

# CẢNH BÁO - DỰ BÁO NƯỚC DƯỚI ĐẤT BẰNG AI-GIS: TỔNG QUAN CẤU TRÚC HIỆN TRẠNG NGHIÊN CỨU, CÁC KHOẢNG TRỐNG VÀ TIỀN ĐỀ TRIỂN KHAI TẠI VIỆT NAM

Hoàng Thị Hạnh, Đặng Đình Phú, Nguyễn Tuấn Linh,  
Phan Quang Thức, Thân Văn Đón

Trung tâm Công nghệ và Dữ liệu tài nguyên nước

Nguyễn Tiếp Tân

Viện Khoa học thủy lợi Việt Nam

Hoàng Nguyệt Minh

Trường Đại học Tài nguyên và Môi trường Hà Nội

Ngô Lê Long

Trường Đại học thủy lợi Hà Nội

**Tóm tắt:** Trong bối cảnh biến đổi khí hậu, khai thác quá mức và xâm nhập mặn gia tăng, nhu cầu cảnh báo và dự báo suy thoái, cạn kiệt, ô nhiễm nguồn nước dưới đất trở thành cấp thiết đối với công tác quản lý tài nguyên nước ở Việt Nam. Sử dụng trí tuệ nhân tạo (AI) khai thác, phân tích và chuyển đổi nguồn dữ liệu thời gian thực thành những kiến thức hữu ích phục vụ cảnh báo, dự báo và quản lý nhà nước về tài nguyên nước là xu thế tất yếu của đổi mới công nghệ và chuyển đổi số trong lĩnh vực tài nguyên nước. Bài báo này tổng hợp, phân tích các nghiên cứu thực hiện ở Việt Nam trong vài năm gần đây về ứng dụng trí tuệ nhân tạo (AI) trong lĩnh vực nước dưới đất, tập trung vào ba hướng chính: (i) lập bản đồ tiềm năng nước dưới đất trên nền GIS tích hợp AI; (ii) dự báo độ mặn và nguy cơ ô nhiễm As bằng các thuật toán học máy (ML); và (iii) phân tích chuỗi dữ liệu quan trắc mực nước bằng mô hình học sâu (DL). Kết quả nghiên cứu bước đầu cho thấy mô hình Extreme Gradient Boosting có tiềm năng nâng cao độ chính xác, giảm chi phí và khối lượng khảo sát thực địa, đặc biệt hữu ích ở các vùng khan hiếm nước; CatBoost và Extra Trees dự báo tốt độ mặn nước dưới đất trong khi Random Forest và Gradient Boosting dự báo tốt nguy cơ ô nhiễm As. Tuy nhiên, các nghiên cứu tại Việt Nam còn hạn chế về quy mô, địa bàn, đối tượng nghiên cứu và dữ liệu sử dụng; chủ yếu dừng ở mức thử nghiệm cho một vùng cụ thể, chưa tạo ra tác động nào đối với công tác quản lý và chưa đáp ứng được yêu cầu thực tiễn về cảnh báo, dự báo tài nguyên nước. Bài báo cũng chỉ ra các khoảng trống về khung pháp lý, chuẩn hóa dữ liệu quan trắc và quy trình nghiên cứu, đồng thời đề xuất định hướng ứng dụng AI cho cảnh báo cạn kiệt, xâm nhập mặn ở vùng ven biển, cảnh báo ô nhiễm tại khu công nghiệp tập trung và tìm kiếm nguồn nước ở vùng sâu, vùng xa, vùng khan hiếm nước. Kết quả nghiên cứu góp phần khẳng định tiềm năng của AI trong quản lý bền vững tài nguyên nước dưới đất và nhấn mạnh sự cần thiết phải chuẩn hóa mạng quan trắc, hoàn thiện khung pháp lý cho cảnh báo, dự báo nguồn nước.

**Từ khóa:** AI-GIS, ML, cảnh báo, dự báo nước dưới đất, cạn kiệt, xâm nhập mặn, khoảng trống nghiên cứu.

**Summary:** The need for warning and forecast groundwater resources degradation, depletion and pollution has become urgent in Vietnam in the context of climate change, groundwater overexploitation and increasing saline intrusion into aquifer systems. Using artificial intelligence (AI) to exploit, analyze and transform raw real-time data sources into useful knowledge for warning, forecasting and state management of water resources is an inevitable trend of technological innovation and digital transformation in water resources sector. This article synthesizes and analyzes various studies conducted in Vietnam in recent years on AI application in the field of groundwater, focusing on three main directions: (i) mapping groundwater potential on AI-integrated GIS platform; (ii) forecasting groundwater salinity and As pollution risk using machine learning (ML) algorithms; and (iii) analyzing groundwater level monitoring data series using deep learning (DL) models. Initial study results show that the Extreme Gradient Boosting model has the potential to improve accuracy, reduce costs and field survey volume, especially useful in water-scarce areas; CatBoost and Extra Trees predict groundwater salinity well while Random Forest and Gradient Boosting predict As pollution risk at very high accuracy. However, research in Vietnam is still limited in scale, location, research subjects and data used, mainly at the experimental level for a specific area, has not created any impact on management work and has not met practical requirements for warning and forecasting water resources. The article also points out gaps in the legal framework, standardization of monitoring data and research procedure, and proposes directions for applying AI for warning of depletion and saline intrusion in coastal areas, warning of pollution in concentrated industrial zones and searching for water sources in remote, isolated and water-scarce areas. The research results contribute to affirming the potential of AI in sustainable management of groundwater resources and emphasize the need to standardize monitoring networks and perfect the legal framework for water sources warning and forecasting.

**Key words:** AI-GIS, ML, groundwater warning forecasting, depletion, saline intrusion, research gaps.

## 1. MỞ ĐẦU

Cảnh báo, dự báo suy thoái, cạn kiệt, xâm

nhập mặn nước dưới đất có vai trò quan trọng trong quản lý, bảo vệ nước dưới đất nói chung và ứng phó chủ động với tình trạng hạn hán, thiếu nước, xâm nhập mặn; giúp cơ quan quản lý, cộng đồng dân cư và doanh nghiệp nắm bắt các nguy cơ và đưa ra các hành động thích hợp. Dự báo dài hạn cho phép xây dựng chiến

