



TUYỂN TẬP BÁO CÁO HỘI NGHỊ TOÀN QUỐC

KHOA HỌC TRÁI ĐẤT VÀ TÀI NGUYÊN VỚI PHÁT TRIỂN BỀN VỮNG (ERSD 2024)

HÀ NỘI 14 - 11 - 2024

ERSD 2024



NHÀ XUẤT BẢN GIAO THÔNG VẬN TẢI

**TIỂU BAN
CÔNG NGHỆ MỚI
TRONG XỬ LÝ MÔI TRƯỜNG**

MỤC LỤC

Bàn về khả năng khai thác nước dưới đất của nhà máy cấp nước khu công nghiệp Sa Đéc tỉnh Đồng Tháp <i>Đỗ Văn Bình, Trần Thị Thanh Thủy, Dương Thị Thanh Xuyên</i>	732
Đánh giá sự phân bố vi nhựa trong trầm tích tại cửa sông Cái (Khánh Hòa) <i>Tạ Lê Đăng Khôi, Nguyễn Thị Thanh Hoài, Lê Hùng Phú, Trần Anh Quân, Hoàng Văn Lương, Vu Ngọc Toan, Trần Thị Thu Hương</i>	737
Đánh giá tính hợp lý của các giếng khai thác nước dưới đất tại trạm cấp nước Dương Nội, Hà Đông <i>Đỗ Văn Bình, Dương Thị Thanh Xuyên, Đỗ Thị Hải, Trần Thị Kim Hà</i>	743
Kết quả khảo sát hoạt độ phóng xạ trong nước biển khu vực Hải Phòng - Quảng Ninh <i>Nguyễn Văn Dũng, Nguyễn Thị Thu Trang, Lê Anh Thơ, Bùi Chí Tiến</i>	749
Đánh giá chất lượng môi trường đất và nước tại một số làng nghề chế tác kim loại và cơ kim khí <i>Nguyễn Mai Hoa</i>	754
Đánh giá lan truyền bụi từ trạm nghiền xi măng Norcem Yên Bình, Lai Châu sử dụng ứng dụng mô hình METI-LIS <i>Trần Anh Quân, Phạm Đức Bình</i>	760
Tối ưu hóa quá trình hòa tách bộ phận nam châm trong ổ cứng đã qua sử dụng để thu hồi kim loại đất hiếm <i>Phạm Khánh Huy</i>	767
Nghiên cứu hiệu quả dự báo chất lượng nước mặt bằng các mô hình học máy: ứng dụng tại sông Ba Chẽ, tỉnh Quảng Ninh <i>Nguyễn Thị Hồng, Nguyễn Thị Thu Huyền, Đào Trung Thành</i>	773
Nghiên cứu vai trò của chủng nấm <i>Curvularia lunata</i> trong kích thích sinh trưởng và xử lý ô nhiễm môi trường <i>Vũ Thị Lan Anh, Nguyễn Thị Hồng, Trần Thị Ngọc, Nguyễn Phương Đông, Nguyễn Thị Nhạn</i>	779
Ảnh hưởng của bộ rễ thực vật đến việc loại bỏ hợp chất 4-Nitrophenon trong nước <i>Nguyễn Hoàng Nam, Nguyễn Hoàng Nam Anh, Nguyễn Mạnh Hà, Phạm Việt Đức</i>	784
Nghiên cứu đánh giá vai trò của chủng vi nấm <i>Trichoderma harzianum</i> <i>Vũ Thị Lan Anh, Nguyễn Văn Dũng, Nguyễn Thị Nhạn</i>	790
Hiệu quả xử lý Benzen, Methyl- <i>tert</i> - butylether trong nước ngầm của hệ thống constructed wetland dòng chảy ngang dưới bề mặt <i>Nguyễn Hoàng Nam, Nguyễn Hoàng Nam Anh, Nguyễn Việt Hùng, Phạm Việt Đức</i>	796
Chuyển đổi số với ngành công nghiệp môi trường ứng dụng với các doanh nghiệp khai khoáng ở Việt Nam <i>Nguyễn Ngọc Bảo, Nông Việt Hùng, Dương Phi Hùng, Nguyễn Thị Thu Hương, Nông Việt Trung, Nguyễn Hồng Thái, Ngô Thái Vinh, Dương Mai Yến</i>	803
Nghiên cứu công nghệ phân tách, thu hồi bụi than, sắt và khoáng silica từ tro bay nhà máy nhiệt điện hướng đến sản xuất nguyên liệu công nghiệp theo mô hình kinh tế tuần hoàn <i>Nông Việt Hùng, Nguyễn Ngọc Bảo, Dương Phi Hùng, Nguyễn Thị Thu Hương, Nông Việt Trung, Vũ Mạnh Anh, Nguyễn Ngọc Trức, Ngô Thái Vinh, Dương Mai Yến</i>	807
Đánh giá khả năng tiếp nhận nước thải và sức chịu tải của sông Ngũ Huyện Khê thuộc địa bàn tỉnh Bắc Ninh <i>Nguyễn Thị Hoà</i>	812

