

DOI:10.22144/ctu.jsi.2019.135

ỨNG DỤNG MẠNG NƠON HỒI QUY TỔNG QUÁT VÀ THUẬT TOÁN NỘI SUY ĐÁNH GIÁ CHẤT LƯỢNG NƯỚC MẶT SÔNG VÀ CÁC CHI LƯU SÔNG ĐỒNG NAI, TỈNH BÌNH DƯƠNG NĂM 2012 - 2018

Nguyễn Thị Thanh Thúy^{1*} và Nguyễn Hiền Thân^{1,2}

¹Khoa Khoa học Quản lý, Trường Đại học Thủ Dầu Một

²Viện Môi trường và Tài nguyên, Đại học Quốc gia Thành phố Hồ Chí Minh

*Người chịu trách nhiệm về bài viết: Nguyễn Thị Thanh Thúy (email: thanhthuy014361@gmail.com)

Thông tin chung:

Ngày nhận bài: 03/07/2019

Ngày nhận bài sửa: 29/08/2019

Ngày duyệt đăng: 16/10/2019

Title:

Application of generalized regression neural network and interpolation algorithm to assess water quality of Dong Nai river and tributaries in Binh Duong from 2012 to 2018

Từ khóa:

Chất lượng nước, GIS, mạng nơon hồi quy, nội suy, sông Đồng Nai

Keywords:

Dong Nai River, generalized regression neural network, GIS, water quality

ABSTRACT

The Dong Nai River is one of the main water supply sources for living and production activities in Binh Duong province. However, the Dong Nai River is currently under pressure by emissions from industries, urban, agriculture and service activities. In this paper, the generalized regression neural network model (GRNN) and the interpolation algorithm was used to evaluate water quality of the Dong Nai river and tributaries. Monitoring data used was for 7 years from 2012 to 2018 at 12 monitoring stations. The results of the study indicated that the GRNN model was a good performance to assess the water quality of the Dong Nai river and tributaries. Comparing the results of the GRNN model with the water quality index showed that the GRNN model was reliable and in close agreement with the actual results with $R^2 = 0.938$, $RMSE = 0.055$, $E = 0.935$. Water quality interpolation map also illustrated that the water quality in the neighbor area was still good. Besides, the GRNN model has a lower cost and less time consumption compared to the WQI method of the Vietnam Environment Administration.

TÓM TẮT

Sông Đồng Nai là một trong những nguồn nước cấp chính cho hoạt động sinh hoạt và sản xuất tỉnh Bình Dương. Tuy nhiên, hiện nay sông Đồng Nai đang chịu nhiều áp lực bởi nguồn phát thải từ các khu công nghiệp, đô thị, nông nghiệp và dịch vụ. Trong bài báo này, mô hình mạng nơon hồi quy tổng quát (GRNN) và thuật toán nội suy được sử dụng để đánh giá chất lượng nước sông Đồng Nai và các chi lưu. Dữ liệu quan trắc được sử dụng trong 7 năm từ năm 2012-2018 tại 12 điểm quan trắc tại Sở Tài nguyên và Môi trường tỉnh Bình Dương. Kết quả nghiên cứu chỉ ra GRNN có thể giúp đánh giá chất lượng nước sông Đồng Nai và các chi lưu với $RMSE = 0,052$ trong huấn luyện và $RMSE = 0,061$ trong kiểm tra mô hình. So sánh kết quả tính chất lượng nước từ mô hình GRNN và chỉ số WQI của Tổng cục Môi trường cho thấy GRNN cho kết quả đánh giá đáng tin cậy và gần với kết quả thực tế với $R^2 = 0,938$, $RMSE = 0,055$, $E = 0,935$. Bên cạnh đó, mô hình GRNN có chi phí đánh giá thấp và thời gian tính toán nhanh hơn so với phương pháp đánh giá WQI của Tổng cục Môi trường.

Trích dẫn: Nguyễn Thị Thanh Thúy và Nguyễn Hiền Thân, 2019. Ứng dụng mạng nơon hồi quy tổng quát và thuật toán nội suy đánh giá chất lượng nước mặt sông và các chi lưu sông Đồng Nai, tỉnh Bình Dương năm 2012 - 2018. Tạp chí Khoa học Trường Đại học Cần Thơ. 55(Số chuyên đề: Môi trường và Biến đổi khí hậu)(2): 85-95.

1 GIỚI THIỆU

Hiện nay, việc đánh giá chất lượng nước có nhiều phương pháp như đánh giá đơn chất, chỉ số chất lượng nước, chỉ số ô nhiễm, đánh giá toàn diện mờ, lý thuyết xám (Xing *et al.*, 2011). Các phương pháp này đều có ưu nhược điểm khác nhau, nhưng có điểm chung là chưa cho biết được mức độ đóng góp của thông số đến chất lượng nước và giải quyết tốt vấn đề phi tuyến (nhiều biến động theo không gian và thời gian của chất lượng nước). Một phương pháp đánh giá mới có thể kết hợp thuận lợi giữa phương pháp truyền thống và trí tuệ nhân tạo là mạng nơ ron nhân tạo (Zulin and Benlin, 2012).