Ngày nhận bài: 20/8/2025

Ngày thông qua phản biện: 10/9/2025

Ngày duyệt đăng: 02/10/2025

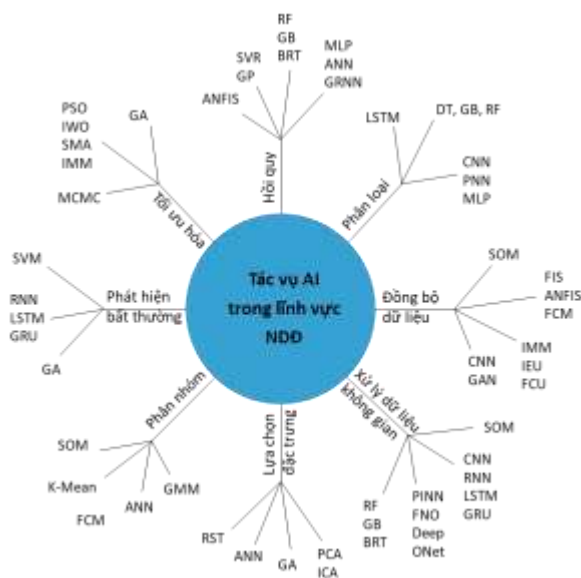
lược phân bổ, khai thác – sử dụng nguồn nước hợp lý và chủ động các biện pháp thích ứng (chuyên đổi cơ cấu cây trồng, xây dựng công trình ngăn mặn, công trình trữ nước ngọt, bổ sung nhân tạo cho tầng chứa nước). Cảnh báo kịp thời giúp giảm thiểu rủi ro cho người dân vùng tổn thương cao, hạn chế thiệt hại về kinh tế – xã hội. Hệ thống cảnh báo sớm thiết lập dựa trên (i) một mạng quan trắc hoàn chỉnh tạo ra dữ liệu quan trắc chi tiết, tin cậy; (ii) các ngưỡng cảnh báo được xác lập rõ ràng; (iii) các công cụ mạnh giúp theo dõi chính xác diễn biến, phát hiện sớm các dấu hiệu nhằm đưa ra cảnh báo kịp thời về nguy cơ suy thoái, cạn kiệt và xâm nhập mặn, giúp cộng đồng chủ động ứng phó.

Quan trắc tài nguyên nước theo thời gian thực sử dụng các hệ thống quan trắc tự động (telemetry system) với công nghệ IoT tạo ra khối lượng dữ liệu khổng lồ; thách thức khả năng xử lý truyền thống của các công cụ thông thường (excel, ArcGIS). Để có thể xử lý, phân tích và chuyên đổi nguồn dữ liệu thời gian thực thô thành những kiến thức hữu ích phục vụ cảnh báo, dự báo và quản lý nhà nước về tài nguyên nước, cần có không chỉ cơ sở hạ tầng tương thích Big Data (cloud computing, HPC) mà cần bản là các công cụ AI khai thác, phân tích dữ liệu và khuôn khổ quản trị dữ liệu mạnh.

Trong vài năm trở lại đây AI đã trở thành công cụ trung tâm trong nghiên cứu, dự báo và quản lý nước dưới đất, với ứng dụng trải rộng từ dự báo mực nước, đánh giá chất lượng, khoanh vùng tiềm năng đến tối ưu hóa khai thác-sử dụng. Các thuật toán ML được nghiên cứu ứng dụng phổ biến trong lĩnh vực nước dưới đất và những tác động mà chúng mang lại đã được xem xét có hệ thống [1], [2]. Giai đoạn 2021-2025 ghi nhận sự gia tăng mạnh mẽ các nghiên cứu ứng dụng AI trong lĩnh vực tài nguyên nước, từ các nghiên cứu tổng quan về dự báo chất lượng nước dưới đất sử dụng ANN, ANFIS, SVM, EA [3] đến các mô hình tích hợp đa nguồn dữ liệu GIS, viễn thám với IoT [4], [5]. Không chỉ dừng ở mô phỏng, nhiều nghiên cứu đã đưa AI vào dự báo mực nước dưới đất phục vụ khai thác mỏ, nông nghiệp, cấp nước đô thị [4], [6] cho thấy ứng dụng thực tiễn rõ nét hơn. Các tác vụ AI cơ bản trong lĩnh vực nước dưới đất và các thuật toán ML tương ứng đã được nghiên cứu ứng dụng được thể hiện trên Hình 1.

Ở Việt Nam, nghiên cứu thử nghiệm AI trong lập bản đồ tiềm năng nước dưới đất lần đầu tiên được thực hiện thí điểm cho tỉnh Đắk

Nông giúp lập bản đồ chính xác hơn với mô hình AI tổ hợp RABANN (Real AdaBoost+ANN) [7]. Các nghiên cứu thử nghiệm AI trong lĩnh vực nước dưới đất ở Việt Nam từ đó đến nay mặc dù chủ yếu vẫn tập trung vào lập bản đồ tiềm năng nước dưới đất; tuy nhiên, đã có sự mở rộng sang dự báo độ mặn nước dưới đất [8], [9] và nguy cơ ô nhiễm Asen trong nước dưới đất [10]. Đặc biệt là nghiên cứu thời gian gần đây đưa CNN (Conv1D), SimpleRNN và GRU vào phân tích chuỗi dữ liệu mực nước cho thấy hướng mở rộng từ bài toán không gian-thời gian sang chuỗi thời gian [11].



Hình 1: Các tác vụ AI trong lĩnh vực nước dưới đất và các thuật toán ML đã được nghiên cứu ứng dụng

Cảnh báo, dự báo nguồn nước là một trong những trụ cột quan trọng của công tác điều tra cơ bản tài nguyên nước. Thông tư số 52/2025/BNNMT ban hành mới đây đề cập các phương pháp cảnh báo, dự báo nguồn nước gồm phương pháp thống kê, giải tích, mô hình; chuỗi thông tin, số liệu phục vụ dự báo nguồn nước phải đảm bảo tối thiểu 20 năm. Tuy nhiên, Thông tư này chỉ liệt kê các bước thực hiện cho dự báo nguồn nước bằng phương pháp thống kê và giải tích; tuyệt đối chưa đề cập đến dự báo nguồn nước bằng phương pháp mô hình. Không chỉ thế, Thông tư này cũng hoàn toàn không đề cập đến hệ thống cảnh báo sớm cũng như các công cụ phục vụ cảnh báo, dự báo nguồn nước.

Nghiên cứu này được nhóm tác giả thực hiện

nhằm: (i) tổng kết tình hình nghiên cứu ứng dụng AI trong lĩnh vực nước dưới đất ở Việt Nam trong mấy năm gần đây; (ii) nhận diện các hạn chế và khoảng trống trong các nghiên cứu; (iii) định hướng cho các nghiên cứu thử nghiệm AI trong lĩnh vực tài nguyên nước dưới đất ở Việt Nam.

## 2. SỐ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Nhóm tác giả đã tiến hành thu thập số liệu từ nguồn sơ cấp (số liệu quan trắc mực nước dưới đất và chất lượng nước dưới đất do Trung tâm Quy hoạch và điều tra tài nguyên nước quốc gia thực hiện), nguồn thứ cấp (các bài báo khoa học công bố trên các tạp chí chuyên ngành trong và ngoài nước, các báo cáo kỹ thuật về nghiên cứu ứng dụng AI trong lĩnh vực nước dưới đất) sau đó tiến hành phân tích-tổng hợp, phân loại-hệ thống và tổng kết kinh nghiệm để có cái nhìn toàn diện về tình hình nghiên cứu, các thách thức chung, các khoảng trống hiện tại và tiềm năng phát triển ở Việt Nam.

