

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT

LÊ TUẤN NGỌC

NGHIÊN CỨU PHÁT TRIỂN MÔ HÌNH TRÍ TUỆ NHÂN
TẠO DỰ BÁO Ô NHIỄM KHÔNG KHÍ THEO THỜI GIAN
THỰC TRONG QUÁ TRÌNH KHAI THÁC MỎ ĐỒNG
SIN QUYỀN, LÀO CAI

TÓM TẮT LUẬN ÁN TIẾN SĨ KỸ THUẬT

Hà Nội - 2026

Công trình hoàn thành tại: **Bộ môn Khai thác lộ thiên**
Khoa Mỏ, Trường Đại học Mỏ - Địa chất

Người hướng dẫn khoa học:

- 1. TS. Nguyễn Hoàng**, Trường Đại học Mỏ - Địa chất
- 2. GS.TS. Bùi Xuân Nam**, Trường Đại học Tôn Đức Thắng

Phản biện 1:

PGS.TS Vũ Đình Hiếu

Phản biện 2:

PGS.TS Đàm Trọng Thắng

Phản biện 3:

PGS.TS Lê Hồng Anh

Luận án sẽ được bảo vệ trước Hội đồng đánh giá luận án cấp Trường, họp
tại.....
vào hồi.....giờ.....ngày.....tháng.....năm 2026.

Có thể tìm thấy luận án tại: **Thư viện Quốc gia Hà Nội**
hoặc **Thư viện Trường Đại học Mỏ - Địa chất**

MỞ ĐẦU

1. Tính cấp thiết của đề tài

Trong chiến lược phát triển kinh tế - xã hội của Việt Nam, công nghiệp khai khoáng luôn giữ vai trò then chốt, cung cấp nguyên nhiên liệu nền tảng cho an ninh năng lượng và phát triển hạ tầng, trong đó có khai thác khoáng sản bằng phương pháp lộ thiên. Đặc thù của công nghệ khai thác lộ thiên là muốn lấy được khoáng sản có ích trong lòng đất thì cần phải bóc đi một lượng lớn đất đá phủ và đất đá gốc bao quanh thân quặng, sử dụng phương pháp khoan-nổ mìn và hệ thống đồng bộ thiết bị khai thác cơ giới hóa công suất lớn. Do đó, các tác động tiêu cực đến môi trường xung quanh là không nhỏ, đặc biệt là môi trường không khí.

Hiện nay, hầu hết các mỏ lộ thiên lớn tại Việt Nam, trong đó có mỏ đồng Sin Quyền, tỉnh Lào Cai, đều đang bước vào giai đoạn khai thác xuống sâu. Đặc điểm hình học của mỏ chuyển dần sang dạng lòng chảo sâu, tạo nên một vùng vi khí hậu đặc thù và phức tạp. Tại đây, các cơ chế thông gió tự nhiên thường xuyên bị triệt tiêu, thay vào đó là hiện tượng nghịch nhiệt và quẩn gió, khiến bụi phát sinh do nổ mìn, và quá trình vận tải không thể khuếch tán, bị nhốt lại ở đáy mỏ và trên các tầng công tác. Tình trạng này không chỉ gây ô nhiễm cục bộ nghiêm trọng, ảnh hưởng trực tiếp đến sức khỏe người lao động và năng suất làm việc của thiết bị, sản lượng khai thác mỏ, mà còn ảnh hưởng đến chất lượng môi trường không khí xung quanh.

Công tác kiểm tra, kiểm soát các thông số môi trường không khí tại các mỏ lộ thiên hiện nay chủ yếu dựa trên quan trắc định kỳ sử dụng các trạm đo di động tại hiện trường, không liên tục với tần suất 1 lần/quý. Cách tiếp cận này mang tính chất hậu kiểm, tức là nhà quản lý chỉ biết được mức độ ô nhiễm khi sự việc đã xảy ra. Trong bối cảnh sản xuất hiện đại, tư duy chạy theo khắc phục sự cố không còn phù hợp. Ngành mỏ đang đứng trước yêu cầu cấp bách phải chuyển đổi sang tư duy dự báo và kiểm soát chủ động. Nghĩa là, phải biết trước được nồng độ bụi, khí độc sẽ vượt ngưỡng vào thời điểm nào, tại vị trí nào để chủ động điều tiết hoạt động nổ mìn, vận tải hay kích hoạt hệ thống dập bụi trước khi các tác động xấu liên quan đến môi trường xảy ra.

Sự phát triển của công nghệ số và cuộc Cách mạng công nghiệp lần thứ tư đã tạo điều kiện thuận lợi cho việc ứng dụng các mô hình trí tuệ nhân tạo trong quản lý và điều hành mỏ theo hướng thông minh, tự động hóa và dự báo thời gian thực. Việc ứng dụng Trí tuệ nhân tạo (AI), đặc biệt là các thuật toán học sâu (Deep learning) kết hợp với các kỹ thuật tối ưu hóa hiện đại, cho phép mô hình hóa và dự báo chính xác diễn biến nồng độ bụi và khí ô nhiễm trong môi trường mỏ theo thời gian thực, điều mà các phương pháp thống kê truyền thống trước đây gặp nhiều hạn chế và mức độ chính xác khi dự báo không cao.

Tại Việt Nam, các nghiên cứu về dự báo chất lượng không khí trong khai thác mỏ lộ thiên bằng mô hình trí tuệ nhân tạo còn tương đối hạn chế, đặc biệt đối với các mỏ khai thác xuống sâu có điều kiện vi khí hậu phức tạp như mỏ đồng Sin Quyền. Xuất phát từ thực tiễn cấp thiết tại mỏ đồng Sin Quyền và xu hướng công nghệ trên thế giới, luận án "Nghiên cứu phát triển mô hình trí tuệ nhân tạo dự báo ô nhiễm không khí theo thời gian thực trong quá trình khai thác mỏ đồng Sin Quyền, Lào Cai" có ý nghĩa khoa học và thực tiễn cao, cung cấp công cụ dự báo ô nhiễm không khí tin cậy nhằm hỗ trợ đắc lực cho công tác điều hành sản xuất mỏ được an toàn, hiệu quả và bền vững.

2. Mục tiêu

Xây dựng và kiểm chứng mô hình trí tuệ nhân tạo dự báo nồng độ bụi mịn $PM_{2.5}$ và

khí CO tại ba vị trí quan trắc đại diện theo cao độ trong mỏ đồng Sin Quyền. Trên cơ sở đó cung cấp thông tin hỗ trợ cán bộ quản lý mỏ lựa chọn biện pháp điều hành sản xuất và kiểm soát môi trường phù hợp. Kết quả được kiểm chứng trong phạm vi dữ liệu thực đo tại mỏ đồng Sin Quyền và có thể được tham khảo, hiệu chỉnh lại khi áp dụng cho các mỏ lộ thiên sâu có điều kiện tương tự.

3. Nội dung nghiên cứu

Luận án tập trung nghiên cứu một số nội dung chính sau:

- Nghiên cứu tổng quan về đặc điểm phát sinh, lan truyền ô nhiễm không khí trong khai thác mỏ lộ thiên sâu; thực trạng công tác quan trắc, dự báo môi trường không khí và xu hướng ứng dụng trí tuệ nhân tạo trong lĩnh vực khai thác mỏ;
- Nghiên cứu cơ sở lý thuyết của các mô hình trí tuệ nhân tạo, các thuật toán học máy, học sâu và thuật toán tối ưu hóa phục vụ bài toán dự báo chất lượng không khí theo thời gian thực;
- Xây dựng cơ sở dữ liệu môi trường không khí, dữ liệu khí tượng và dữ liệu sản xuất tại mỏ đồng Sin Quyền. Phát triển các mô hình trí tuệ nhân tạo dự báo chất lượng không khí phù hợp với điều kiện khai thác lộ thiên xuống sâu;
- Đánh giá, so sánh hiệu suất các mô hình dự báo; phân tích ảnh hưởng của các yếu tố công nghệ và vi khí hậu mỏ đến chất lượng không khí, từ đó đề xuất giải pháp ứng dụng mô hình dự báo trong công tác quản lý môi trường và điều hành sản xuất mỏ.

4. Đối tượng và phạm vi nghiên cứu

4.1. Đối tượng nghiên cứu

Đối tượng nghiên cứu của luận án là:

- Các yếu tố công nghệ khai thác, vi khí hậu và môi trường không khí trong quá trình khai thác xuống sâu tại mỏ đồng Sin Quyền, Lào Cai. Trong phạm vi thực nghiệm của luận án, hai biến đầu ra (output) được lựa chọn là bụi mịn $PM_{2.5}$ và khí CO; các thông số PM_{10} , TSP, NO_x , SO_2 , tiếng ồn và chấn động không phải là biến đầu ra của mô hình và được xác định là hướng mở rộng trong các nghiên cứu tiếp theo.
- Các mô hình trí tuệ nhân tạo và thuật toán tối ưu hóa có khả năng dự báo chất lượng không khí theo thời gian thực trong điều kiện khai thác mỏ lộ thiên sâu.

4.2. Phạm vi nghiên cứu

- Về không gian: Nghiên cứu được thực hiện tại khai trường khai thác lộ thiên mỏ đồng Sin Quyền, tỉnh Lào Cai. Mô hình dự báo chuỗi thời gian tại ba điểm cố định tương ứng các mức cao độ +102 m, +68 m và -58 m. Kết quả không đại diện cho trường nồng độ liên tục trong toàn bộ không gian ba chiều của mỏ, việc ngoại suy đến các góc khuất gió hoặc các tầng mới phải được kiểm chứng bằng số liệu bổ sung.
- Về thời gian: Khảo sát mỏ từ tháng 12/2023 (trúng tuyển NCS). Thu thập dữ liệu thời gian thực từ ngày 10/4/2024 đến 31/3/2025.

5. Phương pháp nghiên cứu

Các phương pháp nghiên cứu chính được thực hiện trong luận án nhằm đạt được các mục tiêu đề ra bao gồm:

- Phương pháp tổng hợp, phân tích và đánh giá các tài liệu, công trình nghiên cứu trong và ngoài nước liên quan đến ô nhiễm không khí, vi khí hậu mỏ và ứng dụng trí tuệ nhân tạo trong dự báo môi trường mỏ;
- Phương pháp khảo sát thực địa, thu thập và xử lý số liệu môi trường không khí, số

liệu khí tượng và dữ liệu sản xuất liên quan đến máy móc, thiết bị sử dụng tại mỏ đồng Sin Quyền, Lào Cai;

- Phương pháp thống kê và phân tích dữ liệu nhằm đánh giá đặc điểm phân bố, xu hướng biến động và mối quan hệ giữa các yếu tố ảnh hưởng đến chất lượng không khí trong mỏ;

- Phương pháp mô hình hóa và ứng dụng trí tuệ nhân tạo, bao gồm các thuật toán học máy, học sâu và mô hình học kết hợp để xây dựng mô hình dự báo chất lượng không khí theo thời gian thực phù hợp điều kiện mỏ đồng Sin Quyền;

- Phương pháp tối ưu hóa siêu tham số bằng các thuật toán tối ưu hiện đại nhằm nâng cao độ chính xác và khả năng tổng quát hóa của các mô hình dự báo;

- Phương pháp đánh giá hiệu suất mô hình thông qua các chỉ tiêu thống kê như MAE, RMSE và hệ số xác định R^2 , kết hợp so sánh giữa các mô hình để lựa chọn cấu hình phù hợp cho điều kiện khai thác mỏ đồng Sin Quyền; kết quả có thể được sử dụng làm cơ sở tham khảo cho các mỏ lộ thiên sâu có điều kiện tương đồng sau khi hiệu chỉnh và kiểm chứng bằng dữ liệu địa phương.

6. Những điểm mới của luận án

CEEMDAN, MPA, XGBoost và Stacking Ensemble đều là các thuật toán/phương pháp đã được công bố trước đây; luận án không coi từng thuật toán riêng lẻ là phát minh mới. Tính mới của nghiên cứu nằm ở khung tích hợp được đồng thiết kế theo cơ chế vật lý - công nghệ của mỏ lộ thiên sâu: từ mạng quan trắc phân tầng, số hóa động lực học phát thải và vi khí hậu bằng biến tương tác/biến trễ/CEEMDAN, tối ưu mô hình bằng MPA, kết hợp mô hình dữ liệu băng - chuỗi thời gian, đến chuyển kết quả dự báo thành thông tin hỗ trợ điều hành. Khung này được kiểm chứng bằng dữ liệu thực đo tại mỏ đồng Sin Quyền.

- Thiết lập được quy trình quan trắc phân tầng theo ba cao độ và quy trình xử lý chuỗi dữ liệu $PM_{2.5}$, CO, khí tượng và hoạt động sản xuất. Trong đó CEEMDAN, biến tương tác và các biến trễ được sử dụng để giảm ảnh hưởng của nhiễu cục bộ và biểu diễn đặc điểm tích tụ ô nhiễm theo thời gian tại mỏ lộ thiên sâu.

- Tối ưu hóa mô hình XGBoost bằng giải thuật Tối ưu hoá kẻ săn mồi đại dương (Marine Predators Algorithm - MPA) và xây dựng mô hình học máy kết hợp Stacking Ensemble. Trong bộ dữ liệu kiểm thử tại mỏ đồng Sin Quyền, MPA-XGBoost cho sai số RMSE, MAE thấp và hệ số xác định R^2 cao hơn các mô hình đối chứng được khảo sát, đồng thời cải thiện khả năng theo sát các đỉnh nồng độ so với mô hình RF khi chưa tối ưu.

