là vuông và khả nghịch thì giả nghịch đảo chính là nghịch đảo. Nghiệm của bài toán tối ưu:

$$w = (XX^T)^+ Xy$$

Hạn chế của Linear Regression:

- Rất nhạy cảm với nhiễu cho nên trước khi thực hiện linear regression thì nhiễu cần phải được loại bỏ (tiền xử lý)
- Không biểu diễn được các mô hình phức tạp.

1.4.2.6. CART - Classification and Regression Trees (cây phân tách và cây hồi quy) [6]

Thuật toán này đã được mô tả lần đầu tiên bởi Breiman L (1984) và được phát triển với các biến mô hình như RF, CART có thể được áp dụng vào vấn đề dự đoán phân tách hoặc hồi quy. Ý tưởng của thuật toán này là một mô hình dự đoán đơn giản nhưng

manh mẽ (robust). Trong phương pháp này, dữ liệu được phân vùng dọc theo các trục (axes) dự đoán thành các tập hợp con (subsets) với giá trị của biến phụ thuộc, nơi mà cây quyết định được tạo ra để đưa ra các dự đoán từ các quan sát (observations) mới. Mô hình CART được trình bày lại như một cây nhị phân (binary tree). Mỗi nút gốc được trình bày một biến dự đoán liên tục (x) và một điểm phân tách trên điểm đó. Từ nút gốc, các nút lá của cây được tạo bằng (generated with) một biến đầu ra (y) được sử dụng để tạo ra dự đoán. Quá trình tiếp tục đi đến dự đoán cuối cùng, nói cách khác, cho đến khi một cây phù hợp được xây dựng.

1.5. Tổng quan về kim loại nặng

1.5.1. Khái niệm về kim loại nặng

Kim loại nặng (KLN) là thuật ngữ dùng để chỉ các nguyên tố kim loại có tỷ trọng lớn hơn 5g/cm^3 [1]. Trong đó các á kim như asen, selen cũng được xem là nguyên tố kim loại nặng. Tỷ trọng của một số kim loại nặng cụ thể: chì (Pb) - $11,3\text{g/cm}^3$, cadimi (Cd) - $8,6\text{g/cm}^3$, thủy ngân (Hg) - $13,5\text{g/cm}^3$, crom (Cr) - $7,19\text{g/cm}^3$, đồng (Cu) - $8,96\text{g/cm}^3$, kẽm (Zn) - $7,1\text{g/cm}^3$, asen (As) - $5,7\text{g/cm}^3$.

1.5.2. Nguồn gốc ô nhiễm và độc tính của một số kim loại nặng

Kim loại là nhóm nguyên tố đã được con người biết đến và ứng dụng từ rất sớm trong đời sống và sản xuất công nghiệp. Một số kim loại đóng vai trò là nguyên tố vi lượng thiết yếu, cần thiết cho hoạt động sống của cơ thể con người và sinh vật. Tuy nhiên, khi tồn tại ở nồng độ cao, các kim loại này có thể trở nên độc hại và gây ảnh hưởng nghiêm trọng đến sức khỏe cũng như môi trường sinh thái. Trong số 118 nguyên tố trong bảng tuần hoàn, có khoảng 90 nguyên tố được xếp vào nhóm kim loại, trong đó khoảng 30 kim loại và bán kim loại đã được nghiên cứu chuyên sâu về mức độ độc hại đối với con người và hệ sinh thái.

Ô nhiễm kim loại nặng có thể bắt nguồn từ cả yếu tố tự nhiên như phong hóa khoáng vật, hoạt động núi lửa, cháy rừng,... và các nguồn do con người gây ra, mà trong đó các hoạt động nhân sinh là nguyên nhân chủ yếu. Các nguồn ô nhiễm kim loại nặng do con người bao gồm:

Hoạt động nông nghiệp sử dụng hóa chất bảo vệ thực vật: Trong hơn một thế kỷ qua, các hợp chất như chì hydro asenat, canxi asenat, đồng sunfat, đồng oxyclorea, đồng hydroxide và thủy ngân sunfat đã được sử dụng để phòng trừ sâu bệnh và nấm hại cây trồng. Dù một số chất hiện nay đã bị hạn chế hoặc cấm sử dụng, dư lượng của chúng có thể vẫn tồn tại trong môi trường, bám lên cây cỏ, sau đó bị rửa trôi vào đất và nguồn

nước, tích lũy lâu dài và gây ảnh hưởng đến hệ sinh thái. Các nguyên tố kim loại nặng này có khả năng tạo phức với các chất hữu cơ trong đất như axit humic, axit fulvic hoặc hấp phụ lên bề mặt các hạt keo, khiến chúng có xu hướng tích tụ lâu dài trong môi trường đất.

Bùn thải từ hệ thống thoát nước và xử lý nước thải: Bùn cồng là sản phẩm phụ của quá trình keo tụ, lắng trong hệ thống xử lý nước thải, hoặc từ quá trình rửa trôi tự nhiên. Loại bùn này thường chứa hàm lượng cao các chất hữu cơ, dưỡng chất như nitơ, photpho và các nguyên tố khoáng vi lượng như kali, canxi, magie. Tuy nhiên, trong bùn cũng tồn tại lượng lớn các kim loại nặng độc hại như cadimi, đồng, niken, crom và kẽm. Việc sử dụng bùn chưa xử lý kỹ làm phân bón hoặc cải tạo đất có thể gây ô nhiễm đất canh tác và nguồn nước. Do đó, cần có quy trình quản lý chặt chẽ để giảm thiểu nguy cơ ô nhiễm môi trường từ nguồn bùn thải này.

Khai thác và chế biến khoáng sản: Các hoạt động khai thác và luyện kim là một trong những nguồn phát thải kim loại nặng lớn nhất ra môi trường. Trong quá trình khai thác mỏ, các công đoạn như nghiền, tuyển quặng tạo ra lượng lớn chất thải chứa kim loại, được thải ra dưới dạng bùn, nước thải hoặc quặng nghèo bị loại bỏ. Những chất thải này có thể gây ô nhiễm đất, nước và trầm tích, ảnh hưởng đến hệ sinh thái và đời sống sinh vật. Ngoài ra, quá trình luyện quặng trong các lò nấu còn tạo ra xỉ thải và hơi kim loại dễ phát tán ra môi trường xung quanh. Các cơ sở luyện kim, mạ kim loại, đúc hợp kim... cũng góp phần phát tán các kim loại nặng như chì, thủy ngân, arsen, niken, đồng, crom, bạc, vàng,... Tại các khu vực gần lò luyện kim, nồng độ kẽm trong lớp đất mặt hữu cơ có thể đạt tới 1,6 – 2% trong phạm vi bán kính chỉ 0,3 km.

Các nguồn phát thải khác: Việc đốt nhiên liệu hóa thạch như than đá, dầu mỏ, khí tự nhiên trong công nghiệp, sản xuất điện, giao thông vận tải, sinh hoạt... tạo ra một lượng lớn kim loại nặng như thủy ngân, arsen, cadimi, chì, crom, niken,... Trong đó, các nguyên tố dễ bay hơi như thủy ngân và cadimi có thể phát tán ở dạng khí, trong khi các kim loại khác thường tồn tại ở dạng hạt mịn trong không khí. Bên cạnh đó, nước rỉ rác từ các bãi chôn lấp chất thải rắn và khí thải từ hoạt động đốt rác (công nghiệp, sinh hoạt, y tế) cũng là nguồn gây ô nhiễm kim loại đáng kể. Trong quá khứ, việc sử dụng xăng pha chì trong phương tiện giao thông cũng là nguyên nhân khiến lượng lớn chì phát tán vào không khí và môi trường đất.

Tóm lại, ô nhiễm kim loại nặng có thể phát sinh từ cả nguồn tự nhiên lẫn các hoạt động của con người. Trong phạm vi nghiên cứu này, tôi sẽ trình bày về nguồn gốc phát

sinh và mức độ độc hại của 7 thông số kim loại nặng chính có ý nghĩa trong đánh giá chất lượng môi trường.

a. Asen [11]

** Nguồn gốc phát sinh*

- Quá trình khai khoáng và nghiền lọc quặng,
- Các chất phế thải trong sản xuất công nghiệp,
- Sử dụng phân bón có chứa Asen.

** Độc tính của Asen*

Asen là một chất độc có tính tích lũy. Asen gây độc cho cơ thể khi chúng xâm nhập vào cơ thể thông qua tiếp xúc với da, qua đường tiêu hóa, qua đường hô hấp. Sau khi được hấp thụ, As vào gan, thận, tim, xương, lông, tóc, da, móng, não,...Asen hữu cơ dễ dàng hấp thụ qua da và màng ruột hơn.

Nhiễm độc cấp tính: Khi cơ thể hấp thụ As qua đường tiêu hóa gây tổn thương mạnh đến hệ tiêu hóa, gây rối loạn thần kinh, khi nồng độ As lên đến 60mg/L có thể gây chết. Khi hít phải không khí, khói, hơi chứa As sẽ gây tổn thương hệ hô hấp như khó thở, ho, gây rối loạn thần kinh như đau đầu; chóng mặt và gây viêm loét da do tiếp xúc với As.

Nhiễm độc mãn tính: Tiếp xúc với As ở nồng độ thấp trong thời gian dài sẽ gây viêm da, nhiễm sắc tố da, tổn thương niêm mạc, viêm kết mạc, kích ứng đường hô hấp trên, rối loạn dạ dày - ruột, rối loạn thần kinh thời gian tiếp xúc càng dài sẽ gây nên ung thư da, ung thư bàng quang và ung thư phổi.

b. Chì

** Nguồn gốc phát sinh*

Chì (Pb) là một trong các kim loại nặng có ảnh hưởng nhiều tới ô nhiễm môi trường, vì nó có khả năng tích lũy lâu dài trong cơ thể và gây nhiễm độc tới người, động vật thủy sinh qua dây chuyền thực phẩm. [10]

Mặc dù Pb xuất hiện trong môi trường tự nhiên ở dạng vết, nồng độ chì trong vỏ trái đất vào khoảng 20ppm, nồng độ Pb trong các nguồn nước thiên nhiên khoảng 5µg/l, nguồn chì phát tán vào môi trường do các hoạt động tự nhiên (núi lửa, cháy rừng, rỉ chảy từ quặng mỏ) chỉ chiếm một phần nhỏ so với nguồn nhân tạo, do các hoạt động của con người. Theo thống kê có tới 360 nghề sử dụng chì và các hợp chất chì vô cơ như đốt nhiên liệu hóa thạch, khai thác khoáng sản, sử dụng ắc quy chì, sơn công nghiệp chứa chì (Pb₃O₄ màu đỏ, PbSO₄ màu trắng, PbCrO₄ màu vàng), sử dụng thuốc trừ sâu có chứa

chì (PbHAsO_4), men đồ gốm (PbO), sử dụng xăng pha chì (tetraetyl chì) một chất chống kích nổ, chất thải, nước thải của một số ngành công nghiệp có sử dụng chì. Nhiều hợp chất của chì không tan, khi thải bỏ vào môi trường sẽ sa lắng vào môi trường bùn đáy.

** Độc tính của chì [11]*

Các hợp chất của chì được hấp thụ chủ yếu qua đường hô hấp và qua da, một phần qua đường tiêu hóa. Các hơi, khói, bụi chì dễ dàng thâm nhập qua đường hô hấp đi vào cơ thể. Chì vô cơ hấp thụ qua da rất ít, chì hữu cơ thì dễ dàng hấp thụ qua da và thành ruột hơn các hợp chất chì vô cơ.

Triệu chứng nhiễm độc cấp tính xuất hiện khi hàm lượng chì được hấp thụ vào cơ thể trên 2000mg/kg thể trọng. Triệu chứng hay gặp là táo bón, nôn mửa, đau bụng trên, trụy tim mạch, trong trường hợp nặng có thể dẫn đến tử vong.

Nhiễm độc mãn tính do tiếp xúc với chì trong một thời gian dài. Biểu hiện ban đầu là mất ngủ, biếng ăn, chân răng có viền đen, nước bọt có vị tanh của kim loại. Trường hợp nhiễm độc nặng sẽ bị thiếu máu, viêm não ở trẻ em, viêm thận mãn tính. Một số trẻ em bị dị tật bẩm sinh như bộ não chậm phát triển, hỏng thận do mẹ tiếp xúc với chì khi mang thai.

c. Cadimi

** Nguồn gốc phát sinh*

Cadimi (Cd) là kim loại màu trắng, dễ kéo dẫn và dát mỏng. Cadimi phân bố rộng rãi trong vỏ trái đất với nồng độ trung bình khoảng $0,1\text{mg/kg}$. Mức độ tích lũy cao nhất của các hợp chất chứa Cd trong môi trường trầm tích khoảng 15mg/kg .

Cadimi trong tự nhiên có lẫn trong quặng Kẽm, là sản phẩm phụ của quá trình khai thác, luyện quặng kẽm, chì. Cadimi còn được giải phóng ra trong quá trình đốt than đá, rác thải và công nghiệp luyện thép. Cadimi được dùng chủ yếu là cực của pin điện, chất tạo màu, tạo độ cứng cho nhựa, men. Trong nông nghiệp, sử dụng rộng rãi phân photphat có lẫn Cd gây ô nhiễm môi trường đất nông nghiệp.

** Độc tính của Cadimi [11]*

Triệu chứng nhiễm độc cấp tính do ăn phải cadimi là nước bọt tiết ra nhiều, buồn nôn và nôn mửa liên tục, chảy máu, choáng váng và ngất.

Nhiễm độc mãn tính: Thường xuyên giảm nhạy khứu giác, xuất hiện viêm mũi, viêm phế quản và nhất là thủng phế quản do tác dụng kích ứng của cadimi trên niêm mạc đường hô hấp. Khi tiếp xúc trong dài với cadimi sẽ gây bệnh nhuyễn xương, tràn khí, suy thận, suy gan.

d. Thủy ngân

**Nguồn gốc phát sinh*

Các hợp chất thủy ngân được ứng dụng rộng rãi trong các ngành kỹ thuật khác nhau (quá trình điện phân, xúc tác, hoá chất bảo vệ thực vật...). Tổng sản lượng thủy ngân trên toàn cầu khoảng 10.10^3 tấn/năm [10].

**Độc tính của Thủy ngân*

Thủy ngân và các hợp chất của thủy ngân chủ yếu hấp thụ qua đường hô hấp, khoảng 80% hơi thủy ngân và metyl thủy ngân trong không khí hấp thụ vào cơ thể qua đường hô hấp, do tính dễ tan trong mỡ nên metyl thủy ngân và phức chất thủy ngân hữu cơ dễ dàng hấp thụ qua da và qua màng ruột. Metyl thủy ngân là hợp chất rất độc, gây nguy hiểm cho não, có khả năng tích lũy trong các mô mỡ, đây là nguyên nhân gây ra bệnh Minamata ở Nhật vào thập niên 1950, do cư dân đã tiêu thụ cá có chứa metyl thủy ngân.

Các biểu hiện của ngộ độc cấp tính là ho, khó thở, thở gấp, sốt, buồn nôn, hôn mê, đau dạ dày và cơ thắt ở vùng ngực. Ngộ độc nặng khi ăn phải lượng lớn thủy ngân thì có thể dẫn đến tử vong.

Tiếp xúc lâu dài với thủy ngân thì có thể gây ra các triệu chứng sau:

- Vàng da do suy yếu chức năng của gan
- Rối loạn tiêu hóa do suy yếu hoạt tính của men tiêu hóa
- Protein niệu
- Viêm lợi do lượng Hg thải qua tuyến nước bọt tích đọng ở chân răng
- Các bệnh liên quan đến não và hệ thần kinh như đau đầu, rối loạn thần kinh dẫn đến nói lắp, run tay, mất cảm giác, co giật... và có thể bị teo vỏ tiểu não

Nhiễm độc mãn tính: Khi tiếp xúc với nồng độ thủy ngân thấp trong thời gian dài, gây ảnh hưởng lên hệ thần kinh trung ương, gây rối loạn tiêu hóa do suy giảm hoạt tính của men tiêu hóa, ở nồng độ $0,01\text{mg/m}^3$ gây ra triệu chứng mất ngủ, kém ăn.

e. Đồng

**Tính chất và nguồn gốc phát sinh*

Các nguồn gây ra ô nhiễm đồng cho môi trường như khai thác mỏ, nghiền, nấu, tuyển quặng. Từ phế thải của quá trình mạ đồng, gia công mỹ nghệ. Trong nông nghiệp các muối và oxit của đồng được dùng để diệt nấm, tảo, sử dụng trong phân bón.

**Độc tính của Đồng*

Phơi nhiễm đồng: Ở nồng độ cao rất độc cho cơ thể, uống nước có chứa đồng lớn hơn 3mg/l sẽ gây ra hội chứng dạ dày - ruột, bao gồm nôn mửa, tiêu chảy. Nuốt phải một lượng lớn đồng, thường là đồng sunfat có thể gây nôn ói, tiêu chảy, có thể gây co giật, hoại tử gan và chết [12].

Trong môi trường sinh thái, đồng có hoạt tính diệt nấm, vì thế hệ sinh vật khác cũng chịu ảnh hưởng từ đồng. Trong hệ sinh thái thủy sinh, cá và giáp xác nói chung rất nhạy cảm với đồng gấp 10-100 lần so với động vật có vú, tảo gấp 1000 lần. Đồng tích tụ trong trầm tích ở nồng độ cao cũng gây ảnh hưởng đến động vật đáy.

f. Kẽm

** Nguồn gốc phát sinh*

Các nguồn phát sinh của kẽm bao gồm mạ điện, luyện kim và chế biến quặng. Hơn nữa, kẽm là một thành phần quan trọng của đúc kim loại, hợp kim khác, sơn, đặc biệt là quá trình mạ kim loại kẽm, để bảo vệ lớp vật liệu bên trong. Hoạt động phát thải kẽm ra môi trường do con người vượt xa lượng phát thải từ các nguồn tự nhiên. Ước tính phát thải 26,76 tấn kẽm vào nước mặt từ 389 cơ sở sản xuất và các cơ sở chế biến tại Mỹ vào năm 2002, chiếm khoảng 0,1% tổng lượng phát thải ra môi trường ước tính [13].

** Độc tính của kẽm*

Kẽm là một nguyên tố dinh dưỡng thiết yếu cho tất cả các sinh vật, là nguyên tố cần thiết cho một lượng lớn enzyme hoạt động trong cơ thể. Tuy nhiên khi tiếp xúc với một lượng lớn sẽ gây ảnh hưởng đến cơ thể. Phần lớn các ảnh hưởng của kẽm đến cơ thể do tiếp xúc qua đường hô hấp, khi hít phải các oxit kẽm gây kích ứng hệ hô hấp, đau và khô cổ họng, đau ngực và khó thở [13]. Ngoài ra, kẽm đi vào cơ thể bằng đường ăn uống với hàm lượng vượt quá 40mg mỗi ngày trong thời gian dài, sẽ gây ra tình trạng ngộ độc với các biểu hiện như sốt, ho, buồn nôn, khó thở, rối loạn tiêu hóa, rối loạn phản ứng miễn dịch.

g. Crom

** Nguồn gốc phát sinh*

Crom có nhiều ứng dụng như sản xuất sắt hợp kim crom và nhiều hợp kim khác chứa coban hay niken, chế tạo thép không gỉ, các cromat khác được sản xuất bằng nấu chảy, nung và chiết tách. Crom thường được dùng trong các sản phẩm nhuộm. Trong công nghiệp, crom được dùng trong sản xuất các hợp kim sắt - crom, tinh luyện quặng, mạ kim loại, sử dụng Cr làm chất bắt và giữ màu trong nhuộm len, da, lông, sắc tố trong

son, mực,... Trong tự nhiên Cr phát tán chủ yếu do hoạt động của núi lửa [12].

** Độc tính của crom*

Crom xâm nhập vào cơ thể qua đường hô hấp, da và đường tiêu hóa. Crom có tính độc cao đối với con người và động vật. Khả năng gây độc của Cr(VI) mạnh hơn nhiều so với Cr(III). Khi xâm nhập vào cơ thể crom sẽ hòa tan vào máu, liên kết với tế bào máu đỏ và được vận chuyển đến các cơ quan khác nhau (gan, phổi, thận,...). Hấp thụ Cr(VI) qua đường ăn uống có thể gây hư hại tiểu cầu và có thể gây chết. Cr(VI) là chất ăn mòn, nó cũng gây ra lở loét da, bệnh chàm da khi tiếp xúc. Sự hít thở phải Cr(VI), như bụi cromat hoặc axit cromic, có thể gây loét màng nhày mũi và làm thủng vách mũi, co thắt phế quản.

CHƯƠNG 2: ĐÁNH GIÁ HIỆN TRẠNG CHẤT LƯỢNG MÔI TRƯỜNG NƯỚC VÀ TRẦM TÍCH MẶT, XÁC ĐỊNH RỦI RO SINH THÁI CỦA MỘT SỐ KIM LOẠI NẶNG VÀ DỰ BÁO SỨC KHỎE MÔI TRƯỜNG BẰNG CÁC MÔ HÌNH HỌC MÁY

2.1. Khảo sát chất lượng môi trường, hiện trạng sức khỏe môi trường và người dân tại khu vực sông Bồng Miêu, xã Tam Lãnh, huyện Phú Ninh, tỉnh Quảng Nam

Phương pháp lập phiếu khảo sát điều tra: khảo sát người dân về sự thay đổi chất lượng môi trường sống và các bệnh tật của người dân sống tại khu vực sông Bồng Miêu. Đánh giá hiện trạng chất lượng môi trường nước và trầm tích, xác định rủi ro sinh thái của một số kim loại nặng và dự báo sức khỏe môi trường bằng các mô hình học máy tại khu vực nghiên cứu.

❖ *Cỡ mẫu khảo sát:*

- Công thức xác định cỡ mẫu theo TT 03/2025/BLĐTBXH [14]:

$$n_0 = \frac{N}{1 + N \times e^2}$$

- Trong đó:

n: số lượng mẫu cần xác định

N: số hộ dân tại khu vực khảo sát (N=380)

e: Sai số cho phép

- Khi tính cỡ mẫu cho toàn ngành thì sai số e cho phép không quá 10%.

- Khi tính cỡ mẫu trong một doanh nghiệp cụ thể của từng ngành thì sai số e cho phép không quá 20% đối với doanh nghiệp có dưới 1.000 người lao động làm nghề, công việc được đánh giá; không quá 10% đối với doanh nghiệp có từ 1.000 người lao động trở lên làm nghề, công việc được đánh giá.

$$n_0 = \frac{380}{1 + 380 \times 0.2^2} = 25 \text{ (người)}$$

Để đảm bảo tính đại diện và đủ cơ sở để kết luận về hiện trạng sức khỏe môi trường, sinh viên tiến hành khảo sát 50 người dân tại thôn Bồng Miêu, xã Tam Lãnh, huyện Phú Ninh, tỉnh Quảng Nam (Mỏ vàng Bồng Miêu). Kết quả thu thập được sẽ mang tính tham khảo và sẽ giúp cung cấp những thông tin sơ bộ, hỗ trợ việc hình thành các đánh giá ban đầu cho nghiên cứu.