## 3. KẾT QUẢ VÀ THẢO LUẬN

### 3.1. Tình hình nghiên cứu ứng dụng AI trong lĩnh vực nước dưới đất

#### 3.1.1. Tích hợp AI trên nền GIS để lập bản đồ tiềm năng nước dưới đất

Tiềm năng nước dưới đất là đối tượng nghiên cứu chủ đạo, chiếm 7 trong số 11 công trình nghiên cứu đăng trên tạp chí chuyên ngành quốc tế (9 bài) và trong nước (2 bài) từ năm 2020 đến tháng 3/2025 được nhóm tác giả khảo cứu (Hình 2); phản ánh sự chuyển dịch từ ứng dụng công cụ GIS thông thường sang GIS tích hợp AI, đặc biệt là các mô hình ML khác nhau.

Ứng dụng AI giúp lập bản đồ và đánh giá chi tiết tiềm năng nước dưới đất có ý nghĩa to lớn đối với công tác tìm kiếm, thăm dò nước dưới đất, góp phần làm giảm khối lượng công tác

điều tra thực địa ở các vùng có điều kiện giao thông kém, đi lại khó khăn, vùng có điều kiện thi công phức tạp, mức độ tài liệu điều tra địa chất, địa chất thủy văn hạn chế. Mọi liên hệ giữa đối tượng nghiên cứu (tức là tiềm năng nước dưới đất) với vùng nghiên cứu được phản ánh rõ nét khi tất cả 7 công trình nghiên cứu này [7], [12-17] đều được triển khai ở các tỉnh Tây Nguyên (Gia Lai, Đắk Nông, Đắk Lắk) là vùng khan hiếm nước.

Việc lựa chọn các yếu tố ảnh hưởng đến khả năng “tìm thấy” nước dưới đất là không đơn giản do bản chất phức tạp và phi tuyến tính của nước dưới đất. Thông thường, các yếu tố được lựa chọn phù hợp với vùng nghiên cứu và theo ý định của nhóm nghiên cứu do không có hướng dẫn chuẩn nào cho việc lựa chọn. Trong các nghiên cứu thử nghiệm ở Việt Nam, các tác giả hầu hết đều lựa chọn 12 yếu tố trong đó luôn bao gồm độ dốc, hướng dốc, độ cao, độ cong, lượng mưa, mật độ sông suối, sử dụng đất, thành tạo địa chất, khoảng cách đến đứt gãy. Ngoài ra, các yếu tố khoảng cách đến sông, chỉ số độ ẩm địa hình (TWI), khoảng cách đến đường giao thông, chỉ số nguồn dòng (SPI), chỉ số vận chuyển trầm tích (STI), chỉ số khác biệt thực vật (NDVI), chỉ số khác biệt độ ẩm (NDMI), chỉ số khác biệt nước (NDWI), chỉ số khác biệt công trình xây dựng (NDBI)... được lựa chọn tùy theo đặc trưng vùng nghiên cứu. Thông thường, các yếu tố địa hình (độ dốc, hướng dốc, độ cao, độ cong) được trích xuất từ mô hình số độ cao DEM ALOS PALSAR với độ phân giải tùy chọn; mật độ sông và khoảng cách đến đường giao thông được thu nhận từ bản đồ địa hình; thông tin về thành tạo địa chất, khoảng cách đến đứt gãy được thu nhận từ bản đồ địa chất; các chỉ số NDVI, NDMI, NDWI, NDBI được tính toán từ ảnh vệ tinh Sentinel và Landsat.

**Bảng 1: Tổng hợp các kỹ thuật AI/thuật toán ML, DL được sử dụng theo bối cảnh dữ liệu trong một số nghiên cứu gần đây về phân vùng tiềm năng nước dưới đất (NDD) ở một số vùng của Việt Nam**

Nguồn tài liệu	Dữ liệu được sử dụng	Kỹ thuật AI/Thuật toán ML được sử dụng
[7]	72 giếng thăm dò NDD	Real AdaBoost+Artificial Neural Networks
[12]	130 giếng thăm dò NDD	AdaBoost-Functional Tree (MRAB-FT), Bagging-Functional Tree (BA-FT), Rotation Forest-Functional Tree (RF-FT)
[13]	134 giếng thăm chiếu địa lý thực địa 2018-2019	Locally Weighted Learning (LWL), AdaBoost - Locally Weighted Learning (ABLWL), Bagging - Locally Weighted Learning (BLWL), Multi Boost - Locally Weighted Learning (MBLWL) and Rotation Forest - Locally Weighted Learning (RFLWL)

Nguồn tài liệu	Dữ liệu được sử dụng	Kỹ thuật AI/Thuật toán ML được sử dụng
[14]	72 giếng thăm dò NĐĐ	AdaBoost-Random Forest (ABRF), Bagging-Random Forest (BRF), LogitBoost-Random Forest (LBRF)
[15]	938 điểm mạch lộ	Deep Neural Network (DeepNN) tối ưu bằng Adaptive Moment Estimation Optimizer (ADAM), SwarmRF tối ưu bằng Harris Hawks Optimizer (HHO)
[16]	181 điểm kiểm kê NĐĐ trong đó 82 điểm không giếng/ không mạch lộ	Deep Neural Network (DNN), CatBoost (CB), extreme Gradient Boosting (XGB)
[17]	501 dữ liệu giếng	Cascade Generalization-Logistics Model Trees (CG-LMT)

Nhìn chung, các mô hình AI/ML được lựa chọn tương thích với bối cảnh dữ liệu: RABANN, MRAB-FT, BA-FT, RF-FT, ABLWL, BLWL, MBLWL, RFLWL cho các tập dữ liệu hạn chế; CB, XGB, DNN, SwarmRF, CG-LMT cho các tập dữ liệu lớn hơn. AI tổ hợp có ưu điểm hơn so với AI đơn lẻ nhờ khả năng nhận dạng tốt các quan hệ phi tuyến và giảm thiên lệch trên bộ dữ liệu nhỏ và vừa. XGB là lựa chọn tốt khi muốn tối ưu độ chính xác và ổn định trong điều kiện dữ liệu trung bình, không quá dồi dào. Deep learning được kỳ vọng là phát huy thế mạnh đối với tập dữ liệu lớn; tuy nhiên nghiên cứu trên 938 dữ liệu mạch lộ nước dưới đất ở Gia Lai cho thấy DeepNN (accuracy 77,9; AUC=0,820) có hiệu năng kém hơn SwarmRF (accuracy 80,2%; AUC=0,854) [15].