Abyaneh (2014) sử dụng hồi quy tương quan và mạng nơ ron nhân tạo (ANN) để dự báo nồng độ COD và BOD từ các thông số TSS, TS, pH và nhiệt độ. Mục tiêu nghiên cứu là tìm ra cấu trúc tối ưu ANN và mô hình hồi quy tương quan để tìm ra mô hình tốt nhất cho dự báo chất lượng nước tại nhà máy xử lý nước thải Ekbatan, Tehran, Iran. Kết quả đánh giá được so sánh giữa hồi quy tương quan và mạng nơ ron perceptron đa lớp cho thấy mô hình mạng nơ ron cho kết quả đánh giá tốt hơn mô hình hồi quy tương quan. Alte (2015) sử dụng pH, Ec, Ca, Mg và Na xác định nồng độ muối natri bằng mạng nơ ron lan truyền ngược. Kết quả cũng cho thấy mạng này xác định tốt nồng độ muối natri. Bên cạnh đó, Juahir *et al.* (2004) đã phát triển mô hình ANN để ước tính chỉ số WQI lưu vực sông Langat, Malaysia thông qua các thông số DO, COD, BOD, SS, AN và pH. Phương pháp mạng nơ ron lan truyền ngược được sử dụng trong công trình này và kết quả cho thấy mô hình mạng với 4 biến đầu vào DO, BOD, SS, AN cho kết quả đánh giá tốt chất lượng nước (WQI). Như vậy, các nghiên cứu đánh giá chất lượng nước bằng mạng nơ ron nhân tạo thường sử dụng mạng truyền thẳng bao gồm thuật toán lan truyền ngược và lan truyền thẳng. Như vậy, các mô hình đánh giá mạng nơ ron chưa thể hiện rõ tính hiệu quả về mặt kinh tế. Các thông số được dự đoán là chủ yếu là DO, COD, BOD... là phổ biến. Một trong những thách thức trong đánh giá chất lượng nước hiện nay là sao để giảm chi phí và thời gian đánh giá là một thách thức đặt ra (Banejad and Olyaei, 2011).

Hệ thống sông ngòi trong tỉnh Bình Dương rất phong phú. Trong đó, sông Đồng Nai là con sông lớn nhất ở miền Đông Nam Bộ, chảy qua địa bàn tỉnh Bình Dương có giá trị lớn về cung cấp nước tưới cho nền nông nghiệp, giao thông thủy và cung cấp nước sinh hoạt cho nhân dân. Tuy nhiên, hiện nay lưu vực sông đoạn qua khu vực tỉnh Bình Dương đang có diễn biến xấu vì chịu ảnh hưởng bởi các quá trình sản xuất công nghiệp và sinh hoạt đô thị. Nghiên cứu đánh giá chất lượng nước các sông rạch

tỉnh Bình Dương đã được nhiều nghiên cứu thực hiện như Huỳnh Thị Như Quỳnh (2011), Nguyễn Phạm Huyền Linh and Bùi Tá Long (2013), Than and Ly (2014). Tuy nhiên, đánh giá chất lượng nước bằng mạng nơ ron hồi quy tổng quát chưa được thực hiện tại khu vực nghiên cứu. Do đó, nghiên cứu này sẽ trình bày mô hình đánh giá chất lượng nước chi phí thấp và tiết kiệm thời gian đánh giá nhưng vẫn đảm bảo độ tin cậy và chính xác nhất định.

2 DỮ LIỆU VÀ PHƯƠNG PHÁP

2.1 Dữ liệu



Hình 1: Bản đồ vị trí quan trắc

Chú thích các điểm quan trắc:

- ĐN1: Cách ngã ba sông ĐN – SB 1 km
- ĐN2: Cầu mới bắc qua cù lao Bạch Đằng
- ĐN3: Họng thu nước nhà máy nước Tân Hiệp
- ĐN4: Họng thu nước nhà máy nước Tân Ba
- ĐN1: Suối Cái tại Cầu Bến Sắn
- ĐN2: Suối Bung Cù tại Cầu Suối Nước
- ĐN3: Suối Ông Đông tại Cầu Tổng Bán.
- ĐN4: Suối Cái tại Cầu Bà Kiên
- ĐN5: Suối Siệp tại cống trên QL 1K
- ĐN6: Rạch Bà Hiệp tại Cầu Bà Hiệp
- ĐN7: Suối Tân Lợi gần KCN Đất Cuốc
- ĐN8: Suối Thợ Ụt tại cầu Thợ Ụt

Dữ liệu quan trắc chất lượng nước được thu thập từ năm 2012 đến năm 2018 tại Sở Tài nguyên và Môi trường tỉnh Bình Dương. Các điểm quan trắc sông và chi lưu sông Đồng Nai, tỉnh Bình Dương bao gồm: 4 điểm sông Đồng Nai (ĐN1, ĐN2, ĐN3, ĐN4) và 8 điểm ở các chi lưu (RĐN1, RĐN2, RĐN3, RĐN4, RĐN5, RĐN6, RĐN7 và RĐN8) (Hình 1). Tần suất lấy mẫu 6 đợt/năm (2 tháng/lần) từ năm 2012 đến 2015 và 12 đợt/năm (1 tháng/lần) từ 2016 – 2018. Tổng số lượng mẫu trong khoảng thời gian 7 năm từ 2012 đến 2018 là 720 mẫu. Lượng mẫu này có thể phản ánh được chất lượng nước các chi lưu và sông Đồng Nai.