2.1.1. Đối tượng và vị trí khảo sát



Hình 2.1. Vị trí lấy phiếu khảo sát hộ dân xung quanh sông Bồng Miêu

- Đối tượng lấy phiếu khảo sát gồm các hộ gia đình sinh sống gần sông Bồng Miêu
- Phạm vi lấy phiếu khảo sát: Khu dân cư gần sông Bồng Miêu như **Hình 2.1**

2.1.2. Thời gian lấy phiếu khảo sát

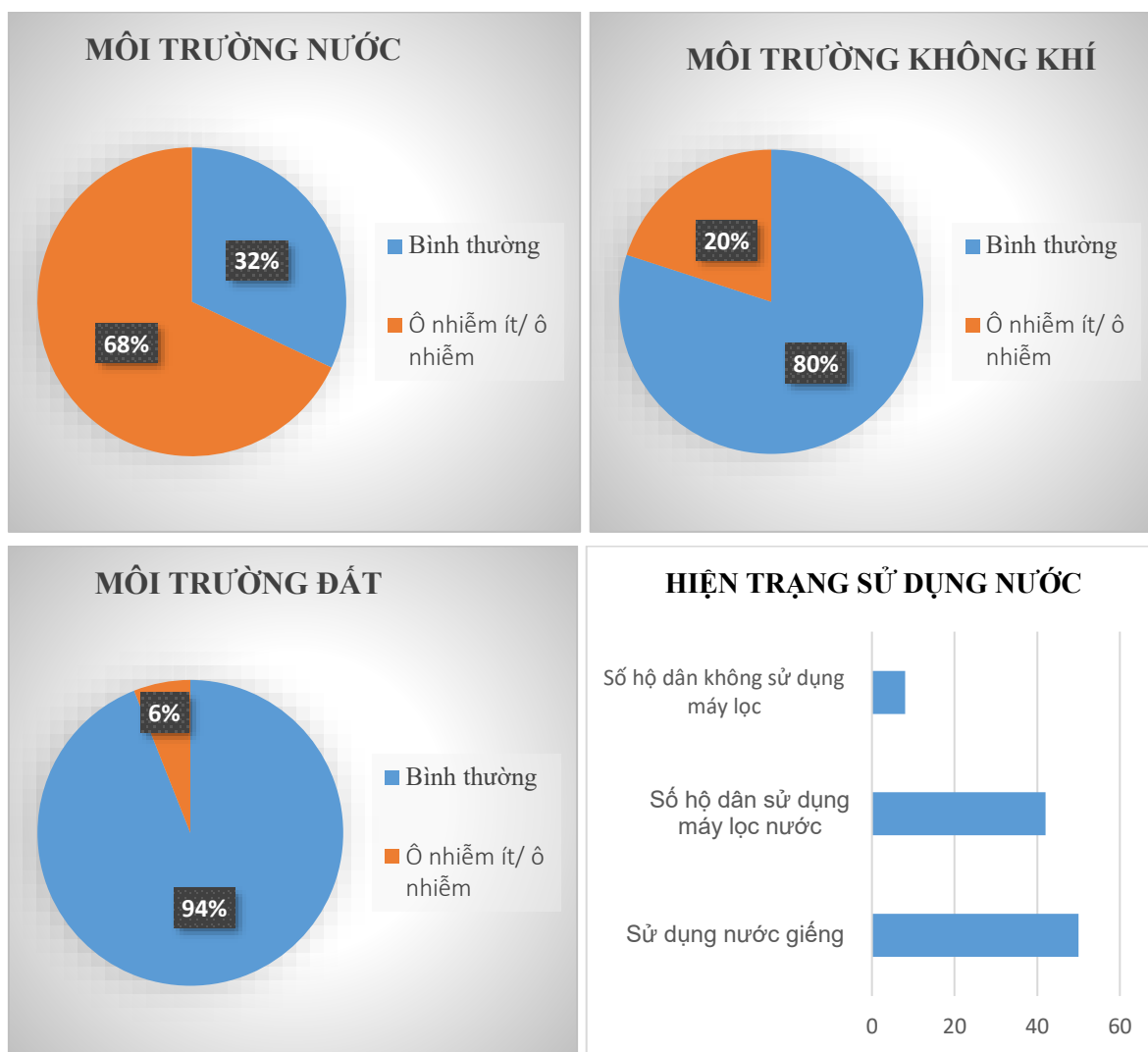
- Thời gian: Từ 27/03/2025 đến ngày 29/03/2025
- Số lượng phiếu khảo sát: 50 phiếu
- Thời tiết:

- Nhiệt độ: 24-27°C
- Hướng gió chủ yếu: Đông Nam
- Độ ẩm tương đối: 72%

- Tình hình chất lượng môi trường tại thời điểm đi khảo sát: trời nắng nhẹ, gió nhẹ, nhiều mây

2.1.3. Phân tích kết quả khảo sát

Phương pháp xử lý số liệu: Kết quả phân tích được thống kê bằng phần mềm Microsoft Excel.



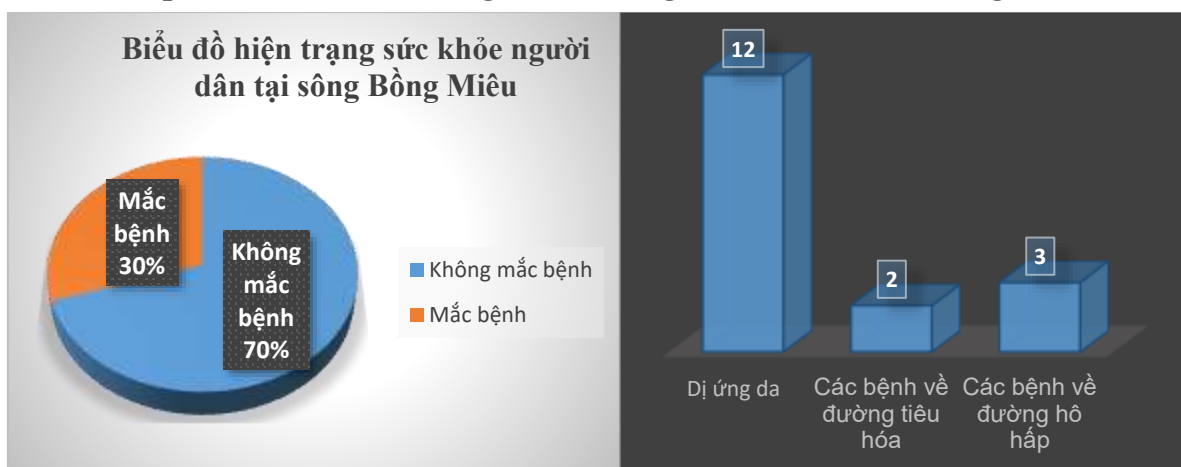
Biểu đồ 2.1: Biểu đồ đánh giá tình trạng môi trường tại khu vực nghiên cứu

Nhận xét:

Trong nghiên cứu này, nhóm khảo sát đã thu thập 50 phiếu khảo sát từ các hộ dân trong tổng số 380 hộ dân tại khu vực nghiên cứu. Để đảm bảo rằng kết quả khảo sát có thể suy rộng ra toàn bộ dân số nghiên cứu với độ tin cậy thống kê, đề tài đã áp dụng công thức xác định cỡ mẫu theo TT 03/2025/BLĐTBXH. Kết quả cho thấy cỡ mẫu tối thiểu cần thiết là khoảng 40 hộ dân. Vì vậy, có thể xem kết quả khảo sát là có ý nghĩa thống kê và có thể suy rộng ra toàn bộ khu vực nghiên cứu với độ tin cậy tương đối cao.

- **Nguồn nước sinh hoạt:** Toàn bộ số hộ dân được khảo sát (100%) sử dụng nước giếng cho mục đích sinh hoạt. Trong đó, 85% số hộ (tương ứng 42/50 hộ) có sử dụng thêm thiết bị lọc nước nhằm cải thiện chất lượng nguồn nước.

- **Chất lượng môi trường nước:** Có 68% phiếu khảo sát (34/50) đánh giá chất lượng nước ở mức ô nhiễm hoặc ít ô nhiễm. Các nguồn gây ô nhiễm chủ yếu được xác định là do hoạt động khai thác vàng và sinh hoạt hằng ngày của người dân. Một số dấu ô nhiễm được người dân phản ánh gồm hiện tượng nước bị đục màu và xuất hiện mùi lạ.
 - **Chất lượng môi trường không khí:** Chỉ có 20% số hộ (10/50 phiếu) đánh giá môi trường không khí bị ô nhiễm hoặc ít ô nhiễm. Nguyên nhân chủ yếu là do hoạt động vận chuyển gỗ bằng xe tải, gây hư hỏng đường giao thông, phát sinh bụi mịn và ảnh hưởng đến chất lượng không khí khu vực.
 - **Chất lượng môi trường đất:** Tỷ lệ hộ dân đánh giá môi trường đất bị ô nhiễm hoặc ít ô nhiễm là 6% (3/50 phiếu khảo sát). Nguyên nhân chính được xác định là do hoạt động khai thác vàng. Biểu hiện của ô nhiễm đất được người dân ghi nhận là hiện tượng giảm năng suất cây trồng, đặc biệt là các loại cây ăn quả.
- **Kết quả khảo sát hiện trạng sức khỏe người dân tại khu vực nghiên cứu**



Biểu đồ 2.2: Biểu đồ kết quả khảo sát sức khỏe người dân sống tại KVNC

Nhận xét:

Kết quả khảo sát cho thấy tỷ lệ người dân mắc các bệnh liên quan đến da liễu, đường tiêu hóa và đường hô hấp chiếm khoảng 30% (tương đương 15/50 phiếu khảo sát). Trong đó, phụ nữ chủ yếu gặp các vấn đề về dị ứng da, ngứa da,... Trong khi đó, nam giới thường mắc các bệnh về đường tiêu hóa và đường hô hấp. Đáng chú ý, một trường hợp được ghi nhận mắc ung thư đường tiêu hóa ở giai đoạn cuối.

* **Khảo sát, điều tra tại Bệnh viện Đa Khoa Quảng Nam**

Bảng 2.1. Thông tin về tình trạng sức khỏe của người dân tại khu vực nghiên cứu do Bệnh viện Đa Khoa tỉnh Quảng Nam cung cấp

| Số liệu người bệnh mắc các bệnh về đường hô hấp, bệnh ngoài da khu vực huyện Phú Ninh khám và điều trị tại BV ĐK Quảng Nam năm 2024 đến nay | | |
|--|----------------------|---------------|
| Nội dung | Bệnh về đường hô hấp | Bệnh ngoài da |
| Điều trị nội trú | 487 | 75 |
| Khám bệnh | 1343 | 209 |
| Tổng | 1830 | 284 |



Hình 2.2. Khảo sát, điều tra thực tế tại Bệnh viện Đa Khoa tỉnh Quảng Nam

*** Nhân xét chung:**

- Kết quả phỏng vấn trực tiếp cho thấy tỷ lệ người mắc bệnh về da cao hơn so với bệnh về đường hô hấp (24% so với 6%). Tuy nhiên, khi so sánh với số liệu từ bệnh viện, số ca bệnh về đường hô hấp lại chiếm tỷ lệ cao hơn (1.830 ca) so với bệnh ngoài da (284 ca). Sự khác biệt này có thể do nhiều nguyên nhân. Thứ nhất, bệnh về da thường có biểu hiện nhẹ, không quá ảnh hưởng nghiêm trọng đến sinh hoạt nên người dân có thể tự điều trị tại nhà hoặc không đi khám, dẫn đến số liệu tại bệnh viện thấp hơn thực tế. Ngược lại, các bệnh về đường hô hấp, đặc biệt là các bệnh mãn tính, hen suyễn, viêm phế quản, thường gây ảnh hưởng trực tiếp, nặng nề đến sức khỏe và sinh hoạt hàng ngày, buộc người dân phải đến cơ sở y tế khám chữa bệnh, nên số liệu ghi nhận tại bệnh viện cao hơn nhiều so với tỷ lệ phản ánh qua phỏng vấn. Ngoài ra, người dân có thể quen với các biểu hiện bệnh ngoài da và ít quan tâm, trong khi bệnh hô hấp dễ gây lo lắng hơn, từ đó tạo ra sự khác biệt về xu hướng giữa hai nguồn dữ liệu.

- Tuy nhiên, ngoài các nguyên nhân trên, có thể giả thiết rằng tâm lý lo ngại, e dè của người dân cũng góp phần làm sai lệch kết quả khảo sát trực tiếp. Khu vực khảo sát là điểm nóng về khai thác vàng trái phép, nhiều hộ dân có tham gia hoặc liên quan đến hoạt động này, do đó họ có thể hạn chế cung cấp thông tin trung thực, sợ ảnh hưởng đến cá nhân hoặc cộng đồng khi làm việc với người lạ, đặc biệt là trong các nội dung nhạy cảm liên quan đến sức khỏe, môi trường.

- Điều này cho thấy việc thu thập thông tin tại các khu vực nhạy cảm cần có sự chuẩn bị kỹ lưỡng về cách tiếp cận, kỹ năng phỏng vấn cũng như cần kết hợp nhiều nguồn dữ liệu (từ cộng đồng, y tế, môi trường, chính quyền địa phương) để đảm bảo tính khách quan, đầy đủ. Đồng thời, kết quả khảo sát cũng cho thấy mức độ ảnh hưởng tiêu cực của môi trường đến sức khỏe người dân là vấn đề đáng lo ngại, cần có sự quan tâm từ các cấp chính quyền, tổ chức y tế, môi trường nhằm tăng cường giám sát, tuyên truyền, hỗ trợ người dân tiếp cận dịch vụ y tế, sử dụng nguồn nước an toàn và giảm thiểu tác động từ hoạt động khai thác khoáng sản (khai thác vàng).

2.2. Khảo sát hiện trạng chất lượng môi trường nước và trầm tích mặt tại Sông Bồng Miêu

a. Mục đích

Nhằm có cái nhìn tổng quan ban đầu về hiện trạng chất lượng nước tại khu vực nghiên cứu, sinh viên đã tiến hành lấy mẫu và phân tích một số chỉ tiêu cơ bản trong nước mặt tại các điểm quan trắc. Việc phân tích này nhằm đánh giá sơ bộ hiện trạng môi trường nước.

b. Đối tượng

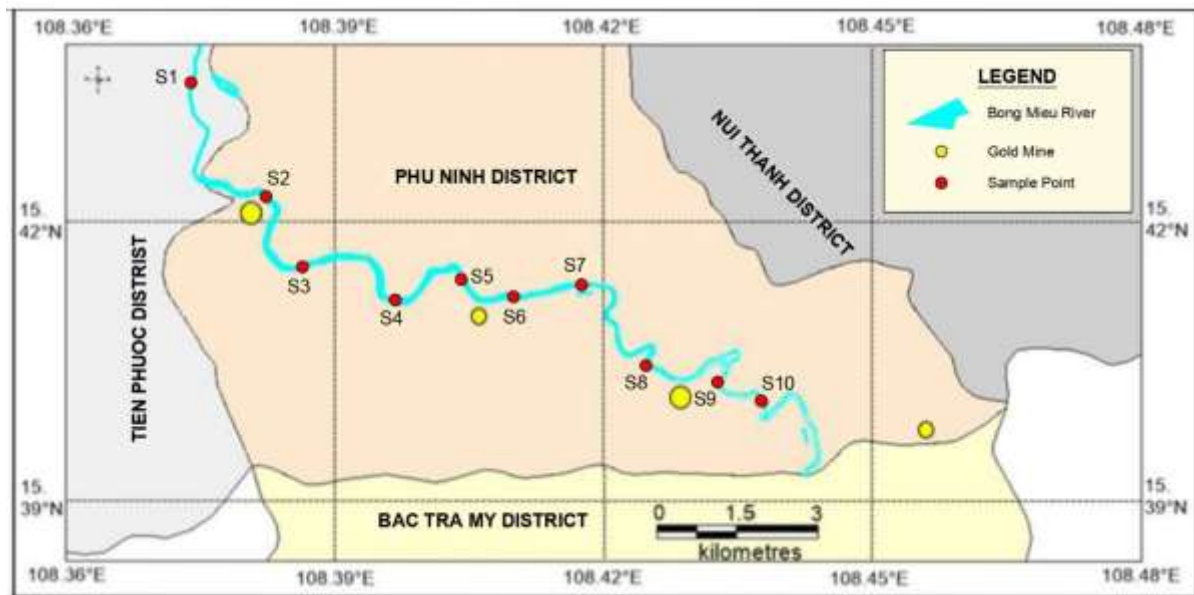
Thông số chất lượng nước/ trầm tích mặt tại Sông Bồng Miêu

c. Phạm vi nghiên cứu

Bảng 2.2. Vị trí lấy mẫu nước sông Bồng Miêu

| STT | Vị trí | Kinh độ | Vĩ độ | Mô tả vị trí lấy mẫu |
|-----|--------|---------------|--------------|---|
| 1 | S1 | 108°22'28.7"E | 15°26'04.7"N | Khu vực hạ lưu sông Bồng Miêu nơi tiếp giáp với sông Tiên |
| 2 | S2 | 108°23'01.2"E | 15°25'20.2"N | Khu vực gần vết tích của khai thác vàng trái phép |

| STT | Vị trí | Kinh độ | Vĩ độ | Mô tả vị trí lấy mẫu |
|------------|---------------|----------------|--------------|--|
| 3 | S3 | 108°23'09.2"E | 15°24'50.3"N | Khu vực thường xuất hiện đãi vàng tự phát |
| 4 | S4 | 108°23'49.4"E | 15°24'43.5"N | Khu vực cách nhà nhà máy khai thác vàng khoảng 1km về hạ lưu |
| 5 | S5 | 108°24'14.5"E | 15°24'49.6"N | Khu vực gần công ty khai thác vàng 6666 về phía hạ lưu |
| 6 | S6 | 108°24'36.1"E | 15°24'39.2"N | Tại cầu Bồng Miêu |
| 7 | S7 | 108°25'03.4"E | 15°24'42.4"N | Khu vực gần công ty khai thác vàng Bồng Miêu về phía hạ lưu |
| 8 | S8 | 108°25'29.2"E | 15°24'16.2"N | Gần khu vực trồng cây lâu năm |
| 9 | S9 | 108°25'58.2"E | 15°24'09.6"N | Gần khu vực khai thác vàng trái phép ở vùng thượng nguồn |
| 10 | S10 | 108°26'10.4"E | 15°24'01.5"N | Gần khu vực khai thác vàng trái phép ở vùng thượng nguồn |



Hình 2.3. Vị trí lấy mẫu nước và trầm tích mặt tại sông Bồng Miêu (chú thích vị trí **Bảng 2.2**)

d. Nội dung

Tiến hành lấy mẫu nước và trầm tích mặt tại các vị trí
Phân tích các thông số chất lượng nước/trầm tích mặt

Mẫu nước

- Phân tích các chỉ tiêu: Nhiệt độ, pH, TSS, BOD₅, COD, N-NH₄⁺, N-NO₂⁻, Cl⁻.
- Quy chuẩn so sánh: QCVN 08:2023/BTNMT - quy chuẩn kỹ thuật quốc gia về chất lượng nước mặt.

Trầm tích

- Phân tích các thông số: Các kim loại nặng (As, Cd, Pb, Zn, Hg, Cr, Cu)
- Quy chuẩn so sánh : QCVN 43 :2017/BTNMT-quy chuẩn kỹ thuật quốc gia về trầm tích

e. Phương pháp và thiết bị quan trắc

- TCVN 6663-1:2011 (ISO 5667-2:2006): Chất lượng nước – Lấy mẫu – Phần 1: Hướng dẫn kỹ thuật lấy mẫu.
- TCVN 6663-6:2018: Chất lượng nước - Lấy mẫu - phần 6: Hướng dẫn lấy mẫu nước sông và suối.
- TCVN 6663-3:2003 (ISO 5667-3:1985): Chất lượng nước – Lấy mẫu – Phần 3: Hướng dẫn bảo quản và xử lý mẫu.

- TCVN 6663-13 : 2000 (ISO 5667-13 : 1997) Phần 13: hướng dẫn lấy mẫu bùn nước, bùn nước thải và bùn liên quan.

- TCVN 6663-15 : 2004 (ISO 5667-15 : 1999) Phần 15: Hướng dẫn bảo quản và xử lý mẫu bùn và trầm tích.

Bảng 2.3. Phương pháp và thiết bị phân tích

| STT | Chỉ tiêu phân tích | Phương pháp | Dụng cụ và thiết bị |
|-----|--|------------------|----------------------------------|
| 1 | Nhiệt độ | Đo nhanh | Bút đo nhanh các thông số cơ bản |
| 2 | pH | Đo nhanh | Bút đo nhanh các thông số cơ bản |
| 3 | Amoni (NH ₄ ⁺) | TCVN 6179-1:1996 | Máy quang phổ LIUV 201 UV/Vis |
| 4 | Chloride (Cl ⁻) | TCVN 6194:1996 | Bộ dụng cụ cơ bản |
| 5 | TSS | TCVN 6625:2000 | Giấy lọc, tủ sấy, cân phân tích |
| 6 | Nitrit (NO ₂ ⁻) | TCVN 6178:1996 | Máy quang phổ LIUV 201 UV/Vis |
| 7 | COD | TCVN 6491:1999 | Bộ phân tích COD VARIO |
| 8 | BOD ₅ | TCVN 6001-1:2008 | Bộ phân tích BOD - FOC225E |

f. Phương pháp phân tích mẫu trầm tích

*** Chiết mẫu trầm tích**

Mẫu được chiết trong dung dịch nước cường toang theo hướng dẫn của ISO 11466:1995. Quy trình chiết mẫu được thực hiện như sau: Cân khoảng 3 gam mẫu trầm tích, chính xác đến 0,001 gam cho vào bình phản ứng dung tích 250 ml. Làm ướt với khoảng từ 0,5 ml đến 1,0 ml nước và vừa trộn vừa cho thêm 21 ml axit clohidric (HCl) 37%, sau đó cho thêm 7 ml axit nitric đậm đặc. Cho 15ml axit nitric (HNO₃) 0,5M, vào bình hấp thụ. Nồi bình hấp thụ và bộ sinh hàn với bình phản ứng và để yên 16 giờ ở nhiệt độ phòng, để quá trình oxi hoá các chất hữu cơ trong mẫu xảy ra từ từ.

Tăng nhiệt độ của hỗn hợp phản ứng từ từ cho đến khi đạt được các điều kiện hồi lưu và duy trì trong 2 giờ, sau đó để nguội. Cho lượng trong bình hấp thụ vào bình phản ứng qua bộ sinh hàn, tráng tiếp bình hấp thụ và bộ sinh hàn bằng 10 ml axit nitric (0,5M).

Lọc dung dịch sau khi phân hủy mẫu qua giấy lọc (đường kính lỗ lọc khoảng 8µm), thu lấy dịch lọc vào bình định mức 100 ml. Lọc tất cả dịch lọc ban đầu qua giấy lọc, sau đó rửa cặn không tan trên giấy lọc với một lượng axit nitric (0,5M) tối thiểu, thu lấy dịch lọc.

*** Phân tích kim loại nặng trong dịch chiết mẫu**

Sau khi chiết mẫu bằng hỗn hợp nước cường toang, thu được dịch chiết. Phân tích nồng độ kim loại nặng trong dịch chiết bằng máy quang phổ hấp thụ nguyên tử AAS contrAA 700 của hãng Analytik Jena, được sản xuất tại Đức. Đối với từng thông số, kỹ thuật phân tích sẽ khác nhau trên máy AAS, cụ thể như sau:

- Nhóm thông số Cd, Pb, Cu, Zn, Cr được phân tích bằng kỹ thuật nguyên tử hóa ngọn lửa (F-AAS) theo hướng dẫn USEPA Method 7000B:2007.

- Thông số As được phân tích bằng kỹ thuật hóa hơi (HG-AAS) sử dụng Argon làm khí mang theo hướng dẫn ISO 20280:2007, thực hiện phản ứng hydride hóa với chất khử natri bohydride (NaBH_4) trong môi trường axit HCl

- Thông số Hg được phân tích bằng kỹ thuật hóa hơi lạnh (CV-AAS) sử dụng Argon làm khí mang và chất khử NaBH_4 trong môi trường axit HCl theo hướng dẫn ISO 16772:2004.



Hình 2.4. Hình ảnh thực tế lấy mẫu nước và trầm tích mặt



Hình 2.5. Bảo quản mẫu tại Phòng thí nghiệm



Hình 2.6. Phân tích các chỉ tiêu nước sông Bồng Miêu tại PTN

Bảng 2.4. Kết quả phân tích mẫu nước sông Bông Miêu (22/4/2025- 29/4/2025)

| Vị trí | Chỉ tiêu | Nhiệt độ | pH | Amoni (NH ₄ ⁺) | Chloride (Cl ⁻) | TSS | Nitrit (NO ₂ ⁻) | BOD ₅ | COD |
|-----------------------|----------|----------|---------------------|---------------------------------------|-----------------------------|-----------------------------------|--|------------------|------|
| | Đơn vị | °C | - | mg/L | mg/L | mg/L | mg/L | mg/L | mg/L |
| | S1 | 26,0 | 7,8 | 0,28 | 6,2 | 17 | 0,044 | 3,0 | 5 |
| | S2 | 25,7 | 7,6 | 0,23 | 4,4 | 25 | 0,055 | 5,2 | 5 |
| | S3 | 25,8 | 7,5 | 0,27 | 1,8 | 55 | 0,036 | 3,6 | 74 |
| | S4 | 26,2 | 8,0 | 0,35 | 4,4 | 20 | 0,030 | 3,8 | 60 |
| | S5 | 25,8 | 7,3 | 0,23 | 3,5 | 24 | 0,027 | 3,1 | 3 |
| | S6 | 26,2 | 7,1 | 0,29 | 7,1 | 29 | 0,023 | 5,2 | 83 |
| | S7 | 26,0 | 7,5 | 0,30 | 3,5 | 42 | 0,017 | 4,1 | 40 |
| | S8 | 26,0 | 7,7 | 0,34 | 3,5 | 67 | 0,051 | 6,1 | 94 |
| | S9 | 26,0 | 7,0 | 0,38 | 6,2 | 84 | 0,031 | 6,4 | 35 |
| | S10 | 26,1 | 7,9 | 0,25 | 7,1 | 48 | 0,030 | 5,2 | 3 |
| QCVN 08:2023/BTNMT | Mức A | - | 6,5 – 8,5 | - | - | ≤ 25 mg/L | - | ≤ 4 | ≤ 10 |
| | Mức B | - | 6,5 – 8,5 | - | - | ≤ 100 mg/L | - | ≤ 6 | ≤ 15 |
| | Mức C | - | 6,5 – 8,5 | - | - | > 100 mg/L và không có rác nổi | - | ≤ 10 | ≤ 20 |
| | Mức D | - | < 6,0 hoặc > 8,5 | - | - | > 100 mg/L và có rác nổi | - | > 10 | > 20 |
| AHSK | | - | - | 0,3 | 250 | - | 0,05 | - | - |

Ghi chú:

- Các mức phân loại đánh giá chất lượng nước được diễn giải cụ thể như sau (theo QCVN 08:2023/BTNMT)

AHSK: Giá trị giới hạn tối đa các thông số ảnh hưởng tới sức khỏe con người.