- Làm rõ mối liên hệ theo thời gian giữa chu kỳ khoan - nổ mìn, xúc bốc - vận tải, điều kiện vi khí hậu và biến thiên $PM_{2.5}$, CO thông qua phân tích chuỗi thời gian, biến trễ và cơ chế chú ý. Kết quả cho thấy việc đưa thông tin các khâu dây chuyền công nghệ khai thác mỏ vào tập đặc trưng giúp mô hình phản ánh tốt hơn các giai đoạn phát thải đột biến và tích tụ kéo dài.

- Đề xuất quy trình chuyển đổi kết quả dự báo trước 1 đến 3 giờ thành các mức theo dõi, cảnh báo và khuyến nghị điều hành có sự phê duyệt của cán bộ quản lý mỏ, qua đó góp phần chuyển phương thức quản lý môi trường từ hậu kiểm sang phòng ngừa trong phạm vi nghiên cứu điển hình tại mỏ đồng Sin Quyền.

7. Các luận điểm bảo vệ

Luận điểm 1:

Trong mỏ lộ thiên sâu, yếu tố hình học mỏ, sự phân tầng nhiệt và các đới quần gió có

thể làm quá trình khuếch tán $PM_{2.5}$ và CO trở nên phi tuyến và có tính lưu giữ theo thời gian. Việc kết hợp phân rã CEEMDAN, biến tương tác vi khí hậu và các biến trễ giúp xây dựng tập dữ liệu phản ánh tốt hơn cơ chế này, giảm ảnh hưởng của nhiễu cục bộ đến mô hình dự báo.

Luận điểm 2:

Trong bộ dữ liệu quan trắc tại mỏ đồng Sin Quyền, mô hình MPA-XGBoost và Stacking Ensemble đạt hiệu suất tốt hơn các mô hình đối chứng được đánh giá dựa trên các chỉ số hiệu suất R^2 , RMSE và MAE. Đặc biệt, MPA-XGBoost giảm hiện tượng dự báo thấp các đỉnh nồng độ so với RF khi chưa tối ưu, qua đó nâng cao giá trị cảnh báo đối với các giai đoạn nổ mìn và vận tải cường độ cao.

Luận điểm 3:

Khung quản lý môi trường tích hợp mô hình dự báo có thể dự báo trước về nồng độ $PM_{2.5}$ và CO tại từng trạm quan trắc, tạo điều kiện để điều chỉnh các hoạt động sản xuất cũng như các giải pháp giảm thiểu ô nhiễm môi trường không khí trong quá trình khai thác.

8. Ý nghĩa khoa học và thực tiễn

8.1. Ý nghĩa khoa học

- Luận án đề xuất cách tiếp cận số hóa đặc điểm phát thải và vi khí hậu mỏ lộ thiên sâu thông qua dữ liệu quan trắc phân tầng, CEEMDAN, biến tương tác và biến trễ. Cách tiếp cận này góp phần bảo toàn xu hướng vật lý của chuỗi dữ liệu $PM_{2.5}$ và CO trước khi huấn luyện mô hình.

- Luận án bổ sung bằng chứng thực nghiệm về khả năng ứng dụng MPA-XGBoost và Stacking Ensemble trong dự báo $PM_{2.5}$, CO tại mỏ lộ thiên sâu thông qua các chỉ số hiệu suất R^2 , RMSE, MAE và khả năng bám sát các cực trị của mô hình.

- Kết quả nghiên cứu cung cấp cơ sở để chuyển từ phương pháp đánh giá tĩnh, hậu kiểm sang phương pháp dự báo, hỗ trợ và chủ động phòng ngừa.

8.2. Ý nghĩa thực tiễn

- Kết quả dự báo chất lượng không khí nói chung và nồng độ $PM_{2.5}$ và CO nói riêng giúp mỏ chủ động điều hành sản xuất đảm bảo an toàn môi trường;

- Việc chuyển đổi từ phun nước dập bụi định kỳ sang dập bụi định lượng có mục tiêu giúp tiết kiệm trực tiếp lượng nước công nghiệp và chi phí nhiên liệu cho hệ thống thiết bị phụ trợ. Việc kiểm soát chặt chẽ nồng độ khí và bụi dưới đáy moong cũng giúp duy trì tầm nhìn an toàn cho thiết bị cơ giới, giảm thiểu rủi ro tai nạn lao động và bệnh nghề nghiệp cho công nhân;

- Khung giải pháp của luận án (từ quy trình quan trắc phân tầng đến lỗi dự báo MPA-XGBoost) có thể được tham khảo và chuyển giao cho các mỏ lộ thiên sâu có điều kiện tương đồng. Khi áp dụng tại mỏ khác cần đánh giá lại hình học moong, khí tượng, cơ cấu nguồn thải và tổ chức sản xuất; đồng thời thu thập dữ liệu địa phương, hiệu chỉnh hoặc huấn luyện lại mô hình và thực hiện kiểm chứng độc lập trước khi đưa vào vận hành.

9. Cơ sở dữ liệu

Cơ sở dữ liệu phục vụ luận án được tổ chức theo các nhóm sau:

- Dữ liệu sơ cấp từ ba trạm quan trắc tự động tại các mức +102 m, +68 m và -58 m, thu thập với bước thời gian 5 phút trong giai đoạn 10/4/2024-31/3/2025. Các trường dữ liệu gồm thời gian, $PM_{2.5}$, CO, nhiệt độ, độ ẩm, áp suất, tốc độ gió và hướng gió.

- Dữ liệu công nghệ và tổ chức sản xuất: thời điểm khoan - nổ mìn, chu kỳ giao ca, hoạt động xúc bốc - vận tải - đổ thải, danh mục và thời gian hoạt động của thiết bị; các dữ liệu này được dùng để đối chiếu sự kiện, không được coi là phép đo phát thải trực tiếp của từng thiết bị.

- Dữ liệu không gian và hồ sơ mỏ: bình đồ hiện trạng, mặt cắt theo cao độ, kết quả khảo sát hiện trường/UAV, kế hoạch khai thác và tài liệu môi trường của mỏ, phục vụ lựa chọn vị trí trạm và diễn giải điều kiện khuếch tán.

- Dữ liệu thứ cấp: các công trình khoa học, báo cáo quan trắc, báo cáo đánh giá tác động môi trường, tiêu chuẩn và quy chuẩn kỹ thuật có liên quan đến chất lượng không khí xung quanh và môi trường lao động.

- Kiểm soát chất lượng dữ liệu: đối chiếu hồ sơ CO, CQ và hồ sơ nghiệm thu thiết bị; đồng bộ dấu thời gian; kiểm tra tính liên tục, khoảng đo hợp lý và tính nhất quán giữa các trạm; gắn cờ dữ liệu thiếu, mất kết nối và ngoại lai trước khi xử lý.

- Giới hạn quyền sở hữu trí tuệ: hệ thống trạm quan trắc là sản phẩm của đề tài KC.05/21-25 do NCS tham gia chủ trì, đóng góp độc lập của luận án tập trung vào thiết kế tập dữ liệu nghiên cứu, kỹ thuật xử lý đặc trưng, tối ưu hóa mô hình dự báo và quy trình hỗ trợ ra quyết định.

10. Cấu trúc luận án

Luận án được trình bày trong 134 trang với 9 bảng và 33 hình minh họa. Ngoài phần mở đầu, kết luận, tài liệu tham khảo và phụ lục, cấu trúc của luận án bao gồm 4 chương, cụ thể như sau:

Mở đầu;

Chương 1: Tổng quan về kiểm soát môi trường không khí trong khai thác mỏ lộ thiên và các mô hình dự báo chất lượng không khí

Chương 2: Cơ sở khoa học và phương pháp nghiên cứu

Chương 3: Thu thập, phân tích dữ liệu và trích xuất đặc trưng tại mỏ đồng Sin Quyền, Lào Cai

Chương 4: Phát triển mô hình trí tuệ nhân tạo dự báo ô nhiễm không khí theo thời gian thực trong quá trình khai thác và đề xuất quy trình điều hành chủ động cho mỏ đồng Sin Quyền, Lào Cai

Kết luận và kiến nghị

Danh mục các công trình đã công bố liên quan đến luận án của NCS.

CHƯƠNG 1: TỔNG QUAN VỀ KIỂM SOÁT MÔI TRƯỜNG KHÔNG KHÍ TRONG KHAI THÁC MỎ LỘ THIÊN VÀ CÁC MÔ HÌNH DỰ BÁO

1.1. Đặc điểm phát thải và khuếch tán ô nhiễm tại mỏ lộ thiên sâu

Nồng độ ô nhiễm tại mỏ lộ thiên biến thiên dưới tác động đồng thời của nguồn phát thải, tổ chức ca kíp, hình học moong và khí tượng. Khoan – nổ mìn tạo nguồn điểm với đỉnh phát thải tức thời; xúc bốc tạo bụi tại gương khai thác; vận tải hình thành nguồn tuyến kéo dài theo đường mỏ và phát sinh cả bụi cuốn lẫn khí xả diesel; đổ thải tạo nguồn diện trên bề mặt bãi thải. Các nguồn này có thể chồng lấn, làm chuỗi dữ liệu có tính chu kỳ, phi tuyến và xuất hiện các cực trị ngắn hạn.

Khi mỏ xuống sâu, bờ tầng và địa hình dạng bậc thang cản trở dòng khí, làm tăng khả năng phân tầng nhiệt và hình thành xoáy quẩn. Bụi mịn có thể duy trì trạng thái lơ lửng, trong khi khí CO tích tụ dọc theo các tuyến vận tải hoặc khu vực thiết bị diesel làm việc với tải lớn.

Cơ chế lưu giữ cho thấy nồng độ tại thời điểm t phụ thuộc cả vào các trạng thái $t-1$, $t-3$, $t-6$ hoặc dài hơn; đây là cơ sở sử dụng biến trễ và mô hình chuỗi thời gian.

Bảng 1. Phân loại nguồn phát thải theo dây chuyền công nghệ khai thác

Khâu công nghệ	Dạng nguồn	Đặc tính chuỗi dữ liệu	Thiết bị/hoạt động chi phối
Khoan – nổ mìn	Nguồn điểm	Đỉnh phát thải đột biến, cường độ cao	Máy khoan, mạng nổ vi sai, loại và lượng thuốc nổ
Xúc bốc	Nguồn điểm	Duy trì nền bụi tại gương khai thác	Máy xúc điện và máy xúc thủy lực
Vận tải	Nguồn tuyến	Nền phát thải liên tục, có đỉnh sau giao ca	Ô tô tự đổ trọng tải lớn, đường dốc
Đổ thải	Nguồn diện	Duy trì nền ô nhiễm trên diện rộng	Ô tô dỡ tải và máy gạt

1.2. Thực trạng công tác quan trắc và dự báo môi trường không khí tại các mỏ lộ thiên sâu

Tổng quan nghiên cứu trong và ngoài nước cho thấy quan trắc thủ công có ba hạn chế chính: không bám sát chu kỳ phát thải; khó nhận diện quá trình lưu giữ bụi do nghịch nhiệt; và chưa cung cấp dữ liệu đủ nhanh cho điều hành. Các nghiên cứu hiện đại chuyển dần sang cảm biến IoT, trạm đa điểm, UAV, mô phỏng CFD, mô hình học máy và học sâu. Tuy nhiên, một số nghiên cứu chỉ sử dụng dữ liệu đô thị hoặc mỏ nông, chưa phản ánh sự phân tầng theo cao độ và sự chồng lấn của các khâu công nghệ trong mỏ sâu.

Tại Việt Nam, yêu cầu quan trắc và kiểm soát môi trường đã được quy định trong các văn bản kỹ thuật, nhưng việc kết nối dữ liệu liên tục với dự báo và quyết định điều hành còn hạn chế. Luận án vì vậy lựa chọn cách tiếp cận tích hợp: đo liên tục, xử lý dữ liệu theo thời gian, phát triển mô hình dự báo và chuyển kết quả thành khuyến nghị có điều kiện.

1.3. Thực trạng ứng dụng Trí tuệ nhân tạo (AI) và vấn đề tối ưu hóa trong quản lý môi trường tại các mỏ lộ thiên sâu

RF và XGBoost phù hợp với dữ liệu bảng, có khả năng biểu diễn quan hệ phi tuyến và tương tác giữa các biến. LSTM, TCN và AttnLSTM có ưu thế xử lý phụ thuộc theo thời gian, trong đó Attention giúp mô hình phân bổ trọng số lớn hơn cho các thời điểm quan trọng. Stacking Ensemble kết hợp dự báo của các mô hình cơ sở thông qua mô hình tổng hợp, nhằm giảm thiên lệch của từng mô hình riêng lẻ. MPA được sử dụng để tìm cấu hình siêu tham số tốt dựa trên RMSE của tập kiểm chứng.

Khoảng trống nghiên cứu được xác định ở bốn điểm: thiếu dữ liệu quan trắc phân tầng tại mỏ lộ thiên sâu; thiếu quy trình kỹ nghệ đặc trưng gắn với cơ chế tích tụ; thiếu đánh giá đồng thời khả năng dự báo trung bình và khả năng bắt đỉnh; và thiếu quy trình chuyển dự báo thành quyết định điều hành. Luận án tập trung giải quyết các khoảng trống này trong phạm vi nghiên cứu điển hình tại Sin Quyền.