Nghiên cứu hiệu quả dự báo chất lượng nước mặt bằng các mô hình học máy: ứng dụng tại sông Ba Chẽ, tỉnh Quảng Ninh

Nguyễn Thị Hồng*, Nguyễn Thị Thu Huyền, Đào Trung Thành
Trường Đại học Mỏ - Địa chất

TÓM TẮT

Việc sử dụng các mô hình học máy trong dự báo chất lượng nước mặt có thể mang lại nhiều lợi ích trong việc quản lý hiệu quả tài nguyên nước và đảm bảo an ninh nguồn nước. Mô hình học máy có khả năng phân tích dữ liệu từ các thông số môi trường như WQI, DO, BOD, pH và các chất ô nhiễm khác để dự báo chất lượng nước mặt trong tương lai. Ngoài ra, mô hình học máy có thể được kết nối với các thiết bị cảm biến để thu thập dữ liệu về chất lượng nước mặt trong thời gian thực. Khi dữ liệu được cập nhật liên tục, mô hình có thể phân tích và dự báo chất lượng nước mặt ngay lập tức, giúp nhà quản lý có thể phát hiện sớm các vấn đề về ô nhiễm nguồn nước và đưa ra biện pháp khắc phục kịp thời. Báo cáo nghiên cứu ứng dụng tại sông Ba Chẽ của huyện Ba Chẽ, đây là huyện miền núi của tỉnh Quảng Ninh với số lượng lớn đồng bào dân tộc thiểu số sinh sống tại nhiều khu vực có địa hình phức tạp. Nhiều năm qua, người dân vẫn sử dụng nguồn nước tự chảy từ các bể chứa, từ khe, suối, nên chất lượng nước khó đảm bảo. Cần thiết phải có biện pháp quản lý hiệu quả tài nguyên nước trên địa bàn để người dân yên tâm sinh hoạt, sản xuất, ổn định đời sống, vươn lên thoát nghèo.

Từ khóa: Học máy, Dự báo chất lượng nước, WQI, Sông Ba Chẽ.

1. Mở đầu

Sông Ba Chẽ bắt nguồn từ đỉnh Thiên Sơn cao hơn 1.200m (phía Nam xã Đồn Đạc, huyện Ba Chẽ) chảy theo hướng Tây Bắc - Đông Nam qua huyện Ba Chẽ và đổ ra Vịnh Bái Tử Long. Đây là một trong bốn con sông lớn của tỉnh Quảng Ninh, không chỉ cung cấp nguồn nước quan trọng cho sinh hoạt, nông nghiệp và công nghiệp, mà còn là một tuyến giao thông thủy quan trọng và có tầm ảnh hưởng lớn đến phát triển kinh tế và văn hóa của khu vực. Dòng sông đã tạo ra một môi trường sống đa dạng và phong phú, góp phần vào sự hình thành và phát triển của các cộng đồng dân cư xung quanh. Tuy nhiên, bên cạnh việc phát triển kinh tế và đô thị hóa một cách ồ ạt của địa phương cũng làm ảnh hưởng đến chất lượng nước sông. Kết quả quan trắc giám sát chất lượng nước cho thấy sông Ba Chẽ đang bị ô nhiễm hữu cơ, giá trị COD, BOD₅ tại hầu hết các điểm quan trắc đều vượt giới hạn cho phép từ 2 đến 2,3 lần. Phương pháp đánh giá hiện tại thường dựa trên kết quả đo nhanh tại chỗ, lấy mẫu và phân tích trong phòng thí nghiệm với các thông số vật lý, hóa học và sinh học trong nước. Phương pháp này tuy có độ chính xác cao nhưng khá tốn thời gian, kinh phí, trang thiết bị và nhân lực kỹ thuật cao.