Mức A: Chất lượng nước tốt. Hệ sinh thái trong môi trường nước có hàm lượng oxy hòa tan (DO) cao. Nước có thể được sử dụng cho mục đích cấp nước sinh hoạt, bơi lội, vui chơi dưới nước sau khi áp dụng các biện pháp xử lý phù hợp.

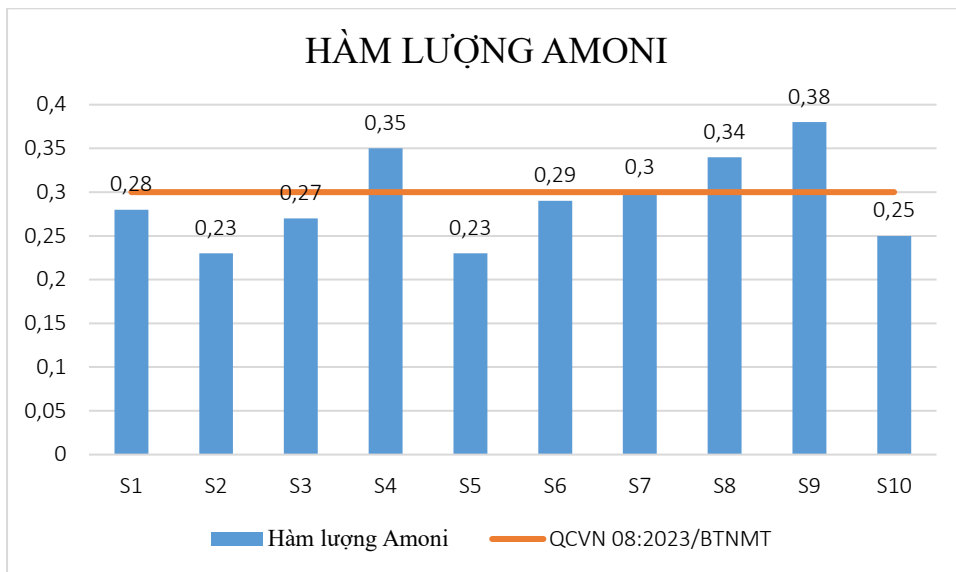
Mức B: Chất lượng nước trung bình. Hệ sinh thái trong nước tiêu thụ nhiều oxy hòa tan do một lượng lớn chất ô nhiễm. Nước có thể sử dụng cho mục đích sản xuất công nghiệp, nông nghiệp sau khi áp dụng các biện pháp xử lý phù hợp.

Mức C: Chất lượng nước xấu. Hệ sinh thái trong nước có lượng oxy hòa tan giảm mạnh do chứa một lượng lớn các chất ô nhiễm. Nước không gây mùi khó chịu, có thể được sử dụng cho các mục đích sản xuất công nghiệp sau khi áp dụng các biện pháp xử lý phù hợp.

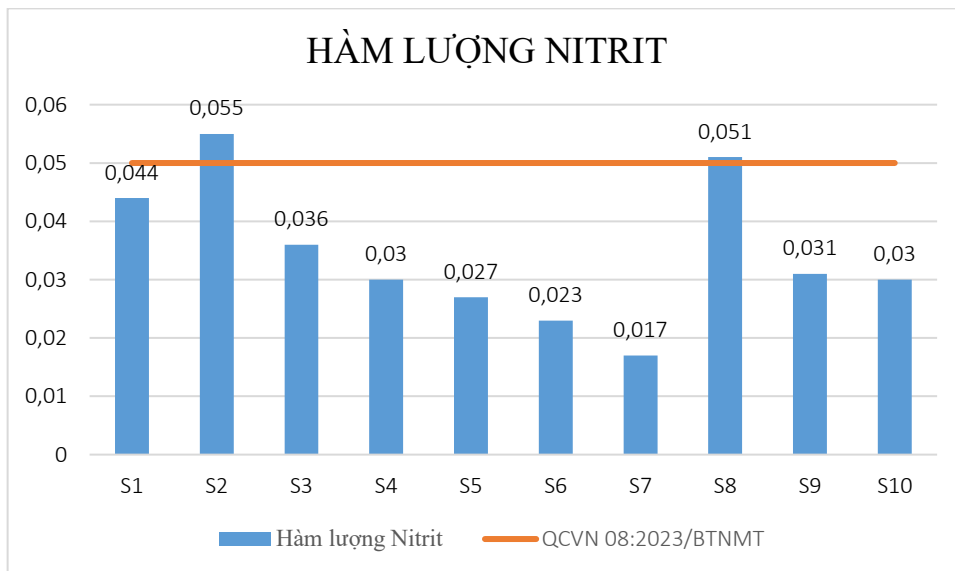
Mức D: Nước có chất lượng rất xấu, có thể gây ảnh hưởng lớn tới cá và các sinh vật sống trong môi trường nước do nồng độ oxy hòa tan thấp, nồng độ chất ô nhiễm cao. Nước có thể được sử dụng cho các mục đích giao thông thủy và các mục đích khác với yêu cầu nước chất lượng thấp [9]

- Các chỉ tiêu được phân tích theo các phương pháp được đưa ra trong Bảng 4, Mục 3, QCVN 08:2023/BTNMT

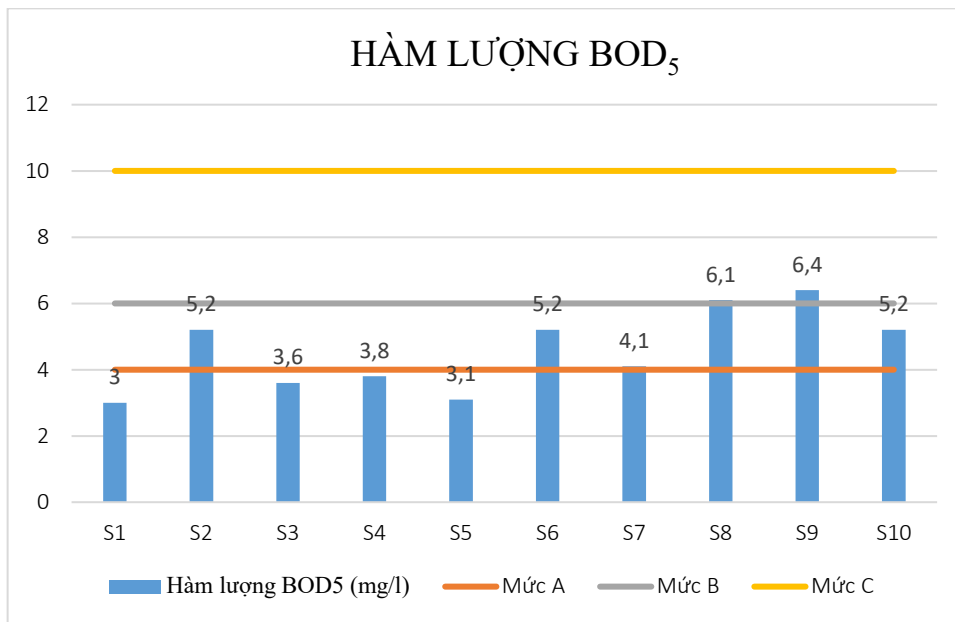
- **Nhận xét:** Kết quả phân tích mẫu nước mặt tại sông Bồng Miêu tại 10 vị trí cho thấy chất lượng nước của sông Bồng Miêu đạt mức tốt, có một số vị trí có chất lượng nước ở mức trung bình: S2 (nồng độ NO_2^- vượt ngưỡng cho phép gây ảnh hưởng đến sức khỏe con người), S4 (nồng độ NH_4^+ vượt ngưỡng cho phép gây ảnh hưởng đến sức khỏe con người), S8 (nồng độ NO_2^- , NH_4^+ vượt ngưỡng cho phép gây ảnh hưởng đến sức khỏe con người, hàm lượng TSS đạt mức B và hàm lượng COD vượt mức D nhiều lần), S9 (nồng độ NH_4^+ vượt ngưỡng cho phép gây ảnh hưởng đến sức khỏe con người, hàm lượng TSS đạt mức B và hàm lượng COD vượt mức D).



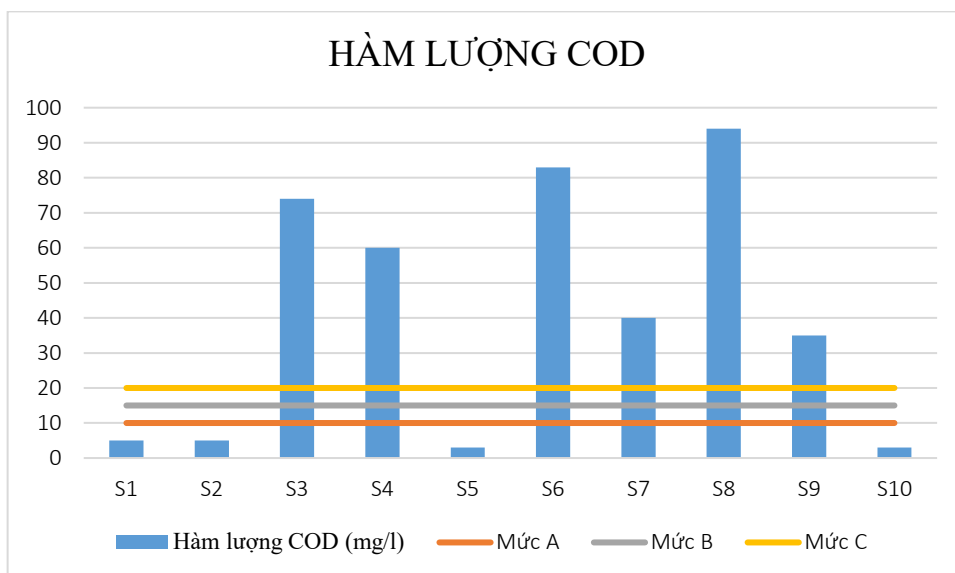
Biểu đồ 2.3: Biểu đồ so sánh nồng độ Amoni (NH_4^+) với QCVN 08:2023/BTNMT



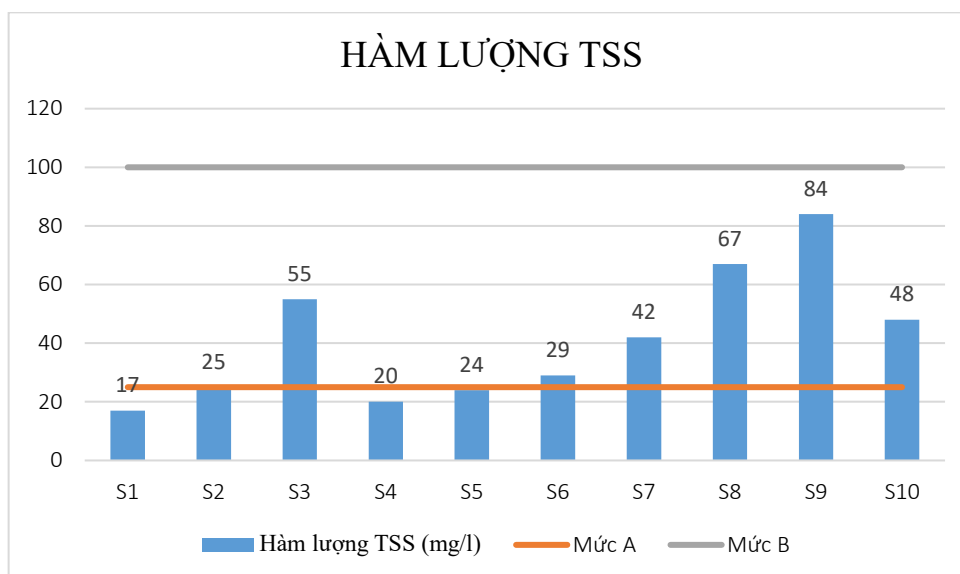
Biểu đồ 2.4: Biểu đồ so sánh nồng độ Nitrit (NO_2^-) với QCVN 08:2023/BTNMT



Biểu đồ 2.5: Biểu đồ so sánh hàm lượng BOD₅ với QCVN 08:2023/BTNMT



Biểu đồ 2.6: Biểu đồ so sánh hàm lượng COD với QCVN 08:2023/BTNMT



Biểu đồ 2.7: Biểu đồ so sánh hàm lượng TSS với QCVN 08:2023/BTNMT

2.3. Xác định chỉ số rủi ro sinh thái RI theo Lars Hakanson (1980).

2.3.1. Tổng quan về chỉ số rủi ro sinh thái RI (Lars Hakanson - 1980)

- Chỉ số rủi ro sinh thái tiềm năng (RI) được đề xuất bởi Lars Hakanson (1980) là một công cụ đánh giá định lượng nhằm phản ánh mức độ nguy cơ tiềm tàng của các kim loại nặng đối với hệ sinh thái thủy sinh, thông qua mức độ ô nhiễm và độc tính của từng kim loại trong trầm tích.
- Phương pháp này cho phép đánh giá tổng hợp các yếu tố: nồng độ thực tế, giá trị nền địa chất, và độc tính sinh thái của từng kim loại. RI là một chỉ số tổng hợp có thể sử dụng để xác định mức độ nguy hiểm sinh thái tiềm năng tại khu vực nghiên cứu.
- Một số nghiên cứu ở Việt Nam và trên thế giới sử dụng chỉ số rủi ro sinh thái (RI). [15], [16]

➤ Các công thức tính [17]

- Hệ số nhiễm bẩn: (Contamination factor - C_f^i)

$$C_f^i = \frac{C_i}{C_n^i}$$

Trong đó:

C_i : nồng độ kim loại i đo được trong mẫu trầm tích

C_n^i : Giá trị nền (background) hoặc nồng độ tham chiếu của kim loại i trong trầm tích chưa bị ảnh hưởng.

- Chỉ số rủi ro sinh thái riêng lẻ (Ecological Risk Factor - E_r^i)

$$E_r^i = T_r^i \times C_f^i = T_r^i \times \frac{C_i}{C_n^i}$$

Trong đó:

T_r^i : Hệ số độc tính của kim loại i – phản ánh độ nguy hại sinh học của kim loại
- **Tổng chỉ số rủi ro sinh thái (RI):**

$$RI = \sum E_r^i = \sum (T_r^i \times \frac{C_i}{C_n^i})$$

Bảng 2.5. Giá trị nền (C_n^i) và hệ số độc tính (T_r^i) của 7 kim loại nặng [17]

| Kim loại nặng | Ký hiệu | Giá trị nền(C_n^i) | Hệ số độc tính(T_r^i) |
|---------------|---------|------------------------|---------------------------|
| Cadimi | Cd | 1,0 | 30 |
| Chì | Pb | 70 | 5 |
| Kẽm | Zn | 175 | 1 |
| Đồng | Cu | 50 | 5 |
| Crôm | Cr | 90 | 2 |
| Thủy ngân | Hg | 0,25 | 40 |
| Asen | As | 15 | 10 |

2.3.2. Xác định chỉ số rủi ro sinh thái (RI) tại khu vực nghiên cứu (sông Bồng Miêu, tỉnh Quảng Nam)

- Tại khu vực sông Bồng Miêu, tỉnh Quảng Nam – nơi chịu ảnh hưởng mạnh mẽ từ hoạt động khai thác vàng, các mẫu trầm tích mặt được thu thập và phân tích để xác định nồng độ các kim loại nặng. Từ đó, áp dụng công thức trên để tính toán chỉ số rủi ro sinh thái RI.

Bảng 2.6. Nồng độ kim loại nặng trong trầm tích và chỉ số rủi ro sinh thái RI (10 vị trí được thực hiện lấy mẫu và phân tích vào 5/2025)

| Vị trí mẫu | Pb | As | Hg | Cd | Cr | Cu | Zn | RI |
|------------|---------|-------|-------|-------|------|------|-------|--------------|
| | (mg/kg) | | | | | | | |
| S1 | 407,2 | 232,8 | 0,064 | 1,187 | 21,7 | 24,8 | 87,0 | 233,6 |
| S2 | 378,3 | 277,3 | 0,248 | 1,161 | 23,6 | 28,7 | 85,9 | 290,3 |
| S3 | 411,6 | 299,6 | 0,033 | 1,524 | 23,0 | 30,6 | 106,9 | 284,3 |
| S4 | 709,9 | 272,4 | 0,049 | 2,713 | 28,9 | 46,6 | 134,2 | 327,6 |
| S5 | 463,3 | 434,8 | 1,431 | 2,614 | 23,9 | 83,2 | 184,2 | 640,2 |
| S6 | 546,6 | 195,6 | 0,877 | 1,882 | 19,3 | 49,1 | 146,5 | 372,4 |

| Vị trí mẫu | Pb | As | Hg | Cd | Cr | Cu | Zn | RI |
|-------------------------------------|-------------|-------------|------------|------------|-----------|------------|------------|--------------|
| | (mg/kg) | | | | | | | |
| S7 | 561,9 | 182,3 | 0,123 | 1,978 | 18,6 | 28,3 | 122,5 | 244,6 |
| S8 | 218,4 | 97,4 | 0,068 | 0,944 | 17,1 | 28,3 | 61,7 | 123,3 |
| S9 | 511,8 | 138,1 | 0,148 | 2,881 | 29,9 | 33,2 | 116,7 | 243,4 |
| S10 | 491,6 | 152,1 | 0,153 | 1,867 | 21,4 | 26,3 | 97,9 | 220,7 |
| QCVN 43:2017/BTNMT* [18] | 91,3 | 17,0 | 0,5 | 3,5 | 90 | 197 | 315 | |

Ghi chú:

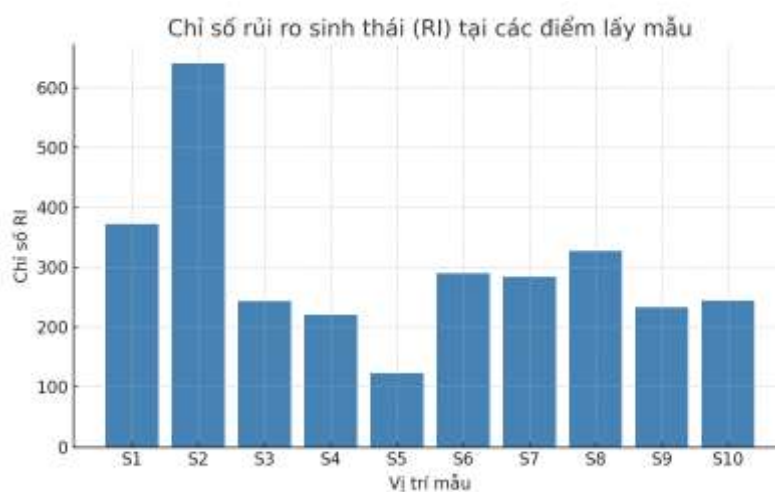
“*”: Áp dụng đối với trầm tích nước ngọt

RI < 150: Rủi ro sinh thái thấp cho hồ/ lưu vực

150 < RI < 300 : Rủi ro sinh thái trung bình cho hồ/ lưu vực

300 < RI < 600: Rủi ro sinh thái đáng kể đối với hồ/ lưu vực

RI >= 600: Rủi ro sinh thái rất cao đối với hồ/ lưu vực



Biểu đồ 2.8: Chỉ số rủi ro sinh thái (RI) của từng vị trí lấy mẫu (đợt 5/2025)

Nhận xét:

a. Đối với Asen

Kết quả phân tích cho thấy nồng độ As trong trầm tích dao động từ 97,4 đến 434,8 mg/kg, với giá trị trung bình là 224,2 mg/kg. Nồng độ cao nhất được ghi nhận tại vị trí S5, trong khi giá trị thấp nhất tại vị trí S8. So sánh với giới hạn cho phép theo Quy chuẩn kỹ thuật quốc gia về chất lượng trầm tích (QCVN 43:2017/BTNMT), với giá trị giới hạn của As là 17,0 mg/kg, cho thấy hàm lượng As trong khu vực nghiên cứu đã vượt từ 5,7

đến 25,5 lần so với tiêu chuẩn, trung bình vượt 13,5 lần.

b. Đối với Chì

Nồng độ Pb trong các mẫu trầm tích dao động từ 218,4 đến 709,9 mg/kg, với giá trị trung bình đạt 470,1 mg/kg. Vị trí có nồng độ Pb cao nhất là S4, khu vực tiếp nhận nguồn thải trực tiếp từ nhà máy khai thác vàng; trong khi đó, vị trí S8 ghi nhận giá trị thấp nhất. So với giới hạn cho phép theo QCVN 43:2017/BTNMT (91,3 mg/kg), nồng độ Pb tại tất cả các vị trí lấy mẫu đều vượt quy chuẩn từ 2,4 đến 7,8 lần, với mức trung bình vượt 4,5 lần.

c. Đối với Thủy ngân

Kết quả phân tích cho thấy nồng độ Hg trong trầm tích dao động từ 0,033 đến 1,431 mg/kg, với giá trị trung bình là 0,319 mg/kg. Nồng độ cao nhất được ghi nhận tại vị trí S5, nằm gần khu vực khai thác vàng của Công ty 6666, trong khi giá trị thấp nhất thuộc về vị trí S3. So với giới hạn cho phép theo QCVN 43:2017/BTNMT (0,5 mg/kg), có hai vị trí (S5 và S6) có hàm lượng Hg vượt ngưỡng tiêu chuẩn, với mức vượt từ 1,3 đến 3 lần.

d. Đối với Cadimi

Nồng độ Cd trong trầm tích dao động từ 0,944 đến 2,881 mg/kg, với giá trị trung bình là 1,87 mg/kg. Tại tất cả các vị trí lấy mẫu, hàm lượng Cd đều nằm dưới ngưỡng giới hạn cho phép theo QCVN 43:2017/BTNMT (3,5 mg/kg), cho thấy mức độ ô nhiễm Cd trong khu vực nghiên cứu chưa vượt quá tiêu chuẩn quy định.

e. Đối với Crom

Nồng độ Cr trong các mẫu trầm tích dao động từ 17,1 đến 29,9 mg/kg, với giá trị trung bình đạt 22,74 mg/kg. So với giới hạn cho phép theo QCVN 43:2017/BTNMT (90 mg/kg), hàm lượng Cr tại tất cả các điểm lấy mẫu đều nằm trong ngưỡng an toàn, không vượt quá tiêu chuẩn quy định.

f. Đối với Đồng

Nồng độ Cu trong trầm tích dao động từ 24,8 đến 83,2 mg/kg, với giá trị trung bình là 37,9 mg/kg. So sánh với giới hạn cho phép theo QCVN 43:2017/BTNMT (197 mg/kg), cho thấy tất cả các mẫu tại các vị trí khảo sát trong cả hai đợt đều có nồng độ Cu nằm trong ngưỡng cho phép, không vượt quá tiêu chuẩn quy định.

g. Đối với Kẽm

Nồng độ Zn trong các mẫu trầm tích dao động từ 61,7 đến 184,2 mg/kg, với giá trị trung bình là 114,2 mg/kg. So sánh với giới hạn cho phép theo QCVN

43:2017/BTNMT (315 mg/kg), cho thấy toàn bộ các mẫu đều có hàm lượng Zn nằm trong ngưỡng an toàn, không vượt quá tiêu chuẩn quy định.

h. Chỉ số RI

Mức độ rủi ro sinh thái tổng thể tại các điểm dao động từ 123,3 (S8) đến 640,2 (S5), phản ánh sự khác biệt rõ rệt về mức độ ô nhiễm kim loại nặng tại từng vị trí. Theo phân loại mức độ nguy cơ sinh thái tiềm tàng của Hakanson (1980):

- Rủi ro thấp → Ghi nhận tại S8 (123,3).
- Rủi ro trung bình → Gồm các điểm S1, S2, S3, S7, S9, S10.
- Rủi ro cao đến rất cao → Ghi nhận tại S4 (372,4), S5 (640,2), S6 (327,6).