Dữ liệu khí tượng bao gồm lượng mưa trung bình tháng, nhiệt độ không khí trung bình tháng, độ ẩm trung bình tháng, số giờ nắng trung bình tháng được thu thập từ Niên giám thống kê tỉnh Bình Dương từ năm 2012 – 2018.

2.2 Phương pháp nghiên cứu

2.2.1 Phương pháp chỉ số chất lượng nước Việt Nam (WQI)

Trong nghiên cứu này chỉ số chất lượng nước được tính toán theo Quyết định số 879 năm 2011 của Tổng cục Môi trường (Tổng Cục Môi Trường, 2011). Các thông số tính toán bao gồm: nhiệt độ, pH, DO, COD, BOD₅, TSS, NH₄⁺, PO₄³⁻, độ đục và Coliform. Công thức tính toán chỉ số chất lượng nước thành phần thể hiện Công thức 1, 2 và 3 như sau:

$$WQI_{SI} = \frac{q_i - q_{i+1}}{BP_{i+1} - BP_i} (BP_{i+1} - C_p) + q_{i+1} \quad (1)$$

$$WQI_{SI} = \frac{q_{i+1} - q_i}{BP_{i+1} - BP_i} (C_p - BP_i) + q_i \quad (2)$$

$$WQI = \frac{WQI_{pH}}{100} \left[\frac{1}{5} \sum_{a=1}^5 WQI_a \times \frac{1}{2} \sum_{b=1}^2 WQI_b \times WQI_c \right]^{1/3} \quad (3)$$

Trong đó:

BP_i: Nồng độ giới hạn dưới của giá trị thông số quan trắc được quy định trong Bảng 1 tương ứng với mức i

BP_{i+1}: Nồng độ giới hạn trên của giá trị thông số quan trắc được quy định trong Bảng 1 tương ứng với mức i+1

q_i: Giá trị WQI ở mức i đã cho trong bảng tương ứng với giá trị BP_i

q_{i+1}: Giá trị WQI ở mức i+1 cho trong bảng tương ứng với giá trị BP_{i+1}

C_p: Giá trị của thông số quan trắc được đưa vào tính toán

WQI_a: Giá trị WQI đã tính toán đối với 05 thông số: DO, BOD₅, COD, N-NH₄, P-PO₄

WQI_b: Giá trị WQI đã tính toán đối với 02 thông số: TSS, độ đục

WQI_c: Giá trị WQI đã tính toán đối với thông số Tổng Coliform

WQI_{pH}: Giá trị WQI đã tính toán đối với thông số pH.

Ghi chú: Nếu 5,5 < giá trị pH < 6 và nếu 20 < giá trị DO% bão hòa < 88 thì WQI_{pH} và WQI_{DO} được tính theo công thức 2.

2.2.2 Mạng nơ ron hồi quy tương quan tổng quát (Generalized regression neural network - GRNN)

Mạng nơ ron hồi quy tổng quát được phát minh bởi Specht (1991). Các nốt GRNN yêu cầu hai hàm chính để tính toán sự khác biệt giữa tất cả các cặp véc tơ mẫu đầu vào và ước tính hàm mật độ xác suất của biến đầu vào. Sự khác biệt giữa vector đầu vào được tính toán sử dụng khoảng cách Euclidean đơn giản giữa giá trị dữ liệu trong không gian thuộc tính. Trọng số khoảng cách được tính toán của mỗi điểm bằng xác suất của các điểm xây ra trong phạm vi mẫu giá trị đầu ra (Palani, S. et al., 2008).

$$E_{y/x}(X) = \int y^* f_{xy}(x, y) dy / \int f_{xy}(x, y) dy \quad (4)$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i^* \exp(-D(x, x_i))}{\sum_{i=1}^n \exp(-D(x, x_i))} \quad (5)$$

Trong đó, x là thông số chất lượng nước, y* là giá trị đầu ra thực tế của mẫu thứ i, n là tổng số mẫu trong tập dữ liệu và D(x, x_i) là hàm chuyển đổi các biến đầu vào được tính như sau:

$$D(x, x_i) = \sum_{j=1}^p (x_j - x_{ij} / \sigma_j)^2 \quad (6)$$

Trong đó, D(x, x_i) x là thông số đầu vào, x_i là mẫu thứ i, x_j là giá trị dữ liệu thứ j trong vector đầu vào, x_{ij} là giá trị dữ liệu thứ j trong mẫu thứ i, và σ_j là yếu tố làm trơn.