1.4. Kết luận chương 1

Chương 1 xác lập cơ sở lựa chọn $PM_{2.5}$ và CO, phân loại nguồn phát thải theo khâu công nghệ, làm rõ vai trò của hình học moong và vi khí hậu, đồng thời chỉ ra nhu cầu kết hợp quan trắc liên tục với mô hình AI. Kết quả tổng quan là nền tảng xây dựng cơ sở khoa học và quy trình nghiên cứu ở Chương 2.

CHƯƠNG 2: CƠ SỞ KHOA HỌC VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đặc tính phát thải và cơ chế vi khí hậu tại mỏ đồng Sin Quyền

Động lực học nguồn thải được biểu diễn theo chu kỳ công nghệ. Nổ mìn tạo xung $PM_{2.5}$ có biên độ lớn; xúc bốc và vận tải duy trì nền phát thải, trong đó CO gắn chủ yếu với đội xe diesel; đổ thải đóng góp bụi trên diện rộng. Khi các khâu hoạt động đồng thời, việc quy một giá trị nồng độ cho duy nhất một thiết bị là không phù hợp. Luận án sử dụng dấu thời gian sự kiện sản xuất để đối chiếu và giải thích đỉnh nồng độ, nhưng không coi đó là phép đo phát thải trực tiếp của từng thiết bị.

Nghịch nhiệt được hiểu là trạng thái nhiệt độ lớp không khí phía trên cao hơn lớp phía dưới, làm suy yếu đối lưu thẳng đứng. Tại mặt mỏ +102 m, gió tự nhiên giúp khuếch tán nhanh; tại tầng +68 m, dòng khí chịu ảnh hưởng của sườn tầng và tuyến vận tải; tại đáy -58 m, không gian hẹp và quán gió làm tăng thời gian lưu giữ. Sự khác biệt này được chuyển thành yêu cầu lựa chọn biến trễ khác nhau và huấn luyện mô hình riêng cho từng trạm.

Bảng 2. Liên hệ giữa phân tầng không gian mỏ và yêu cầu xử lý chuỗi thời gian

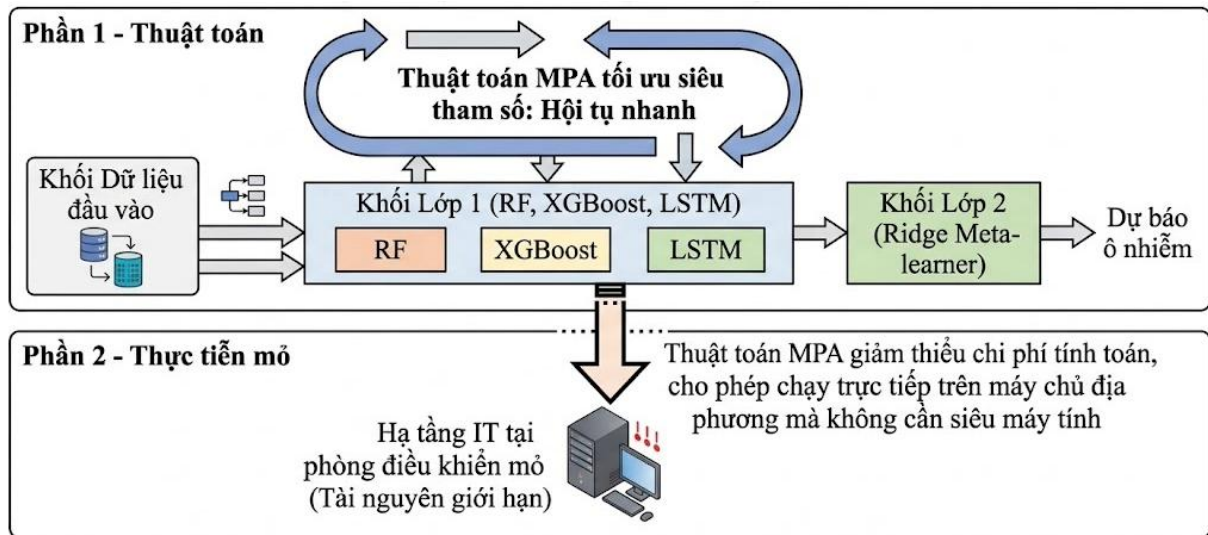
Cao độ	Đặc tính vi khí hậu	Khả năng khuếch tán	Dải thông tin quá khứ cần xem xét
+102 m	Thông gió tự nhiên tương đối ổn định	Nhanh, thời gian tồn lưu ngắn	Biến trễ ngắn 1–3 giờ
+68 m	Giao thoa dòng khí, quán gió cục bộ	Trung bình, có tích tụ ngắn hạn	Biến trễ 3–6 giờ
-58 m	Không gian sâu, nghịch nhiệt và quán gió	Thấp, có khả năng cộng dồn qua ca	Biến trễ dài hơn; đánh giá 6–24 giờ

2.2. Kỹ nghệ đặc trưng và cơ sở lựa chọn mô hình

Dữ liệu môi trường mỏ chứa dao động cục bộ do xe đi qua trạm, hoạt động thiết bị gần cảm biến, gián đoạn truyền tin và nhiễu đo. CEEMDAN được sử dụng để phân rã chuỗi thành các thành phần dao động nội tại ở nhiều thang thời gian, hỗ trợ tách nhiễu tần số cao khỏi xu hướng. CEEMDAN không phải lựa chọn duy nhất; các phương pháp EMD, EEMD, STL hoặc bộ lọc thông kê có thể thay thế, nhưng CEEMDAN được chọn vì phù hợp với chuỗi phi tuyến, không dừng và giảm hiện tượng trộn mode.

Biến tương tác khí tượng được hình thành từ vận tốc gió, hướng gió và định hướng tuyến vận tải/sườn tầng. Một vận tốc gió lớn không luôn làm giảm ô nhiễm: nếu gió thổi ngang vào bờ tầng, dòng khí có thể bị bóc tách và tạo xoáy quán. Biến tương tác cho phép mô hình học được sự khác biệt giữa thông gió thuận và thông gió nghịch. Các biến trễ 1, 3 và 6 giờ phản ánh xung phát thải ngắn, chu kỳ sau nổ mìn và tích tụ trong nửa ca; các cửa sổ dài hơn được xem xét đối với đáy moong.

RF và XGBoost được dùng làm mô hình dữ liệu bảng; LSTM và TCN mô hình hóa phụ thuộc theo thời gian; AttnLSTM tập trung vào các thời điểm có ảnh hưởng lớn; Stacking kết hợp các dự báo cơ sở bằng Ridge meta-learner. MPA cập nhật quần thể cấu hình siêu tham số thông qua chuyển động Brown và Lévy, cân bằng giữa tìm kiếm rộng và tinh chỉnh cục bộ. Trong luận án, “Prey” chỉ một cấu hình ứng viên, “Elite” là cấu hình tốt nhất tạm thời; các thuật ngữ sinh học không phải đại lượng vật lý của mỏ.



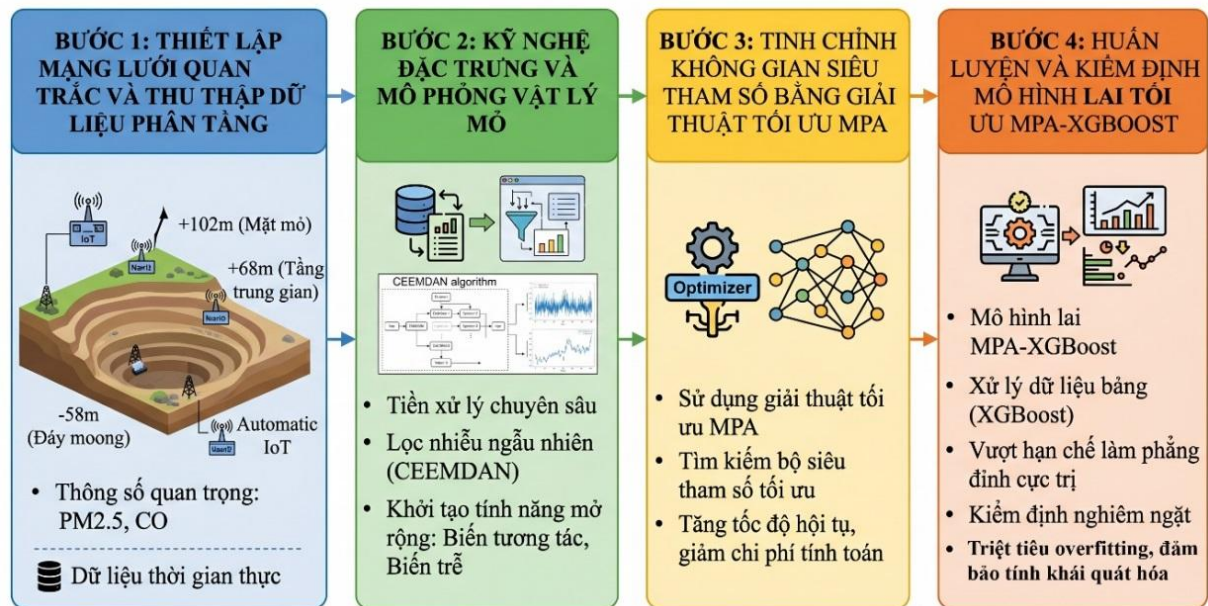
Hình 1. Cấu trúc Stacking Ensemble và cơ chế tối ưu hóa siêu tham số bằng MPA

2.3. Quy trình thực hiện nghiên cứu

Quy trình gồm bốn bước:

- (1) thiết lập mạng quan trắc và thu thập dữ liệu phân tầng;
- (2) tiền xử lý, phân rã, tạo biến tương tác và biến trễ;
- (3) tối ưu không gian siêu tham số bằng MPA trên tập kiểm chứng;
- (4) huấn luyện, kiểm thử và chuyển kết quả dự báo thành thông tin hỗ trợ điều hành.

Quy trình được thiết kế để bảo đảm tính lặp lại, kiểm soát rò rỉ dữ liệu và xác định rõ phạm vi hiệu lực của mô hình.



Hình 2. Quy trình thực hiện nghiên cứu của luận án

2.4. Kết luận chương 2

Chương 2 xây dựng cầu nối giữa cơ chế phát thải – vi khí hậu và cấu trúc dữ liệu – thuật toán. Cơ sở khoa học này định hướng cách bố trí trạm, lựa chọn biến đầu vào, thiết lập cửa sổ thời gian và đánh giá mô hình trong Chương 3 và Chương 4.

CHƯƠNG 3. THU THẬP, TIỀN XỬ LÝ VÀ CHUẨN HÓA, PHÂN TÍCH DỮ LIỆU VÀ TRÍCH XUẤT ĐẶC TRƯNG TẠI MỎ ĐỒNG SIN QUYỀN, LÀO CAI

3.1. Thiết kế mạng lưới quan trắc phân tầng

Về điều kiện địa hình và quy mô hình học, mỏ đồng Sin Quyền nằm trong khu vực đồi núi trung bình (cao độ tự nhiên khoảng +100 m đến +500 m), địa hình bị phân cắt mạnh với sườn dốc phổ biến 30° - 50° . Theo Giấy phép khai thác số 116/GP-BNNMT ngày 15/4/2026, khai trường tiếp tục xuống sâu đến mức -307 m; chiều dài khai trường theo phương khoảng 1.700 m, chiều rộng khoảng 1.400 m, trong đó moong khai thác quặng dài khoảng 600 m và rộng khoảng 100 m. Chênh cao lớn cùng cấu trúc bậc thang - lòng chảo làm điều kiện thông gió, thời gian lưu giữ và khả năng phát tán chất ô nhiễm thay đổi rõ theo cao độ. Vì vậy, quy mô và hình học moong là một trong các cơ sở trực tiếp để lựa chọn các tầng quan trắc đại diện.

Mỏ đồng Sin Quyền có địa hình dạng bậc thang, đáy moong sâu và tuyến vận tải kéo dài. Ba trạm được bố trí theo nguyên tắc đại diện cho ba vùng khí động học: Trạm 1 tại +102 m phản ánh môi trường nền và điều kiện thông gió tốt; Trạm 2 tại +68 m đại diện vùng giao thoa dòng khí và hoạt động vận tải; Trạm 3 tại -58 m đại diện đáy moong, nơi dễ xuất hiện nghịch nhiệt, quẩn gió và đỉnh phát thải. Vị trí được lựa chọn trên cơ sở bình đồ, mặt cắt, khảo sát hiện trường, an toàn lắp đặt, khả năng cấp nguồn và truyền dữ liệu.