Vị trí S5 có chỉ số RI cao nhất (640,2), tiệm cận ngưỡng rủi ro sinh thái rất cao, chủ yếu do hàm lượng As và Hg vượt trội so với các điểm khác. Điều này cho thấy khu vực này có thể đang chịu ảnh hưởng mạnh từ nguồn thải kim loại nặng – có khả năng liên quan đến hoạt động khai thác vàng hoặc tích tụ từ dòng chảy mặt.

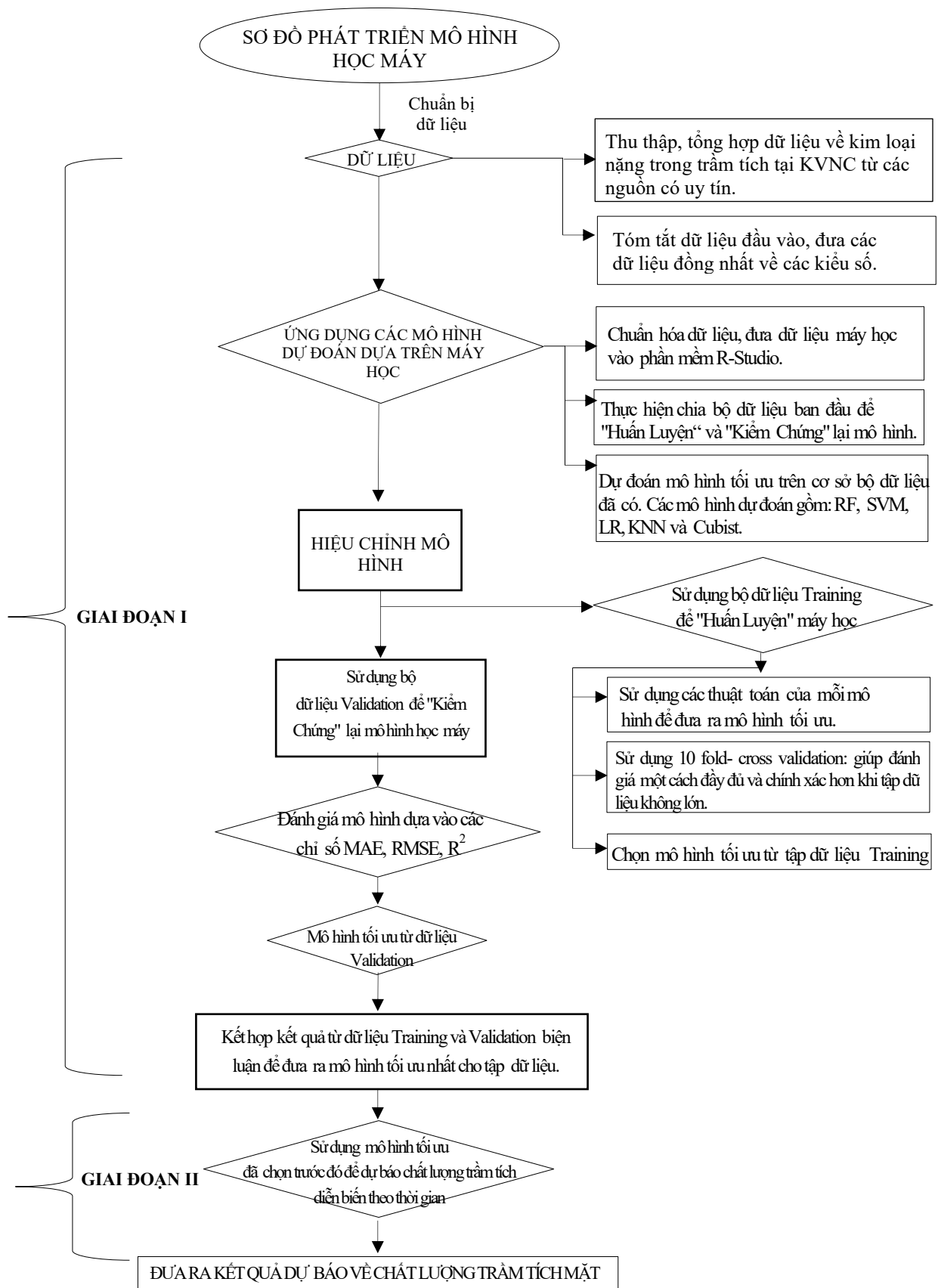
Các kim loại như Cd, As và Hg là những thành phần góp phần lớn nhất vào RI tại đa số các vị trí do có hệ số độc tính cao (T_i), đặc biệt là Cd ($T_i = 30$) và Hg ($T_i = 40$), dù nồng độ không quá lớn nhưng vẫn gây ảnh hưởng mạnh đến chỉ số tổng.

Những điểm như S1, S8, R10 có RI thấp hơn tương đối, cho thấy mức độ ô nhiễm kim loại ở mức trung bình hoặc cục bộ, có thể là khu vực ít bị tác động bởi các nguồn ô nhiễm chính.

- *Nội dung tính toán chi tiết, đầy đủ dữ liệu xin được trình bày tại Phụ lục 1 vì giới hạn số trang của Đồ án tốt nghiệp*

2.4. Xác định mô hình học máy phù hợp để dự báo chất lượng môi trường trầm tích mặt tại khu vực Sông Bồng Miêu, Tỉnh Quảng Nam

2.4.1. Sơ đồ thực hiện mô hình học máy



2.2.2. Thu thập số liệu

Để chạy mô hình học máy đánh giá và dự báo chất lượng trầm tích mặt tại khu vực khảo sát, cần có bộ dữ liệu đáng tin cậy về các thông số nước mặt và trầm tích. Do quy mô và kinh phí làm đồ án tốt nghiệp còn nhiều hạn chế nên số liệu về chất lượng trầm tích mặt để chạy mô hình dự báo được thu thập, tham khảo từ các nguồn khác nhau. Bộ dữ liệu thu thập được gồm 210 dữ liệu trầm tích tại khu vực sông Bồng Miêu, tỉnh Quảng Nam từ năm 2024 – 2025.

Trong đó:

- **Trung tâm Quan trắc và Phân tích môi trường Quảng Nam (TP. Tam Kỳ) cung cấp:** 200 dữ liệu và 10 mẫu được thực hiện lấy mẫu trực tiếp trong thời gian thực hiện nghiên cứu và được phân tích tại trung tâm này.

Trong đồ án tốt nghiệp này, sinh viên sử dụng phần mềm R-Studio để thực hiện chạy mô hình học máy dự báo chất lượng trầm tích mặt tại khu vực nghiên cứu.

2.2.3. Giới thiệu về ngôn ngữ R

Trong phân tích số liệu với khối lượng dữ liệu lớn, các phần mềm thống kê như SPSS, SAS, Stata, S-Plus thường được sử dụng rộng rãi tại các trường đại học và trung tâm nghiên cứu trên thế giới. Tuy nhiên, vì chi phí bản quyền cao, việc sử dụng lâu dài các phần mềm này là trở ngại đối với nhiều trường đại học ở các quốc gia đang phát triển.

Trước thực tế đó, vào năm 1996, hai nhà thống kê học Ross Ihaka và Robert Gentleman thuộc Đại học Auckland (New Zealand) đã giới thiệu một ngôn ngữ lập trình thống kê mới mang tên R. Đây là phần mềm mã nguồn mở và hoàn toàn miễn phí, được cộng đồng các nhà nghiên cứu thống kê toàn cầu chung tay phát triển và cải tiến không ngừng.

Hiện nay, R đã trở thành một trong những công cụ phổ biến và mạnh mẽ nhất trong lĩnh vực phân tích thống kê và khoa học dữ liệu. Đặc biệt, giao diện RStudio được phát triển nhằm hỗ trợ người dùng tương tác với R dễ dàng và trực quan hơn. Về bản chất, R và RStudio không chỉ là phần mềm thống kê, mà còn là một ngôn ngữ lập trình đa năng, có khả năng thực hiện từ các phép tính đơn giản đến những phân tích dữ liệu phức tạp và tùy biến cao.

Với tính linh hoạt, mạnh mẽ và hoàn toàn miễn phí, R/RStudio đang ngày càng được ưa chuộng trong giảng dạy, nghiên cứu và ứng dụng thực tế.

2.2.4. Tải R – Studio và cài đặt

Để sử dụng R-Studio, việc đầu tiên là chúng ta phải cài đặt R- Studio trong máy tính của mình. Để làm việc này, ta cần truy cập vào website có tên là “Comprehensive R Archive Network” (CRAN), sau đây: <https://cran.r-project.org/>.

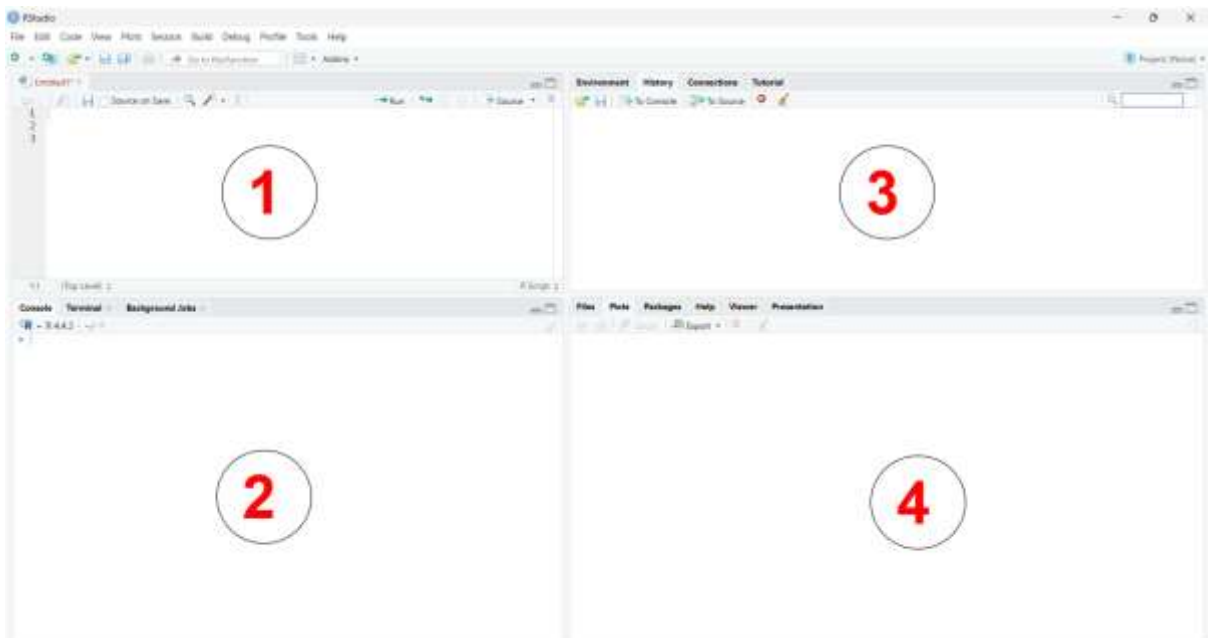
Chọn file tài liệu cần tải về để cài đặt phần mềm, tùy theo phiên bản, nhưng thường có tên bắt đầu bằng mẫu tự R và số phiên bản. Ví dụ phiên bản sinh viên sử dụng được cập nhật lại vào ngày 28/02/2025 và file tài liệu có tên là R- 4.4.3.tar.gz.

Khi tải file tài liệu xuống máy tính, bước tiếp theo là cài đặt vào máy tính. Để làm việc này, chúng ta cần nhấn chuột vào file tài liệu đã tải về và làm theo các hướng dẫn cách cài đặt trên màn hình. Sau khi hoàn tất việc cài đặt, trên màn hình desktop đã xuất hiện:



Hình 2.7. Biểu tượng của phần mềm R trên desktop

Sau khi cài đặt hoàn tất, nhấn chuột vào biểu tượng ở trên **Hình 2.7**, chúng ta sẽ có một cửa sổ làm việc trên R-Studio như sau:



Hình 2.8. Màn hình cửa sổ của R- Studio

Màn hình cửa sổ làm việc của R-Studio gồm có 4 khung:

Bên trái phía trên **(1)** là khung “nguồn”, khung này dùng để soạn các câu lệnh để xử lý dữ liệu và thực hiện mô hình;

Bên trái phía dưới **(2)** là khung “tương tác”, khung này thể hiện các câu lệnh và kết quả thực hiện câu lệnh. Khung này giống với cửa sổ làm việc của R khi không sử dụng R-Studio.

Bên phải phía trên **(3)** là khung "đối tượng" chứa ba thẻ, thẻ Environment dùng để thể hiện thông tin về các đối tượng sử dụng trong khi thực hiện các câu lệnh, thẻ History dùng để thể hiện tất cả các câu lệnh đã soạn kể từ một thời điểm nào đấy, thẻ Connection thể hiện các kết nối.

Bên phải phía dưới **(4)** là khung "đa chức năng" chứa 5 thẻ:

Files: chứa các thư mục và tập tin có liên quan;

Plots: thể hiện các biểu đồ là kết quả của những lệnh vẽ;

Packages: thể hiện các phụ kiện;

Help: thể hiện các trợ giúp;

Viewer: dùng để xem.

Bây giờ ta có thể bắt đầu thực hiện các mô hình học máy trên R-Studio.

2.2.5. Các bước chạy mô hình học máy

Trước khi tiến hành thực hiện các bước trong mô hình, cần làm sạch dữ liệu bằng cách xóa bỏ các câu lệnh và kết quả thực hiện trước đó (khi sử dụng lần 2 trở đi) bằng lệnh `rm()` (xóa bỏ các câu lệnh, kết quả đã thực hiện trước đó), lệnh `graphics()` (xóa bỏ các biểu đồ) nhằm đảm bảo không bị nhầm lẫn kết quả khi thực hiện lại mô hình từ dữ liệu khác.

Bước 1: Làm sạch bộ dữ liệu cũ

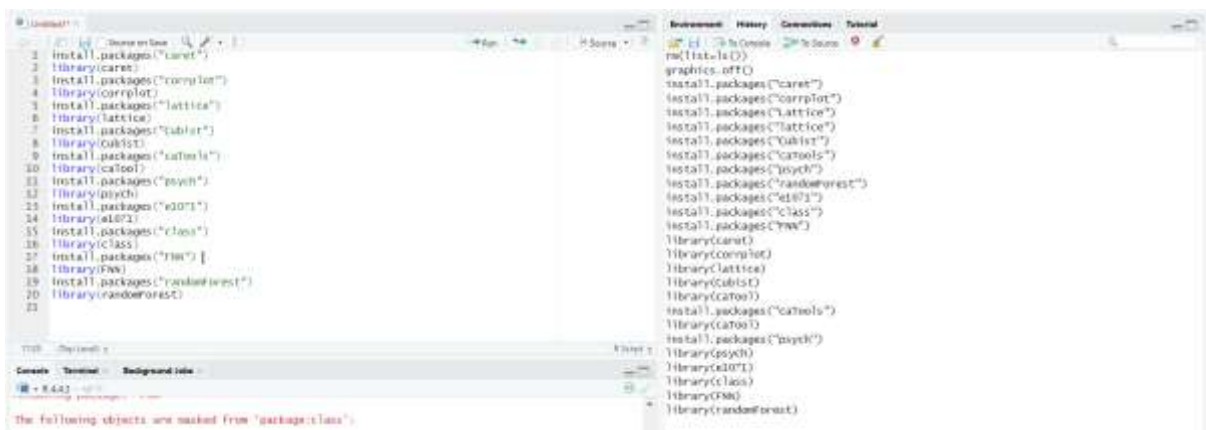
```
# lam sach, xoa bo du lieu cu
rm(list=ls())
graphics.off()
```

Để thực hiện được việc chạy mô hình, cần tải các gói dữ liệu và thư viện chứa các thuật toán trong các mô hình.

Bước 2: Tải và Nạp các gói dữ liệu

```
# Tai va nap cac goi du lieu
-> Tải: install.packages("")
-> Nạp: library()
install.packages("caret")
```

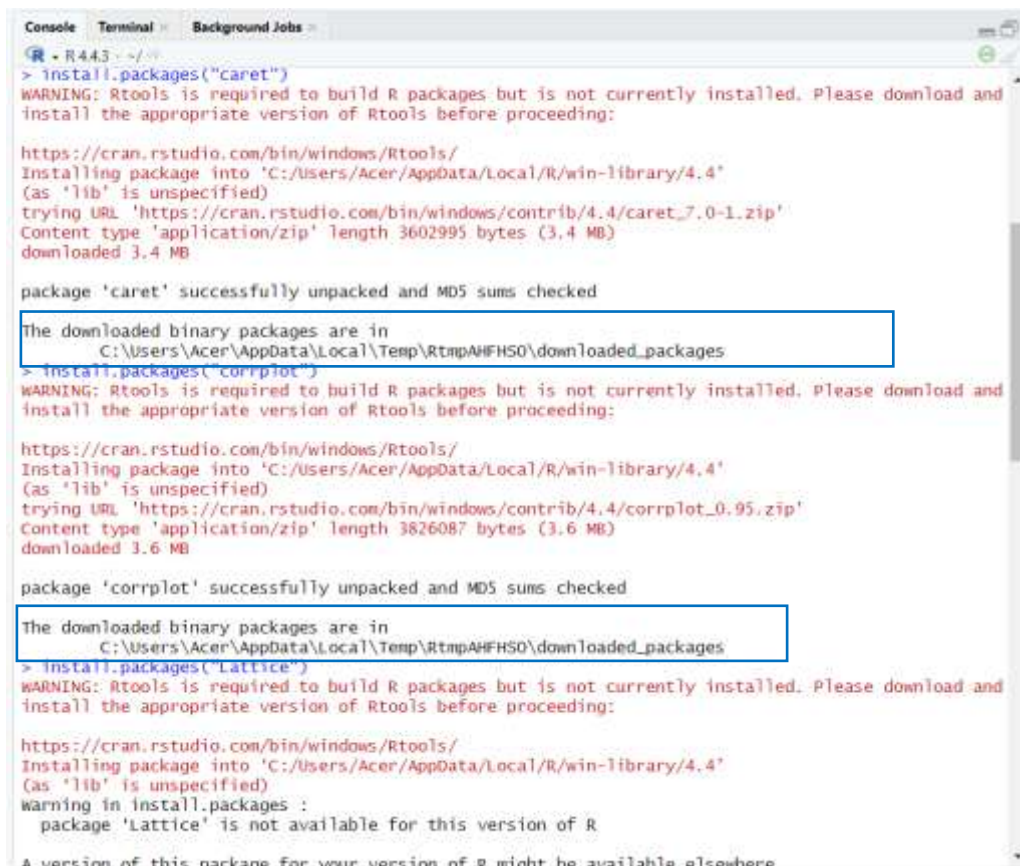
```
library(caret)
install.packages("corrplot")
library(corrplot)
install.packages("lattice")
library(lattice)
install.packages("Cubist")
library(Cubist)
install.packages("caTools")
library(caTool)
install.packages("psych")
library(psych)
install.packages("e1071")
library(e1071)
install.packages("class")
library(class)
install.packages("FNN")
(Fast Neighbor, hỗ trợ nhanh, KNN hồi quy)
library(FNN)
install.packages("randomForest")
library(randomForest)
```



```
1 install.packages("caret")
2 library(caret)
3 install.packages("corrplot")
4 library(corrplot)
5 install.packages("lattice")
6 library(lattice)
7 install.packages("Cubist")
8 library(Cubist)
9 install.packages("caTools")
10 library(caTool)
11 install.packages("psych")
12 library(psych)
13 install.packages("e1071")
14 library(e1071)
15 install.packages("class")
16 library(class)
17 install.packages("FNN")
18 library(FNN)
19 install.packages("randomForest")
20 library(randomForest)
21
```

The following objects are masked from 'package:class':

```
rm(list=ls())
graphics.off()
install.packages("caret")
install.packages("corrplot")
install.packages("lattice")
install.packages("lattice")
install.packages("Cubist")
install.packages("caTools")
install.packages("psych")
install.packages("randomForest")
install.packages("e1071")
install.packages("class")
library("caret")
library(corrplot)
library(lattice)
library(Cubist)
library(caTools)
install.packages("caTools")
library(caTool)
install.packages("psych")
library(psych)
library(e1071)
library(class)
library(FNN)
library(randomForest)
```



```
Console Terminal Background Jobs
R 4.4.3 ~/>
> install.packages("caret")
WARNING: Rtools is required to build R packages but is not currently installed. Please download and
install the appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/Acer/AppData/Local/R/win-library/4.4'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.4/caret_7.0-1.zip'
Content type 'application/zip' length 3602995 bytes (3.4 MB)
downloaded 3.4 MB

package 'caret' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:/Users/Acer/AppData/Local/Temp/RtmpAHFHS0/downloaded_packages
> install.packages("corrplot")
WARNING: Rtools is required to build R packages but is not currently installed. Please download and
install the appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/Acer/AppData/Local/R/win-library/4.4'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.4/corrplot_0.95.zip'
Content type 'application/zip' length 3826087 bytes (3.6 MB)
downloaded 3.6 MB

package 'corrplot' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:/Users/Acer/AppData/Local/Temp/RtmpAHFHS0/downloaded_packages
> install.packages("Lattice")
WARNING: Rtools is required to build R packages but is not currently installed. Please download and
install the appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/Acer/AppData/Local/R/win-library/4.4'
(as 'lib' is unspecified)
warning in install.packages :
  package 'Lattice' is not available for this version of R
A version of this package for your version of R might be available elsewhere
```

Hình 2.9. Tải và nạp thành công các gói dữ liệu (Packages)

Chọn đường dẫn và đọc dữ liệu vào học máy: Sử dụng lệnh `setwd()`: đường dẫn đến nơi mà R-studio lấy dữ liệu vào. Sau đó đưa file dữ liệu "heavymetalsDATA.csv" vào đường dẫn đó để phần mềm đọc dữ liệu.

```
getwd()
heavymetals <- read.csv("heavymetalsDATA.csv")
```



```
Console Terminal Background Jobs
R 4.4.3 ~/>
> getwd()
[1] "C:/Users/Acer/Documents"
> data <- read.csv("heavymetalsDATA.csv")
> heavymetals <- read.csv("heavymetalsDATA.csv")
>
```

Hình 2.10. Thể hiện đường dẫn và đọc dữ liệu kim loại nặng vào học máy

```
# Kiem tra bo du lieu
Heavymetals
```

```
str(heavymetals)
> str(heavymetals)
'data.frame': 210 obs. of 8 variables:
 $ Pb: num 547 463 512 492 218 ...
 $ As: num 195.6 434.8 138.1 152.1 97.4 ...
 $ Hg: num 0.877 1.431 0.148 0.153 0.068 ...
 $ Cd: num 1.882 2.614 2.881 1.867 0.944 ...
 $ Cr: num 19.3 23.9 29.9 21.4 17.1 23.6 23 28.9 21.7 18.6 ...
 $ Cu: num 49.1 83.2 33.2 26.3 28.3 28.7 30.6 46.6 24.8 28.3 ...
 $ Zn: num 146.5 184.2 116.7 97.9 61.7 ...
 $ RI: num 372 640 243 221 123 ...
```

Hình 2.11. Kết quả sau khi kiểm tra bộ dữ liệu

Từ hình ảnh, ta thấy dữ liệu gồm 210 dòng, có 8 biến: Pb, As, Hg, Cd, Cr, Cu, Zn và RI. Nhìn vào phần kết quả xuất ra từ lệnh `Str()` từ **Hình 2.11**, ta thấy các biến đã đồng nhất về 1 kiểu dữ liệu. Nếu có biến chưa cùng dữ liệu về dạng “numeric” thì thực hiện chuyển đổi dạng dữ liệu như sau:

Ví dụ: Nếu biến Pb ở dạng “Character” chưa cùng dạng dữ liệu thì dùng cú pháp:

```
heavymetals$Pb <- as.numeric(as.character(heavymetals$Pb))
```

Tóm tắt bộ dữ liệu đầu vào bằng lệnh `Summary()`:

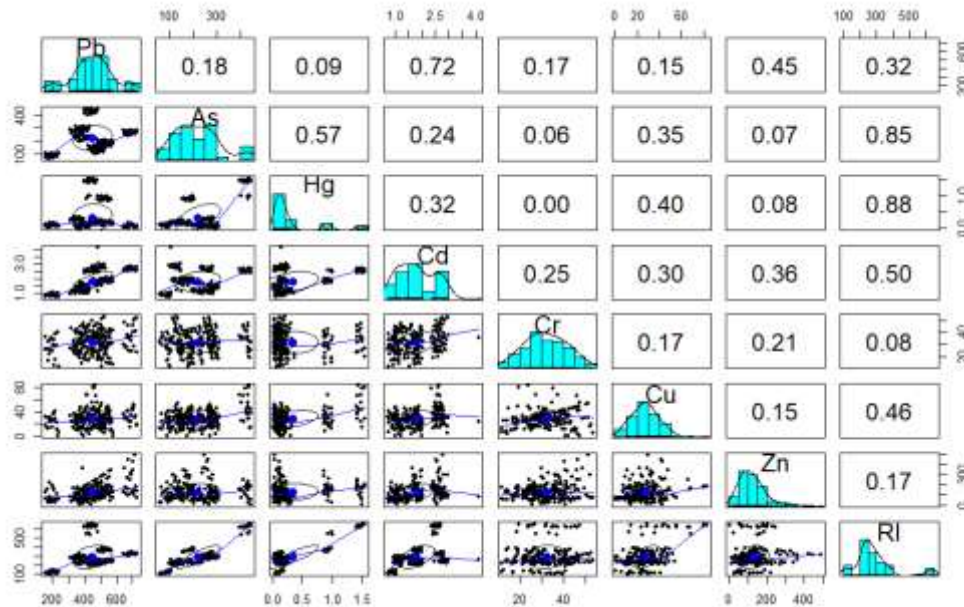
```
# Tóm tắt dữ liệu
summary(Heavymetals)
> summary(heavymetals)
      Pb           As           Hg           Cd
Min.   :153.8   Min.   : 59.1   Min.   :0.0100   Min.   :0.711
1st Qu.:376.9   1st Qu.:148.0   1st Qu.:0.0935   1st Qu.:1.177
Median :453.6   Median :203.4   Median :0.1460   Median :1.757
Mean   :449.0   Mean   :220.9   Mean   :0.3361   Mean   :1.798
3rd Qu.:518.8   3rd Qu.:274.4   3rd Qu.:0.2655   3rd Qu.:2.487
Max.   :728.6   Max.   :444.6   Max.   :1.5490   Max.   :4.126
      Cr           Cu           Zn           RI
Min.   :11.80   Min.   : 0.50   Min.   : 2.40   Min.   :101.2
1st Qu.:25.40   1st Qu.:18.57   1st Qu.: 73.25   1st Qu.:227.1
Median :31.65   Median :27.35   Median :113.05   Median :263.0
Mean   :31.95   Mean   :28.15   Mean   :129.36   Mean   :291.3
3rd Qu.:38.98   3rd Qu.:34.80   3rd Qu.:162.55   3rd Qu.:320.3
Max.   :53.40   Max.   :83.20   Max.   :491.20   Max.   :657.6
```

Hình 2.12. Kết quả sau khi tóm tắt dữ liệu đầu vào

Nhìn vào bảng kết quả được trình bày ở **Hình 2.12** ta biết được kết quả giá trị lớn nhất, giá trị nhỏ nhất, giá trị trung vị, trung bình, 25% dữ liệu và 75% dữ liệu của từng biến số.