### 3.1.2. Sử dụng Machine Learning dự báo độ mặn nước dưới đất và nguy cơ ô nhiễm Asen

Nghiên cứu thử nghiệm các thuật toán ML dự báo độ mặn nước dưới đất và ô nhiễm Asen nước dưới đất ở Việt Nam mới ở bước khởi đầu với số lượng rất ít công trình được công bố. Mô hình CB cho thấy hiệu năng dự báo hàm lượng  $\text{Cl}^-$  tốt nhất trong cả pha huấn luyện (RMSE=29,90;  $R^2=0,99$ ) và pha kiểm tra (RMSE=205,96;  $R^2=0,84$ ) [8]. Giá trị hàm lượng  $\text{Cl}^-$  dự báo bằng mô hình CB được sử dụng để lập bản đồ độ mặn nước dưới đất; từ đó tính toán ra diện tích vùng và số dân bị ảnh hưởng mặn. Trong một nghiên cứu mới đây, 9 mô hình ML tổ hợp đã được so sánh, sử dụng tối ưu Bayes và thực hiện hoán đổi nhiều lần tập dữ liệu huấn luyện/kiểm định để kiểm tra độ ổn định và xác định mô hình tối ưu cho từng chỉ thị độ mặn ( $\text{Cl}^-$ ,  $\text{HCO}_3^-$ , pH); xác định ET ổn định nhất cho  $\text{Cl}^-$  ( $R^2>0,93$ ; RMSE < 1050 mg/L); RF ổn định cho pH ( $R^2=0,67$ ; RMSE=0,43); BA, RF ổn định nhất cho  $\text{HCO}_3^-$  ( $R^2=0,78/0,72$ ; RMSE=82–88) [9].

Nguy cơ ô nhiễm As trong nước dưới đất được dự báo tốt (AUC-ROC tới 0,96) bằng mô hình RF và GB và xác định Eh và Fe là hai biến quan trọng nhất chi phối hàm lượng As trong nước dưới đất [10]. Giá trị Eh thấp biểu thị điều kiện môi trường khử, tạo điều kiện cho quá trình hòa tan các oxit Fe và giải phóng As từ trầm tích vào trong nước dưới đất.

### 3.1.3. Sử dụng Deep Learning xử lý các chuỗi dữ liệu nước dưới đất thời gian thực

Việc vận hành mạng quan trắc nước dưới đất trong thời gian dài, đặc biệt là khi hệ thống quan trắc bán tự động và tự động được triển khai lắp đặt phổ biến tại các công trình quan trắc, mang lại khối lượng dữ liệu quan trắc đồ sộ và kéo theo yêu cầu xử lý hiệu quả, tối ưu nguồn dữ liệu này để phục vụ cho công tác cảnh báo, dự báo nước dưới đất. Mặc dù các thuật toán ML đã được nghiên cứu thử nghiệm mạnh mẽ và áp dụng rộng rãi cho quan trắc nước dưới đất ở nhiều nước trên thế giới từ những năm 2010 [2]; tuy nhiên, ở Việt Nam, chủ đề này vẫn chưa được khai thác nhiều. Kết quả nghiên cứu ban đầu đã cho thấy ANN dự báo mực nước dưới đất ở các giếng quan trắc (QT08 và QT16) tại Cần Thơ với EI và  $R^2 > 0,9$  cho cả giai đoạn hiệu chỉnh (2004–2012) và kiểm định (2013–2015); RMSE rất thấp (0,06–0,22 m) [18]. Dạng đường dự báo bám sát với đường quan trắc, phù hợp về xu hướng trung bình tháng và thể hiện sự suy giảm mực nước theo thời gian do khai thác mạnh; tuy nhiên vẫn có sai số cục bộ thể hiện ở các peak được dự báo quá thấp/quá cao.

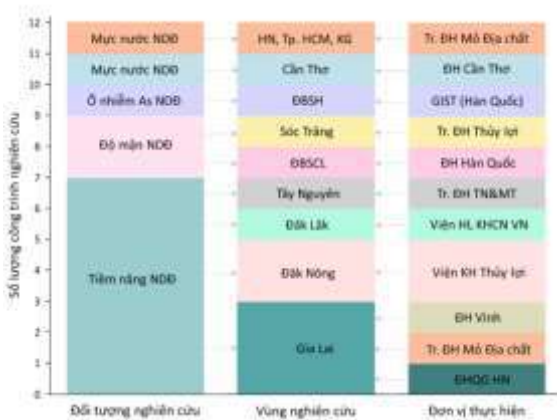
Thử nghiệm dự báo mực nước dưới đất sử dụng ba mô hình DL khác nhau (Convolutional Neural Networks biệt hóa cho tập dữ liệu một chiều (Conv1D), Simple Recurrent Neural Networks (SimpleRNN), Gated Recurrent Unit (GRU) cho dữ liệu quan trắc nước dưới đất tại 3 trạm quan trắc quốc

gia cho thấy CNN (Conv1D) cho kết quả tốt nhất ở cả 3 trạm quan trắc và đặc biệt nổi trội với dữ liệu có biên động vừa phải hoặc có thiếu hụt (ví dụ trạm Q080810 – HCM), với RMSE rất thấp (0,06 mm). SimpleRNN hoạt động rất tốt với dữ liệu có độ phân giải cao và ít biên động (trạm Q55M1 – Hà Nội, dữ liệu theo giờ). Tuy nhiên, khi dữ liệu biên động lớn (trạm Q40101T – Kiên Giang) thì độ chính xác giảm mạnh, sai số tăng. GRU theo lý thuyết có khả năng xử lý tốt quan hệ dài hạn, nhưng trong thử nghiệm lại kém nhất. Ở Q40101T, sai số RMSE lên đến 7,39 mm, lớn hơn hẳn hai mô hình kia [11].

**3.2. Các hạn chế và khoảng trống**

**3.2.1. Quy mô, đối tượng, địa bàn và đơn vị thực hiện nghiên cứu**

Có thể thấy các nghiên cứu thử nghiệm AI/ML trong lĩnh vực nước dưới đất ở Việt Nam còn hạn chế về quy mô, đối tượng và địa bàn nghiên cứu. Tổng số công trình nghiên cứu chỉ dừng lại ở mức 12, phản ánh thực tế rằng ứng dụng AI/ML trong lĩnh vực nước dưới đất ở Việt Nam vẫn còn ở giai đoạn thử nghiệm, chưa hình thành một hướng nghiên cứu mạnh mẽ, liên tục. Phần lớn công trình nghiên cứu (7/12 công trình) tập trung vào tiềm năng nước dưới đất (Hình 2) trong khi các vấn đề chuyên biệt, cấp bách hơn như suy giảm mực nước, gia tăng độ mặn, nguy cơ ô nhiễm lại ít được khai thác (chỉ 1-2 công trình). Điều này cho thấy sự mất cân đối giữa nhu cầu thực tiễn và hướng nghiên cứu.