Các bước xây dựng mô hình đánh giá chất lượng nước bằng GRNN như sau:

Tiền xử lý số liệu

Dữ liệu chất lượng nước sau khi thu thập được kiểm tra và xử lý trị bất thường (outliers) bằng phương pháp Grub Test (Quang et al., 2019). Các giá trị bất thường được loại bỏ và tái tạo bằng phương pháp phục dựng dữ liệu khuyết (imputation missing analysis) (OECD, 2008).

Dữ liệu sau khi loại bỏ trị bất thường được chuẩn hóa để giảm bớt ảnh hưởng biến động của dữ liệu đến kết quả chất lượng nước. Vì các dữ liệu chất lượng nước có các đơn vị khác nhau và có sự biến động lớn (hiện tượng phi tuyến) do đó tác động đến kết quả. Nghiên cứu này sử dụng phương pháp chuẩn hóa tỷ lệ tuyến tính (linear scaling) được sử dụng trong nhiều công trình nghiên cứu mạng nơ ron (Ostad-Ali-Askari *et al.*, 2017, Hanrahan, 2011). Công thức chuẩn hóa theo công thức 7:

$$z_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (7)$$

Trong đó, z_{ij} là thông số đơn lẻ trong ma trận chuẩn hóa Z, x_{ij} giá trị thông số chất lượng nước thứ i mẫu chất lượng nước thứ j đo được, min và max là giá trị nhỏ nhất và lớn nhất của thông số i.

Lựa chọn biến

Trong nghiên cứu này, biến đầu vào bao gồm các thông số : pH, DO, độ đục, TSS, COD, BOD₅, NH₄⁺, PO₄³⁻, Coliform, TDS, NaCl, EC và các yếu tố khí tượng: nhiệt độ không khí trung bình (M1), lượng mưa (M2), độ ẩm trung bình (M3) và số giờ nắng trung bình (M4). Phân tích hệ số tương quan Pearson và phân tích hồi quy tương quan (Quang *et al.*, 2019) được sử dụng để lựa chọn biến tham gia vào mô hình.

Cấu trúc mạng

Các biến đầu vào là các biến được lựa chọn từ bước lựa chọn biến và biến đầu ra là chỉ số chất lượng nước WQI theo hướng dẫn của Tổng cục Môi trường Việt Nam.

Huấn luyện mạng

Để huấn luyện mạng, dữ liệu được phân chia thành các tập dữ liệu huấn luyện và tập đánh giá mô hình. Theo nhiều nghiên cứu, tỷ lệ dữ liệu huấn luyện và đánh giá mô hình dao động từ 60-80% và 20-40% tập dữ liệu (Han Yan and Wang, 2010; Heddam, 2016) nhằm tránh hiện tượng quá tải trong mô hình mạng ANN

Mô hình đánh giá chất lượng nước sông Đồng Nai bằng mạng GRNN được thiết lập theo tỷ lệ 70% dữ liệu huấn luyện mô hình (470 mẫu), 30% dữ liệu (191 mẫu) dùng để kiểm tra mô hình huấn luyện. Dữ liệu năm 2018 dùng để đánh giá kết quả thực nghiệm của mô hình. Mô hình được huấn luyện bằng phần mềm DecisionTools Suite 7.6 Industrial (2019).

2.2.3 Tiêu chí đánh giá mô hình

Kết quả mô hình đánh giá chất lượng nước bằng mạng GRNN được kiểm chứng và đánh giá thông

qua sai số bình phương trung bình gốc (root mean square error – RMSE), hệ số xác định (R²) (Csábrági *et al.*, 2017) và chỉ số Nash–Sutcliffe (E) (Han Yan and Wang, 2010). Sai số RMSE và chỉ số Nash–Sutcliffe càng gần về 0 mô hình đánh giá các có độ chính xác cao. Hệ số xác định R² càng gần 1 cho biết mô hình huấn luyện có sự liên hệ chặt chẽ với chỉ số chất lượng nước thực. Công thức tính toán RMSE và E được thể hiện như sau:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (8)$$

$$E = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (9)$$

$$R^2 = \frac{\left[\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}}) \right]^2}{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2} \quad (10)$$

Trong đó: y là giá trị WQI tính toán theo hướng dẫn của Tổng Cục Môi trường tại mẫu thứ I, \bar{y} là giá trị trung bình chỉ số WQI tính toán, \hat{y} là chỉ số chất lượng nước từ mô hình GRNN, $\bar{\hat{y}}$ là giá trị trung bình chỉ số chất lượng nước từ mô hình GRNN và n là số lượng mẫu.