Hiện nay, việc áp dụng học máy (Machine Learning) vào nhiều lĩnh vực khoa học khác nhau đang trở thành xu thế nhằm tăng độ chính xác của các phép toán dự báo và phân loại. Học máy học các mối quan hệ thống kê bậc cao, do đó thường vượt trội so với các phương pháp mô hình hóa và dự báo truyền thống. Trong dự báo chất lượng nước, các thuật toán học máy có thể ước lượng và phân tích các thông số, dựa trên mối quan hệ tuyến tính và phi tuyến tính. Các kỹ thuật học máy cũng đóng một vai trò quan trọng trong việc dự báo chất lượng nước sông nhờ mạng nơ ron nhân tạo (ANN), vectơ hỗ trợ (SVM), rừng ngẫu nhiên (RF), cây quyết định (DT) và học sâu (DL) (Sanja Cojbasic và nkk, 2023). Học máy làm tăng hiệu quả đánh giá chất lượng nước mặt bằng Chỉ số chất lượng nước (WQI) thông qua các thuật toán có giám sát như dự báo đa lớp đối với thuật toán AdaBoost (AB), thuật toán K-nearest neighbors (KNN), quy trình Gaussian (GP) và hồi quy tăng cường độ dốc cực đại (XGBR) (Masood và nkk, 2023; Anita và nkk, 2023). Ngoài ra, các phương pháp tiếp cận dựa trên máy học như rừng ngẫu nhiên đã thành công trong việc phân loại chất lượng nước sông thành các loại ô nhiễm hoặc ô nhiễm nhẹ với độ chính xác cao, nhấn mạnh tầm quan trọng của thông số DO trong dự đoán chất lượng nước (Nasaruddin và nkk, 2023). Hơn nữa, việc tích hợp các thuật toán học sâu như mạng tích chập đồ thị (Graph Convolutional Network - GCN) và mạng thần

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

kinh bộ nhớ ngắn hạn dài (Long Short-Term Memory - LSTM) đã cải thiện đáng kể độ chính xác của các mô hình dự báo chất lượng nước, hỗ trợ quản lý môi trường nước hiệu quả (Juan Huan và nnk, 2023).

Ở Việt Nam, nhiều nghiên cứu thường dùng chỉ số chất lượng nước (WQI) để đánh giá chất lượng nước mặt và khả năng sử dụng của nguồn nước đối với từng mục đích khác nhau. Tuy nhiên, việc tính toán WQI từ các thông số quan trắc chất lượng nước còn gặp nhiều hạn chế do cần nhiều thông số quan trắc và tính toán còn tương đối phức tạp. Trong những năm gần đây, nhiều nghiên cứu tại Việt Nam áp dụng thuật toán học máy để dự báo chỉ số chất lượng nước mặt, nâng cao độ chính xác và giảm chi phí giám sát. Các thuật toán học máy, đặc biệt là tăng cường độ dốc (GBoost), được ứng dụng hiệu quả trong ước tính chỉ số chất lượng nước mặt ở đồng bằng sông Hồng, Việt Nam, nâng cao độ chính xác dự báo và hiệu quả chi phí (Nguyen Quoc Son và nnk, 2023). Kể cả với môi trường nước biển, các thuật toán cây quyết định (DT), rừng ngẫu nhiên (RF), hồi quy tăng cường độ dốc (GBR) và hồi quy tăng cường độ dốc Ada (ABR) sử dụng ảnh viễn thám Sentinel-2 đối với khu vực ven biển tỉnh Bình Định. Nghiên cứu cho thấy thuật toán rừng ngẫu nhiên (RF) dự báo các thông số chất lượng nước biển chính xác nhất so với dữ liệu đo tại hiện trường ($R^2 = 0,59 \div 0,74$) (Quang Nguyen Hong và nnk, 2023). Việc tìm ra thuật toán để dự báo hiệu quả WQI là rất quan trọng và cần thiết nhằm phục vụ tốt hơn cho công tác đánh giá chất lượng nước mặt cho vùng nghiên cứu. Nghiên cứu tại bán đảo Cà Mau đã áp dụng phương pháp Bayes (BMA) để lựa chọn các thông số chất lượng nước tối ưu (pH, BOD5, PO4 và Coliform) để tính toán WQI. Kết quả cho thấy sử dụng các thuật toán tăng cường độ dốc (GBoost) có kết quả dự báo chính xác nhất ($R^2 = 0,973$), tăng cường độ dốc cực đại (XGBoost) có $R^2 = 0,966$, cây quyết định có $R^2 = 0,944$ và tăng cường độ dốc nhẹ (LGBoost) có $R^2 = 0,928$ (Nguyễn Đức Phong và Hà Hải Dương, 2023).