Hình 2: Các công trình nghiên cứu ứng dụng AI trong lĩnh vực nước dưới đất ở Việt Nam thống kê theo đối tượng nghiên cứu, vùng nghiên cứu và đơn vị thực hiện

Các nghiên cứu mới chỉ được thực hiện ở một số khu vực cụ thể như Tây Nguyên, ĐBSH,

ĐBSCL trong khi nhiều tỉnh/thành khác với tình trạng khai thác nước dưới đất phức tạp (ví dụ: các khu công nghiệp tập trung, vùng ven biển ĐBSH, vùng ven biển miền Trung) lại chưa được đề cập.

Các đơn vị có nền tảng khoa học kỹ thuật và kinh nghiệm nghiên cứu thực địa như Viện Hàn lâm Khoa học công nghệ Việt Nam, Viện Khoa học thủy lợi, Đại học Quốc gia Hà Nội, Trường Đại học Mỏ-Địa chất đều đã tham gia nghiên cứu và được kỳ vọng tiếp tục đóng vai trò nòng cốt. Sự tham gia của đối tác Hàn Quốc như Viện Khoa học Công nghệ Gwangju (GIST) và Đại học Hàn Quốc dù với số lượng công trình nghiên cứu còn rất hạn chế đã cho thấy tiềm năng mở rộng mạng lưới hợp tác và tiếp thu tri thức, công nghệ tiên tiến. Cần thúc đẩy hợp tác quốc tế, mở rộng đối tác và chú trọng xây dựng đề cương, lập dự án nghiên cứu chi tiết bài bản, định hình rõ mục tiêu nghiên cứu và vai trò của Việt Nam; tránh tình trạng chỉ cung cấp dữ liệu cho đối tác nước ngoài thực hiện nghiên cứu và công bố kết quả.

NAWAPI là đơn vị có bề dày kinh nghiệm quản lý vận hành mạng quan trắc tài nguyên nước quốc gia, nắm giữ dữ liệu quan trắc tài nguyên nước và có chức năng nhiệm vụ thực hiện cảnh báo, dự báo tài nguyên nước. Do vậy, có thể kỳ vọng sự tham gia tích cực, mạnh mẽ của đơn vị này trong các nghiên cứu thử nghiệm, ứng dụng AI/ML trong quan trắc, cảnh báo, dự báo nước dưới đất trong thời gian tới.

**3.2.2. Dữ liệu được sử dụng**

Các nghiên cứu đã thực hiện về dự báo độ mặn và ô nhiễm As nước dưới đất hầu hết mới chỉ sử dụng các tập dữ liệu tại một thời điểm hoặc gắn với một chiến dịch điều tra khảo sát trong một khoảng thời gian nhất định, thường là vài năm. Khả năng tổng quát hóa về mặt không gian-thời gian của nghiên cứu là cực kỳ hạn chế; kết quả của nghiên cứu không thể áp dụng cho giai đoạn khác hoặc địa phương khác. Việc mở rộng sang dự báo chuỗi thời gian bằng Conv1D, SimpleRNN và GRU [11] mới ở giai đoạn thử nghiệm với phạm vi dữ liệu hạn chế, chỉ gồm 3 trạm quan trắc: Q55M1 (Hà Nội; 2020-2023), Q080810 (TP.HCM; 1992-2016), Q40101T (Kiên Giang; 2000-2016) với thời gian và đặc điểm dữ liệu khác nhau nên khó khái quát cho toàn bộ Việt Nam. Các yếu tố khí tượng, thủy văn, khai thác ... vốn ảnh hưởng mạnh đến động thái nước dưới đất lại chưa được xem xét. Trong khi đó, Thông tư số 52/2025/BNNMT quy định về

chuỗi thông tin, số liệu phục vụ dự báo nguồn nước phải đảm bảo tối thiểu 20 năm.

### 3.2.3. Các bước tiến hành nghiên cứu lựa chọn mô hình

Nhìn chung, kết quả nghiên cứu thử nghiệm AI/ML trong lĩnh vực nước dưới đất ở Việt Nam chưa được các nhóm tác giả trình bày thống nhất và rõ ràng theo các bước thực hiện,

gây cản trở đối với việc (i) hiểu đúng kết quả của nghiên cứu, cụ thể là hiệu năng dự báo/dự đoán của mô hình và (ii) lặp lại quy trình/áp dụng thử nghiệm mô hình đó trên tập dữ liệu mới. Các nghiên cứu đã thực hiện ở Việt Nam sẽ được xem xét đánh giá theo các bước tiêu chuẩn như tóm tắt trong Hình 3 để làm rõ các hạn chế nếu có.



Hình 3: Các bước tiêu chuẩn trong nghiên cứu lựa chọn mô hình AI/ML cho dự báo nước dưới đất

#### a. Chuẩn bị dữ liệu

Các nghiên cứu đều tiến hành làm sạch dữ liệu, loại các giá trị ngoại lai, đồng bộ dữ liệu về độ phân giải, tọa độ; dán nhãn/mã hóa biến phụ thuộc theo ngưỡng xác định (ví dụ: lưu lượng 2 L/s để phân biệt có/không có nước dưới đất; hàm lượng As 10 µg/l để phân biệt có/không ô nhiễm). Hầu hết các nghiên cứu đều tiến hành chia dữ liệu ngẫu nhiên hoặc chia theo giai đoạn thời gian (ví dụ: 2004-2012 vs 2013-2015) thành tập dữ liệu *train* và tập dữ liệu *test* theo tỉ lệ 70/30. Dữ liệu *test* cần được giữ nguyên không can thiệp, ngoại trừ làm sạch, đồng bộ và gán nhãn cho các biến mà không làm rò rỉ thông tin. Lưu ý nguy cơ rò rỉ tương quan không gian khi các điểm gần nhau có đặc điểm tương đồng được phân chia ngẫu nhiên vào cả tập *train* và tập *test*. Cần đảm bảo duy trì tỉ lệ dương tính của tập dữ liệu trong quá trình phân chia. Tuy nhiên, một số nghiên cứu đã thực hiện chuẩn hóa dữ liệu [9] (nhưng không trình bày rõ chuẩn hóa như thế nào); gán dữ liệu bị thiếu cho một số tham số (T, HCO<sub>3</sub>, DO...) bằng KNN imputer [10] trước khi chia tập *train/test*.

Một số nghiên cứu đã sử dụng các kỹ thuật/công cụ Feature importance [8], One-R feature selection [12], Correlation Attribute Evaluation [13] để lựa chọn các yếu tố đầu vào.