Sau khi tiến hành xử lý, hiệu chỉnh dữ liệu và xác định các thuộc tính của từng biến đầu vào, nhằm hiểu rõ hơn về đặc điểm cũng như mối quan hệ giữa các biến, biểu đồ tương quan được xây dựng thông qua hàm `pairs.panels()` để trực quan hóa ma trận tương quan giữa các biến.

```
# Vẽ biểu đồ tương quan giữa các biến  
pairs.panels(heavymetals)
```



Hình 2.13. Biểu đồ tương quan giữa các biến trong bộ dữ liệu

Biểu đồ tương quan giữa các kim loại nặng và chỉ số rủi ro sinh thái (RI) cho thấy một số yếu tố như **Hg**, **As**, và **Cd** có hệ số tương quan tương đối cao với chỉ số RI ($r > 0,6$), trong đó đặc biệt **Hg** có tương quan rất mạnh ($r = 0,88$). Tuy nhiên, các kim loại khác như **Cr**, **Zn** và **Pb** lại có tương quan thấp hơn, phản ánh đặc tính không tuyến tính hoặc phức hợp trong cách các kim loại ảnh hưởng đến tổng chỉ số RI.

```
# Tương quan ma trận giữa các biến
```

```
cor(heavymetals[,c("Pb", "As", "Hg", "Cd", "Cr", "Cu", "Zn", "RI")])
```

```
> cor(heavymetals[,c("Pb", "As", "Hg", "Cd", "Cr", "Cu", "Zn", "RI")])
      Pb      As      Hg      Cd      Cr      Cu      Zn
Pb 1.00000000 0.18021909 0.08519125 0.7161734 0.17181666 0.1510865 0.45056245
As 0.18021909 1.00000000 0.56973001 0.2410895 0.05726176 0.3498412 0.06536216
Hg 0.08519125 0.56973001 1.00000000 0.3182833 -0.00266994 0.3999069 0.08110828
Cd 0.71617340 0.24108946 0.31828326 1.0000000 0.25175575 0.3019158 0.35624331
Cr 0.17181666 0.05726176 -0.00266994 0.2517557 1.00000000 0.1714420 0.21455480
Cu 0.15108645 0.34984121 0.39990687 0.3019158 0.17144205 1.0000000 0.15497805
Zn 0.45056245 0.06536216 0.08110828 0.3562433 0.21455480 0.1549780 1.00000000
RI 0.32255267 0.85378940 0.87791806 0.5036610 0.08316543 0.4572628 0.16977669
```

Hình 2.14. Mối tương quan ma trận giữa các biến

Sinh viên tiến hành chia bộ dữ liệu ban đầu sau khi hiệu chỉnh thành 2 phần gồm: Training và Validation. Sử dụng lệnh `Split()` để thực hiện chia dữ liệu. Vì đây là phép toán chia một cách ngẫu nhiên nên dùng lệnh `set.seed()` để kết quả không đổi sau khi thực hiện lại mô hình lần 2.

```
# Thực hiện chia bộ dữ liệu
set.seed(22)
heavymetals$ID <- 1:210
split <- sample.split(heavymetals$ID, SplitRatio = 0.70)
split
table(split)
```

Dòng đầu tiên, vì đây là phân chia một cách ngẫu nhiên nên ta cần mặc định cho 1 số bất kỳ để kết quả không đổi khi sinh viên chạy lại lần 2.

Thêm biến ID để dễ dàng chia bộ dữ liệu ban đầu, tiến hành chia bộ dữ liệu bằng lệnh `Split()` để chia bộ dữ liệu dựa vào biến ID, chia tỷ lệ Training và Validation từ dữ liệu ban đầu theo tỷ lệ là (70:30)

```
> split <- sample.split(heavymetals$ID, SplitRatio = 0.70)
> split
 [1] TRUE TRUE FALSE TRUE FALSE FALSE TRUE FALSE TRUE TRUE FALSE TRUE TRUE FALSE
[15] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE FALSE TRUE TRUE FALSE
[29] TRUE TRUE TRUE TRUE TRUE FALSE TRUE FALSE TRUE FALSE TRUE TRUE TRUE TRUE TRUE
[43] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE FALSE TRUE TRUE TRUE TRUE
[57] FALSE TRUE TRUE TRUE FALSE TRUE TRUE TRUE TRUE FALSE FALSE TRUE TRUE TRUE TRUE
[71] TRUE FALSE TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE
[85] FALSE FALSE FALSE TRUE FALSE TRUE FALSE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE
[99] FALSE TRUE TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE TRUE
[113] FALSE FALSE TRUE FALSE FALSE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE FALSE
[127] TRUE FALSE TRUE TRUE FALSE FALSE TRUE FALSE FALSE TRUE TRUE TRUE TRUE TRUE FALSE
[141] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE FALSE TRUE FALSE
[155] TRUE FALSE FALSE TRUE FALSE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE FALSE FALSE
[169] FALSE TRUE TRUE FALSE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE
[183] TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE
[197] TRUE TRUE TRUE FALSE TRUE TRUE TRUE FALSE TRUE FALSE FALSE FALSE TRUE TRUE
```

Hình 2.15. Các dòng dữ liệu sau khi chia ngẫu nhiên bộ dữ liệu

Tóm tắt dữ liệu sau khi phân chia:

```
> table(split)
split
FALSE  TRUE
  63   147
```

Gắn tập dữ liệu Training gồm có 147 quan sát (TRUE), Validation gồm 63 quan sát (FALSE)

```
# Gan tap du lieu
```

```
Training <- subset(heavymetals, split==TRUE)
```

```
Validation <- subset(heavymetals, split==FALSE)
```

Sau khi hoàn tất quá trình phân chia dữ liệu, bước tiếp theo là chuẩn bị khai thác dữ liệu nhằm đánh giá hiệu quả của các thuật toán thông qua các chỉ số đánh giá gồm: MAE (Mean Absolute Error - sai số tuyệt đối trung bình), RMSE (Root Mean Squared Error - sai số bình phương trung bình), và R^2 (Độ tin cậy). Quá trình đánh giá được thực hiện bằng phương pháp xác thực chéo 10 lần (10-fold cross-validation) trên tập dữ liệu huấn luyện, với mỗi lần xác thực được lặp lại 3 lần để đảm bảo tính ổn định và độ tin cậy của kết quả.

```
# Chuan bi khai thac du lieu bang xac thuc cheo 10 lan, lap lai 3 lan
```

```
trainControl <- trainControl(method="repeatedcv", number=10, repeats=3)
```

```
metric <- "RMSE"
```

Nghiên cứu đã lựa chọn nhiều thuật toán khác nhau phù hợp với bài toán hồi quy để xây dựng mô hình học máy, bao gồm: Rừng ngẫu nhiên (Random Forest - RF), Máy vector hỗ trợ (Support Vector Machine - SVM), K-Điểm gần nhất (K-Nearest Neighbors - KNN), Hồi quy tuyến tính (Linear Regression - LR) và thuật toán Cubist (mô hình cây lập thể).

Thuật toán của 5 mô hình học máy từ tập dữ liệu Training:

```
#RF
```

```
set.seed(22)
```

```
fit.rf1 <- train(RI ~ ., data = Training, method = "rf", metric = metric,
preProc = c("center", "scale"), trControl = trainControl)
```

```
#SVM
```

```
set.seed(22)
```

```
fit.svm1 <- train(RI ~ ., data = Training, method = "svmRadial", metric = metric,
preProc = c("center", "scale"), trControl = trainControl)
```

```
#KNN
```

```

set.seed(22)
fit.knn1 <- train(RI ~ ., data = Training, method = "knn", metric = metric,
  preProc = c("center", "scale"), trControl = trainControl)
#LR
set.seed(22)
fit.lr1 <- train(RI ~ ., data = Training, method = "lm", metric = metric,
  preProc = c("center", "scale"), trControl = trainControl)
# cubist
set.seed(22)
fit.cubist1 <- train(RI ~ ., data = Training, method = "cubist", metric = metric,
  preProc = c("center", "scale"), trControl = trainControl)

```

Tất cả các thuật toán đều sử dụng các tham số điều chỉnh mặc định. Sinh viên sẽ sử dụng một bảng kết quả đơn giản để biết nhanh về hiệu suất của các mô hình, sử dụng biểu đồ để hiển thị mức độ tin cậy là 95% cho các chỉ số ước tính.

```

# ket qua
results1 <- resamples(list(RF1=fit.rf1, SVM1=fit.svm1, KNN1=fit.knn1,
  LR=fit.lr1, CUBIST1=fit.cubist1 ))
summary(results1)

```

```

> results1 <- resamples(list(RF1=fit.rf1, SVM1=fit.svm1, KNN1=fit.knn1, LR=fit.lr1, CUBIST1=fi
t.cubist1 ))
> summary(results1)

Call:
summary.resamples(object = results1)

Models: RF1, SVM1, KNN1, LR, CUBIST1
Number of resamples: 30

MAE
      Min.      1st Qu.      Median      Mean      3rd Qu.      Max. NA's
RF1    6.13566500  7.68375093  9.13680639  9.39675506 10.4062151 15.82359190  0
SVM1   9.97558653 13.53702167 16.94136014 17.63539967 21.8210345 30.05743974  0
KNN1  15.62933333 19.30538462 21.78790476 21.58477411 24.5005357 27.60285714  0
LR     0.01763346 0.02294039 0.02523258 0.02532284 0.0281545 0.03226899  0
CUBIST1 0.12785950 0.17958578 0.20283624 0.21171516 0.2318625 0.32280986  0

RMSE
      Min.      1st Qu.      Median      Mean      3rd Qu.      Max. NA's
RF1    7.74464752 10.12874109 12.01143416 13.06584062 13.86833488 25.46332587  0
SVM1  14.58390172 18.20050135 25.14268473 26.71886824 31.45498612 46.73480006  0
KNN1  18.44942781 24.49779810 27.56951101 27.75374278 31.82216113 37.69414960  0
LR     0.02163368 0.02663038 0.02993619 0.02978832 0.03262337 0.03795947  0
CUBIST1 0.17524658 0.22356688 0.26357190 0.27712909 0.31465145 0.44798712  0

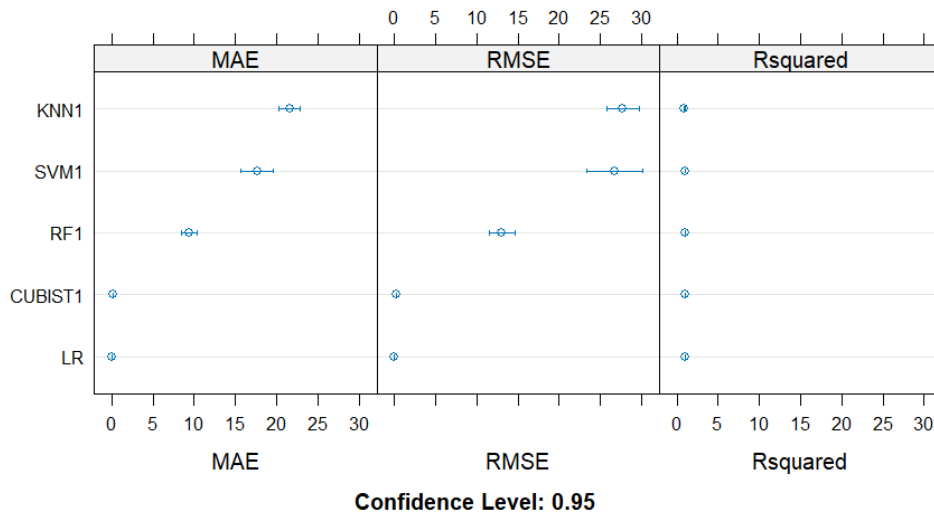
Rsquared
      Min.      1st Qu.      Median      Mean      3rd Qu.      Max. NA's
RF1    0.9320942 0.9833033 0.9916943 0.9867220 0.9949711 0.9984038  0
SVM1   0.8962202 0.9446860 0.9663746 0.9616850 0.9821753 0.9906211  0
KNN1   0.8022730 0.9342538 0.9516140 0.9428882 0.9717883 0.9923671  0
LR     0.9999998 0.9999999 0.9999999 0.9999999 1.0000000 1.0000000  0
CUBIST1 0.9999691 0.9999940 0.9999963 0.9999942 0.9999978 0.9999989  0

```

Hình 2.16. Kết quả sau khi thực hiện 5 mô hình học máy từ tập dữ liệu Training

Vẽ biểu đồ các mô hình với độ tin cậy 95%

`dotplot(results1)`



Biểu đồ 2.9: Kết quả từ dữ liệu Training

Từ kết quả **Hình 2.16** và **Biểu đồ 2.9**, ta thấy mô hình LR và Cubist có giá trị $R^2 \sim 1$ và chỉ số RMSE, MAE là thấp nhất.

Để kiểm chứng lại kết quả đã thu được, sinh viên tiến hành thực hiện kiểm chứng các mô hình trên tập dữ liệu Validation và không sử dụng xác thực chéo 10 lần.

Cú pháp thực hiện:

```
## Validation
pred.rf <- predict(fit.rf1, newdata = Validation)
pred.svm <- predict(fit.svm1, newdata = Validation)
pred.knn <- predict(fit.knn1, newdata = Validation)
pred.lr <- predict(fit.lr1, newdata = Validation)
pred.cubist <- predict(fit.cubist1, newdata = Validation)

# RF
rmse_rf <- rmse(Validation$RI, pred.rf)
mae_rf <- mae(Validation$RI, pred.rf)
r2_rf <- R2(pred.rf, Validation$RI)

# SVM
rmse_svm <- rmse(Validation$RI, pred.svm)
mae_svm <- mae(Validation$RI, pred.svm)
r2_svm <- R2(pred.svm, Validation$RI)
```

```
# KNN
rmse_knn <- rmse(Validation$RI, pred.knn)
mae_knn <- mae(Validation$RI, pred.knn)
r2_knn <- R2(pred.knn, Validation$RI)

# LR
rmse_lr <- rmse(Validation$RI, pred.lr)
mae_lr <- mae(Validation$RI, pred.lr)
r2_lr <- R2(pred.lr, Validation$RI)

# Cubist
rmse_cubist <- rmse(Validation$RI, pred.cubist)
mae_cubist <- mae(Validation$RI, pred.cubist)
r2_cubist <- R2(pred.cubist, Validation$RI)
```

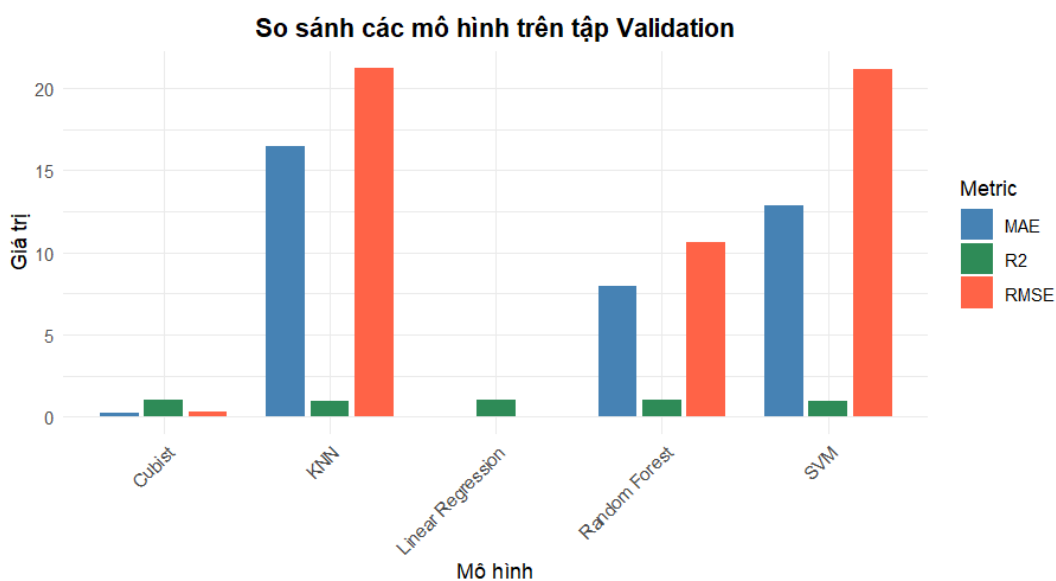
Kết quả và biểu đồ thu được:

```
results <- data.frame(
  Model = c("Random Forest", "SVM", "KNN", "Linear Regression", "Cubist"),
  RMSE = c(rmse_rf, rmse_svm, rmse_knn, rmse_lr, rmse_cubist),
  MAE = c(mae_rf, mae_svm, mae_knn, mae_lr, mae_cubist),
  R2 = c(r2_rf, r2_svm, r2_knn, r2_lr, r2_cubist))
print(results)

> print(results)
```

| | Model | RMSE | MAE | R2 |
|---|-------------------|-------------|-------------|-----------|
| 1 | Random Forest | 10.60325158 | 7.93536624 | 0.9925763 |
| 2 | SVM | 21.14199587 | 12.87324583 | 0.9724774 |
| 3 | KNN | 21.20896052 | 16.48222222 | 0.9714673 |
| 4 | Linear Regression | 0.03162783 | 0.02794912 | 0.9999999 |
| 5 | Cubist | 0.26970565 | 0.20011953 | 0.9999962 |

Hình 2.17. Kết quả chạy mô hình từ tập dữ liệu Validation



Biểu đồ 2.10: Kết quả từ dữ liệu Validation

Tương tự thực hiện chia bộ dữ liệu cho tỷ lệ (71:29), (72:28) đến (85:15) thu được bảng số liệu sau:

Bảng 2.7. Kết quả chạy mô hình từ dữ liệu Training

| Split Ratio | TRAINING | | | | | | | | | | | | | | |
|-------------|----------|-------|-------|------|--------|------|-------|-------|------|--------|----------------|------|------|------|--------|
| | RMSE | | | | | MAE | | | | | R ² | | | | |
| | RF | SVM | KNN | LR | CUBIST | RF | SVM | KNN | LR | CUBIST | RF | SVM | KNN | LR | CUBIST |
| 70:30 | 13,07 | 26,72 | 27,58 | 0,03 | 0,28 | 9,4 | 17,63 | 21,58 | 0,03 | 0,21 | 0,99 | 0,96 | 0,94 | 1,00 | 1,00 |
| 71:29 | 12,8 | 24,35 | 26,53 | 0,03 | 0,27 | 9,3 | 15,5 | 20,3 | 0,02 | 0,21 | 0,98 | 0,96 | 0,94 | 0,99 | 0,99 |
| 72:28 | 12,7 | 25,5 | 25,03 | 0,03 | 0,26 | 9,23 | 15,04 | 19,8 | 0,03 | 0,26 | 0,98 | 0,97 | 0,97 | 1,00 | 0,99 |
| 73:27 | 12,67 | 25,13 | 24,35 | 0,03 | 0,27 | 9,18 | 16,60 | 18,65 | 0,02 | 0,21 | 0,99 | 0,98 | 0,97 | 1,00 | 0,99 |
| 74:26 | 13,01 | 26,25 | 24,43 | 0,03 | 0,27 | 9,39 | 16,86 | 18,83 | 0,02 | 0,21 | 0,99 | 0,98 | 0,97 | 1,00 | 0,99 |
| 75:25 | 12,79 | 28,05 | 24,56 | 0,03 | 0,27 | 9,14 | 17,74 | 19,05 | 0,02 | 0,21 | 0,99 | 0,98 | 0,97 | 1,00 | 0,99 |
| 76:24 | 12,56 | 22,02 | 24,86 | 0,03 | 0,26 | 9,10 | 14,05 | 19,26 | 0,02 | 0,19 | 0,98 | 0,97 | 0,94 | 0,99 | 0,99 |
| 77:23 | 12,71 | 21,64 | 25,42 | 0,03 | 0,26 | 9,11 | 13,71 | 19,34 | 0,02 | 0,19 | 0,98 | 0,97 | 0,94 | 0,99 | 0,99 |
| 78:22 | 12,10 | 20,88 | 24,73 | 0,03 | 0,25 | 9,06 | 13,33 | 19,14 | 0,02 | 0,19 | 0,98 | 0,97 | 0,94 | 0,99 | 0,99 |
| 79:21 | 12,58 | 19,82 | 25,20 | 0,03 | 0,25 | 9,08 | 12,85 | 19,07 | 0,02 | 0,19 | 0,98 | 0,97 | 0,93 | 0,99 | 0,99 |
| 80:20 | 12,40 | 25,28 | 24,91 | 0,02 | 0,25 | 8,99 | 16,45 | 18,77 | 0,02 | 0,18 | 0,99 | 0,96 | 0,95 | 0,99 | 0,99 |
| 85:15 | 11,93 | 21,18 | 25,96 | 0,03 | 0,24 | 8,67 | 13,79 | 19,62 | 0,02 | 0,18 | 0,98 | 0,97 | 0,95 | 0,99 | 0,99 |