Như vậy, các nghiên cứu trước đây về ứng dụng học máy trong dự báo chất lượng nước chủ yếu tập trung vào dự báo các thông số chất lượng nước và tính toán chỉ số chất lượng nước (WQI). Một số nghiên cứu còn kết hợp thuật toán học máy với mạng giám sát thời gian thực và cho kết quả rất khả quan. Tuy nhiên, các nghiên cứu này chủ yếu trên thế giới, còn tại Việt Nam, có rất ít nghiên cứu đánh giá tiềm năng của thuật toán học máy trong dự báo chỉ số chất lượng nước mặt (WQI) dựa trên dữ liệu đầu vào nhằm giảm chi phí giám sát chất lượng nước mặt. Vì vậy, việc nghiên cứu ứng dụng các mô hình học máy để dự báo chỉ số chất lượng nước mặt tại sông Ba Chẽ là cần thiết và góp phần đưa ra các phương pháp tính toán chỉ số chất lượng nước mặt hiệu quả và tiết kiệm hơn.

2. Thu thập mẫu và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu và lấy mẫu

Vị trí quan trắc lấy mẫu được thống kê trong Bảng 1 và thể hiện trên Hình 1 trong 2 đợt khảo sát mùa khô và mùa mưa giai đoạn 2017÷2022. Mẫu được lấy theo TCVN 6663-6-2018 (ISO 5667-6-2014) trong lấy mẫu nước sông và TCVN 6663-3-2016 (ISO 5667-3-2012) trong bảo quản và xử lý mẫu nước.



Hình 1. Các điểm lấy mẫu nước trên sông Ba Chẽ.

Bảng 1. Tọa độ điểm khảo sát thu thập mẫu nước mặt trên sông Ba Chẽ.

SHM	Vị trí	Tọa độ VN2000 (107°45', múi chiều 3°)	
		X	Y
NM01	Chân cầu ngầm xã Thanh Sơn	2356132	441516
NM02	Phía sau điểm hợp lưu với sông Làng Cống tại xã Đồn Đạc	2353555	448881
NM03	Điểm lấy nước về nhà máy nước Ba Chẽ	2353478	450465
NM04	Đập tràn thị trấn Ba Chẽ	2353426	451224
NM05	Chân cầu Ba Chẽ 2	2353207	452971

2.2. Phương pháp thu thập dữ liệu

Dữ liệu cho bài báo này chủ yếu tập trung vào việc thu thập thông số chất lượng nước như pH, TSS, NTU, BOD₅, COD, DO, NO₃⁻, NH₄⁺, NO₂⁻, PO₄³⁻, Zn, Fe, Cu, Pb, Hg, và Coliform.

2.3. Xử lý số liệu

- Đánh giá mẫu nước theo QCVN 08:2023/BTNMT, mức A - Quy chuẩn kỹ thuật Quốc gia về chất lượng nước mặt phục vụ mục đích cấp nước sinh hoạt.

- Quy trình tính toán chỉ số WQI theo hướng dẫn trong Quyết định số 1460/QĐ-TCMT (Tổng cục môi trường, 2019) và so sánh đánh giá chất lượng nước như trong Bảng 2, đối với thủy vực cần chú ý ô nhiễm chất hữu cơ thì tính theo công thức như sau:

$$WQI = \frac{WQI_I}{100} \times \frac{(\prod_{i=1}^m WQI_{III})^{1/m}}{100} \times [(\frac{1}{k} \sum_{i=1}^k WQI_{IV})^2 \times \frac{1}{l} \sum_{i=1}^l WQI_V]^{1/3} \quad (1)$$

Trong đó:

WQI_I: Kết quả tính toán đối với thông số nhóm I (Thông số pH);

WQI_{III}: Kết quả tính toán đối với thông số nhóm III (Các thông số As, Cd, Pb, Cr⁶⁺, Cu, Zn, Hg);

WQI_{IV}: Kết quả tính toán đối với thông số nhóm IV (Các thông số DO, BOD₅, COD, TOC, N-NH₄⁺, N-NO₃⁻, N-NO₂⁻, P-PO₄³⁻);

WQI_V: Kết quả tính toán đối với thông số nhóm V (Các thông số E.coli, Coliform).

Bảng 2. Bảng đánh giá chất lượng nước theo WQI.