2.2.4 Phương pháp nội suy không gian theo khoảng cách IDW

Để phân vùng chất lượng nước và đánh giá lan truyền ô nhiễm chất lượng nước sông Đồng Nai và các chi lưu, nhóm nghiên cứu thực hiện xây dựng bản đồ chất lượng nước và nội suy không gian bằng phương pháp IDW (phương pháp nội suy khoảng cách) (Nguyễn Thị Thu Hiền *et al.*, 2018; Nông Thị Hải Yến và Nguyễn Hải Hòa, 2017). Bản đồ nền tỉnh Bình Dương được sử dụng trong nghiên cứu được thu thập từ Sở Tài nguyên và Môi trường tỉnh Bình Dương. Phần mềm ArcGis version 10.2.2 được sử dụng để xây dựng bản đồ.

3 KẾT QUẢ

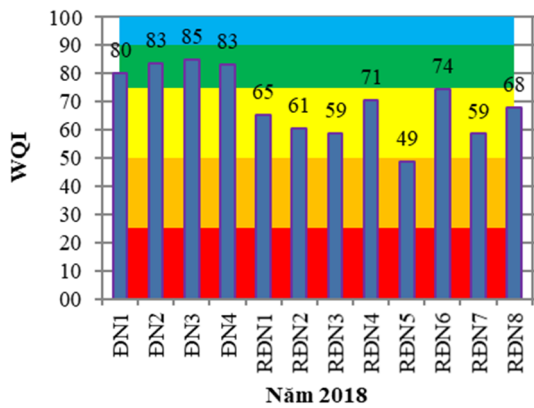
3.1 Đánh giá chất lượng nước mặt sông Đồng Nai và chi lưu bằng phương pháp chỉ số đánh giá chất lượng nước

3.1.1 Hiện trạng chất lượng nước sông Đồng Nai và các chi lưu năm 2018

Theo Hình 2, chất lượng nước tại 4 điểm sông Đồng Nai (ĐN1, ĐN2, ĐN3 và ĐN4) khá tương đồng dao động từ 80,1 – 85, mức sử dụng cho mục đích cấp nước sinh hoạt nhưng cần các biện pháp xử

lý phù hợp. Chất lượng nước tại 7 điểm chi lưu sông Đồng Nai (RĐN1, RĐN2, RĐN3, RĐN4, RĐN6, RĐN7, RĐN8) dao động từ 60,6 – 74,5, mức sử dụng cho mục đích tưới tiêu và các mục đích tương đương khác. Tuy nhiên, riêng chất lượng nước tại điểm RĐN5 còn khá thấp với giá trị là 48,0 mức sử dụng cho giao thông thủy.

Theo Hình 3, chất lượng nước thành phần sông Đồng Nai của các thông số có thay đổi theo hướng tích cực. Trong đó, thông số chất lượng nước phosphat (PO_4^{3-}), Coliform, TSS cải thiện đáng kể. Ngược lại, chất lượng nước pH, DO, BOD₅, COD

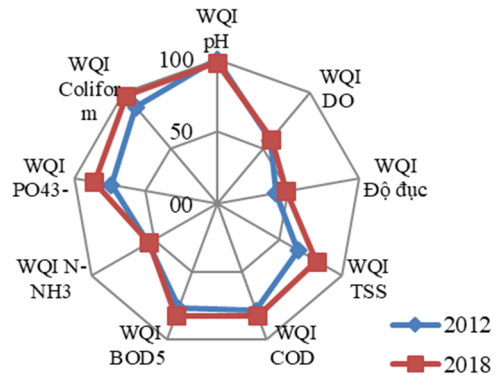


Hình 2: Trung bình chất lượng nước sông Đồng Nai và các chi lưu năm 2018

3.1.2 *Diễn biến chất lượng nước trung bình năm sông Đồng Nai*

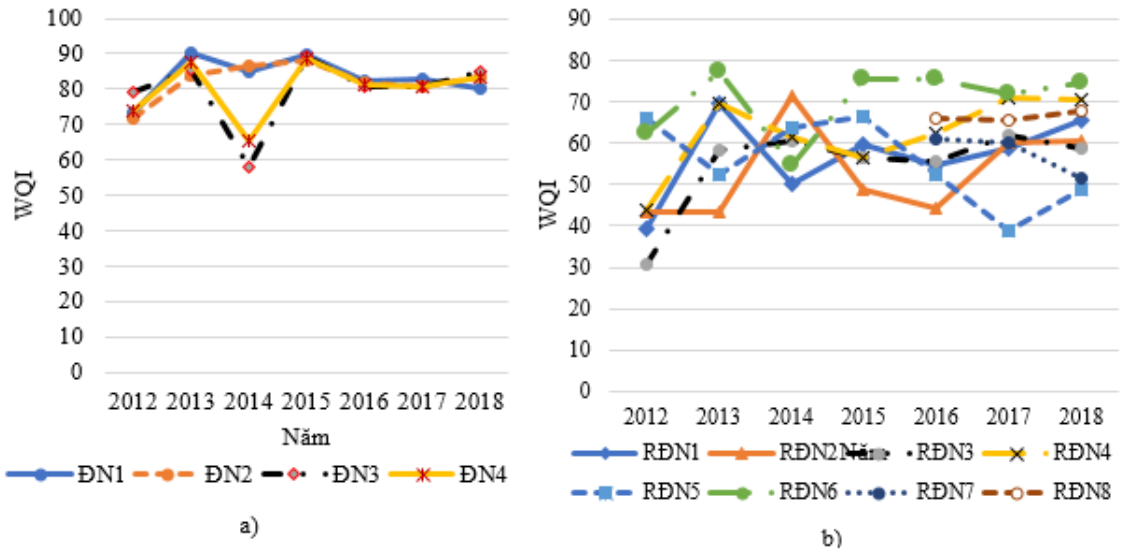
Hình 4 cho thấy chất lượng nước sông Đồng Nai diễn biến tích cực qua các năm từ 2012 đến 2018. Tại sông Đồng Nai, chỉ số WQI có sự biến động từ 2012 – 2015 dao động từ 60 – 90. Chất lượng nước trong giai đoạn này có xu hướng cải thiện. Riêng điểm ĐN3 và ĐN4 có nhiều biến động bởi các điểm này là nơi hợp lưu các nhiều kênh rạch thuộc chi lưu sông Đồng Nai có chất lượng nước thấp. Giai đoạn có sự biến động ít từ 2015 đến 2018 WQI dao động từ 80 – 88. Chất lượng nước có thể sử dụng cho mục đích sinh hoạt nhưng cần có biện pháp xử lý. Tại các