Hình 3. Mỏ đồng Sin Quyền và vị trí ba trạm quan trắc tự động

Bảng 3. Đặc điểm kỹ thuật và chức năng của các vị trí quan trắc

Trạm	Cao độ	Đặc điểm khí động học	Nguồn phát thải đặc trưng	Vai trò trong mô hình
1	+102 m	Thông gió tự nhiên tương đối ổn định	Môi trường nền, ảnh hưởng sau nổ mìn	Tham chiếu nồng độ cơ sở
2	+68 m	Giao thoa dòng khí và tuyến vận tải	Bụi cuốn, khí xả vận tải	Đặc trưng nguồn tuyến
3	-58 m	Nghịch nhiệt, quần gió cục bộ	Đỉnh bụi và khí xả thiết bị	Biến trở và dự báo cực trị

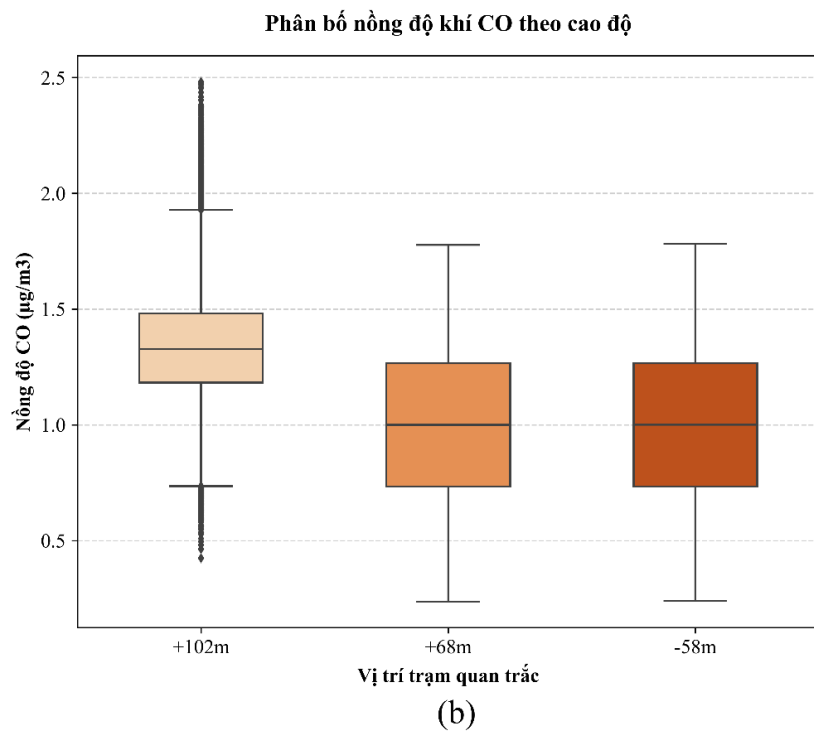
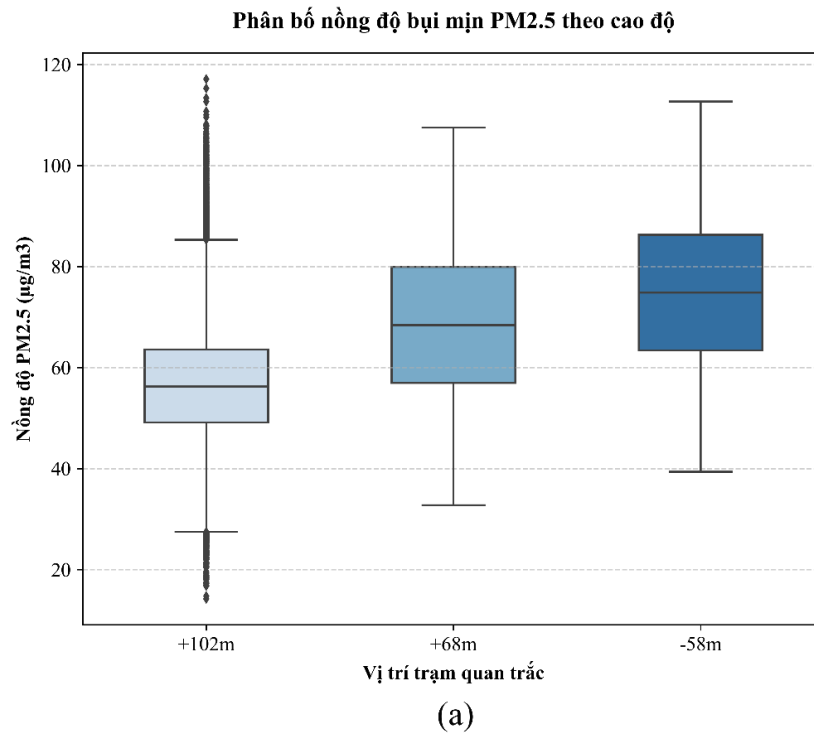
3.2. Phân tích thống kê chuỗi dữ liệu

Dữ liệu được ghi với bước 5 phút. Thống kê mô tả cho thấy PM_{2.5} trung bình tăng từ 56,69 µg/m³ tại +102 m lên 68,60 µg/m³ tại +68 m và 75,06 µg/m³ tại -58 m. Giá trị PM_{2.5} tối thiểu tại đáy moong cao hơn đáng kể so với mặt mô, thể hiện nền ô nhiễm được duy trì. Nhiệt độ trung bình tăng theo chiều sâu trong bộ dữ liệu khảo sát, trong khi tốc độ gió tại các tầng sâu thấp hơn; kết quả phù hợp với nhận định về khả năng lưu giữ ô nhiễm.

CO có giá trị trung bình khoảng 1,00–1,34 mg/m³, với độ lệch chuẩn lớn hơn tương đối tại hai trạm sâu. Sự phân bố CO khác PM_{2.5} vì CO tập trung theo tuyến vận tải và khu vực động cơ diesel, trong khi PM_{2.5} còn chịu ảnh hưởng của bụi vách, nổ mìn và đổ thải. Do đó, mô hình được huấn luyện riêng cho từng biến đích và từng trạm.

Bảng 4. Thống kê mô tả các thông số môi trường tại ba mức cao độ

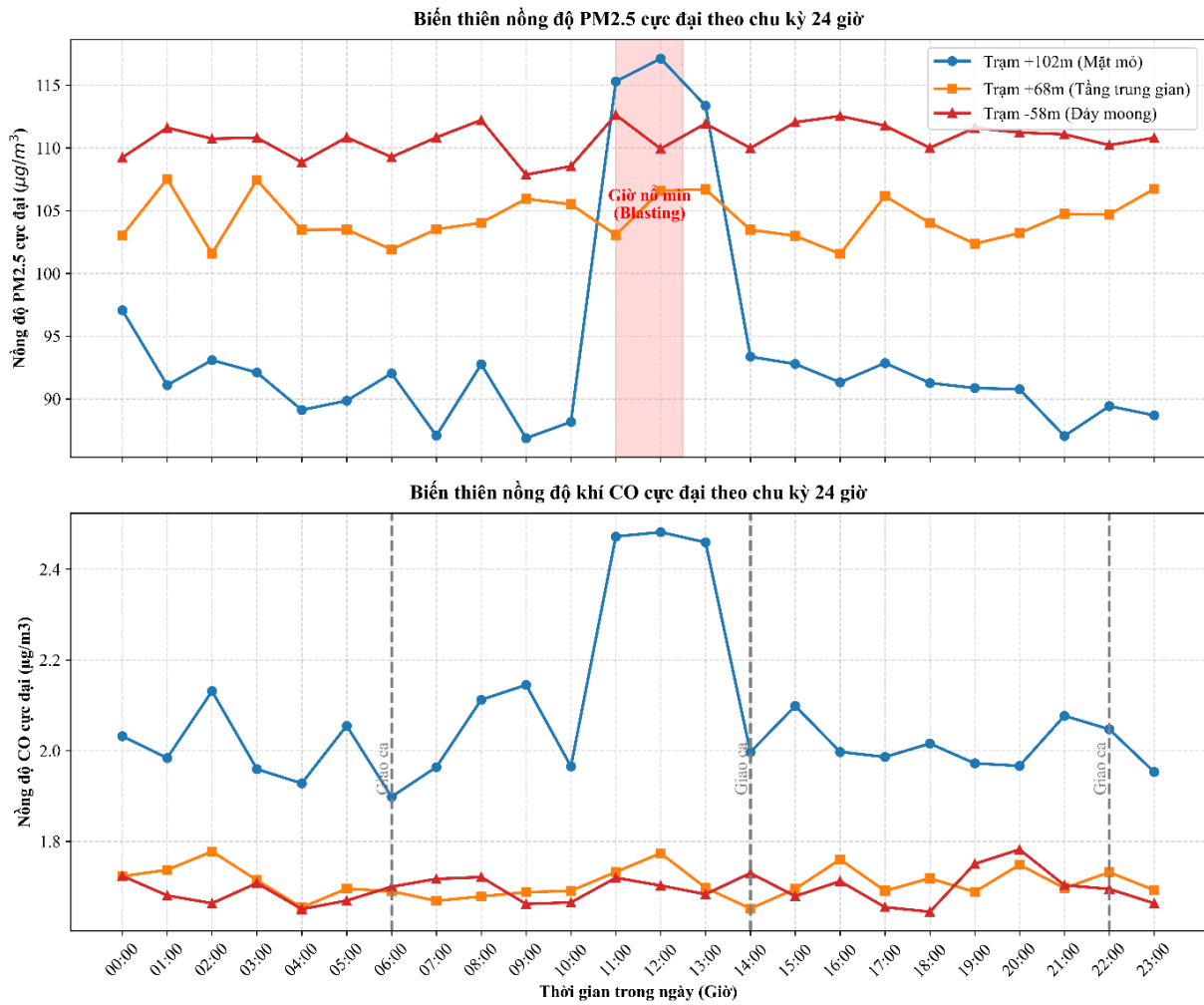
Thông số	Thống kê	+102 m	+68 m	-58 m
PM _{2.5} (µg/m ³)	Trung bình	56,69	68,60	75,06
PM _{2.5} (µg/m ³)	Độ lệch chuẩn	11,26	13,58	13,67
PM _{2.5} (µg/m ³)	Lớn nhất	117,12	107,53	112,64
PM _{2.5} (µg/m ³)	Nhỏ nhất	14,21	32,77	39,34
CO (mg/m ³)	Trung bình	1,34	1,00	1,00
CO (mg/m ³)	Độ lệch chuẩn	0,23	0,30	0,30
Nhiệt độ (°C)	Trung bình	17,69	19,72	21,20
Độ ẩm (%)	Trung bình	76,60	80,00	80,00
Tốc độ gió (m/s)	Trung bình	1,51	1,32	1,32



Hình 4. Phân bố PM_{2.5} và CO theo cao độ quan trắc

3.2.2. Nhận diện sự kiện phát thải đột biến theo chu kỳ ca kíp

Đối chiếu dấu thời gian cho thấy nồng độ PM_{2.5} thường tăng mạnh quanh khung giờ nổ mìn; tại +102 m, đỉnh giảm nhanh hơn nhờ thông gió, trong khi tại +68 m và -58 m thời gian suy giảm kéo dài. Sau giao ca, việc khởi động đồng loạt xe tải và máy xúc diesel có thể tạo các đỉnh phụ của CO và PM_{2.5}. Việc gán nguyên nhân được thực hiện bằng đối chiếu nhật ký sản xuất, điều kiện khí tượng và diễn biến đồng thời tại các trạm; luận án không khẳng định mọi đỉnh đều do duy nhất một khâu công nghệ.



Hình 5. Biến thiên cực đại của PM_{2.5} và CO theo chu kỳ 24 giờ

3.3. Xử lý dữ liệu và cấu trúc tập đặc trưng

Bước xử lý gồm kiểm tra tính liên tục, gán cờ dữ liệu mất kết nối, loại hoặc hiệu chỉnh giá trị ngoài khoảng hợp lý, làm mượt bằng cửa sổ trượt, phân rã CEEMDAN và tạo biến trễ. Trung bình trượt 3 giờ mô tả xu hướng sau nổ mìn; cửa sổ 6 giờ phản ánh tích tụ trong nửa ca. Chuỗi PM_{2.5} và CO đều được xử lý tương tự, nhưng tham số được lựa chọn riêng để tránh làm mất các đỉnh có ý nghĩa.

Biến trễ được tạo từ chính PM_{2.5}/CO quá khứ và các biến khí tượng. Dữ liệu được sắp xếp theo thời gian, chia 70% huấn luyện, 15% kiểm chứng và 15% kiểm thử. MinMaxScaler chỉ được fit trên tập huấn luyện rồi áp dụng cho hai tập còn lại. Quy trình này ngăn thông tin tương lai đi vào huấn luyện và bảo đảm kết quả kiểm thử phản ánh đúng khả năng tổng quát hóa.

3.4. Kết luận chương 3

Chương 3 xây dựng được bộ dữ liệu thực chứng phân tầng, làm rõ sự khác biệt giữa ba cao độ và mối liên hệ giữa dữ liệu sản xuất với chuỗi PM_{2.5}, CO. Bộ dữ liệu sau tiền xử lý là đầu vào cho phát triển mô hình tại Chương 4. Giới hạn quan trọng là dữ liệu đại diện cho ba điểm đo, chưa thay thế bản đồ nồng độ 3D của toàn bộ moong.

CHƯƠNG 4. PHÁT TRIỂN MÔ HÌNH TRÍ TUỆ NHÂN TẠO DỰ BÁO Ô NHIỄM KHÔNG KHÍ THEO THỜI GIAN THỰC VÀ ĐỀ XUẤT QUY TRÌNH ĐIỀU HÀNH CHỦ ĐỘNG CHO MỎ ĐỒNG SİN QUYỀN

4.1. Thiết lập cấu trúc hệ thống mô hình dự báo

Hệ thống gồm nhóm mô hình dữ liệu bảng RF và XGBoost; nhóm mô hình chuỗi thời gian LSTM, TCN, AttnLSTM; hai mô hình tối ưu MPA-RF và MPA-XGBoost; và Stacking Ensemble. Mỗi mô hình được huấn luyện riêng cho từng trạm và từng biến đích. RF cung cấp mô hình đối chứng dễ diễn giải; XGBoost khai thác tương tác phi tuyến; LSTM/TCN/AttnLSTM xử lý phụ thuộc thời gian; Stacking kết hợp dự báo của nhiều cấu trúc.