Giá trị WQI	Mức đánh giá chất lượng nước	Màu
91÷100	Sử dụng tốt cho mục đích cấp nước sinh hoạt	Xanh nước biển
76÷90	Sử dụng cho mục đích cấp nước sinh hoạt nhưng cần các biện pháp xử lý phù hợp	Xanh lá cây
51÷75	Sử dụng cho mục đích tưới tiêu và các mục đích tương đương khác	Vàng
26÷50	Sử dụng cho giao thông thủy và các mục đích tương đương khác	Da cam
10÷25	Nước ô nhiễm nặng, cần các biện pháp xử lý trong tương lai	Đỏ
<10	Nước nhiễm độc, cần có biện pháp khắc phục, xử lý	Nâu

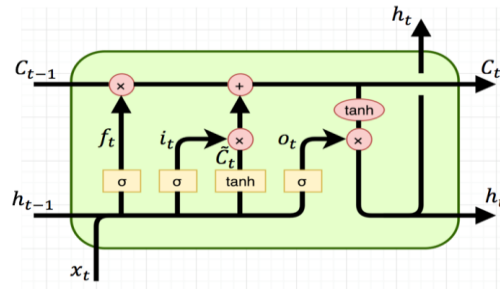
2.4. Phương pháp học máy

- Mạng nơ ron hồi quy (RNN) là một loại thuật toán học máy phù hợp với dữ liệu liên tục, đa biến có thể xử lý dữ liệu tuần tự và đã được chứng minh hiệu quả trong dự báo chất lượng nước (Abba và nnk, 2020).

- Mạng nơ ron bộ nhớ ngắn-dài (LSTM) kế thừa thuật toán của RNN để xử lý các chuỗi dữ liệu theo thời gian (sequence/ time-series), phù hợp trong việc dự báo chất lượng nước thời gian thực. LSTM vừa học được các thông tin từ state gần nó (short term memory) giống RNN và các thông tin từ state ở trước nó rất xa (long term memory) điều mà RNN không làm được (Hochreiter, 1997; Ye và nnk, 2019). Ở state thứ t của mô hình LSTM:

+ Output: c_t, h_t , trong đó c là cell state, h là hidden state.

+ Input: c_{t-1}, h_{t-1}, x_t , trong đó x_t là input ở state thứ t của mô hình, c_{t-1}, h_{t-1} là output của layer trước.



Hình 2. Cấu trúc của mô hình LSTM.

$$\text{Forget gate: } f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f), \quad \tilde{c}_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c). \quad (2)$$

$$\text{Input gate: } i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i). \quad (3)$$

$$\text{Output gate: } o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o). \quad (4)$$

$$\text{Long memory: } c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t. \quad (5)$$

$$\text{Short memory: } h_t = o_t \times \tanh(c_t). \quad (6)$$

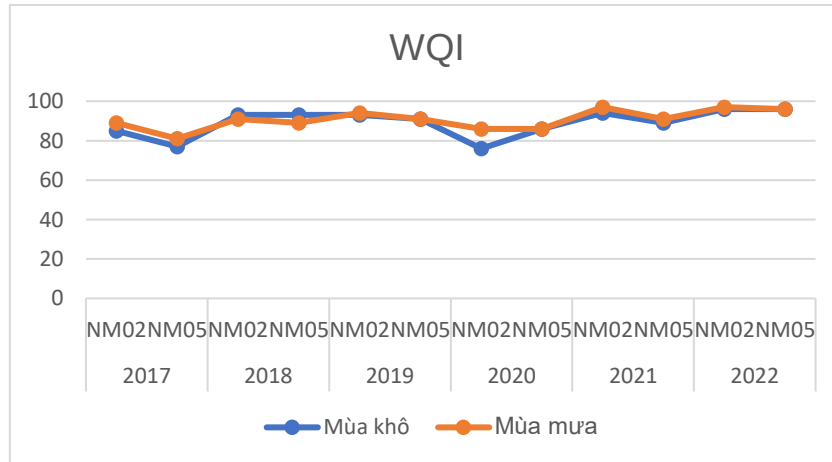
Ghi chú: σ , $\tanh$ là bước dùng sigma, tanh activation function; $\times$, $+$ là phép nhân, phép cộng ma trận. f_t , i_t , và o_t tương ứng với forget gate, input gate và output gate.

- Việc đánh giá độ chính xác của mô hình trong dự báo chỉ số chất lượng nước tại khu vực nghiên cứu bằng các chỉ số như: Sai số tuyệt đối trung bình (MAE), Sai số bình phương trung bình (MSE), Căn bậc hai của giá trị trung bình của sai số bình phương (RMSE), R-square (R^2).