và amoni không có sự thay đổi đáng kể trong 6 năm qua. Trong các năm qua, tỉnh Bình Dương thực hiện nhiều biện pháp bảo vệ môi trường như xây dựng hệ thống thu gom nước thải sinh hoạt ở các đô thị lớn, bê tông hóa đường giao thông, di dời các xí nghiệp nằm ngoài khu công nghiệp, nghiêm chỉnh yêu cầu xử lý nước thải đối với các xí nghiệp. Do đó, TSS (phần lớn do ảnh hưởng nước mưa chảy tràn) và PO_4^{3-} (hoạt động công nghiệp), Coliform (sinh hoạt) có sự cải thiện. Các thông số chất lượng nước pH, DO, BOD₅, COD và amoni còn chịu ảnh hưởng nhiều của hoạt động đô thị chưa có nhiều thay đổi.



Hình 3: Trung bình chung chỉ số chất lượng nước thành phần sông và chi lưu Đồng Nai năm 2012 và 2018

chi lưu, chất lượng nước có sự biến động liên tục qua các năm theo xu hướng tích cực. Khác với chất lượng nước tại sông chính, các chi lưu sông Đồng Nai có mức độ ô nhiễm cao hơn biên độ dao động chất lượng nước từ 31 – 77. Hầu hết các điểm có chất lượng nước cải thiện trong giai đoạn 2012 - 2018, riêng RĐN5 và RĐN7 có xu hướng suy giảm chất lượng nước. Nguyên nhân dẫn đến chất lượng nước tại sông Đồng Nai và các chi lưu có xu hướng cải thiện là do hiệu quả tích cực thực việc triển khai quyết liệt công tác bảo vệ môi trường nước mặt của tỉnh Bình Dương và các tỉnh lân cận như Đồng Nai, Bình Phước,...



Hình 4: Diễn biến chất lượng nước trung bình năm tại các điểm quan trắc: a) sông Đồng Nai; b) các chi lưu sông Đồng Nai

3.2 Đánh giá chất lượng nước sông Đồng Nai và các chi lưu bằng GRNN

3.2.1 Lựa chọn biến trong mô hình GRNN

Kết quả phân tích tương quan Pearson cho biết các biến đều có sự liên hệ với chỉ số chất lượng

Bảng 1: Tóm tắt lựa chọn biến từ phân tích hồi quy tương quan

Mô hình	Tham số trong mô hình	Biến vào/ra	Hiện trạng	RMSE	R ²	R ² _{adj}
1	TSS	TSS	Vào	0,170	0,408	0,407
2	pH / TSS	pH	Vào	0,114	0,735	0,734
3	pH / TSS / EC	EC	Vào	0,104	0,779	0,778
4	pH / TSS / Độ đục / EC	Độ đục	Vào	0,097	0,808	0,807
5	M1 / pH / TSS / Độ đục / EC	M1	Vào	0,095	0,817	0,816
6	M1 / pH / COD / TSS / Độ đục / EC	COD	Vào	0,093	0,824	0,823
7	M1 / M3 / pH / COD / TSS / Độ đục / EC	M3	Vào	0,093	0,826	0,825
8	M1 / M4 / M3 / pH / COD / TSS / Độ đục / EC	M4	Vào	0,090	0,836	0,834
7	M4 / M3 / pH / COD / TSS / Độ đục / EC	M1	Ra	0,090	0,836	0,834
8	M4 / M2 / M3 / pH / COD / TSS / Độ đục / EC	M2	Vào	0,090	0,837	0,836
9	M4 / M2 / M3 / pH / COD / TSS / Amoni / Độ đục / EC	Amoni	Vào	0,090	0,838	0,836
10	M4 / M2 / M3 / pH / DO / COD / TSS / Amoni / Độ đục / EC	DO	Vào	0,089	0,839	0,838