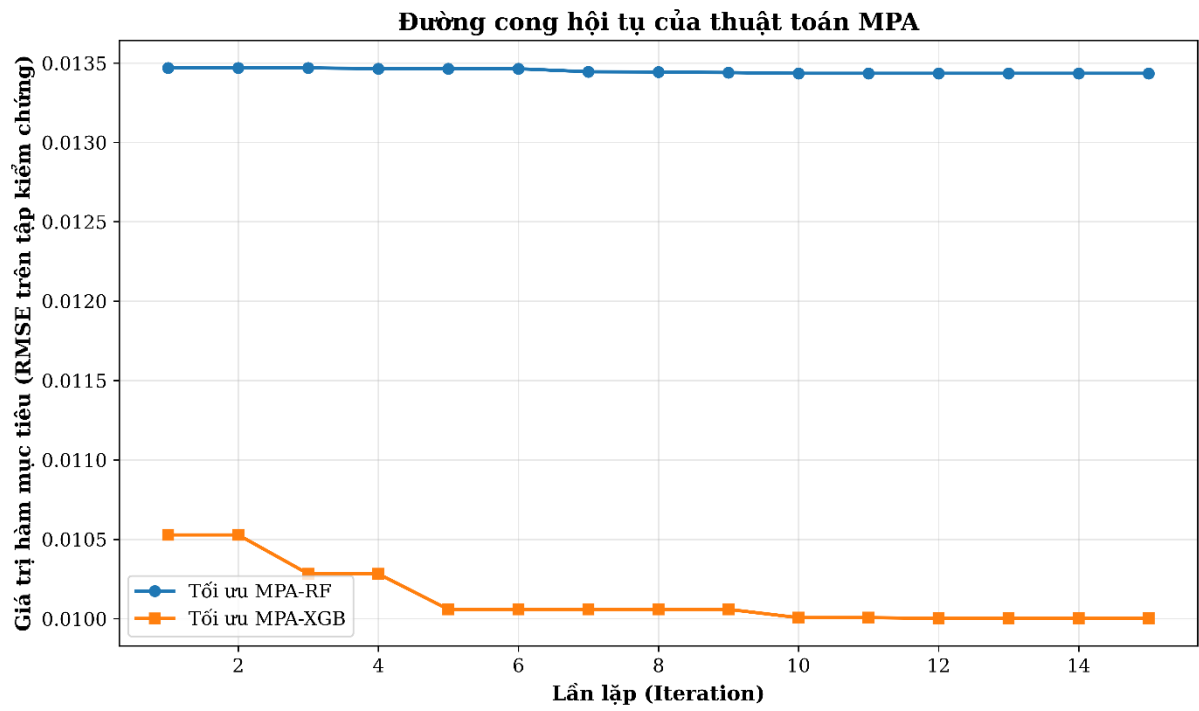
Bảng 5. Vai trò của các mô hình được phát triển và so sánh

Mô hình	Vai trò chính	Điểm mạnh trong bài toán	Hạn chế cần kiểm soát
RF	Đối chứng học máy	Ổn định, xử lý dữ liệu bảng, đánh giá tầm quan trọng biến	Có thể làm phẳng đỉnh và quá khớp
XGBoost	Mô hình tăng cường cây	Sai số thấp, xử lý phi tuyến và tương tác	Nhạy với siêu tham số
LSTM	Chuỗi thời gian	Ghi nhớ phụ thuộc dài hạn	Hiệu quả phụ thuộc lượng dữ liệu và cấu hình
TCN	Tích chập thời gian	Trường tiếp nhận rộng, huấn luyện song song	Có thể kém ổn định tại dữ liệu nhiều nhiễu
AttnLSTM	LSTM + Attention	Tập trung vào thời điểm quan trọng	Cấu trúc phức tạp, chi phí huấn luyện cao
MPA-RF/MPA-XGB	Mô hình tối ưu	Tìm siêu tham số theo RMSE kiểm chứng	Không bảo đảm mọi chỉ tiêu kiểm thử đều tốt hơn mô hình gốc
Stacking	Học kết hợp	Bù trừ sai số của các mô hình cơ sở	Phụ thuộc tính đa dạng của mô hình thành phần

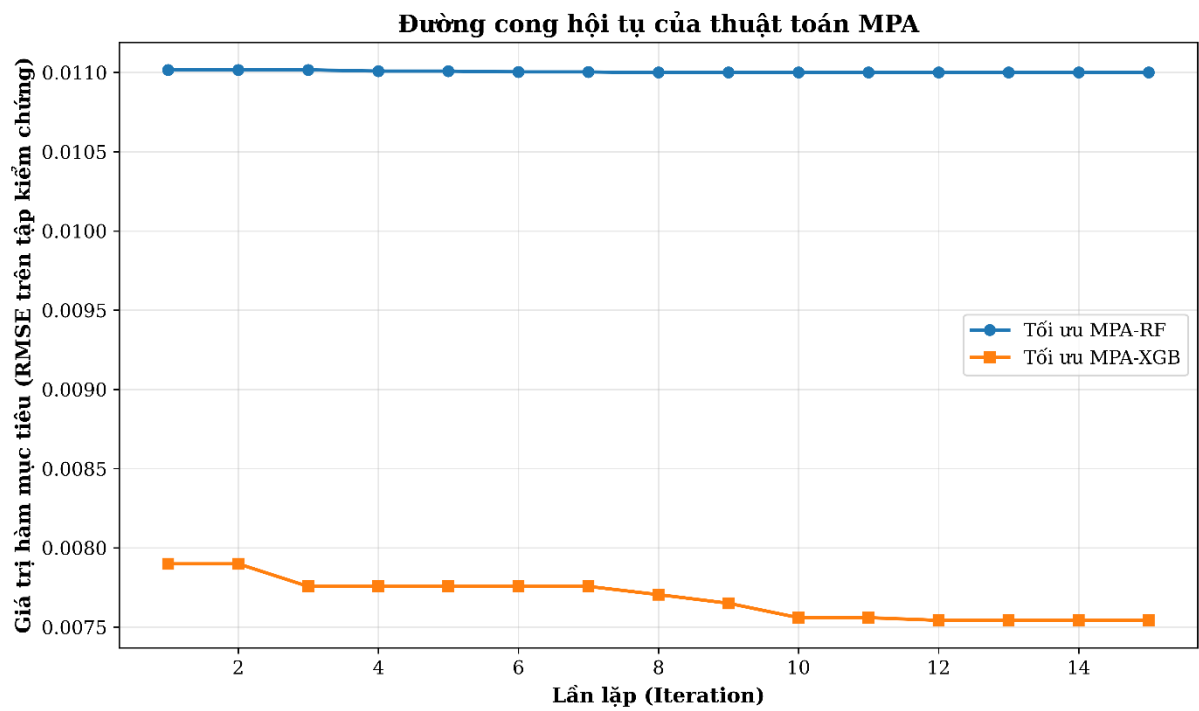
4.2. Tối ưu hóa bằng MPA và phát triển mô hình lai

Không gian tìm kiếm của RF/XGBoost bao gồm số cây, độ sâu cây, tỷ lệ học, tỷ lệ lấy mẫu và các tham số điều chuẩn. MPA khởi tạo nhiều cấu hình ứng viên, đánh giá bằng RMSE trên tập kiểm chứng và cập nhật theo ba giai đoạn tìm kiếm. Đường cong hội tụ cho thấy hàm mục tiêu giảm nhanh trong các vòng đầu và ổn định ở giai đoạn sau. Sau khi chọn cấu hình, mô hình được huấn luyện lại trên dữ liệu huấn luyện kết hợp kiểm chứng và đánh giá một lần trên tập kiểm thử.

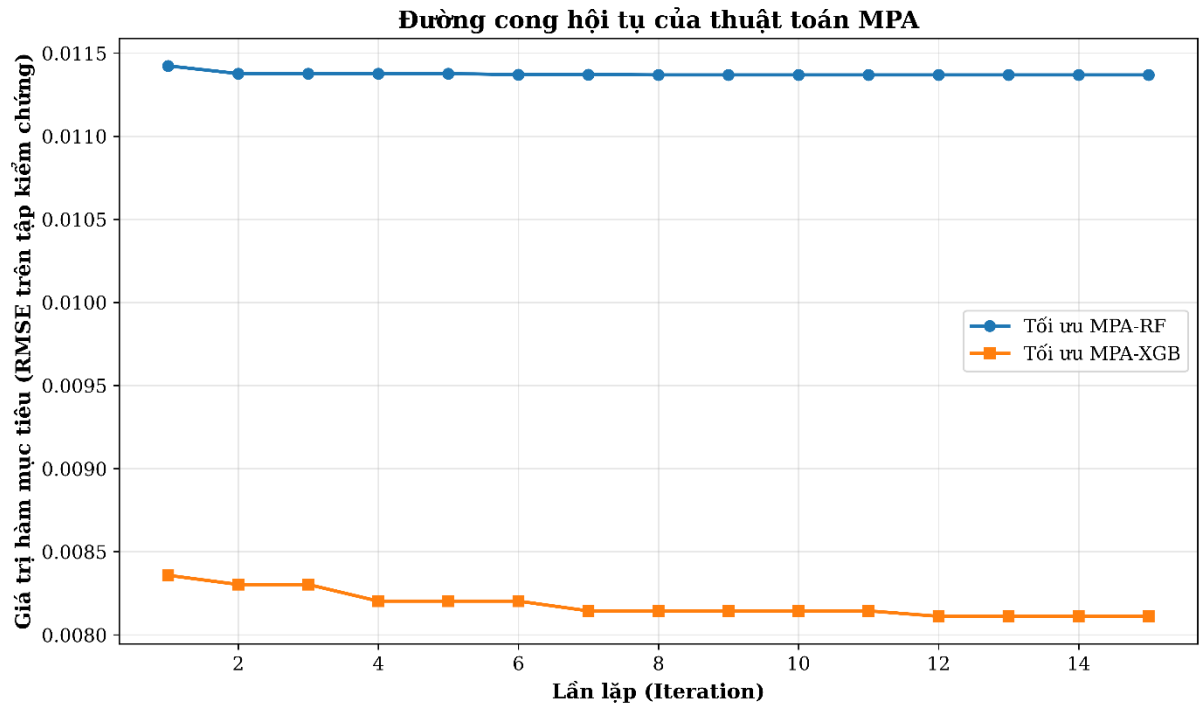
Kết quả cho thấy XGBoost và MPA-XGBoost là nhóm mô hình ổn định nhất trên cả ba trạm. MPA-XGBoost có lợi thế rõ hơn ở giai đoạn tối ưu/kiểm chứng và khả năng theo sát cực trị; trên tập kiểm thử, chênh lệch giữa XGBoost và MPA-XGBoost ở một số trường hợp rất nhỏ, thậm chí XGBoost có RMSE hoặc MAE thấp hơn nhẹ. Vì vậy, đóng góp của MPA được diễn giải là cải thiện quy trình tìm kiếm cấu hình, độ ổn định và năng lực bắt đỉnh, không phải bảo đảm tuyệt đối mọi chỉ tiêu đều tốt hơn mô hình gốc.



(a) Với dữ liệu tại trạm +102m



(b) Với dữ liệu tại trạm +68m

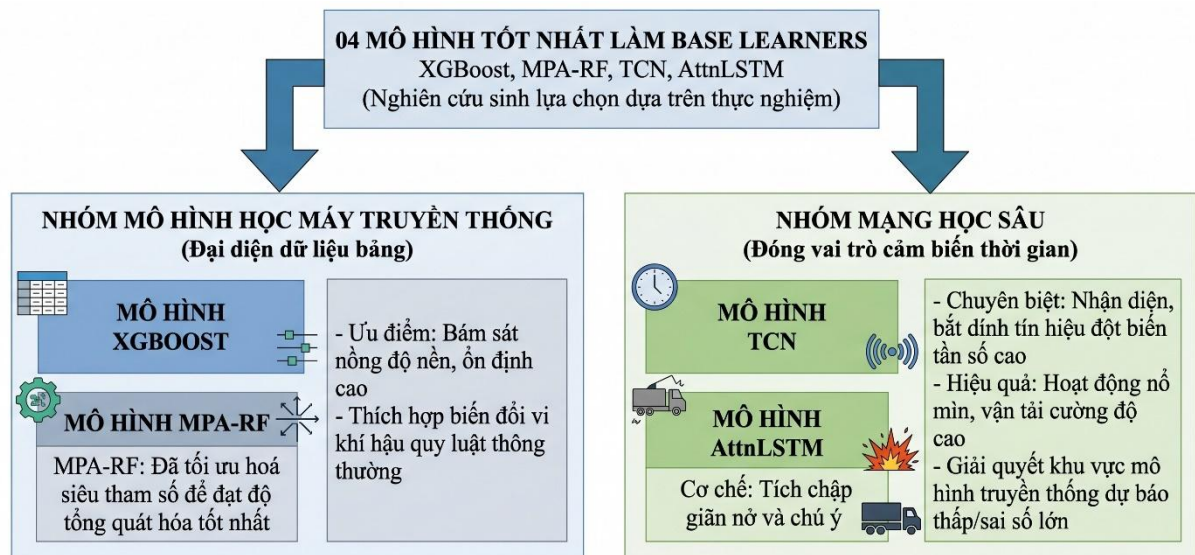


(c) Với dữ liệu tại trạm -58m

Hình 6. Biểu đồ đường cong hội tụ của giải thuật tối ưu hóa MPA khi tối ưu mô hình RF và XGBoost

4.3. Phát triển Stacking Ensemble

Stacking sử dụng các dự báo ngoài mẫu của RF, XGBoost, TCN và AttnLSTM làm đầu vào cho Ridge meta-learner. Cấu trúc này giúp tổng hợp ưu thế của mô hình cây và mạng chuỗi thời gian. Stacking đạt R^2 cao đối với $PM_{2.5}$ ở cả ba trạm và thể hiện độ ổn định tốt, nhưng không luôn có sai số thấp nhất. Kết quả khẳng định cần lựa chọn mô hình theo nhiều chỉ tiêu và mục tiêu vận hành, thay vì chỉ dựa vào một hệ số.