Bảng 2.8. Kết quả chạy mô hình từ dữ liệu Validation

| VALIDATION | | | | | | | | | | | | | | | |
|--------------|-------|-------|-------|------|--------|-------|-------|-------|------|--------|----------------|------|------|------|--------|
| Split Ratio | RMSE | | | | | MAE | | | | | R ² | | | | |
| | RF | SVM | KNN | LR | CUBIST | RF | SVM | KNN | LR | CUBIST | RF | SVM | KNN | LR | CUBIST |
| 70.00 | 13,07 | 26,72 | 27,58 | 0,03 | 0,28 | 9,4 | 17,63 | 21,58 | 0,03 | 0,21 | 0,99 | 0,96 | 0,94 | 1,00 | 1,00 |
| 71.00 | 12,41 | 39,73 | 25,44 | 0,03 | 0,24 | 9,62 | 20,86 | 19,78 | 0,02 | 0,18 | 0,99 | 0,94 | 0,97 | 1,00 | 0,99 |
| 72.00 | 12,09 | 41,99 | 24,95 | 0,02 | 0,27 | 9,51 | 22,41 | 19,63 | 0,02 | 0,21 | 0,99 | 0,94 | 0,97 | 1,00 | 0,99 |
| 73.00 | 13,18 | 43,53 | 25,33 | 0,02 | 0,25 | 10,31 | 23,57 | 20,04 | 0,02 | 0,2 | 0,99 | 0,94 | 0,97 | 1,00 | 0,99 |
| 74.00 | 13,68 | 44,93 | 24,70 | 0,02 | 0,25 | 10,43 | 24,34 | 19,42 | 0,02 | 0,19 | 0,99 | 0,93 | 0,97 | 1,00 | 0,99 |
| 75.00 | 12,33 | 47,61 | 23,26 | 0,02 | 0,25 | 9,58 | 25,76 | 19,70 | 0,02 | 0,19 | 0,99 | 0,93 | 0,97 | 1,00 | 0,99 |
| 76.00 | 12,33 | 42,30 | 24,50 | 0,02 | 0,27 | 9,47 | 22,97 | 19,06 | 0,02 | 0,19 | 0,99 | 0,94 | 0,97 | 1,00 | 0,99 |
| 77.00 | 12,78 | 39,36 | 24,92 | 0,02 | 0,27 | 9,76 | 21,26 | 19,61 | 0,02 | 0,19 | 0,99 | 0,96 | 0,97 | 1,00 | 0,99 |
| 78.00 | 13,21 | 38,96 | 23,77 | 0,02 | 0,30 | 9,69 | 21,01 | 18,44 | 0,02 | 0,23 | 0,99 | 0,96 | 0,97 | 1,00 | 0,99 |
| 79.00 | 12,67 | 37,12 | 24,36 | 0,02 | 0,32 | 9,40 | 19,84 | 18,99 | 0,02 | 0,25 | 0,99 | 0,96 | 0,97 | 1,00 | 0,99 |
| 80.00 | 13,07 | 39,76 | 23,24 | 0,02 | 0,30 | 9,46 | 21,73 | 18,06 | 0,02 | 0,22 | 0,99 | 0,95 | 0,97 | 1,00 | 0,99 |
| 85.00 | 10,77 | 29,76 | 23,03 | 0,02 | 0,28 | 8,67 | 14,90 | 17,73 | 0,02 | 0,23 | 0,99 | 0,97 | 0,97 | 1,00 | 0,99 |

Bảng 2.9. So sánh hiệu suất của từng mô hình giữa tập Training và Validation

| Model | RMSE | | MAE | | R ² | | Độ lệch | Nhận xét |
|---------------|--------------|-------------|-------------|-------------|----------------|-------------|-----------------|-------------------|
| | Training | Validation | Training | Validation | Training | Validation | | |
| RF1 | 13,07 | 10,6 | 9,40 | 7,93 | 0,99 | 0,99 | Giảm nhẹ | Mô hình ổn |
| SVM1 | 26,72 | 21,14 | 17,63 | 12,97 | 0,98 | 0,97 | Giảm | Chấp nhận được |
| KNN1 | 27,58 | 21,21 | 21,58 | 16,48 | 0,98 | 0,97 | Giảm | Trung bình |
| LR | 0,03 | 0,03 | 0,03 | 0,03 | 1,00 | 1,00 | Gần bằng 0 | Quá Khớp |
| Cubist | 0,28 | 0,27 | 0,21 | 0,21 | 1,00 | 1,00 | Gần bằng 0 | Quá Khớp |

Nhận xét:

Sau khi kiểm chứng lại các mô hình từ tập dữ liệu Validation thì 2 mô hình LR và Cubist có hiệu suất cao nhất với độ tin cậy $R^2 \sim 1$ và các chỉ số MAE và RMSE rất thấp (tiệm cận bằng 0). Từ đây đặt ra một giả thiết, có thể vì kích cỡ của tập dữ liệu đầu vào quá nhỏ nên mô hình đã học thuộc lòng tập Training hoặc dữ liệu bị rò rỉ và mô hình đã học được nên không còn khả năng tổng quát hóa. Vì vậy mà chúng ta thấy được dấu hiệu rất rõ của “Overfitting” (Quá khớp) đối với 2 mô hình này.

Ngược lại, đối với mô hình RF (Random Forest) cho ra kết quả từ tập dữ liệu Training và Validation đều có hiệu suất ở mức tốt với độ tin cậy R^2 gần bằng 1, chỉ số MAE và RMSE ở mức ổn chỉ xếp sau 2 mô hình LR và Cubist và kết quả sau khi kiểm chứng lại với tập Validation thể hiện sự cân bằng giữa độ chính xác và khả năng tổng quát hóa (thể hiện ở **bảng 2.7**, **bảng 2.8** và **bảng 2.9**)

Bảng 2.10. Xếp hạng các mô hình tối ưu hóa dữ liệu nhất

| Xếp hạng | Mô hình tốt nhất |
|---------------------------|--|
| Random Forest (RF) | Cân bằng giữa độ chính xác và khả năng tổng quát hóa |
| SVM | Hiệu suất không quá tệ, nhưng không tốt bằng RF |
| LR & Cubist | Overfitting nặng – không nên dùng nếu không xử lý lại dữ liệu |

Kết Luận:

Khi hoàn thiện việc xây dựng các mô hình học máy, dữ liệu ban đầu được chia theo tỷ lệ từ (70:30) đến (85:15) cho tập huấn luyện và tập kiểm chứng. Mô hình được lựa chọn là mô hình tối ưu cho việc dự báo chất lượng trầm tích mặt tại khu vực khảo sát sẽ dựa trên các chỉ số đánh giá như sai số tuyệt đối trung bình (MAE), sai số bình phương trung bình (RMSE) và hệ số xác định (R^2).

Trong quá trình phân tích, hai mô hình **Linear Regression (LR)** và **Cubist** cho kết quả dự báo gần như hoàn hảo trên cả hai tập huấn luyện và kiểm chứng, với sai số rất thấp và giá trị R^2 tiệm cận 1. Tuy nhiên, chính hiệu suất quá cao này và gần như không có sự chênh lệch giữa hai tập Training và Validation khiến cho khả năng “**overfitting**” trở nên đáng nghi ngờ – tức là mô hình đã học quá sát với dữ liệu huấn luyện, có thể làm giảm khả năng tổng quát hóa đối với các dữ liệu mới.

Ngược lại, mô hình **Random Forest (RF)** tuy không đạt hiệu suất cao nhất theo các chỉ số đánh giá, nhưng lại thể hiện sự ổn định và cân bằng giữa hai tập dữ liệu. Bên cạnh đó, RF là thuật toán vốn được biết đến với khả năng chống **overfitting** tốt, nhờ cơ chế lấy mẫu ngẫu nhiên (bootstrap) và tổng hợp nhiều cây quyết định (bagging).

Vì vậy, mặc dù không phải là mô hình có chỉ số tốt nhất, Random Forest vẫn được lựa chọn là mô hình phù hợp nhất trong nghiên cứu này – như một giải pháp an toàn, đáng tin cậy và có khả năng ứng dụng thực tế cao, đặc biệt trong điều kiện dữ liệu còn hạn chế.

CHƯƠNG 3: DỰ BÁO DIỄN BIẾN CHẤT LƯỢNG MÔI TRƯỜNG THEO THỜI GIAN VÀ ĐỀ XUẤT CÁC GIẢI PHÁP QUẢN LÝ - KỸ THUẬT

3.1. Dự báo hiện trạng môi trường diễn biến theo thời gian dựa trên mô hình tối ưu đã lựa chọn.

3.1.1. Dự báo chỉ số rủi ro sinh thái tại 10 vị trí nghiên cứu đến tháng 1/2026

Các bước thực hiện dự báo:

Bước 1: Cài đặt và tải các thư viện như ở **Chương 2**

Bước 2: Đọc và xử lý dữ liệu

```
# Các cột: Time, Location, As, Zn, Pb, Cr, Cu, Hg, Cd, RI
```

```
# Chuyển cột Time từ dạng 'YYYY-M' về số (ví dụ: "2023-9" -> 2023.9)
```

```
heavymetals$Time <- as.numeric(gsub("-", ".", heavymetals$Time))
```

Bước 3: Huấn luyện mô hình Random Forest để dự báo RI

```
set.seed(123)
```

```
trainControl <- trainControl(method = "repeatedcv", number = 10, repeats = 3)
```

```
fit.rf <- train(
```

```
RI ~ Time + Location + As + Zn + Pb + Cr + Cu + Hg + Cd,
```

```
data = heavymetals,
```

```
method = "rf",
```

```
trControl = trainControl,
```

```
preProc = c("center", "scale")
```

```
)
```

Bước 4: Tạo dữ liệu dự báo cho tháng 1/2026

```
future_dates <- c("2026-1")
```

```
future_data <- expand.grid(
```

```
Time = future_dates,
```

```
Location = unique(heavymetals$Location)
```

```
)
```

```
# Chuyển cột Time từ "2026-1" thành số 2026.1 (nếu dữ liệu đầu vào chưa chuyển)
```

```
future_data$Time <- as.numeric(gsub("-", ".", future_data$Time))
```

```
metal_means <- heavymetals %>%
```

```
group_by(Location) %>%
```

```
summarise(across(c(As, Zn, Pb, Cr, Cu, Hg, Cd), mean, na.rm = TRUE))
```

```
future_data <- merge(future_data, metal_means, by = "Location")
```

#Dự báo

```
future_data$RI_predicted <- predict(fit.rf, newdata = future_data)
```

Bước 5: Xuất kết quả

```
print(future_data[, c("Time", "Location", "RI_predicted")])
# Tùy chọn: xuất ra file CSV
write.csv(future_data, "RI_du_bao_2026.csv", row.names = FALSE)
# Tùy chọn: vẽ biểu đồ (nếu cần)
ggplot(future_data, aes(x = Time, y = RI_predicted, color = Location, group =
  Location)) +
  geom_line(size = 1.2) +
  geom_point() +
  theme_minimal() +
  labs(title = "Dự báo chỉ số RI theo thời gian (1/2026)",
  x = "Thời gian (năm-tháng)", y = "Chỉ số RI dự báo") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
> print(future_data[, c("Time", "Location", "RI_predicted")])
  Time Location RI_predicted
1  2026.1     S1      226.4023
2  2026.1    S10      215.1000
3  2026.1     S2      282.7633
4  2026.1     S3      279.5357
5  2026.1     S4      319.0563
6  2026.1     S5      625.6487
7  2026.1     S6      362.0046
8  2026.1     S7      240.2300
9  2026.1     S8      115.7690
10 2026.1     S9      241.3365
```

Hình 3.1. Kết quả dự báo chỉ số rủi ro sinh thái (RI) cho tháng 1/2026

3.1.2. Dự báo nồng độ của từng kim loại nặng tại 10 vị trí nghiên cứu đến tháng 1/2026

Bước 1: Huấn luyện mô hình cho từng kim loại

```
# Thiết lập xác thực chéo
train_control <- trainControl(method = "repeatedcv", number = 10, repeats = 3)
# Huấn luyện mô hình cho kim loại As
fit.As <- train(
  As ~ Time + Location,
  data = heavymetals,
```

```
        method = "rf",
    preProcess = c("center", "scale"),
    trControl = train_control,
    metric = "RMSE"
    )
# Huấn luyện mô hình cho kim loại Pb
fit.Pb <- train(
  Pb ~ Time + Location,
  data = heavymetals,
  method = "rf",
  preProcess = c("center", "scale"),
  trControl = train_control,
  metric = "RMSE"
  )
# Huấn luyện mô hình cho kim loại Hg
fit.Hg <- train(
  Hg ~ Time + Location,
  data = heavymetals,
  method = "rf",
  preProcess = c("center", "scale"),
  trControl = train_control,
  metric = "RMSE"
  )
# Huấn luyện mô hình cho kim loại Cu
fit.Cu <- train(
  Cu ~ Time + Location,
  data = heavymetals,
  method = "rf",
  preProcess = c("center", "scale"),
  trControl = train_control,
  metric = "RMSE"
  )
# Huấn luyện mô hình cho kim loại Cr
```

```
fit.Cr <- train(
  Cr ~ Time + Location,
  data = heavymetals,
  method = "rf",
  preProcess = c("center", "scale"),
  trControl = train_control,
  metric = "RMSE"
)
# Huấn luyện mô hình cho kim loại Cd
fit.Cd <- train(
  Cd ~ Time + Location,
  data = heavymetals,
  method = "rf",
  preProcess = c("center", "scale"),
  trControl = train_control,
  metric = "RMSE"
)
# Huấn luyện mô hình cho kim loại Zn
fit.Zn <- train(
  Zn ~ Time + Location,
  data = heavymetals,
  method = "rf",
  preProcess = c("center", "scale"),
  trControl = train_control,
  metric = "RMSE"
)
```

Bước 2: Tạo khung dữ liệu mới cho tháng 1/2026 (10 vị trí S1-S10)

```
future_month <- data.frame(
  Time = rep("2026-1", times = 10),
  Location = paste0("S", 1:10)
)
#Chuyển dạng ký tự
library(zoo)
```

```
future_month$Time <- as.numeric(as.yearmon(future_month$Time, "%Y-%m"))
```

Bước 3: Dự báo nồng độ từng kim loại nặng

```
#Dự báo nồng độ của từng kim loại nặng
pred.As <- predict(fit.As, newdata = future_month)
pred.Zn <- predict(fit.Zn, newdata = future_month)
pred.Pb <- predict(fit.Pb, newdata = future_month)
pred.Cr <- predict(fit.Cr, newdata = future_month)
pred.Cu <- predict(fit.Cu, newdata = future_month)
pred.Hg <- predict(fit.Hg, newdata = future_month)
pred.Cd <- predict(fit.Cd, newdata = future_month)

# Gộp các dữ liệu kim loại nặng thành 1 bảng
future_results <- cbind(
  future_month,
  As = pred.As,
  Zn = pred.Zn,
  Pb = pred.Pb,
  Cr = pred.Cr,
  Cu = pred.Cu,
  Hg = pred.Hg,
  Cd = pred.Cd
)
```

Bước 4: Xuất kết quả

```
print(future_results)

write.csv(future_results, "du_bao_kim_loai_tung_vi_tri.csv", row.names = FALSE)
> print(future_results)
```

| | Time | Location | As | Zn | Pb | Cr | Cu | Hg | Cd |
|----|------|----------|-----------|-----------|----------|----------|----------|------------|-----------|
| 1 | 2026 | S1 | 227.01617 | 102.59067 | 410.0322 | 21.68470 | 27.23168 | 0.06907447 | 1.2834231 |
| 2 | 2026 | S2 | 271.17560 | 85.45001 | 363.5953 | 22.28897 | 27.57371 | 0.29427420 | 1.2959852 |
| 3 | 2026 | S3 | 294.50167 | 104.60164 | 399.8656 | 21.94626 | 28.98766 | 0.06212110 | 1.4633172 |
| 4 | 2026 | S4 | 266.05447 | 217.61687 | 694.0941 | 26.08566 | 42.57336 | 0.06588440 | 2.6743979 |
| 5 | 2026 | S5 | 429.77174 | 161.12208 | 447.4077 | 22.41663 | 77.56820 | 1.44344760 | 2.5369593 |
| 6 | 2026 | S6 | 194.07708 | 154.28526 | 531.3542 | 20.37227 | 45.35751 | 0.92483347 | 1.8260730 |
| 7 | 2026 | S7 | 180.52744 | 141.42671 | 546.4471 | 19.69181 | 27.58220 | 0.16236860 | 1.8792795 |
| 8 | 2026 | S8 | 91.92829 | 75.12253 | 202.6763 | 18.36655 | 27.50529 | 0.06950610 | 0.9127865 |
| 9 | 2026 | S9 | 133.81815 | 120.05459 | 495.2550 | 26.90243 | 30.93611 | 0.18859850 | 2.9226447 |
| 10 | 2026 | S10 | 149.19276 | 103.39395 | 475.2088 | 21.42132 | 26.98365 | 0.19317607 | 1.6187493 |

Hình 3.2. Kết quả dự báo nồng độ các kim loại nặng tháng 1/2026

Bảng 3.1. Tổng hợp dữ liệu sau dự báo (1/2026)

| Vị trí mẫu | Thời gian | As | Zn | Pb | Cr | Cu | Hg | Cd | RI (dự báo) | RI (tính toán) |
|---------------------------------|-----------|-------------|------------|-------------|-----------|------------|------------|------------|-------------|----------------|
| | | (mg/kg) | | | | | | | | |
| S1 | 1/2026 | 225,4 | 102,3 | 386,1 | 30,5 | 22,5 | 0,089 | 1,131 | 226,4 | 229,6 |
| S2 | 1/2026 | 269,9 | 77,3 | 357,2 | 32,8 | 24,5 | 0,273 | 1,122 | 282,7 | 286,5 |
| S3 | 1/2026 | 292,2 | 95,6 | 390,5 | 32,3 | 24,9 | 0,058 | 1,459 | 279,5 | 279,6 |
| S4 | 1/2026 | 265,1 | 224,9 | 688,8 | 36,6 | 33,0 | 0,074 | 2,642 | 319,0 | 322,5 |
| S5 | 1/2026 | 427,5 | 142,7 | 442,2 | 33,7 | 46,8 | 1,371 | 2,524 | 625,6 | 618,0 |
| S6 | 1/2026 | 188,3 | 157,9 | 525,5 | 29,3 | 29,3 | 0,902 | 1,814 | 362,0 | 366,3 |
| S7 | 1/2026 | 174,9 | 156,3 | 540,8 | 28,1 | 23,2 | 0,148 | 1,855 | 240,2 | 238,5 |
| S8 | 1/2026 | 90,1 | 91,2 | 197,3 | 28,1 | 22,9 | 0,093 | 0,875 | 115,7 | 118,8 |
| S9 | 1/2026 | 130,8 | 131,8 | 490,7 | 36,9 | 29,7 | 0,173 | 2,847 | 241,3 | 239,9 |
| S10 | 1/2026 | 144,8 | 113,3 | 470,5 | 31,1 | 24,8 | 0,178 | 1,709 | 215,1 | 213,7 |
| QCVN 43:2017/BTNMT* [18] | | 17,0 | 315 | 91,3 | 90 | 197 | 0,5 | 3,5 | | |

Nhận xét:

So với những tháng gần nhất, nồng độ các kim loại nặng được dự báo bằng mô hình Random Forest (RF) có xu hướng giảm nhẹ. Tuy nhiên, tại các vị trí, nồng độ của các kim loại nặng như Pb và As vẫn vượt ngưỡng cho phép theo QCVN 43:2017/BTNMT, và không có sự thay đổi đáng kể. Chỉ số rủi ro sinh thái (RI) dự báo từ mô hình cho thấy các vị trí S5 và S6 tiếp tục là hai điểm có giá trị RI cao nhất. Khi so sánh kết quả dự báo với chỉ số RI được tính theo công thức của Lars Hakanson, độ sai số dao động trong khoảng 0,01–0,03, cho thấy sai số rất nhỏ. Điều này chứng tỏ mô hình RF có độ chính xác cao và hiệu suất dự báo tốt.

3.2. Các giải pháp quản lý - kỹ thuật

Khu vực sông Bông Miêu (tỉnh Quảng Nam) đang chịu ảnh hưởng nghiêm trọng bởi các hoạt động khai thác vàng trái phép diễn ra trong thời gian dài. Các hoạt động này không những gây mất an ninh trật tự địa phương, phá vỡ cấu trúc tự nhiên mà còn để lại hậu quả nghiêm trọng về môi trường, đặc biệt là ô nhiễm kim loại nặng (thủy ngân, chì, asen, kẽm, đồng, Crom, Cadimi...) trong trầm tích, nước và sinh vật thủy sinh.

Để hạn chế và ngăn chặn tình trạng này, cần phải áp dụng đồng thời cả biện pháp quản lý hành chính lẫn biện pháp kỹ thuật một cách quyết liệt và bền vững.



Hình 3.3. Hiện trạng khai thác vàng và môi trường tại sông Bồng Miêu

3.2.1 Biện pháp quản lý

3.2.1.1 Tăng cường giám sát và xử lý hoạt động khai thác trái phép

- **Thành lập lực lượng liên ngành kiểm tra, truy quét thường xuyên:** Thiết lập các tổ công tác liên ngành gồm công an, kiểm lâm, và đại diện chính quyền địa phương để phối hợp kiểm tra định kỳ và đột xuất tại các khu vực thường xuyên xảy ra hoạt động khai thác vàng trái phép. Đặc biệt chú trọng các điểm nóng tại khu vực rừng núi, nơi khó kiểm soát.
- **Ứng dụng công nghệ cao trong giám sát:** Tận dụng máy bay không người lái (drone), ảnh vệ tinh và hệ thống định vị GPS để giám sát diện rộng, phát hiện sớm các dấu hiệu khai thác trái phép. Công nghệ này đặc biệt hữu ích tại những khu vực địa hình phức tạp, nơi lực lượng chức năng khó tiếp cận.

- **Tăng cường chế tài xử phạt nghiêm minh:** Áp dụng nghiêm khắc các hình phạt theo quy định của Luật Bảo vệ Môi trường và Luật Khoáng sản đối với cá nhân hoặc tổ chức vi phạm. Các hình thức xử lý có thể bao gồm xử phạt hành chính, buộc bồi thường thiệt hại môi trường, và xử lý hình sự nếu có đủ căn cứ. [19], [20]

3.2.1.2 Quản lý quy hoạch vùng nhạy cảm

- Khoanh vùng và **cấm mốc rõ ràng các khu vực nghiêm cấm khai thác**, đặc biệt là thượng nguồn sông Bồng Miêu và các vùng có hệ sinh thái nhạy cảm.
- Xây dựng và áp dụng **quy trình đánh giá tác động môi trường nghiêm ngặt** nếu có dự án khai thác hợp pháp, nhằm tránh ảnh hưởng đến nước sông và đời sống người dân.

3.2.1.3 Nâng cao nhận thức cộng đồng

- **Tổ chức các buổi truyền thông cộng đồng**, phổ biến kiến thức về tác hại của kim loại nặng đối với sức khỏe (gây ung thư, tổn thương hệ thần kinh, ảnh hưởng hệ sinh sản...).
- Kết hợp **giáo dục môi trường trong trường học**, tạo ý thức bảo vệ tài nguyên cho thế hệ trẻ.
- **Hỗ trợ chuyển đổi sinh kế** cho người dân từng tham gia khai thác trái phép sang các ngành nghề bền vững như nông nghiệp hữu cơ, du lịch sinh thái, trồng rừng.

3.2.2 Biện pháp kỹ thuật

3.2.2.1 Xử lý nước thải có chứa kim loại nặng

- Xây dựng **hệ thống xử lý nước thải tập trung** tại khu vực mỏ vàng cũ, gồm các bước:
 - **Kết tủa hóa học:** sử dụng NaOH, Ca(OH)₂, hoặc sulfide để kết tủa các ion kim loại như Pb²⁺, Hg²⁺, Cd²⁺.
 - **Hấp phụ:** dùng vật liệu hấp phụ như than hoạt tính, zeolit, biochar từ phế phẩm nông nghiệp để loại bỏ kim loại nặng trong nước. [21]
 - **Lọc màng:** áp dụng công nghệ lọc nano hoặc màng RO cho các nguồn nước cần phục hồi ở cấp độ cao hơn. [22]

3.2.2.2 Quản lý bùn thải và chất thải nguy hại

- Xây dựng **hồ lắng nhiều ngăn**, thu gom bùn thải có chứa cyanua hoặc thủy ngân để xử lý riêng theo quy định về chất thải nguy hại.
- Tuyệt đối **không để bùn thải, nước thải chưa qua xử lý chảy vào sông**, đặc biệt vào mùa mưa khi dòng chảy mạnh và dễ phát tán.