3. Kết quả và thảo luận

3.1. Kết quả đánh giá chất lượng nước mặt sông Ba Chẽ bằng chỉ số WQI

Kết quả tính toán giá trị WQI của các mẫu phục vụ huấn luyện mô hình được thể hiện trong Hình 3 cho thấy diễn biến chất lượng nước sông Ba Chẽ có sự biến động theo mùa, chất lượng nước mùa khô thấp hơn mùa mưa. Trong giai đoạn giám sát chất lượng nước khá ổn định ở mức tốt đến rất tốt với giá trị WQI dao động trong khoảng 76 đến 97, đáp ứng tiêu chuẩn cấp nước sinh hoạt. Trong đó, tại các điểm quan trắc NM-02 là điểm trước khu vực lấy nước về nhà máy nước sạch Ba Chẽ có chất lượng tốt hơn ở các điểm phía hạ nguồn khi chưa chịu ảnh hưởng nhiều của việc xả nước thải sinh hoạt, dịch vụ và công nghiệp trên địa bàn.



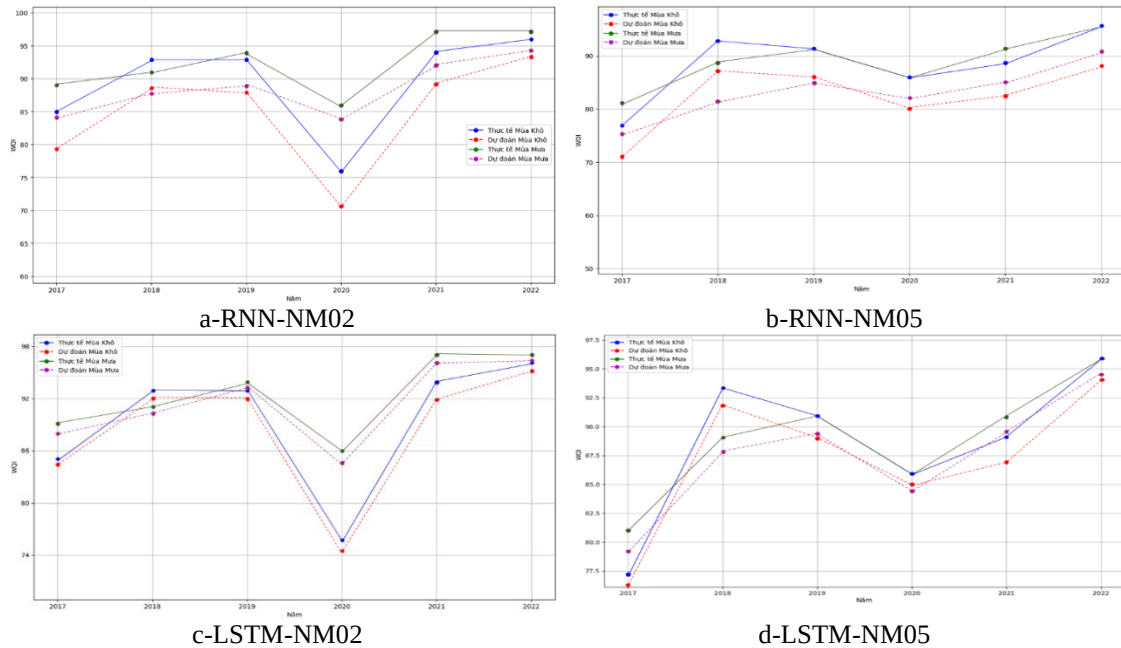
Hình 3. Chỉ số chất lượng nước (WQI) của sông Ba Chẽ giai đoạn 2017-2022.

3.2. Nghiên cứu dự báo chỉ số chất lượng nước mặt bằng học máy

3.2.1 Kết quả tính toán chỉ số chất lượng nước mặt

Dựa trên kết quả thu được, lựa chọn các thông số để xây dựng mô hình học máy. Nghiên cứu đã xây dựng mô hình từ các thông số (BOD_5 , NH_4^+ , PO_4^{3-} , TSS, NTU và Coliform) gồm dữ liệu đào tạo (24 mẫu) và thử nghiệm (36 mẫu) để dự báo chất lượng nước mặt (WQI) theo các mô hình là Mạng nơ ron hồi quy (RNN) và Mạng nơ ron bộ nhớ ngắn-dài (LSTM).

Kết quả báo cáo WQI và biểu đồ so sánh giữa giá trị WQI dự đoán và đo được đối với tập dữ liệu thực nghiệm theo các mô hình học máy được trình bày trên Hình 4.



Hình 4. So sánh giữa WQI dự báo và thực tế đối với dữ liệu thử nghiệm theo các mô hình học máy.