Hình 7. Sơ đồ phân nhóm và vai trò của 4 mô hình cơ sở được lựa chọn để phát triển mô hình stacking ensemble dự báo chất lượng không khí cho mỏ đồng Sin Quyền

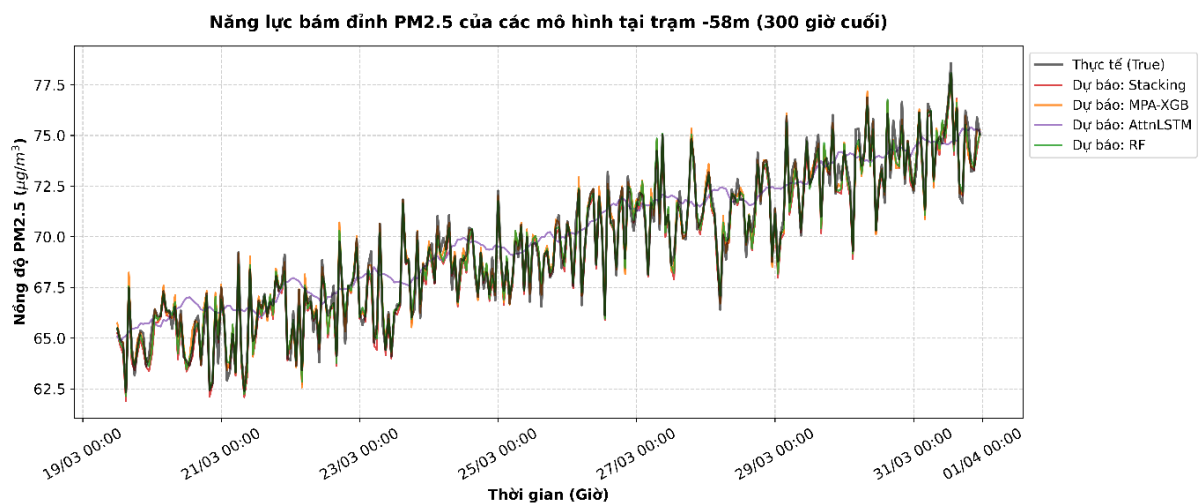
4.4. Đánh giá hiệu năng và độ tin cậy

Tại trạm +102 m, XGBoost và MPA-XGBoost đạt R^2 kiểm thử 0,996 đối với $PM_{2.5}$ và 0,997 đối với CO. Tại +68 m, hai mô hình đạt R^2 khoảng 0,997 đối với $PM_{2.5}$ và xấp xỉ 1,000 đối với CO. Tại -58 m, XGBoost/MPA-XGBoost đạt R^2 khoảng 0,997 đối với $PM_{2.5}$ và 1,000 đối với CO. Các mạng học sâu đơn lẻ có kết quả thấp hơn đối với $PM_{2.5}$ tại tầng trung gian và đáy moong, đặc biệt TCN giảm mạnh ở -58 m; điều này cho thấy dữ liệu bằng đã được kỹ nghệ đặc trưng phù hợp có thể hiệu quả hơn cấu trúc học sâu phức tạp trong bộ dữ liệu hiện tại.

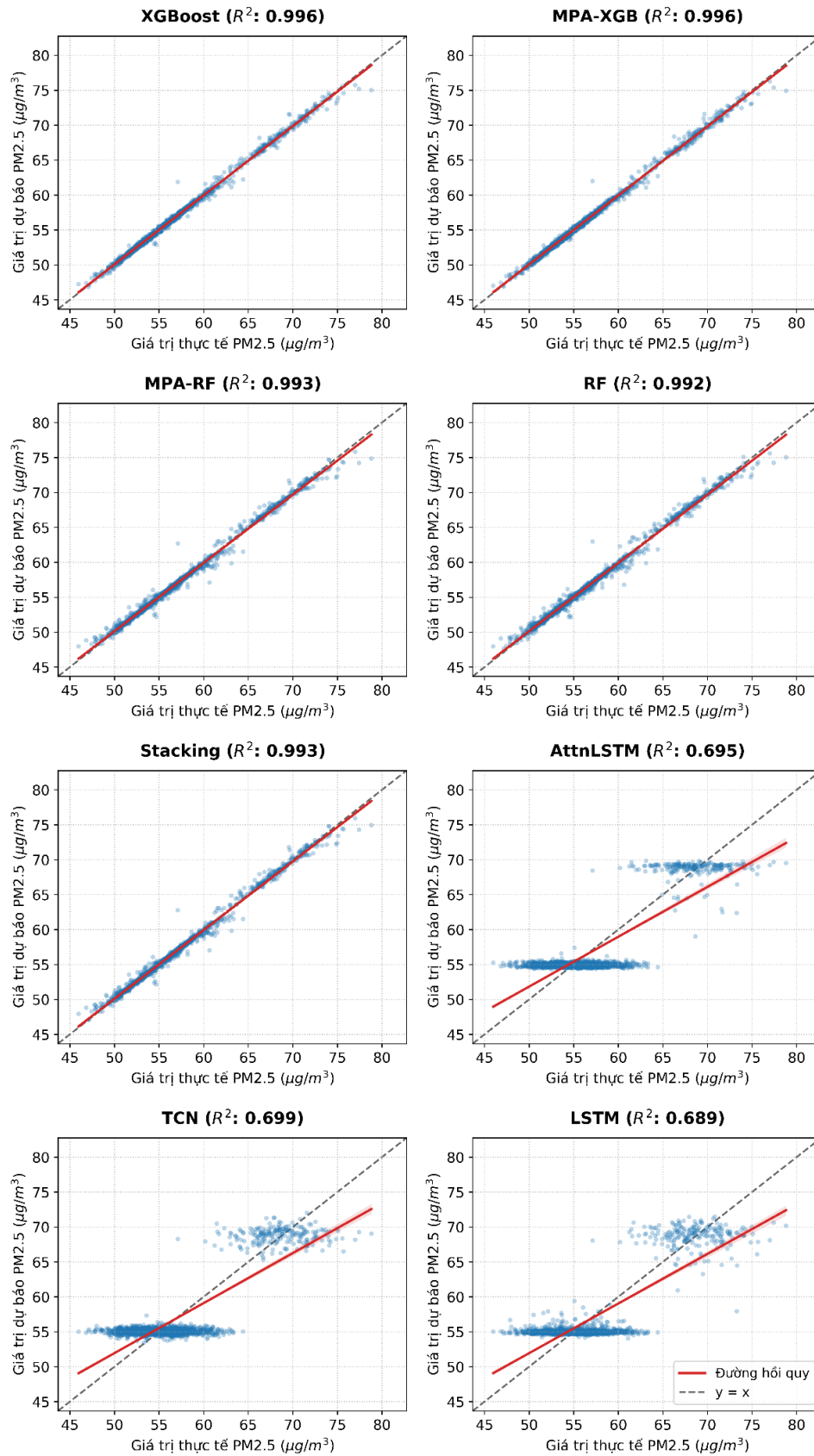
R^2 chỉ phản ánh tỷ lệ biến thiên được giải thích và phải được đọc cùng RMSE, MAE. Phân tích tương quan giá trị thực – dự báo và lát cắt 300 giờ cuối cho thấy MPA-XGBoost và XGBoost theo sát phần lớn dao động; RF và AttnLSTM có xu hướng làm mượt ở một số đoạn. Khả năng bắt đỉnh có ý nghĩa đặc biệt đối với cảnh báo sau nổ mìn và giai đoạn vận tải cường độ cao.

Bảng 6. Kết quả kiểm thử tiêu biểu của nhóm mô hình có hiệu suất cao

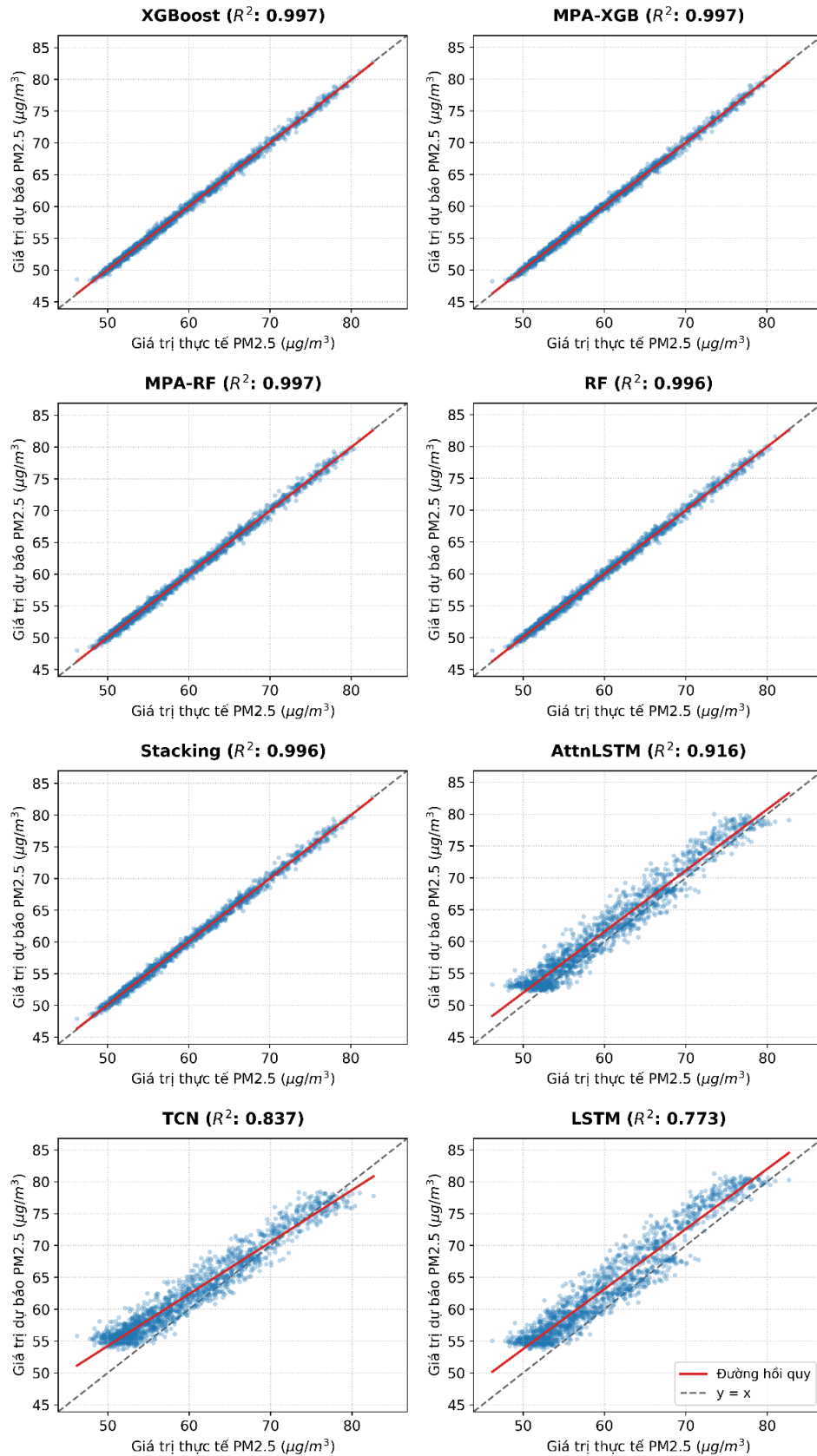
Trạm/biến	Nhóm mô hình có hiệu suất cao	RMSE kiểm thử tiêu biểu	MAE kiểm thử tiêu biểu	R^2 kiểm thử
+102 m – $PM_{2.5}$	XGBoost / MPA-XGBoost	0,346	0,210–0,217	0,996
+102 m – CO	XGBoost / MPA-XGBoost	0,007	0,004–0,005	0,997
+68 m – $PM_{2.5}$	XGBoost / MPA-XGBoost	0,414–0,416	0,325–0,328	0,997
+68 m – CO	XGBoost / MPA-XGBoost	0,005	0,004	$\approx 1,000$
-58 m – $PM_{2.5}$	XGBoost / MPA-XGBoost	0,428–0,440	0,342–0,353	0,997
-58 m – CO	XGBoost / MPA-XGBoost	0,006	0,004	$\approx 1,000$