3.2.2.3. Bổ sung hệ thống cấp nước sạch

Tại khu vực sông Bồng Miêu, hiện nay gần như 100% hộ dân vẫn đang sử dụng nguồn nước giếng khoan hoặc giếng đào, trong khi nước sạch từ các công trình cấp nước tập trung vẫn chưa được đưa vào sử dụng hoặc bị bỏ hoang sau đầu tư. Đây là một thực trạng đáng lo ngại, đặc biệt khi các mẫu quan trắc tại khu vực đã ghi nhận sự tồn lưu kim loại nặng như thủy ngân, asen, chì trong môi trường đất và nước — những chất có thể thấm thấu và tích tụ trong nguồn nước ngầm.

Việc không có nguồn nước sạch thay thế khiến người dân phải sử dụng nước tiềm ẩn nguy cơ cao gây bệnh tật mãn tính, như tổn thương gan, thận, hệ thần kinh, hoặc thậm chí là ung thư. Do đó, cần:

- Rà soát lại các công trình cấp nước sạch đã đầu tư tại địa phương, xác định nguyên nhân chưa thể đưa vào hoạt động (ví dụ: thiếu vốn vận hành, chưa có đường ống kết nối...).
- Khẩn trương khôi phục và đưa các công trình cấp nước vào sử dụng, ưu tiên khu vực dân cư gần sông Bồng Miêu và vùng từng bị ảnh hưởng khai thác vàng trái phép.
- Tăng cường lấy mẫu giếng ngầm để kiểm tra chất lượng nước định kỳ, cảnh báo sớm nếu vượt ngưỡng cho phép.

Việc đảm bảo tiếp cận nguồn nước sạch không chỉ là một giải pháp kỹ thuật mà còn là một biện pháp bảo vệ sức khỏe cộng đồng lâu dài và thể hiện vai trò quản lý của chính quyền địa phương.

KẾT LUẬN

Từ kết quả nghiên cứu cho thấy sự hiện diện đáng kể của một số kim loại nặng như chì (Pb), asen (As) và cadimi (Cd) trong trầm tích bề mặt tại khu vực sông Bồng Miêu, tỉnh Quảng Nam. Nhiều vị trí lấy mẫu ghi nhận nồng độ vượt ngưỡng cho phép theo quy chuẩn QCVN 43:2017/BTNMT, đặc biệt tại vị trí S4, S5 có nồng độ KLN vượt ngưỡng cho phép nhphả, phản ánh tình trạng ô nhiễm kim loại nặng nghiêm trọng tại khu vực nghiên cứu. Chỉ số rủi ro sinh thái (RI) được tính toán dao động từ mức trung bình đến rất cao, với giá trị RI lớn nhất lên tới 640,2 (S5) – cho thấy nguy cơ ảnh hưởng nghiêm trọng đến hệ sinh thái thủy sinh và sức khỏe cộng đồng.

Dữ liệu khảo sát ý kiến người dân địa phương cũng phản ánh thực trạng đáng lo ngại về sức khỏe, với tỷ lệ đáng kể người dân báo cáo các triệu chứng bệnh lý liên quan đến da liễu, hô hấp và tiêu hóa. Những dấu hiệu này có khả năng liên quan đến ô nhiễm môi trường do sự tích tụ kim loại nặng kéo dài trong nước và trầm tích.

Về mặt phương pháp, nghiên cứu đã ứng dụng thành công các mô hình học máy, trong đó Random Forest (RF) được xác định là mô hình tối ưu với độ chính xác cao trong dự báo chỉ số RI. So sánh kết quả dự báo từ mô hình với chỉ số tính theo công thức của Lars Hakanson cho thấy sai số rất nhỏ (0,01–0,03), chứng minh hiệu quả mô hình trong thực tiễn. Việc áp dụng các mô hình học máy hiện đại như RF, SVM và XGBoost không chỉ nâng cao độ tin cậy trong đánh giá và dự báo chất lượng môi trường, mà còn góp phần giảm thiểu thời gian, chi phí và công sức cho công tác quan trắc truyền thống.

Bên cạnh đó, hướng tiếp cận tích hợp giữa dữ liệu môi trường, dữ liệu sức khỏe cộng đồng và mô hình hóa bằng học máy đã được chứng minh là khả thi và hiệu quả. Đây là một định hướng nghiên cứu tiềm năng, góp phần hỗ trợ công tác giám sát môi trường, cảnh báo sớm rủi ro sinh thái và đề xuất các giải pháp can thiệp kịp thời. Kết quả nghiên cứu không chỉ có ý nghĩa khoa học mà còn có giá trị thực tiễn cao trong việc xây dựng chiến lược quản lý môi trường bền vững, đặc biệt tại các khu vực chịu tác động mạnh từ hoạt động khai thác khoáng sản.

KIẾN NGHỊ

❖ Đối với hiện trạng sức khỏe môi trường

Trước thực trạng ô nhiễm kim loại nặng trong trầm tích và mức độ rủi ro sinh thái cao tại khu vực sông Bồng Miêu, nghiên cứu kiến nghị các nội dung sau:

- ✚ Tăng cường quan trắc môi trường định kỳ tại các điểm có nguy cơ cao (đặc biệt là vị trí S5 và S6), với phạm vi mở rộng bao gồm cả nước mặt, trầm tích và sinh vật thủy sinh, nhằm đánh giá chính xác diễn biến ô nhiễm theo thời gian.
- ✚ Áp dụng các mô hình học máy như Random Forest vào hệ thống quản lý môi trường địa phương, phục vụ mục đích cảnh báo sớm và hỗ trợ ra quyết định trong công tác kiểm soát ô nhiễm, đặc biệt tại các khu vực chịu ảnh hưởng từ khai thác khoáng sản.
- ✚ Thực hiện các nghiên cứu chuyên sâu về mối liên hệ giữa ô nhiễm môi trường và sức khỏe cộng đồng, từ đó xây dựng bộ chỉ thị môi trường – sức khỏe phù hợp với điều kiện địa phương, phục vụ công tác giám sát và phản ứng nhanh với các nguy cơ môi trường.
- ✚ Đề xuất các biện pháp cải tạo và phục hồi môi trường, như xử lý trầm tích ô nhiễm, kiểm soát nguồn thải từ hoạt động khai thác khoáng sản, đồng thời tăng cường phổ biến kiến thức môi trường cho cộng đồng dân cư để nâng cao nhận thức và khả năng tự bảo vệ sức khỏe.

❖ Đối với mô hình học máy

Từ quá trình thử nghiệm và so sánh các mô hình học máy khác nhau như K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Cubist và Linear Regression (LR), nghiên cứu nhận thấy rằng mô hình Random Forest (RF) cho hiệu suất dự báo ổn định và chính xác hơn trong điều kiện dữ liệu đầu vào còn hạn chế. Trong khi các mô hình Cubist và LR có dấu hiệu “overfitting”, làm giảm độ tin cậy của kết quả, thì mô hình RF đã chứng minh khả năng xử lý tốt dữ liệu phi tuyến và không yêu cầu nhiều giả định về phân phối dữ liệu.

Từ đó, nghiên cứu kiến nghị các cơ quan quản lý môi trường nên cân nhắc tích hợp các mô hình học máy, đặc biệt là Random Forest, vào hệ thống quản lý và giám sát chất lượng môi trường. Việc này không chỉ hỗ trợ hiệu quả trong việc dự báo xu hướng ô nhiễm và đánh giá rủi ro sinh thái mà còn giúp giảm thiểu chi phí, thời gian lấy mẫu và phân tích thực địa – một lợi thế quan trọng trong bối cảnh nguồn lực còn hạn chế và yêu

cầu cảnh báo sớm ngày càng cao. Đồng thời, cũng cần đầu tư mở rộng cơ sở dữ liệu môi trường, vì chất lượng và độ tin cậy của mô hình học máy phụ thuộc chặt chẽ vào quy mô và độ đa dạng của dữ liệu đầu vào.

TÀI LIỆU THAM KHẢO

- [1] *Từ điển hóa học Anh - Việt 2002*. Nhà xuất bản khoa học kỹ thuật
- [2] Lê Huy Bá, 2004. *Độc học Môi trường cơ bản*. Nhà xuất bản Đại học quốc gia Hồ Chí Minh
- [3] Lê Hồng Thía, Nguyễn Văn Phương, Tạp chí Khoa học và Công nghệ, 2022. *Đánh giá ô nhiễm kim loại nặng trong trầm tích hạ lưu sông Vàm Thuật*
- [5] *HIEN TRANG MT QUANG NAM*.
- [6] V. H. Tiệp. *Machine Learning cơ bản*.
- [7] Le Phuoc Cuong, Xuan Cuong Nguyen, Thi Thanh Huyen Nguyen, Journal of Environmental Management 2021 *Developing a new approach for design support of subsurface constructed wetland using machine learning algorithms*.
- [8] Nguyễn Hoàng, 2020. *Nghiên cứu một số mô hình trí tuệ nhân tạo dự báo chấn động nổi mìn trong khai thác mỏ lộ thiên*
- [9] "QCVN 08:2023/BTNMT: Quy chuẩn kỹ thuật Quốc gia về Chất lượng môi trường nước mặt".
- [10] Đặng Kim Chi, 2002. *Hóa học Môi trường*. Nxb Giáo dục.
- [11] Lê Phước Cường, 2018. *Giáo trình Độc học Môi trường*. Nhà xuất bản Đại học Quốc gia Hà Nội.
- [12] Nguyễn Đức Huệ, 2010. *Giáo trình Độc học Môi trường*. Nhà xuất bản Đại học Quốc gia Hà Nội
- [13] U.S. Environmental Protection Agency, 2005. *Toxicological review of Zinc and compounds*
- [14] Thông tư 03/2025/BLĐTBXH: Quy định tiêu chuẩn phân loại lao động theo điều kiện lao động
- [15] Siriporn Pradit, J. Mar. Sci. Eng, 2022. *Heavy Metal Contamination and Ecological Risk Assessment in the Sediment Cores of the Wetlands in Southern Thailand*
- [16] Lê Văn Hào, Tạp chí Khoa học và Công nghệ, 2023. *Đánh giá mức độ tích lũy và rủi ro sinh thái của một số kim loại nặng trong trầm tích tầng mặt tại sông Bông Miêu, tỉnh Quảng Nam*
- [17] Lars Hakanson, A Sedimentological Approach, 1980. *AN ECOLOGICAL RISK INDEX FOR AQUATIC POLLUTION CONTROL*.
- [18] QCVN 43:2017/BTNMT: Quy chuẩn kỹ thuật Quốc gia về chất lượng trầm tích

- [19] Nghị định số 36/2020/NĐ-CP: Quy định về xử phạt hành chính trong lĩnh vực tài nguyên nước và khoáng sản
- [20] Bộ Luật hình sự 2015 (sửa đổi, bổ sung 2017), Điều 227
- [21] Yizhuo Wang, He Li, Shaohua Lin, Water 2022. *Advances in the Study of Heavy Metal Adsorption from Water and Soil by Modified Biochar*
- [22] Cristina Ileana Covaliu-Mierlă, Oana Păunescu, Horia Iovu, Membranes 2023. *Recent Advances in Membranes Used for Nanofiltration to Remove Heavy Metals from Wastewater: A Review*

DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC

- [1]. Lê Duy Lâm, Tô Viết Đức Trí, **Trần Minh Hùng**, Lê Năng Định & Lê Phước Cường (2023). “Đánh giá mức độ tích lũy và rủi ro sinh thái một số PCB và OCP trong trầm tích mặt tại cửa An Hòa thuộc sông Trường Giang, Núi Thành, Quảng Nam”, *Tạp chí Khoa học và Công nghệ - Đại học Đà Nẵng*, Vol. 21(3), tr 63–68.

PHỤ LỤC 1
BẢNG DỮ LIỆU NỒNG ĐỘ CÁC KIM LOẠI NẶNG VÀ
CHỈ SỐ RỦI RO SINH THÁI (RI)

| Thời gian | Vị trí mẫu | Pb | As | Hg | Cd | Cr | Cu | Zn | RI |
|---------------------|-------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Tháng 5/2025 | S1 | 407.2 | 232.8 | 0.064 | 1.187 | 21.7 | 24.8 | 87 | 233.6 |
| | S2 | 378.3 | 277.3 | 0.248 | 1.161 | 23.6 | 28.7 | 85.9 | 290.3 |
| | S3 | 411.6 | 299.6 | 0.033 | 1.524 | 23 | 30.6 | 106.9 | 284.3 |
| | S4 | 709.9 | 272.4 | 0.049 | 2.713 | 28.9 | 46.6 | 134.2 | 327.6 |
| | S5 | 463.3 | 434.8 | 1.431 | 2.614 | 23.9 | 83.2 | 184.2 | 640.2 |
| | S6 | 546.6 | 195.6 | 0.877 | 1.882 | 19.3 | 49.1 | 146.5 | 372.4 |
| | S7 | 561.9 | 182.3 | 0.123 | 1.978 | 18.6 | 28.3 | 122.5 | 244.6 |
| | S8 | 218.4 | 97.4 | 0.068 | 0.944 | 17.1 | 28.3 | 61.7 | 123.3 |
| | S9 | 511.8 | 138.1 | 0.148 | 2.881 | 29.9 | 33.2 | 116.7 | 243.4 |
| | S10 | 491.6 | 152.1 | 0.153 | 1.867 | 21.4 | 26.3 | 97.9 | 220.7 |
| Tháng 4/2025 | S1 | 397.7 | 234.6 | 0.175 | 1.161 | 17.5 | 21.7 | 81.7 | 250.7 |
| | S2 | 368.8 | 279.1 | 0.359 | 1.765 | 19.4 | 25.6 | 80.6 | 326.3 |
| | S3 | 402.1 | 301.4 | 0.144 | 1.501 | 18.8 | 27.5 | 101.6 | 301.5 |
| | S4 | 700.4 | 274.2 | 0.16 | 2.691 | 24.7 | 43.5 | 128.9 | 344.8 |
| | S5 | 453.8 | 436.6 | 1.542 | 2.598 | 19.7 | 80.1 | 178.9 | 657.6 |
| | S6 | 537.1 | 197.4 | 0.988 | 1.861 | 15.1 | 46 | 141.2 | 389.6 |
| | S7 | 552.4 | 184.1 | 0.234 | 1.921 | 14.4 | 25.2 | 117.2 | 260.8 |
| | S8 | 208.9 | 99.2 | 0.179 | 0.924 | 12.9 | 25.2 | 56.4 | 140.5 |
| | S9 | 502.3 | 139.9 | 0.259 | 2.864 | 25.7 | 30.1 | 111.4 | 260.7 |
| | S10 | 482.1 | 153.9 | 0.264 | 1.115 | 17.2 | 23.2 | 92.6 | 216.0 |
| Tháng 3/2025 | S1 | 402.6 | 219.5 | 0.133 | 1.037 | 33.1 | 20.9 | 122.2 | 231.0 |
| | S2 | 373.7 | 264 | 0.317 | 1.018 | 35 | 24.8 | 93.3 | 287.7 |
| | S3 | 407 | 286.3 | 0.102 | 1.367 | 34.4 | 26.7 | 126.6 | 281.4 |
| | S4 | 705.3 | 259.1 | 0.118 | 2.568 | 40.3 | 42.7 | 424.9 | 326.6 |
| | S5 | 458.7 | 421.5 | 1.5 | 2.345 | 35.3 | 79.3 | 178.3 | 633.8 |
| | S6 | 542 | 182.3 | 0.946 | 1.738 | 30.7 | 45.2 | 261.6 | 370.4 |
| | S7 | 557.3 | 169 | 0.192 | 1.745 | 30 | 24.4 | 276.9 | 240.2 |
| | S8 | 213.8 | 84.1 | 0.137 | 0.801 | 28.5 | 24.4 | 66.6 | 120.7 |
| | S9 | 507.2 | 124.8 | 0.217 | 2.664 | 41.3 | 29.3 | 226.8 | 239.2 |
| | S10 | 487 | 138.8 | 0.222 | 1.738 | 32.8 | 22.4 | 206.6 | 219.1 |
| Tháng 2/2025 | S1 | 387.4 | 242.6 | 0.131 | 1.656 | 30 | 21.8 | 361 | 265.0 |
| | S2 | 358.5 | 287.1 | 0.315 | 1.264 | 31.9 | 17.6 | 32.5 | 308.0 |
| | S3 | 391.8 | 309.4 | 0.1 | 1.624 | 31.3 | 19.6 | 80 | 302.1 |
| | S4 | 690.1 | 282.2 | 0.116 | 2.814 | 37.2 | 24 | 299.1 | 345.3 |

| Thời gian | Vị trí mẫu | Pb | As | Hg | Cd | Cr | Cu | Zn | RI |
|----------------------|-------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | S5 | 443.5 | 444.6 | 1.498 | 2.714 | 32.2 | 44.6 | 52.5 | 654.7 |
| | S6 | 526.8 | 205.4 | 0.944 | 1.984 | 27.6 | 17.4 | 135.8 | 388.3 |
| | S7 | 542.1 | 192.1 | 0.19 | 2.044 | 26.9 | 26.4 | 151.1 | 262.6 |
| | S8 | 198.6 | 107.2 | 0.135 | 1.047 | 25.4 | 25.4 | 192.4 | 142.9 |
| | S9 | 492 | 147.9 | 0.215 | 2.984 | 38.2 | 30.5 | 101 | 262.1 |
| | S10 | 471.8 | 161.9 | 0.22 | 1.536 | 29.7 | 29.1 | 80.8 | 226.9 |
| Tháng 1/2025 | S1 | 414.3 | 209.4 | 0.053 | 1.251 | 43.9 | 7.7 | 188.5 | 218.0 |
| | S2 | 385.4 | 253.9 | 0.237 | 1.231 | 45.8 | 11.6 | 159.6 | 274.7 |
| | S3 | 418.7 | 276.2 | 0.022 | 1.591 | 45.2 | 13.5 | 192.9 | 268.7 |
| | S4 | 717 | 249 | 0.038 | 2.781 | 51.1 | 29.5 | 491.2 | 313.6 |
| | S5 | 470.4 | 411.4 | 1.42 | 2.681 | 46.1 | 66.1 | 244.6 | 624.5 |
| | S6 | 553.7 | 172.2 | 0.866 | 1.951 | 41.5 | 32 | 327.9 | 357.4 |
| | S7 | 569 | 158.9 | 0.112 | 2.011 | 40.8 | 11.2 | 343.2 | 228.8 |
| | S8 | 225.5 | 74 | 0.057 | 1.014 | 39.3 | 11.2 | 111 | 107.6 |
| | S9 | 518.9 | 114.7 | 0.137 | 2.951 | 52.1 | 16.1 | 293.1 | 228.4 |
| | S10 | 498.7 | 128.7 | 0.142 | 1.951 | 43.6 | 9.2 | 272.9 | 206.1 |
| Tháng 12/2024 | S1 | 361.1 | 228.6 | 0.041 | 1.113 | 21.6 | 25.7 | 65.2 | 221.6 |
| | S2 | 332.2 | 273.1 | 0.225 | 1.093 | 23.5 | 21.5 | 64.1 | 277.6 |
| | S3 | 365.5 | 295.4 | 0.01 | 1.453 | 22.9 | 23.5 | 85.1 | 271.6 |
| | S4 | 663.8 | 268.2 | 0.026 | 2.643 | 28.8 | 27.9 | 112.4 | 313.7 |
| | S5 | 417.2 | 430.6 | 1.408 | 2.543 | 23.8 | 48.5 | 162.4 | 624.7 |
| | S6 | 500.5 | 191.4 | 0.854 | 1.813 | 19.2 | 21.3 | 124.7 | 357.6 |
| | S7 | 515.8 | 178.1 | 0.1 | 1.873 | 18.5 | 30.3 | 100.7 | 231.8 |
| | S8 | 172.3 | 93.2 | 0.045 | 0.876 | 17 | 29.3 | 39.9 | 111.5 |
| | S9 | 465.7 | 133.9 | 0.125 | 2.813 | 29.8 | 34.4 | 94.9 | 231.6 |
| | S10 | 445.5 | 147.9 | 0.13 | 1.813 | 21.3 | 33 | 76.1 | 209.8 |
| Tháng 11/2024 | S1 | 423.7 | 228.1 | 0.118 | 1.25 | 45.2 | 9.4 | 62.8 | 241.0 |
| | S2 | 394.8 | 272.6 | 0.302 | 1.23 | 47.1 | 0.9 | 61.8 | 296.6 |
| | S3 | 428.1 | 294.9 | 0.087 | 1.59 | 46.5 | 2.8 | 75.2 | 290.5 |
| | S4 | 726.4 | 267.7 | 0.103 | 2.78 | 52.4 | 18.8 | 71.2 | 333.7 |
| | S5 | 479.8 | 430.1 | 1.485 | 2.68 | 47.4 | 55.4 | 67.2 | 646.0 |
| | S6 | 563.1 | 190.9 | 0.931 | 1.95 | 42.8 | 21.3 | 40.4 | 378.3 |
| | S7 | 578.4 | 177.6 | 0.177 | 2.01 | 42.1 | 0.5 | 74.6 | 249.7 |
| | S8 | 234.9 | 92.7 | 0.122 | 1.013 | 40.6 | 0.5 | 53 | 129.7 |
| | S9 | 528.3 | 133.4 | 0.202 | 2.95 | 53.4 | 5.4 | 72 | 249.6 |
| | S10 | 508.1 | 147.4 | 0.207 | 1.95 | 44.9 | 7.8 | 62.3 | 228.3 |
| Tháng 10/2024 | S1 | 363.8 | 194.5 | 0.11 | 1.202 | 34.1 | 9 | 53.5 | 211.3 |

| Thời gian | Vị trí mẫu | Pb | As | Hg | Cd | Cr | Cu | Zn | RI |
|---------------------|---------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | S2 | 334.9 | 239 | 0.294 | 1.182 | 32.7 | 12.9 | 52.4 | 268.1 |
| | S3 | 368.2 | 261.3 | 0.079 | 1.542 | 32 | 14.8 | 73.4 | 262.0 |
| | S4 | 666.5 | 234.1 | 0.095 | 2.732 | 35.6 | 30.8 | 100.7 | 305.3 |
| | S5 | 419.9 | 396.5 | 1.477 | 2.632 | 37.3 | 67.4 | 150.7 | 618.0 |
| | S6 | 503.2 | 157.3 | 0.923 | 1.902 | 28.7 | 33.3 | 113 | 350.2 |
| | S7 | 518.5 | 144 | 0.169 | 1.962 | 31.8 | 12.5 | 89 | 221.4 |
| | S8 | 175 | 59.1 | 0.114 | 0.965 | 32.1 | 12.5 | 28.2 | 101.2 |
| | S9 | 468.4 | 99.8 | 0.194 | 2.902 | 33.9 | 17.4 | 83.2 | 221.1 |
| | S10 | 448.2 | 113.8 | 0.199 | 1.902 | 32.9 | 10.5 | 64.4 | 198.9 |
| | Tháng 9/2024 | S1 | 355.4 | 230 | 0.044 | 1.165 | 25.4 | 9.9 | 162.6 |
| S2 | | 326.5 | 274.5 | 0.228 | 1.119 | 25.6 | 13.8 | 161.6 | 279.2 |
| S3 | | 359.8 | 296.8 | 0.013 | 1.479 | 29.2 | 15.7 | 175 | 273.2 |
| S4 | | 658.1 | 269.6 | 0.029 | 2.669 | 27.7 | 31.7 | 171 | 316.2 |
| S5 | | 411.5 | 432 | 0.92 | 2.569 | 27.5 | 68.3 | 167 | 550.1 |
| S6 | | 494.8 | 192.8 | 0.857 | 1.839 | 30.9 | 34.2 | 140.2 | 361.1 |
| S7 | | 510.1 | 179.5 | 0.103 | 1.543 | 11.8 | 13.4 | 174.4 | 221.5 |
| S8 | | 166.6 | 94.6 | 0.047 | 0.902 | 26.3 | 13.4 | 152.8 | 112.3 |
| S9 | | 460 | 135.3 | 0.124 | 2.839 | 26.5 | 18.3 | 171.8 | 231.5 |
| S10 | | 439.8 | 149.3 | 0.131 | 1.839 | 25.7 | 11.4 | 162.1 | 209.7 |
| Tháng 8/2024 | S1 | 378 | 202 | 0.062 | 1.262 | 25.4 | 13.4 | 53.6 | 211.7 |
| | S2 | 349.1 | 246.5 | 0.246 | 1.242 | 39.5 | 27.5 | 52.5 | 269.8 |
| | S3 | 382.4 | 268.8 | 0.031 | 1.602 | 38 | 26 | 73.5 | 263.4 |
| | S4 | 680.7 | 241.6 | 0.047 | 2.792 | 35.7 | 23.7 | 100.8 | 304.7 |
| | S5 | 434.1 | 404 | 0.998 | 2.423 | 36.8 | 24.8 | 150.8 | 536.9 |
| | S6 | 517.4 | 164.8 | 0.875 | 1.987 | 37.8 | 25.8 | 113.1 | 350.5 |
| | S7 | 532.7 | 151.5 | 0.121 | 2.022 | 33.2 | 21.2 | 89.1 | 222.4 |
| | S8 | 189.2 | 66.6 | 0.066 | 1.025 | 35.9 | 23.9 | 28.3 | 102.6 |
| | S9 | 482.6 | 107.3 | 0.146 | 4.126 | 36 | 24 | 83.3 | 256.8 |
| | S10 | 462.4 | 121.3 | 0.151 | 1.962 | 36.6 | 24.6 | 64.5 | 200.6 |
| Tháng 7/2024 | S1 | 342.6 | 241.5 | 0.059 | 0.948 | 28.4 | 34.3 | 84.6 | 227.9 |
| | S2 | 313.7 | 286 | 0.243 | 0.928 | 27 | 32.9 | 83.5 | 284.2 |
| | S3 | 347 | 308.3 | 0.028 | 1.427 | 26.3 | 32.2 | 104.5 | 282.0 |
| | S4 | 645.3 | 281.1 | 0.044 | 2.477 | 29.9 | 35.8 | 131.8 | 319.8 |
| | S5 | 398.7 | 443.5 | 1.426 | 2.378 | 31.6 | 37.5 | 181.8 | 629.1 |
| | S6 | 482 | 204.3 | 0.872 | 1.648 | 23 | 28.9 | 144.1 | 363.8 |
| | S7 | 497.3 | 191 | 0.118 | 1.708 | 26.1 | 32 | 120.1 | 237.4 |
| | S8 | 153.8 | 106.1 | 0.063 | 0.711 | 26.4 | 32.3 | 59.3 | 117.3 |