Có thể dùng permutation importance (PI) và/hoặc SHAP để loại bớt những yếu tố nhiễu, chỉ giữ lại những đặc trưng quan trọng. Xử lý đồng biến tuyến tính, đa cộng tuyến bằng Variance Inflation Factor (VIF) là cần thiết đối với các tập dữ liệu có

nhiều biến tương quan cao.

Việc lấy mẫu âm để cân bằng dữ liệu cho tập *train* và *validation* cần thận trọng, đảm bảo có kiểm soát không gian để tránh overfitting và thực hiện sau khi chia tập *train/validation/test*. Việc lấy thêm các điểm không mạch lộ với số lượng bằng số điểm mạch lộ ban đầu trước khi chia dữ liệu thành tập *train* và *validation* (không chia tập *test*) [15] dẫn đến nguy cơ rò rỉ không gian và khiến cho các tiêu chí đánh giá hiệu năng dự báo của mô hình không khách quan.

#### b. Xây dựng mô hình

Phần lớn các nghiên cứu chưa mô tả quy trình tinh chỉnh siêu tham số một cách có hệ thống. Hàm grid search 5-fold CV và random search 10-fold CV lần lượt được áp dụng để tinh chỉnh siêu tham số trong các nghiên cứu [8], [10]. Trong một nghiên cứu khác [9], hàm tối ưu Bayes được sử dụng kết hợp với thực hiện hoán đổi 300 lần tập *train/validation* để kiểm tra độ ổn định của mô hình khi dữ liệu thay đổi. Tuy nhiên, việc giữ nguyên các siêu tham số như được xác định qua lần đánh giá đầu tiên trong các lần hoán đổi này vô hình chung làm mất ý nghĩa tinh chỉnh siêu tham số.

Trong trường hợp một mô hình cho độ chính xác vượt trội hẳn so với các mô hình khác hoặc có sự khác biệt lớn giữa *training/validation*, cần thực hiện kiểm định chéo không gian (block-CV/k-fold CV), chia dữ liệu theo khối lưới, phân vị địa chất hoặc

tiểu lưu vực và tiến hành kiểm định chéo nhiều lần cho các nhóm.

c. Đánh giá mô hình

Đa số các nghiên cứu mới dừng ở việc báo cáo giá trị RMSE,  $R^2$  đơn thuần cùng với EI [18]; AIC, BIC và Taylor diagram [8]; PR-AUC và AUC-ROC [10]; MAE [11].

Nên tính toán và báo cáo khoảng tin cậy (CI) cho AUC/ACC và giá trị PR-AUC khi dữ liệu mất cân bằng; ví dụ: dữ liệu quan trắc mực nước dưới đất dài kỳ, trong đó đa số là giá trị mực nước không biến động, một số ít ngày có mực nước dâng cao/hạ thấp bất thường do mưa lớn, xả lũ/khai thác tầng đột biến; hoặc dữ liệu quan trắc chất lượng nước trong một vùng với phần lớn các điểm là không ô nhiễm, chỉ một số ít các điểm bị ô nhiễm.

d. Lựa chọn mô hình

Các nghiên cứu chọn mô hình tối ưu theo RMSE và  $R^2$  hoặc AUC-ROC tốt nhất trên tập *test* và thứ hạng ổn định qua các lần kiểm định chéo lặp lại.

e. Kiểm chứng và triển khai

Các nghiên cứu mới chỉ dừng ở đánh giá hiệu năng dự báo của mô hình trên tập *test*; chưa thực hiện kiểm định độc lập trên tập dữ liệu ở vùng khác, hoặc ở khoảng thời gian khác; do vậy, chưa đánh giá được khả năng khái quát hóa của mô hình để áp dụng rộng rãi.

### 3.2.4. Các thuật toán ML được ứng dụng

Nhìn chung, trong các nghiên cứu đã được thực hiện ở Việt Nam, các tác giả đã lựa chọn các mô hình AI/ML tương thích với bối cảnh dữ liệu. Conv1D là mô hình nổi trội nhất cho dữ liệu mực nước dưới đất ở Việt Nam hiện tại. Tuy nhiên, để áp dụng rộng rãi mô hình này cho dữ liệu thời gian thực, cần mở rộng tập dữ liệu, tích hợp thêm yếu tố ngoại sinh (mưa, chỉ số ENSO, mực nước biển, khoảng cách tới biển/sông, địa tầng, độ sâu tầng chứa nước, sử dụng đất).

Cần thử nghiệm các kiến trúc tiên tiến hơn (LSTM, LSTM autoencoder, hybrid CNN-RNN, Transformer) với khả năng xử lý dữ liệu dài hạn, nhiều chiều và cập nhật online; phù hợp với hệ thống quản lý tài nguyên nước thông minh.

## 3.3. Định hướng cho nghiên cứu ứng dụng AI trong lĩnh vực nước dưới đất của Việt Nam

Ứng dụng AI trong lĩnh vực nước dưới đất ở Việt Nam cần được định hướng phù hợp với đặc trưng từng vùng và nhu cầu thực tiễn. Về cơ bản, cần chú trọng công tác cảnh báo, dự báo suy thoái, cạn kiệt, ô nhiễm nước dưới đất

ở những vùng có tính tổn thương cao và tìm kiếm, thăm dò nguồn nước ở những vùng khan hiếm nước.

### 3.3.1. Ứng dụng AI cho cảnh báo, dự báo cạn kiệt, xâm nhập mặn nước dưới đất vùng ven biển

Ở vùng ven biển, đặc trưng nổi bật là khai thác nước dưới đất quy mô lớn phục vụ sinh hoạt, sản xuất nông nghiệp và nuôi trồng thủy sản, trong khi nguồn bổ cập hạn chế và chịu ảnh hưởng mạnh của thủy triều, mực nước biển dâng. Nhu cầu cấp thiết là dự báo cạn kiệt, cảnh báo xâm nhập mặn theo không gian – thời gian để điều chỉnh kế hoạch khai thác và bảo vệ tầng chứa nước. Do đó, dữ liệu bắt buộc phải bao gồm mực nước, EC, T (tần suất 1 lần/giờ) tại các giếng quan trắc; lưu lượng khai thác, chế độ khai thác, các thông tin theo giấy phép khai thác cho từng giếng, dữ liệu thủy văn–hải văn (mưa, bốc thoát hơi, mực nước sông/kênh, mực nước biển, độ mặn nước sông/cửa sông), dữ liệu địa chất thủy văn (độ sâu mái, đáy tầng chứa nước, hệ số thấm, chiều sâu ống lọc), dữ liệu viễn thám (DEM, LULC, NDVI, NDWI, NDMI), GRACE, khoảng cách tới biển/sông, công trình ngăn mặn.