Bảng 1 cho thấy các biến pH, TSS, Độ đục và EC là các biến có mức độ liên hệ cao với chỉ số chất lượng nước được thể hiện trong mô hình từ 1 – 4. Thông số khí tượng nhiệt độ trung bình tháng (M1) và độ ẩm trung bình tháng (M3) là hai trong bốn thông số khí tượng có sự liên hệ cao nhất đối với chất lượng nước tại Sông Đồng Nai đoạn chảy qua tỉnh Bình Dương thể hiện trong mô hình hồi quy 5 và 7. So sánh giữa các mô hình, mô hình 5 gồm các biến M1, pH, TSS, Độ đục và EC có số lượng biến ít nhưng đảm bảo sai số RMSE thấp 0,095, hệ số xác định R² và R²_{adj} hiệu chỉnh cao lần lượt 0.817 và 0.816. Độ ẩm không khí là một trong biến khí tượng

nước. Sau đó, các biến được sàng lọc qua phân tích hồi quy tương quan đa biến bằng kỹ thuật Stepwise, kết quả thu được 10 mô hình hồi quy thể hiện trong Bảng 1.

được khá nhiều nghiên cứu sử dụng khi đánh giá chất lượng nước như Lee and Resdi (2016). Mục tiêu trong nghiên cứu này là xây dựng mô hình đánh giá chất lượng nước nhanh và chi phí thấp nhất nhưng vẫn đảm bảo độ tin cậy và chính xác. Do đó, các biến trong mô hình 5 biến và biến độ ẩm trung bình tháng được lựa chọn để tiến hành xây dựng mô hình đánh giá chất lượng nước bằng mạng GRNN.

Mô hình hồi quy tương quan 6 biến thu được phương trình sau (công thức 11):

$$WQI = -0,05427 \times M1 + 0,01352 \times M3 + 0,37502 \times pH + 0,38799 \times TSS + 0,22813 \times \text{Độ đục} - 0,34032 \times EC \quad (11)$$

$R^2 = 0,842$, $RMSE = 0,086$, $P = 0$, $Dustin\ waston = 1,535$. Kết quả cho thấy mô hình hồi quy

trương quan đa biến đánh giá chất lượng nước rất có ý nghĩa thống kê. Kiểm định các biến tham gia vào mô hình cho thấy 5/6 biến là rất có ý nghĩa thống kê ($p < 0,01$). Biến M3 (độ âm trung bình tháng) có $p = 0,33 > 0,05$, điều này có nghĩa M3 không có ý nghĩa thống kê trong mô hình hồi quy tương quan đánh giá chất lượng nước (Bảng 2).

Bảng 2: Kiểm định thông số trong mô hình hồi quy tương quan

Thông số	Giá trị	Sai số chuẩn	t	Pr > t	Chặn dưới (95%)	Chặn trên (95%)
Intercept	0,0000					
M1	-0,0543	0,0196	-2,7633	0,0059	-0,0929	-0,0157
M3	0,0135	0,0139	0,9752	0,3300	-0,0137	0,0408
Ph	0,3750	0,0179	20,9175	< 0,0001	0,3398	0,4102
TSS	0,3880	0,0183	21,1963	< 0,0001	0,3520	0,4240
Độ đục	0,2281	0,0199	11,4354	< 0,0001	0,1889	0,2673
EC	-0,3403	0,0314	-10,8239	< 0,0001	-0,4021	-0,2785

3.2.2 Mô hình chất lượng nước bằng mạng GRNN

Cấu trúc mạng GRNN H_6IO_1 thu được kết quả mô hình như sau:

Mô hình đánh giá chất lượng nước GRNN thu được kết quả huấn luyện mạng rất tốt $R = 0,974$ (giải thích 97,4% biến trong mô hình đánh giá chất lượng

nước), $R^2 = 0,948$, $RMSE = 0,052$ và chỉ số Nash–Sutcliffe = 0,945 (gần 1) và kết quả kiểm tra mô hình huấn luyện thu được $R = 0,957$, $R^2 = 0,916$, $RMSE = 0,061$ và chỉ số Nash–Sutcliffe = 0,913. Kết quả huấn luyện và đánh giá mô hình cho thấy mô hình đánh giá chất lượng nước bằng GRNN có độ chính xác cao và có độ liên chặt chẽ với kết quả đánh giá chất lượng nước theo quy định của Việt Nam.