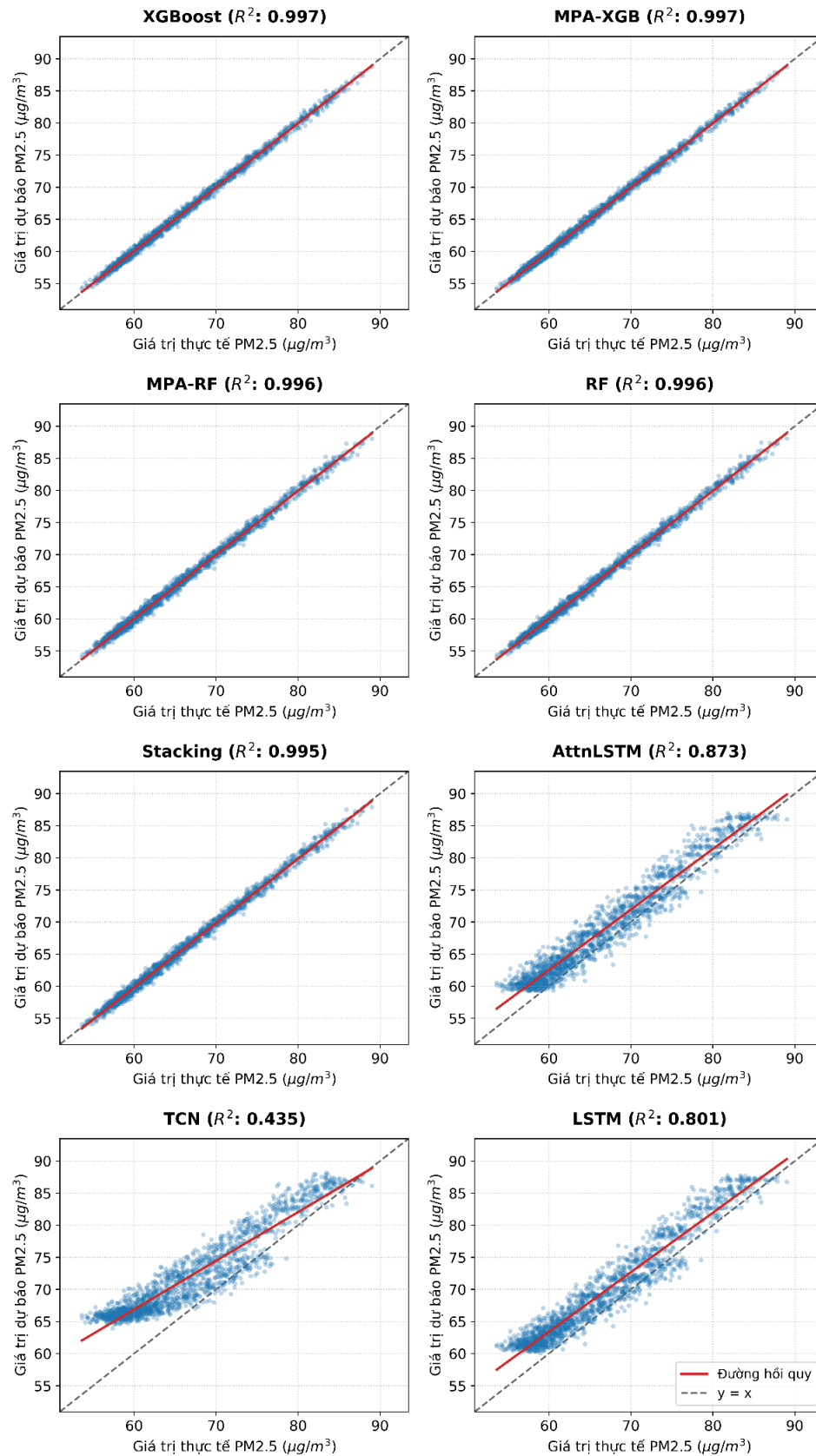
Hình 8. Khả năng theo sát các đỉnh $PM_{2.5}$ tại trạm -58 m trong 300 giờ kiểm thử cuối



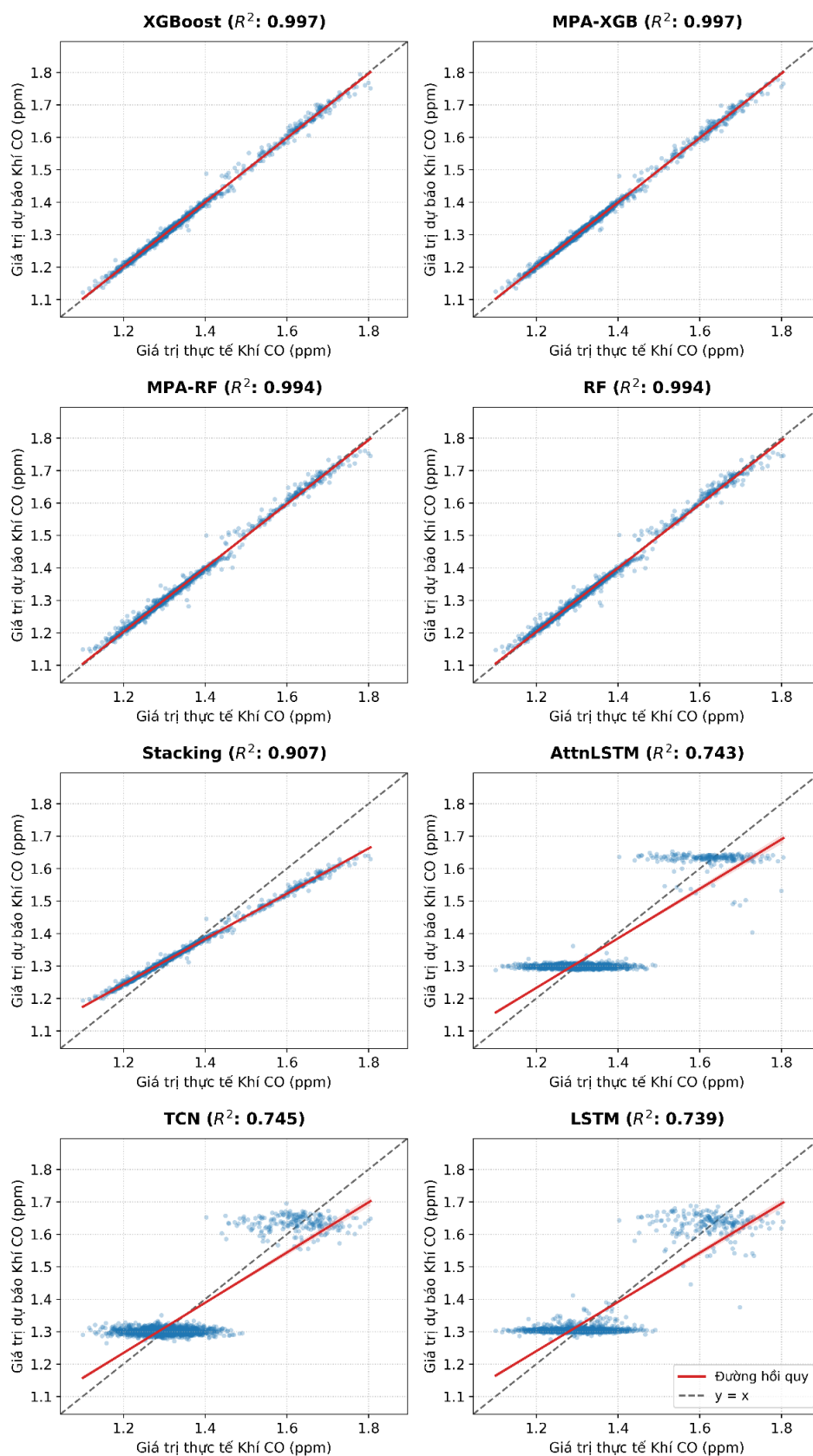
(a) Tương quan giữa giá trị thực tế và giá trị dự báo của PM_{2.5} tại trạm +102m



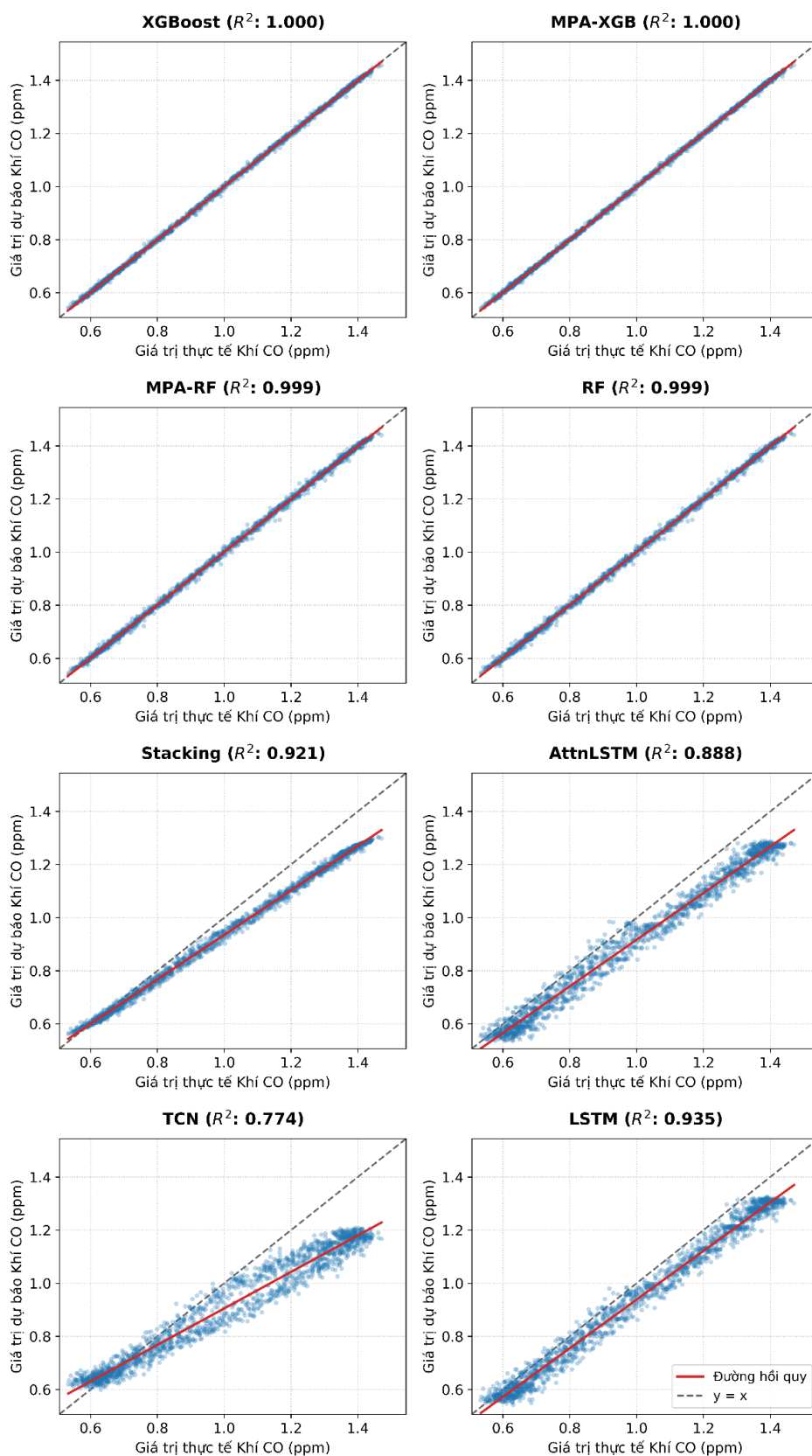
(b) Tương quan giữa giá trị thực tế và giá trị dự báo của PM_{2.5} tại trạm +68m



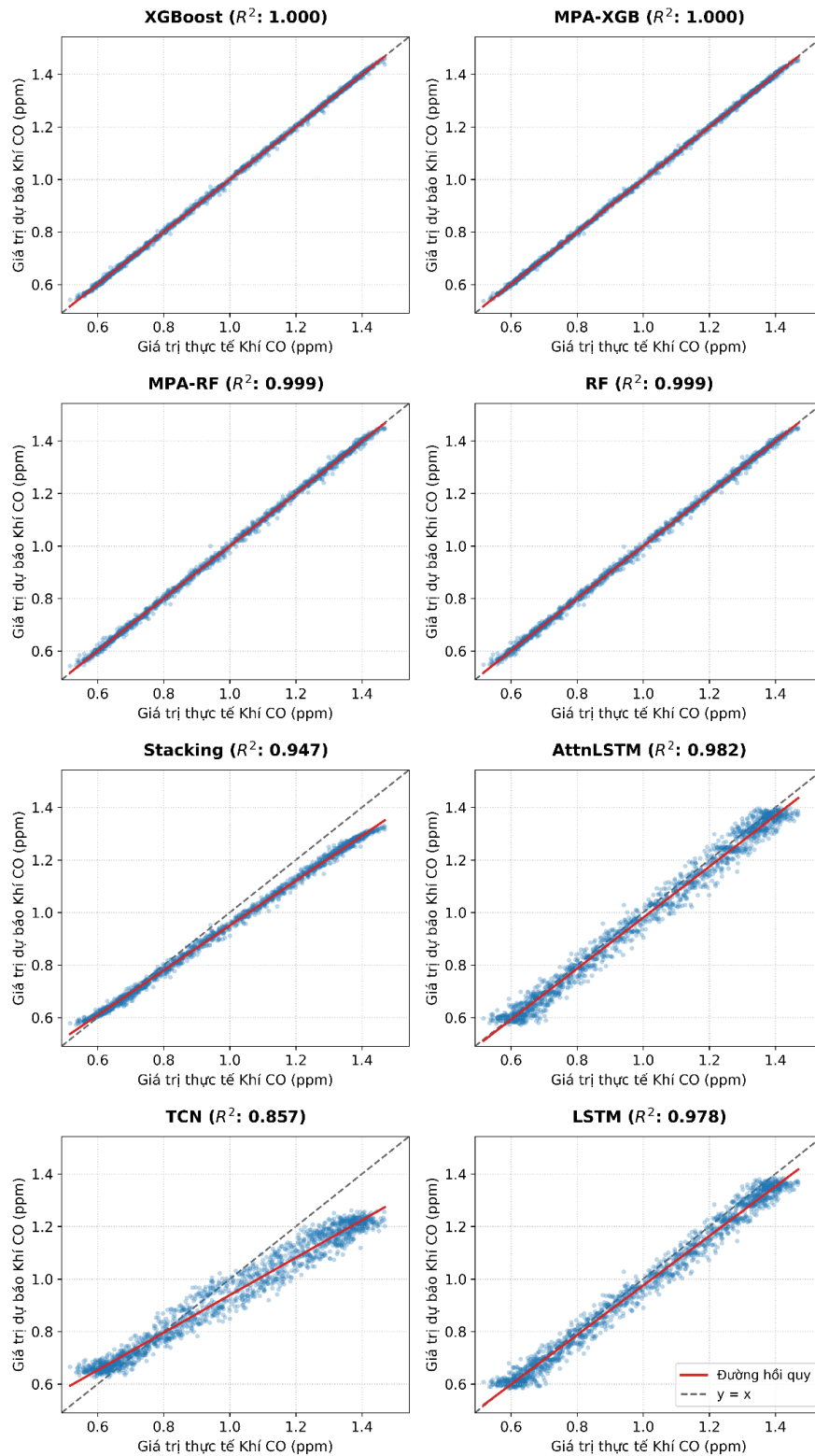
(c) Tương quan giữa giá trị thực tế và giá trị dự báo của PM_{2.5} tại trạm -58m