Các mô hình DL theo chuỗi thời gian (LSTM, GRU, Temporal Fusion Transformer) kết hợp với mô hình không gian–thời gian (Graph Neural Network, ConvLSTM) thích hợp để dự báo biến động. Độ chính xác tối thiểu cần đạt là sai số mực nước  $\leq 0,3-0,5$  m và AUC-ROC  $\geq 0,85$  trong cảnh báo.

### 3.3.1. Ứng dụng AI cho cảnh báo, dự báo suy thoái, cạn kiệt, ô nhiễm nước dưới đất tại các khu công nghiệp tập trung

Ở các khu công nghiệp tập trung, thách thức lớn là suy thoái, cạn kiệt nước dưới đất do khai thác quá mức và nguy cơ ô nhiễm từ nước thải công nghiệp. Nhu cầu thực tiễn là thiết lập hệ thống cảnh báo sớm để phát hiện dị thường mực nước hoặc nồng độ ô nhiễm cũng như dự báo lan truyền chất ô nhiễm để phục vụ quản lý rủi ro. Do đó, yêu cầu dữ liệu không chỉ là mực nước dưới đất và các chỉ tiêu hóa lý cơ bản (pH, T, EC,  $\text{NO}_3^-$ ,  $\text{Cl}^-$ ) mà còn phải có thông tin định kỳ về hàm lượng kim loại nặng (As, Pb, Cd, Cr), hợp chất hữu cơ dễ bay hơi (VOC, BTEX), bản đồ phân bố tầng chứa nước, hướng dòng chảy nước dưới đất, hệ số thấm, loại hình sản xuất, lưu lượng xả thải, sơ đồ công trình ngầm, sơ đồ hệ thống thoát nước mưa/nước thải.

Các mô hình Isolation Forest, LSTM Autoencoder, One-Class SVM, Bayesian

Change-Point cho phát hiện dị thường; XGBoost hoặc LightGBM cho phân loại nguồn ô nhiễm là phù hợp để sử dụng. Độ chính xác mong muốn PR-AUC  $\geq 0,85$  và thời gian phát hiện sự cố  $\leq 7$  ngày.

### 3.3.2. Ứng dụng AI cho tìm kiếm, thăm dò nước dưới đất ở các vùng sâu, vùng xa, miền núi và hải đảo

Đối với vùng sâu, vùng xa, miền núi và hải đảo, thách thức là thiếu mạng lưới quan trắc và nhu cầu bức thiết là tìm kiếm, thăm dò nước dưới đất phục vụ cấp nước sinh hoạt tập trung. Do vậy, dữ liệu cần bao gồm các bản đồ địa chất - địa chất thủy văn, các phiếu lỗ khoan đã thực hiện, số liệu địa vật lý (ERT, VES, CSAMT), dữ liệu viễn thám (DEM, NDVI, NDWI, NDMI, lineament, LULC), lượng mưa, bốc hơi, giếng khai thác đang hoạt động, giếng khô. Các dữ liệu này phải được tích hợp thành hệ thống không gian thống nhất, có metadata về vị trí, cao độ và thời gian đo.

Các mô hình RF, XGB, LGB, CB kết hợp với mạng U-Net để lập bản đồ khoanh vùng tiềm năng nước dưới đất là lựa chọn thích hợp. Mục tiêu là dự báo được vị trí khoan giếng có nước với độ chính xác  $\geq 80\%$  và chiều sâu khoan giếng với sai số  $\leq 20\%$ .

## 4. KẾT LUẬN

Trong những năm gần đây, các nghiên cứu ứng dụng AI trong lĩnh vực nước dưới đất tại Việt Nam đã bước đầu được triển khai, phản ánh xu thế tất yếu của đổi mới công nghệ và chuyển đổi số trong quản lý tài nguyên nước. Tuy nhiên, số lượng công trình còn hạn chế và chất lượng chưa đồng đều, chủ yếu dừng ở mức thử nghiệm cho một vùng nhất định, chưa tạo được tác động rõ rệt trong thực tiễn quản lý. Đáng chú ý, hệ thống khung pháp lý về cảnh báo và dự báo nguồn nước dưới đất vẫn chưa hoàn thiện; Thông tư số 52/2025/TT-BNNMT vừa ban hành vẫn chưa đáp ứng được yêu cầu thực tiễn đối với công tác cảnh báo, dự báo nguồn nước dưới đất trong bối cảnh biến đổi khí hậu, gia tăng nhu cầu khai thác

và chuyển đổi số.

Một rào cản lớn khác là sự thiếu chuẩn hóa trong mạng lưới quan trắc và công tác thu thập dữ liệu. Hiện nay, dữ liệu quan trắc nước dưới đất tại Việt Nam chưa đảm bảo tính liên tục, đồng bộ và chất lượng cao – đây là yếu tố cốt lõi để AI có thể phát huy hiệu quả trong phân tích, cảnh báo và dự báo. Do đó, việc chuẩn hóa mạng quan trắc và công tác vận hành cũng như nâng cao chất lượng dữ liệu là điều kiện tiên quyết để các nghiên cứu ứng dụng AI có thể chuyển sang giai đoạn triển khai thực tế.

Trong tương lai, cần định hướng phát triển các nghiên cứu ứng dụng AI ở những khu vực có áp lực khai thác lớn và có nguy cơ cạn kiệt hoặc ô nhiễm nguồn nước dưới đất, chẳng hạn như các vùng phát triển công nghiệp tập trung, vùng ven biển ĐBSCL, vùng ven biển ĐBSH. Đây là những nơi vừa có nhu cầu quản lý cấp thiết, vừa có tiềm năng dữ liệu để triển khai thí điểm.

Tóm lại, việc ứng dụng AI trong cảnh báo và dự báo nước dưới đất ở Việt Nam mới ở giai đoạn khởi đầu, còn nhiều hạn chế về nghiên cứu, dữ liệu và chính sách. Cần có sự kết hợp đồng bộ giữa đầu tư nghiên cứu, hoàn thiện khung pháp lý và đặc biệt là chuẩn hóa mạng quan trắc nhằm tạo nền tảng dữ liệu vững chắc cho các giải pháp AI; sự phối hợp giữa các viện nghiên cứu, trường đại học với các cơ quan quản lý để biến kết quả nghiên cứu ứng dụng AI thành công cụ dự báo, cảnh báo, phục vụ quản lý bền vững tài nguyên nước dưới đất.

**LỜI CẢM ƠN:** Công trình nghiên cứu được nhóm tác giả thực hiện trong khuôn khổ Đề tài “Nghiên cứu ứng dụng trí tuệ nhân tạo (AI) và GIS cảnh báo và dự báo suy thoái, cạn kiệt và xâm nhập mặn nguồn nước dưới đất, áp dụng thử nghiệm tại các tỉnh ven biển đồng bằng Bắc Bộ” (Mã số đề tài: CS.2025.24), tổ chức chủ trì: Trung tâm Công nghệ và Dữ liệu tài nguyên nước (WATEDA).