3.2.2. Kết quả đánh giá mô hình

Kết quả đánh giá mô hình học máy dự vào 4 tiêu chí (MAE, MSE và RMSE và R^2) để dự đoán chỉ số chất lượng nước mặt tại khu vực nghiên cứu được thể hiện tại Bảng 3.

Bảng 3. Đánh giá các mô hình dự báo chỉ số chất lượng nước mặt khu vực nghiên cứu.

Học máy	Thông số đầu vào	Đầu ra	Đánh giá			
			MAE	MSE	RMSE	R^2
RNN	BOD ₅ , NH ₄ ⁺ , PO ₄ ³⁻ , TSS, NTU, Coliform	WQI	5,7	78,6	8,85	0,82
LSTM	BOD ₅ , NH ₄ ⁺ , PO ₄ ³⁻ , TSS, NTU, Coliform	WQI	5,5	73,0	8,52	0,86

Theo kết quả tính toán, mô hình học máy LSTM có độ chính xác cao hơn mô hình RNN. Trong đó, mô hình mô hình RNN có R^2 là 0,82; giá trị lỗi là 5,70; 78,6; 8,85 và mô hình LSTM có R^2 là 0,86; giá trị lỗi là 5,50; 73,0; và 8,52. Có thể thấy, các mô hình học máy áp dụng trong bài báo đều có thể dự đoán tốt WQI cho khu vực nghiên cứu (R^2 đều lớn hơn 80%). Do đó, việc ứng dụng các mô hình học máy trong tính toán WQI cho sông Ba Chẽ là phù hợp và tiết kiệm chi phí. Từ đó, giúp tăng hiệu quả trong công tác giám sát chất lượng nước sông phục vụ cấp nước sinh hoạt, ổn định đời sống và phát triển kinh tế-xã hội trên địa bàn huyện. Đồng thời, nâng cao công tác quản lý tài nguyên nước, đặc biệt là tại các vùng có điều kiện khó khăn như huyện miền núi Ba Chẽ.

4. Kết luận

Kết quả nghiên cứu đã đưa ra cơ sở khoa học cho việc tính toán chỉ số chất lượng nước mặt bằng học máy và đã phát huy hiệu quả trong dự báo chỉ số chất lượng nước (WQI) tại khu vực nghiên cứu. Báo cáo đã xây dựng mô hình học máy dự báo WQI với 5 tham số được chọn, bao gồm: BOD₅, NH₄⁺, PO₄³⁻, TSS, NTU và Coliform. Theo kết quả tính toán dự báo WQI của các mô hình học máy cho thấy mô hình LSTM có độ chính xác hơn mô hình RNN và cả hai mô hình đều có độ chính xác cao, dao động từ 82÷86%.

Tuy nhiên, bộ dữ liệu sử dụng trong bài báo còn hạn chế chưa đạt được độ chính xác cao như mong muốn. Vì vậy, cần tiếp tục bổ sung thêm bộ dữ liệu về nguồn xả thải, thủy động lực học dòng chảy và đặc biệt là những yếu tố nhân tạo gây ảnh hưởng đến chất lượng nước sông. Từ đó, giúp nâng cao độ chính xác hơn nữa của mô hình học máy trong công tác dự báo và nâng cao hiệu quả quản lý môi trường sông để đảm bảo chất lượng nước phục vụ mục đích cấp nước sinh hoạt bền vững trên địa bàn huyện Ba Chẽ.

Lời cảm ơn

Kết quả bài báo được hoàn thành dưới sự hỗ trợ kinh phí và số liệu của đề tài “Ứng dụng mô hình LSTM trong dự báo chất lượng nước sông Ba Chẽ phục vụ cấp nước sinh hoạt cho huyện Ba Chẽ, tỉnh Quảng Ninh” - Mã số T24-24. Các tác giả xin trân trọng cảm ơn.