(d) Tương quan giữa giá trị thực tế và giá trị dự báo của khí CO tại trạm +102m



(e) Tương quan giữa giá trị thực tế và giá trị dự báo của khí CO tại trạm +68m



(f) Tương quan giữa giá trị thực tế và giá trị dự báo của khí CO tại trạm -58m

Hình 9. Tương quan giữa giá trị thực tế và giá trị dự báo của bụi $PM_{2.5}$ và khí CO tại các trạm

4.5. Quy trình hỗ trợ ra quyết định điều hành chủ động

Kết quả dự báo không được sử dụng để điều khiển thiết bị hoàn toàn tự động. Hệ thống thực hiện bốn bước: kiểm tra chất lượng dữ liệu; dự báo nồng độ theo chân trời 1–3 giờ; so sánh với ngưỡng phù hợp về thời gian trung bình và phạm vi áp dụng; phát hành khuyến nghị để cán bộ có thẩm quyền xem xét. Khi dự báo tăng tại +68 m, ưu tiên tăng dập bụi và

giãn mật độ xe trên tuyến liên quan; khi đáy -58 m có xu hướng tích tụ, cần hạn chế người vào vùng ảnh hưởng, điều tiết xúc bốc – vận tải và rà soát hoạt động phát thải cao; trước nổ mìn, kết quả dự báo được kết hợp với khí tượng để quyết định thời điểm và phạm vi kiểm soát.

Bảng 7. Ma trận mức cảnh báo và khuyến nghị điều hành

Mức	Điều kiện kích hoạt	Khuyến nghị điều hành	Trách nhiệm
Bình thường	Dữ liệu hợp lệ; dự báo dưới ngưỡng tiền cảnh báo nội bộ	Duy trì sản xuất, tiếp tục quan trắc và ghi nhật ký	Phòng điều hành giám sát
Theo dõi	Nồng độ tiến gần ngưỡng, xu thế tăng hoặc gió bất lợi	Tăng tần suất kiểm tra; tăng dập bụi tại đúng tầng/tuyến; giãn mật độ xe,...	An toàn – môi trường phối hợp điều hành
Cảnh báo	Dự báo đạt/vượt ngưỡng sau khi quy đổi đúng thời gian trung bình	Kích hoạt thông gió cưỡng bức; Tăng biện pháp xử lý và điều hành sản xuất	Lãnh đạo ca hoặc người được ủy quyền phê duyệt
Lỗi dữ liệu	Mất kết nối, cảm biến bất thường hoặc mô hình ngoài phạm vi hiệu lực	Không dùng dự báo để điều khiển; chuyển sang thiết bị dự phòng/đo thủ công; khắc phục và xác nhận dữ liệu	Kỹ thuật đo lường và an toàn - môi trường

Để duy trì hệ thống, luận án đề xuất quy trình dự phòng: khoảng mất dữ liệu ngắn được gán cờ và chỉ diễn khi không qua sự kiện nổ mìn; mất kết nối kéo dài phải dừng khuyến nghị tự động và chuyển sang đo dự phòng; khi hình học mỏ hoặc đội thiết bị thay đổi cần đánh giá trôi dữ liệu, huấn luyện lại và kiểm thử độc lập. Mỗi phiên bản mô hình phải lưu tập dữ liệu, cấu hình, phiên bản thư viện, chỉ tiêu kiểm thử và người phê duyệt triển khai.

4.6. Kết luận chương 4

Chương 4 đã phát triển và kiểm chứng các mô hình dự báo trên dữ liệu thực đo tại ba mức +102 m, +68 m và -58 m của mỏ đồng Sin Quyền. Nhóm mô hình cây, đặc biệt XGBoost và MPA-XGBoost, đạt R^2 kiểm thử cao và sai số nhỏ ở nhiều tổ hợp $PM_{2.5}/CO$; một số mô hình học sâu cho hiệu suất thấp hơn tại các chuỗi phức tạp. MPA-XGBoost được lựa chọn làm mô hình cốt lõi dựa trên tổng hợp R^2 , RMSE, MAE, độ ổn định giữa các tập dữ liệu và khả năng theo sát cực trị, thay vì chỉ dựa vào một chỉ số riêng lẻ.

Trong lát cắt kiểm thử, MPA-XGBoost làm giảm hiện tượng dự báo thấp/làm phẳng một số đỉnh nồng độ so với RF chưa tối ưu, đồng thời duy trì chuỗi dự báo ổn định ở vùng nồng độ nền. Kết quả cho thấy mô hình có tiềm năng cung cấp thông tin dự báo cho khung hỗ trợ điều hành chủ động tại mỏ đồng Sin Quyền. Mức độ tin cậy khi áp dụng ngoài phạm vi dữ liệu nghiên cứu phải được đánh giá lại bằng dữ liệu mới và kiểm chứng độc lập.

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết luận

Luận án đã xây dựng và kiểm chứng một khung dự báo ô nhiễm không khí theo thời gian tại mỏ lộ thiên sâu, nghiên cứu điển hình tại mỏ đồng Sin Quyền. Khung nghiên cứu liên kết điều kiện vi khí hậu và công nghệ khai thác với quan trắc phân tầng, kỹ nghệ đặc trưng, tối ưu siêu tham số và mô hình dự báo. Những kết quả chính gồm:

- Kết quả thực nghiệm cho thấy MPA là một phương án khả thi để tinh chỉnh siêu tham số RF/XGBoost trên tập kiểm chứng. Trong bộ dữ liệu Sin Quyền, MPA-XGBoost duy trì hiệu suất dự báo cao và làm giảm hiện tượng dự báo thấp ở nhiều đỉnh nồng độ so với RF chưa tối ưu; tuy nhiên, XGBoost gốc vẫn có thể cho RMSE/MAE kiểm thử tương đương hoặc thấp hơn nhẹ ở một số trường hợp.

- Luận án làm rõ sự khác biệt của chuỗi $PM_{2.5}/CO$ giữa ba cao độ quan trắc và cho thấy điều kiện thông gió bất lợi tại các tầng sâu làm tăng độ phức tạp của bài toán dự báo. Việc kết hợp CEEMDAN, biến tương tác và biến trễ giúp biểu diễn tốt hơn nhiều động cục bộ và sự phụ thuộc thời gian trong dữ liệu; hiệu quả được đánh giá trong phạm vi ba trạm của nghiên cứu, không được khái quát cho mọi điều kiện địa hình.

- Điểm mới của luận án là khung tích hợp “vật lý mô - dữ liệu - đặc trưng - mô hình - quyết định”: CEEMDAN, MPA, XGBoost và Stacking là các phương pháp đã có, nhưng được tổ chức thành một quy trình thống nhất theo điều kiện phát thải, vi khí hậu và tổ chức sản xuất của mỏ đồng Sin Quyền, đồng thời tuân thủ nguyên tắc phân chia dữ liệu theo thời gian để hạn chế rò rỉ thông tin.

Luận án có ý nghĩa quản lý và thực tiễn như sau:

- Kết quả của luận án cung cấp thêm một công cụ dự báo hỗ trợ quản trị rủi ro môi trường, góp phần chuyển phương thức quản lý từ hậu kiểm sang chủ động phòng ngừa tại các vị trí quan trắc được nghiên cứu.

- Khung MPA-XGBoost được thiết lập để cung cấp dự báo $PM_{2.5}$ và CO trước 1-3 giờ. Trong bộ dữ liệu kiểm thử, mô hình theo sát tốt nhiều cực trị liên quan đến các giai đoạn phát thải mạnh; khoảng thời gian dự báo này tạo thêm cơ sở để cán bộ an toàn - môi trường và điều hành sản xuất xem xét biện pháp ứng phó trước khi nồng độ có nguy cơ tăng cao.

- Thông tin dự báo có thể hỗ trợ phân bổ biện pháp giảm bụi và điều chỉnh hoạt động sản xuất theo vị trí, thời điểm và mức cảnh báo (ví dụ tăng cường phun nước tại khu vực có nguy cơ vượt ngưỡng). Luận án chưa thực hiện phân tích kinh tế đầy đủ nên kết quả được diễn giải là góp phần nâng cao hiệu quả sử dụng nguồn lực và hỗ trợ bảo vệ sức khỏe người lao động, không khẳng định mức tiết kiệm chi phí cụ thể.

2. Kiến nghị

- Kết hợp mô hình lai MPA-XGBoost với các thuật toán nội suy địa thống kê (Kriging) hoặc mạng nơ-ron đồ thị (Graph Neural Networks - GNN) để thiết lập bản đồ dự báo ô nhiễm dạng lưới (Grid-map) cho toàn bộ không gian mỏ lộ thiên.

- Thu thập thêm dữ liệu để huấn luyện AI dự báo song song các khí độc khác sinh ra từ thuốc nổ công nghiệp (NO_x , SO_2), hướng tới một hệ thống cảnh báo đa chỉ tiêu.

- Phát triển kết quả của luận án thành một phần mềm độc lập (Web-based hoặc Desktop application) có giao diện trực quan, kết nối API trực tiếp với hệ thống IoT của mỏ đồng Sin Quyền để các kỹ sư điều hành có thể sử dụng dễ dàng mà không cần kiến thức chuyên sâu về lập trình.

- Khi tham khảo hoặc chuyển giao khung mô hình sang mỏ lộ thiên sâu khác, cần thu thập dữ liệu tại mỏ đích, đánh giá lại vị trí mạng quan trắc, hiệu chỉnh/huấn luyện lại mô hình và kiểm chứng ngoại mẫu trước khi sử dụng cho cảnh báo hoặc điều hành./.

DANH MỤC CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN

1. **Ngoc, L. T.**, Nguyen, H., Bui, X. N., & Tran, T. T. (2026). Dynamic weighting and genetic algorithm-based hybrid SVR-Ridge model for improving PM_{2.5} forecasting in open-pit copper mines. *International Journal of Mining, Reclamation and Environment*, 40(4), 349–379. <https://doi.org/10.1080/17480930.2025.2492086>
2. **Le Tuan Ngoc**, Hoang Nguyen, Xuan-Nam Bui, Le Thi Thu Hoa (2025). RF-PM₁₀Hybrid Model for Real-Time PM₁₀ Forecasting in Open-Pit Copper Mine: A Case Study at the Sin Quyen Copper Mine. *Journal of Mining and Earth Sciences*, 66(4), 48–69. [https://doi.org/10.46326/JMES.2025.66\(4\).04](https://doi.org/10.46326/JMES.2025.66(4).04)
3. **Tuan-Ngoc Le**, Nguyen Truc Anh, Hoang Nguyen, Xuan-Nam Bui (2025). Predicting PM_{2.5} Concentrations in an Open-Pit Copper Mine Using Advanced AI Techniques: Insights from the Sin Quyen Copper Mine, Vietnam. *Proceedings of the International Workshop on Advances in Surface Mining for Environmental Protection and Sustainable Development*, 17 October 2025, Ha Noi, Viet Nam, 249–268.
4. Hoang Nguyen, Tran Dinh Bao, Xuan-Nam Bui, Pham Van Viet, Dinh-An Nguyen, Ngoc-Hoan Do, Le Thi Thu Hoa, Qui-Thao Le, **Tuan-Ngoc Le** (2024). Measuring and Predicting Blast-Induced Flyrock Using Unmanned Aerial Vehicles and Lévy Flight Technique-Based Jaya Optimization Algorithm Integrated with Adaptive Neuro-Fuzzy Inference System. *Natural Resources Research*, 34(9). <https://doi.org/10.1007/s11053-025-10455-4>.