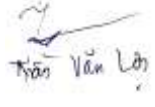



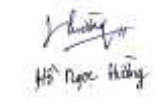
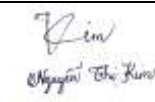
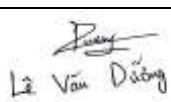
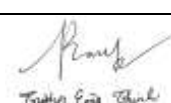
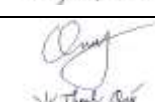
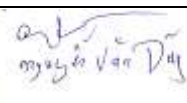
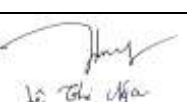
| Thời gian | Vị trí mẫu | Pb | As | Hg | Cd | Cr | Cu | Zn | RI |
|---------------------|-------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | S9 | 447.2 | 146.8 | 0.143 | 2.648 | 28.2 | 34.1 | 114.3 | 236.8 |
| | S10 | 427 | 160.8 | 0.148 | 1.535 | 27.2 | 33.1 | 95.5 | 211.9 |
| Tháng 6/2024 | S1 | 372.1 | 224.8 | 0.044 | 0.99 | 39 | 34.5 | 142.6 | 218.3 |
| | S2 | 343.2 | 269.3 | 0.228 | 0.97 | 40.9 | 36.4 | 113.7 | 274.8 |
| | S3 | 376.5 | 291.6 | 0.013 | 1.33 | 40.3 | 35.8 | 147 | 268.6 |
| | S4 | 674.8 | 264.4 | 0.029 | 2.52 | 46.2 | 41.7 | 445.3 | 312.4 |
| | S5 | 428.2 | 426.8 | 1.411 | 2.42 | 41.2 | 36.7 | 198.7 | 619.2 |
| | S6 | 511.5 | 187.6 | 0.857 | 1.69 | 36.6 | 32.1 | 282 | 355.1 |
| | S7 | 526.8 | 174.3 | 0.103 | 1.75 | 35.9 | 31.4 | 297.3 | 228.4 |
| | S8 | 183.3 | 89.4 | 0.048 | 0.753 | 34.4 | 29.9 | 65.1 | 107.1 |
| | S9 | 476.7 | 130.1 | 0.128 | 2.69 | 47.2 | 42.7 | 247.2 | 228.7 |
| | S10 | 456.5 | 144.1 | 0.133 | 1.69 | 38.7 | 34.2 | 227 | 206.2 |
| Tháng 5/2024 | S1 | 369 | 235.7 | 0.182 | 1.02 | 23.8 | 25.7 | 108.1 | 246.9 |
| | S2 | 340.1 | 280.2 | 0.366 | 1 | 22.4 | 24.3 | 107 | 303.2 |
| | S3 | 373.4 | 302.5 | 0.151 | 1.36 | 21.7 | 23.6 | 128 | 296.9 |
| | S4 | 671.7 | 275.3 | 0.167 | 2.55 | 25.3 | 27.2 | 155.3 | 338.9 |
| | S5 | 425.1 | 437.7 | 1.549 | 2.45 | 27 | 28.9 | 205.3 | 648.2 |
| | S6 | 508.4 | 198.5 | 0.995 | 1.72 | 18.4 | 20.3 | 167.6 | 382.8 |
| | S7 | 523.7 | 185.2 | 0.241 | 1.78 | 21.5 | 23.4 | 143.6 | 256.5 |
| | S8 | 180.2 | 100.3 | 0.186 | 0.783 | 21.8 | 23.7 | 82.8 | 136.3 |
| | S9 | 473.6 | 141 | 0.266 | 2.72 | 23.6 | 25.5 | 137.8 | 255.9 |
| | S10 | 453.4 | 155 | 0.271 | 1.72 | 22.6 | 24.5 | 119 | 234.3 |
| Tháng 4/2024 | S1 | 390.4 | 202.6 | 0.119 | 1.06 | 31.7 | 32.3 | 12 | 217.8 |
| | S2 | 361.5 | 247.1 | 0.303 | 1.04 | 30.3 | 28.1 | 16.9 | 273.8 |
| | S3 | 394.8 | 269.4 | 0.088 | 1.4 | 29.6 | 30.1 | 16.4 | 267.6 |
| | S4 | 693.1 | 242.2 | 0.104 | 2.59 | 33.2 | 34.5 | 314.7 | 311.3 |
| | S5 | 446.5 | 404.6 | 1.486 | 2.49 | 34.9 | 55.1 | 68.1 | 620.8 |
| | S6 | 529.8 | 165.4 | 0.932 | 1.76 | 26.3 | 27.9 | 151.4 | 354.3 |
| | S7 | 545.1 | 152.1 | 0.178 | 1.82 | 29.4 | 36.9 | 166.7 | 228.7 |
| | S8 | 201.6 | 67.2 | 0.123 | 0.823 | 29.7 | 35.9 | 176.8 | 108.8 |
| | S9 | 495 | 107.9 | 0.203 | 2.76 | 31.5 | 41 | 116.6 | 228.0 |
| | S10 | 474.8 | 121.9 | 0.208 | 1.76 | 30.5 | 39.6 | 96.4 | 206.4 |
| Tháng 3/2024 | S1 | 345.5 | 213.3 | 0.108 | 1.134 | 41.3 | 21.6 | 59.8 | 221.6 |
| | S2 | 316.6 | 257.8 | 0.292 | 1.114 | 43.2 | 23.5 | 30.9 | 278.1 |
| | S3 | 349.9 | 280.1 | 0.077 | 1.474 | 42.6 | 22.9 | 64.2 | 271.9 |
| | S4 | 648.2 | 252.9 | 0.093 | 2.651 | 48.5 | 28.8 | 362.5 | 315.3 |
| | S5 | 401.6 | 415.3 | 1.475 | 2.567 | 43.5 | 23.8 | 115.9 | 622.6 |





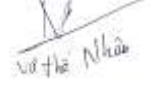
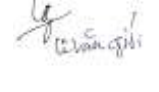

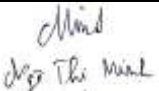
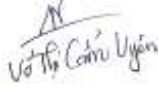

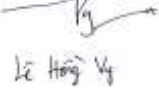
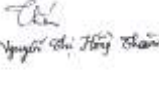
| Thời gian | Vị trí mẫu | Pb | As | Hg | Cd | Cr | Cu | Zn | RI |
|----------------------|-------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | S6 | 484.9 | 176.1 | 0.921 | 1.834 | 38.9 | 19.2 | 199.2 | 358.3 |
| | S7 | 500.2 | 162.8 | 0.167 | 1.894 | 38.2 | 18.5 | 214.5 | 231.7 |
| | S8 | 156.7 | 77.9 | 0.112 | 0.897 | 36.7 | 17 | 129 | 111.2 |
| | S9 | 450.1 | 118.6 | 0.192 | 2.834 | 49.5 | 29.8 | 164.4 | 232.0 |
| | S10 | 429.9 | 132.6 | 0.197 | 1.175 | 41 | 21.3 | 144.2 | 189.7 |
| Tháng 2/2024 | S1 | 383.9 | 241.9 | 0.133 | 1.131 | 14.3 | 3.3 | 53.7 | 244.9 |
| | S2 | 355 | 286.4 | 0.317 | 1.111 | 28.4 | 17.4 | 82.6 | 303.2 |
| | S3 | 388.3 | 308.7 | 0.102 | 1.471 | 26.9 | 15.9 | 49.5 | 296.5 |
| | S4 | 686.6 | 281.5 | 0.118 | 2.661 | 24.6 | 13.6 | 249 | 338.7 |
| | S5 | 440 | 443.9 | 1.5 | 2.561 | 25.7 | 14.7 | 2.4 | 646.2 |
| | S6 | 523.3 | 204.7 | 0.946 | 1.845 | 26.7 | 15.7 | 85.7 | 383.2 |
| | S7 | 538.6 | 191.4 | 0.192 | 1.891 | 22.1 | 11.1 | 101 | 255.7 |
| | S8 | 195.1 | 106.5 | 0.137 | 0.894 | 24.8 | 13.8 | 42.5 | 135.8 |
| | S9 | 488.5 | 147.2 | 0.217 | 2.472 | 24.9 | 13.9 | 50.9 | 244.1 |
| | S10 | 468.3 | 161.2 | 0.222 | 1.831 | 25.5 | 14.5 | 30.7 | 233.6 |
| Tháng 1/2024 | S1 | 418.5 | 229 | 0.057 | 0.994 | 34.2 | 26.7 | 136 | 225.7 |
| | S2 | 389.6 | 273.5 | 0.241 | 0.974 | 36.1 | 28.6 | 134.9 | 282.4 |
| | S3 | 422.9 | 295.8 | 0.026 | 1.334 | 35.5 | 28 | 155.9 | 276.1 |
| | S4 | 721.2 | 268.6 | 0.042 | 2.524 | 41.4 | 33.9 | 183.2 | 318.4 |
| | S5 | 474.6 | 431 | 1.424 | 2.424 | 36.4 | 28.9 | 233.2 | 626.8 |
| | S6 | 557.9 | 191.8 | 0.87 | 1.694 | 31.8 | 24.3 | 195.5 | 362.0 |
| | S7 | 573.2 | 178.5 | 0.116 | 1.754 | 31.1 | 23.6 | 171.5 | 235.2 |
| | S8 | 229.7 | 93.6 | 0.061 | 0.757 | 29.6 | 22.1 | 110.7 | 114.8 |
| | S9 | 523.1 | 134.3 | 0.141 | 2.694 | 42.4 | 34.9 | 165.7 | 235.7 |
| | S10 | 502.9 | 148.3 | 0.146 | 1.644 | 33.9 | 26.4 | 146.9 | 211.7 |
| Tháng 12/2023 | S1 | 404.9 | 231.5 | 0.053 | 1.01 | 43 | 45.2 | 53.3 | 227.8 |
| | S2 | 376 | 276 | 0.237 | 0.99 | 44.9 | 47.1 | 52.2 | 284.5 |
| | S3 | 409.3 | 298.3 | 0.022 | 1.35 | 44.3 | 46.5 | 73.2 | 278.2 |
| | S4 | 707.6 | 271.1 | 0.038 | 2.54 | 50.2 | 52.4 | 100.5 | 320.5 |
| | S5 | 461 | 433.5 | 1.42 | 2.44 | 45.2 | 47.4 | 150.5 | 628.9 |
| | S6 | 544.3 | 194.3 | 0.866 | 1.71 | 40.6 | 42.8 | 112.8 | 364.1 |
| | S7 | 559.6 | 181 | 0.112 | 1.77 | 39.9 | 42.1 | 88.8 | 237.3 |
| | S8 | 216.1 | 96.1 | 0.057 | 0.773 | 38.4 | 40.6 | 28 | 116.9 |
| | S9 | 509.5 | 136.8 | 0.137 | 2.71 | 51.2 | 53.4 | 83 | 237.8 |
| | S10 | 489.3 | 150.8 | 0.142 | 1.57 | 42.7 | 44.9 | 64.2 | 211.1 |
| Tháng 11/2023 | S1 | 399.3 | 223.3 | 0.078 | 1.09 | 38.7 | 40.5 | 40.1 | 227.7 |
| | S2 | 370.4 | 267.8 | 0.262 | 1.07 | 40.6 | 42.4 | 11.2 | 284.2 |

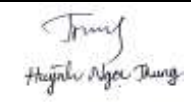

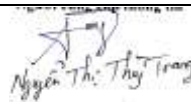
| Thời gian | Vị trí mẫu | Pb | As | Hg | Cd | Cr | Cu | Zn | RI |
|----------------------|-------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| | S3 | 403.7 | 290.1 | 0.047 | 1.43 | 40 | 41.8 | 44.5 | 278.0 |
| | S4 | 702 | 262.9 | 0.063 | 2.62 | 45.9 | 47.7 | 342.8 | 321.8 |
| | S5 | 455.4 | 425.3 | 1.445 | 2.52 | 40.9 | 42.7 | 96.2 | 628.6 |
| | S6 | 538.7 | 186.1 | 0.891 | 1.79 | 36.3 | 38.1 | 179.5 | 364.4 |
| | S7 | 554 | 172.8 | 0.137 | 1.85 | 35.6 | 37.4 | 194.8 | 237.8 |
| | S8 | 210.5 | 87.9 | 0.082 | 0.853 | 34.1 | 35.9 | 148.7 | 117.5 |
| | S9 | 503.9 | 128.6 | 0.162 | 2.79 | 46.9 | 48.7 | 144.7 | 238.1 |
| | S10 | 483.7 | 142.6 | 0.167 | 1.79 | 38.4 | 40.2 | 124.5 | 215.6 |
| Tháng 10/2023 | S1 | 425.9 | 239.6 | 0.058 | 1 | 16.6 | 10.3 | 87.3 | 231.3 |
| | S2 | 397 | 284.1 | 0.242 | 0.98 | 18.5 | 12.2 | 86.2 | 288.0 |
| | S3 | 430.3 | 306.4 | 0.027 | 1.34 | 17.9 | 11.6 | 107.2 | 281.7 |
| | S4 | 728.6 | 279.2 | 0.043 | 2.53 | 23.8 | 17.5 | 134.5 | 324.0 |
| | S5 | 482 | 441.6 | 0.994 | 2.43 | 18.8 | 12.5 | 184.5 | 563.5 |
| | S6 | 565.3 | 202.4 | 0.871 | 1.7 | 14.2 | 7.9 | 146.8 | 367.6 |
| | S7 | 580.6 | 189.1 | 0.117 | 1.76 | 13.5 | 7.2 | 122.8 | 240.8 |
| | S8 | 237.1 | 104.2 | 0.062 | 0.763 | 12 | 5.7 | 62 | 120.4 |
| | S9 | 530.5 | 144.9 | 0.142 | 2.7 | 24.8 | 18.5 | 117 | 241.3 |
| | S10 | 510.3 | 158.9 | 0.147 | 1.7 | 16.3 | 10 | 98.2 | 218.8 |
| Tháng 9/2023 | S1 | 365.7 | 229.6 | 0.052 | 1.11 | 31.1 | 34 | 131.9 | 225.7 |
| | S2 | 336.8 | 274.1 | 0.236 | 1.09 | 33 | 35.9 | 60.8 | 281.9 |
| | S3 | 370.1 | 296.4 | 0.021 | 1.45 | 32.4 | 35.3 | 27.5 | 275.3 |
| | S4 | 668.4 | 269.2 | 0.037 | 2.64 | 38.3 | 41.2 | 270.8 | 318.8 |
| | S5 | 421.8 | 431.6 | 0.988 | 2.54 | 33.3 | 36.2 | 24.2 | 556.6 |
| | S6 | 505.1 | 192.4 | 0.865 | 1.81 | 28.7 | 31.6 | 107.5 | 361.5 |
| | S7 | 520.4 | 179.1 | 0.111 | 1.87 | 28 | 30.9 | 122.8 | 234.8 |
| | S8 | 176.9 | 94.2 | 0.056 | 0.873 | 26.5 | 29.4 | 220.7 | 115.4 |
| | S9 | 470.3 | 134.9 | 0.136 | 2.81 | 39.3 | 42.2 | 72.7 | 235.1 |
| | S10 | 450.1 | 148.9 | 0.141 | 1.81 | 30.8 | 33.7 | 52.5 | 212.6 |

PHỤ LỤC 2
DANH SÁCH NGƯỜI DÂN THAM GIA THỰC HIỆN PHIẾU KHẢO SÁT
KHU VỰC XUNG QUANH SÔNG BÔNG MIÊU, TỈNH QUẢNG NAM

| ST T | Họ và Tên | Giới tính | Tuổi | Địa chỉ | SĐT | Chữ Ký | Ghi chú |
|------|-------------------|-----------|------|----------------|------------|---|---------|
| 1 | Bùi Thị Nga | Nữ | 37 | Thôn Bông Miêu | | | |
| 2 | Võ Thị Phương | Nữ | 32 | Thôn Bông Miêu | 0236382446 |  | |
| 3 | Hồ Ngọc Hương | Nữ | 38 | Thôn Bông Miêu | 037182524 |  | |
| 4 | Trương Quang Sang | Nam | 31 | Thôn Bông Miêu | 0354279005 |  | |
| 5 | Huỳnh Thị Dung | Nữ | 43 | Thôn An Lâu 1 | 0905184132 |  | |
| 6 | Nguyễn Đình Trung | Nam | 47 | Thôn Bông Miêu | 0987377641 |  | |
| 7 | Nguyễn Thị Nhung | Nữ | 34 | Thôn Bông Miêu | 0384414372 |  | |
| 8 | Phạm Văn Hòa | Nam | 35 | Thôn Bông Miêu | 0354627404 |  | |
| 9 | Cao Thanh Lực | Nam | 42 | Thôn Trung Sơn | 0921376385 |  | |
| 10 | Trịnh Đức Thanh | Nam | 53 | Thôn Bông Miêu | 0376445173 |  | |
| 11 | Nguyễn Tấn Luận | Nam | 46 | Thôn An Lâu 1 | 0354274521 |  | |
| 12 | Trần Khả Hành | Nam | 28 | Thôn Trung Sơn | 0236456741 |  | |
| 13 | Nguyễn Viết Cảnh | Nam | 64 | Thôn Bông Miêu | 0905404218 |  | |
| 14 | Trần Văn Lê | Nam | 65 | Thôn Bông Miêu | 0394472849 |  | |

| ST T | Họ và Tên | Giới tính | Tuổi | Địa chỉ | SĐT | Chữ Ký | Ghi chú |
|------|------------------------|-----------|------|----------------|------------|---|---------|
| 15 | Phạm Ngọc Hòe | Nam | 41 | Thôn Bồng Miêu | | | |
| 16 | Trần Văn Lợi | Nam | 28 | Thôn Bồng Miêu | 0384464132 |  | |
| 17 | Trương Ngọc Bích | Nam | 53 | Thôn Bồng Miêu | |  | |
| 18 | Phạm Thị Thúy Vũ | Nữ | 28 | Thôn Bồng Miêu | 0376432674 |  | |
| 19 | Nguyễn Quang Duy | Nam | 41 | Thôn Bồng Miêu | 0376007125 |  | |
| 20 | Nguyễn Văn Đông | Nam | 45 | Thôn Bồng Miêu | 0326467474 |  | |
| 21 | Nguyễn Thị Tuyết Nhung | Nữ | 35 | Thôn Bồng Miêu | | | |
| 22 | Hồ Ngọc Hương | Nữ | 36 | Thôn Bồng Miêu | 0935484314 |  | |
| 23 | Nguyễn Thị Kim | Nữ | 41 | Thôn Bồng Miêu | 0905436412 |  | |
| 24 | Lê Văn Dương | Nam | 47 | Thôn Bồng Miêu | 0372593761 |  | |
| 25 | Trương Công Thịnh | Nam | 42 | Thôn Bồng Miêu | 0338005335 |  | |
| 26 | Võ Thanh Quý | Nam | 28 | Thôn Bồng Miêu | 0325148372 |  | |
| 27 | Phạm Thị Hoa | Nữ | 38 | Thôn Bồng Miêu | | | |
| 28 | Nguyễn Văn Dũng | Nam | 50 | Thôn Bồng Miêu | 0987261622 |  | |
| 29 | Lê Thị Nga | Nữ | 51 | Thôn Bồng Miêu | 0374985117 |  | |

| ST T | Họ và Tên | Giới tính | Tuổi | Địa chỉ | SĐT | Chữ Ký | Ghi chú |
|------|----------------------|-----------|------|----------------|------------|---|---------|
| 30 | Phan Văn Sâm | Nam | 55 | Thôn Bồng Miêu | |  Phan Văn Sâm | |
| 31 | Phạm Thị Thúy Anh | Nữ | 32 | Thôn Bồng Miêu | 0376405102 |  Phạm Thị Thúy Anh | |
| 32 | Nguyễn Thị Cúc | Nữ | 41 | Thôn Bồng Miêu | | | |
| 33 | Nguyễn Văn Trí | Nam | 40 | Thôn Bồng Miêu | 0987625479 |  Nguyễn Văn Trí | |
| 34 | Nguyễn Thế Kỷ | Nam | 65 | Thôn Bồng Miêu | 0335466625 |  Nguyễn Thế Kỷ | |
| 35 | Vũ Thế Nhân | Nam | 32 | Thôn Trung Sơn | 0905322466 |  Vũ Thế Nhân | |
| 36 | Lê Văn Giới | Nam | 63 | Thôn Bồng Miêu | 0394874702 |  Lê Văn Giới | |
| 37 | Đình Tiên Hoàng | Nam | 61 | Thôn Bồng Miêu | | | |
| 38 | Phạm Thị Lan | Nữ | 47 | Thôn Bồng Miêu | 0375751811 |  Phạm Thị Lan | |
| 39 | Ngô Thị Minh | Nữ | 43 | Thôn Bồng Miêu | 0355461421 |  Ngô Thị Minh | |
| 40 | Võ Thị Cẩm Uyên | Nữ | 32 | Thôn Bồng Miêu | 0964309246 |  Võ Thị Cẩm Uyên | |
| 41 | Vũ Tiến Vĩnh | Nam | 38 | Thôn Bồng Miêu | 0946530576 |  Vũ Tiến Vĩnh | |
| 42 | Huỳnh Ngọc Bang | Nam | 56 | Thôn Bồng Miêu | | | |
| 43 | Lê Hồng Vy | Nữ | 43 | Thôn Bồng Miêu | 0323456412 |  Lê Hồng Vy | |
| 44 | Nguyễn Thị Hồng Thắm | Nữ | 36 | Thôn Bồng Miêu | 0322684174 |  Nguyễn Thị Hồng Thắm | |

| ST T | Họ và Tên | Giới tính | Tuổi | Địa chỉ | SĐT | Chữ Ký | Ghi chú |
|------|-----------------------|-----------|------|----------------|------------|---|---------|
| 45 | Huỳnh Ngọc Trung | Nam | 37 | Thôn Bông Miêu | 0336597621 |  | |
| 46 | Nguyễn Hà | Nam | 52 | Thôn An Lâu 1 | |  | |
| 47 | Trần Thị Đào | Nữ | 53 | Thôn Bông Miêu | |  | |
| 48 | Nguyễn Thị Thùy Trang | Nữ | 29 | Thôn Bông Miêu | 0935462386 |  | |
| 49 | Cô Bông | Nữ | 46 | Thôn Bông Miêu | | | |
| 50 | Nguyễn Thị Anh | Nữ | 46 | Thôn Bông Miêu | 0376748352 |  | |