Tài liệu tham khảo

- Sanja Cojbasic, Sonja Dmitrasinovic, Marija Kostic, Maja Turk Sekulic, Jelena Radonic, Ana Dodig, Milan Stojkovic, (2023). Application of machine learning in river water quality management: a review. *Water Science & Technology*, 88 (9): 2297–2308. doi: 10.2166/wst.2023.331
- Masood, A.; Niazkar, M.; Zakwan, M.; Piraei, R. (2023). A Machine Learning-Based Framework for Water Quality Index Estimation in the Southern Bug River. *Water*, 15(20), 3543. doi: 10.3390/w15203543
- Anita M., M. G. Dinesh, C. Lakshmi Priya, V. Sharmila, A. Muthuram, Siva Subramanian R, (2023). Water Quality Prediction using Machine Learning: A Comparative Study. *Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, 348-353. doi: 10.1109/ICAISS58487.2023.10250743.
- Nasaruddin, N.; Ahmad, A.; Zakaria, S. F.; Ul-Saufie, A. Z.; & Osman, M. S. (2023). Predicting Kereh River's Water Quality: A comparative study of machine learning models. *Environment-Behaviour Proceedings Journal*, 8(SI15), 213–219. <https://doi.org/10.21834/e-bpj.v8iSI15.5097>
- Juan Huan, Wenjie Liao, Yongchun Zheng, Xiangen Xu, Hao Zhang, Bing Shi, (2023). A deep learning model with spatio-temporal graph convolutional networks for river water quality prediction. *Water Science & Technology: Water Supply*, 23(7), 2940-2957. doi: 10.2166/ws.2023.164
- Nguyen Quoc Son, Nguyen Cam Linh, Le Thi Phuong Quynh, and Le Phuong Thu, (2023). Application of Machine Learning Algorithms in Studying Water Quality in the Red River System. *VNU Journal Of Science: Earth And Environmental Sciences*, 39(2), 52-63.
- Quang Nguyen Hong, Nguyen Tran Dinh, Nguyen Tran Dien, and Le Thanh Son, (2023). Calibration of Sentinel-2 Surface Reflectance for Water Quality Modelling in Binh Dinh's Coastal Zone of Vietnam. *Sustainability*, 15(2), 1410.
- Nguyễn Đức Phong, Hà Hải Dương (2023). Nghiên cứu ứng dụng các mô hình học máy để dự báo chỉ số chất lượng nước mặt vùng bán đảo Cà Mau. *Tạp chí Khoa học và Công nghệ Thủy lợi* 76.
- Tổng cục môi trường, 2019. *Quyết định số 1460/QĐ-TCMT ngày 12 tháng 11 năm 2019 của Tổng cục môi trường ban hành Hướng dẫn Kỹ thuật tính toán và Công bố chỉ số chất lượng nước Việt Nam (VN_WQI)*.
- Abba, Sani Isah; Pham, Quoc Bao; Saini, Gaurav; Linh, Nguyen Thi Thuy; Ahmed, Ali Najah; Mohajane, Meriam; Khaledian, Mohammadreza; Abdulkadir, Rabi'u Aliyu; Bach, Quang-Vu (2020). Implementation of data intelligence models coupled with ensemble machine learning for prediction of water quality index. *Environmental Science and Pollution Research* 27, 41524 - 41539.
- S. Hochreiter, J. Schmidhuber (1997). Long short-term memory. *Neural Comput* 9, 1735-1780.
- Ye, Qiangqiang; Yang, Xueqin; Chen, Chaobo; Wang, Jingcheng (2019). *River Water Quality Parameters Prediction Method Based on LSTM-RNN Model*. *Chinese Control And Decision Conference (CCDC)*, 3024–3028.

ABSTRACT

Research on surface water quality prediction using machine learning: application at the Ba Che River, Quang Ninh province

Dao Trung Thanh, Nguyen Thi Hong, Nguyen Thi Thu Huyen
Hanoi University of Mining and Geology

Applications of machine learning models in predicting surface water quality have brought efficiency to water resource management. Machine learning models can predict surface water quality through data analysis from physical, chemical, and biological parameters (WQI, DO, BOD, pH, and other pollutants). In addition, it can be connected to sensor devices to collect data on surface water quality in real-time. These models can immediately analyze and predict surface water quality, so managers can detect water pollution problems early and take timely corrective measures. The models used in the article are RNN and LSTM, which have been used by many scientists and have proven their effectiveness. The scope of the article's research is the Ba Che river area of the Ba Che mountainous district with a large number of ethnic minorities living in many areas with complex terrain. The scope of the article's research is the Ba Che river area of the Ba Che mountainous district with complex terrain, sparse population, and difficult life. In recent years, people have used water from reservoirs, streams and streams, so water quality is difficult to ensure. It is necessary to have measures to effectively manage water resources so that people can feel secure in living, producing, stabilizing their lives, and escaping poverty.

Keywords: Machine learning; Water quality forecasting; WQI; Ba Che River.

KHOA HỌC TRÁI ĐẤT VÀ TÀI NGUYÊN VỚI PHÁT TRIỂN BỀN VỮNG (ERSD 2024)



ISBN: 978-604-76-3040-0



9 786047 630400

SÁCH KHÔNG BÁN