## TÀI LIỆU THAM KHẢO

- [1] Wissal Ed-Dehbi, Mustapha Ahlaqqach and Jamal Benhra. 2025. Artificial intelligence for optimal water resource management: a literature review. Eng. Proc., 2025, 97,52.
- [2] Mrunmayee Dhapre, Shrikant Jadhav, Debanjana Das, Jehanzeb Khan, Youngsoo Kim, Sen Chiao, Thomas Danielson. 2025. A systematic review of machine learning in groundwater monitoring. Environmental Modelling and Software, 192, 106549
- [3] Nur Farahin Che Nordin, Nuruol Syuhadaa Mohd, Suhana Koting, Zubaidah Ismail, Mohsen Sherif, Ahmed El-Shafie. 2021. Groundwater quality forecasting modelling using artificial intelligence: a review. Groundwater for Sus. Development, 14, 100643.
- [4] Siva Rama Krishnan, M. K. Nallakaruppan, Rajeswari Chengoden, Srinivas Koppu, M. Iyapparaja,

- Jayakumar Sadhasivam and Sankaran SethuramanKrishnan. 2022. Smart water resources management using Artificial Intelligence-A review. *Sustainability*, 14, 13384.
- [5] Habtamu Tamiru, Meseret Wagari and Bona Tadese. 2022. An integrated Artificial Intelligence and GIS spatial analyst tools for delineation of groundwater potential zones in complex terrain: Fincha catchment, Abay Basi, Ethiopia. *Air, Soil and Water Research*, 15, 1-15.
- [6] Amirhossein Najafabadipour, Gholamreza Kamali and Hossein Nezamabadi-pour. 2022. Application of Artificial Intelligence techniques for the determination of groundwater level using spatio-temporal parameters. *ACS Omega*, 7, 10751-10764.
- [7] Nguyen Phong Tung, Duong Hai Ha, Abolfazl Jaafari, Nguyen Huu Duy, Tran Van Phong, Nadhir Al-Ansari, Indra Prakash, Le Van Hiep and Pham Thai Binh. 2020. Groundwater potential mapping combining Artificial Neural Network and Real AdaBoost Ensemble Technique: The DakNong Province case-study, Vietnam. *Int J Environ Res Public Health*. 17(7):2473. doi:10.3390/ijerph17072473
- [8] Tran Dang An, Maki Tsujimura, Ha Nan Thang, Nguyen Van Tam, Doan Van Binh, Dang Thanh Duc, Doan Quang Van, Bui Dieu Tien, Trieu Anh Ngoc, Le Vo Phu, Pham Thi Bich Thuc, Pham Tien Dat. 2021. Evaluating the predictive power of different machine learning algorithms for groundwater salinity prediction of multi-layer coastal aquifers in the Mekong Delta, Vietnam. *Ecological Indicators*, 127, 107790. <https://doi.org/10.1016/j.ecolind.2021.107790>
- [9] Heewon Jeong, Ather Abbas, Hyo Gyeom Kim, Hoang Van Hoan, Pham Van Tuan, Phan Thang Long, Eunhee Lee, Kyung Hwa Cho. 2024. Spatial prediction of groundwater salinity in multiple aquifers of the Mekong Delta region using explainable machine learning models. *Water Research*, 266, 122404. <https://doi.org/10.1016/j.watres.2024.122404>
- [10] Zheina J. Ottong, Reta L. Puspasari, Daeung Yoon, Kyoung-Woong Kim. 2022. Predicting As contamination risk in Red River Delta using machine learning algorithms. *Economic and Environmental Geology*, 55(2), 127–135. <https://doi.org/10.9719/EEG.2022.55.2.127>
- [11] Nguyen Gia Trong, Thanh, P. T., Tinh, L. D., Thao, N. B., & Elshewy, M. A. 2025. Application of deep learning models for groundwater data analysis: A comparative study of CNN (Conv1D), SimpleRNN, and Gated Recurrent Unit (GRU) models. *International Journal of Geoinformatics*, 21(3).
- [12] Tran Van Phong, Pham Thai Binh, Phan Trong Trinh, Ly Hai Bang, Vu Quoc Hung, Ho Si Lanh, Le Van Hiep, Lai Hop Phong, Mohammadtaghi Avand, Indra Prakash. 2021. Groundwater potential mapping using GIS-based hybrid Artificial Intelligence methods. *Groundwater*. 59, 745-760. doi:10.1111/gwat.13094
- [13] Hoang Phan Hai Yen, Pham Thai Binh, Tran Van Phong, Duong Hai Ha, Romulus Costache, Le Van Hiep, Nguyen Huu Duy, Mahdis Amiri, Nguyen Van Tao, Indra Prakash. 2021. Locally weighted learning based hybrid intelligence models for groundwater potential mapping and modeling: A case study at Gia Lai province, Vietnam. *Geoscience Frontiers*. 12:101154. doi:10.1016/j.gsf.2021.101154
- [14] Duong Hai Ha, Ngo Thi Thanh Huong, Tran Van Phong, Nguyen Duc Dam, Mohammadtaghi Avand, Nguyen Huu Duy, Mahdis Amiri, Le Van Hiep, Indra Prakash, Pham Thai Binh. 2022. Development and application of hybrid artificial Intelligence models for groundwater potential mapping and assessment. *Vietnam J Earth Sci*. 44(3), 410-429.
- [15] Nhu Viet Ha, Pham Viet Hoa, Laura Melgar-Garcia, Bui Dieu Tien. 2023. Comparative analysis of deep learning and Swarm-Optimized Random Forest for groundwater spring potential identification in tropical regions. *Remote Sensing*. 15(19), 4761. doi:10.3390/rs15194761
- [16] Nguyen Huu Duy, Giang Van Trong, Truong Quang Hai, Gheorghe Serban, Alexandra-Ionut Petrisor. 2024. Groundwater potential assessment in Gia Lai Province (Vietnam) using machine learning, remote sensing and GIS. *Geographia Technica*. 19(2), 13–32. doi:10.21163/GT\_2024.192.02
- [17] Tran Xuan Bien, Pham The Trinh, Luu Thuy Duong, Tran Van Phong, Vuong Hong Nhat, Le Van Hiep, Nguyen Duc Dam, Indra Prakash, Pham Thanh Tam, Pham Thai Binh. 2024. Groundwater potential zoning using Logistics Model Trees based novel ensemble machine learning model. *Vietnam J Earth Sci*. 46(2), 272-281.
- [18] Tran Van Ty, Le Van Phat, Huynh Van Hiep. 2018. Groundwater level prediction using artificial neural networks: A case study in Tra Noc Industrial Zone, Can Tho City, Vietnam. *Journal of Water Resource and Protection*, 10, 870–883. <https://doi.org/10.4236/jwarp.2018.109050>