Bảng 3: Thống kê mô hình đánh giá chất lượng nước bằng mô hình GRNN và hồi quy tương quan

Phương pháp	Tập dữ liệu	R	R ²	RMSE	E
GRNN (6 biến)	Huấn luyện	0,974	0,948	0,052	0,945
	Kiểm tra	0,957	0,916	0,061	0,913
	Ước tính	0,968	0,938	0,055	0,937
Hồi quy tương quan (6 biến)	Huấn luyện	0,918	0,842	0,086	0,840
	Kiểm tra	0,895	0,802	0,109	0,792
	Ước tính	0,872	0,760	0,107	0,750

* Các biến tham gia trong mô hình này chính là các biến được sử dụng tính toán chỉ số WQI

Mô hình hồi quy tương quan thu được kết quả khá tốt (mô hình huấn luyện $R = 0,918$, $R^2 = 0,842$, $RMSE = 0,086$ và chỉ số Nash–Sutcliffe = 0,840), tập đánh giá kết quả mô hình $R = 0,895$, $R^2 = 0,802$, $RMSE = 0,109$ và chỉ số Nash–Sutcliffe = 0,840. Qua Bảng 3, mô hình GRNN có kết quả hoạt động tốt hơn ở tất cả các chỉ tiêu khi so sánh với mô hình hồi quy, đặc biệt sai số RMSE thấp hơn đáng kể (huấn luyện 0,052 < 0,086, kiểm tra 0,061 < 0,109). Kết quả nghiên cứu chỉ ra rằng mạng nơ ron nhân tạo có hoạt động tốt rất nhiều so với mô hình hồi quy, điều này cũng được chứng minh trong nhiều loại mạng khác trong các nghiên cứu trước đây như Banejad và Olyaie (2011), Nasir et al. (2012).

3.2.3 Kết quả đánh giá chất lượng nước bằng mô hình GRNN

Sử dụng mô hình GRNN đã xây dựng, chỉ số chất lượng nước sông Đồng Nai và các chỉ lưu năm

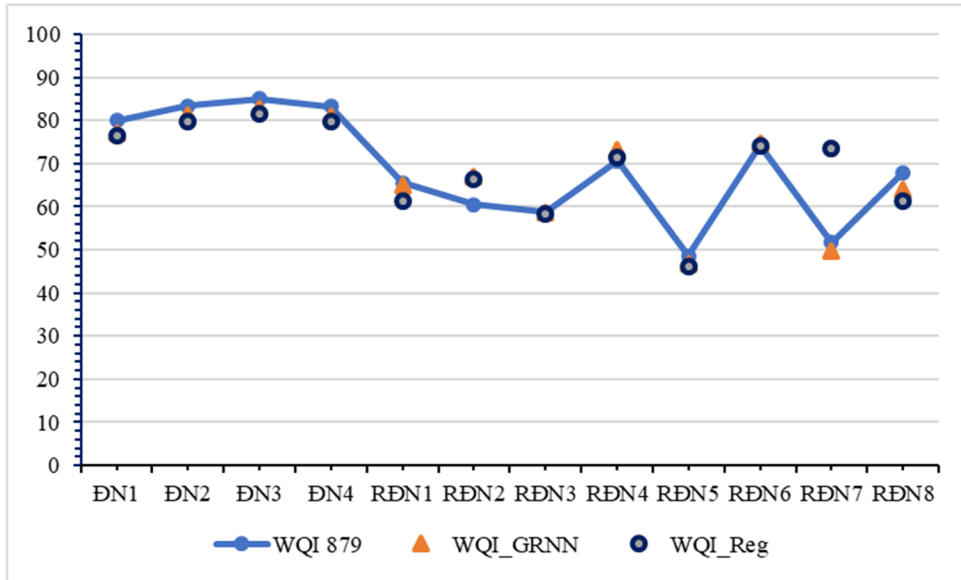
2018 khu vực tỉnh Bình Dương được tính toán. Kết quả tính toán mô hình thu được kết quả rất tốt: hệ số xác định cao ($R^2 = 0,938$, $RMSE = 0,055$, $E = 0,937$). So sánh kết quả tính toán WQI giữa mô hình GRNN và hồi quy tương quan chứng minh rằng mô hình GRNN hoạt động tốt trong điều kiện dữ liệu khác nhau. Kết quả tính toán WQI từ mô hình hồi quy cho hệ số xác định có độ tương quan khá $R^2 = 0,760$ và sai số $RMSE = 0,107$, $E = 0,750$ thấp hơn đáng kể so với mô hình GRNN. Hình 5 cho thấy chỉ số chất lượng nước trung bình tại các điểm quan trắc của mô hình mạng GRNN (WQI_{GRNN}) có sự liên hệ rất chặt chẽ với chỉ số WQI 879. Bên cạnh đó, kết quả đánh giá chất lượng nước sông Đồng Nai năm 2018 bằng mạng GRNN có chất lượng nước còn tốt, dao động từ 80,2 – 82,9, chênh lệch giữa 4 vị trí không nhiều. Kết quả đánh giá chất lượng nước các chỉ lưu sông Đồng Nai năm 2018 bằng mạng GRNN

có chất lượng nước ở mức trung bình dao động từ 52,6 – 75,8.

Kiểm định sai số kết quả đánh giá chất lượng sông Đồng Nai và các chi lưu so với phương pháp đánh giá chất lượng nước theo chỉ số chất lượng nước của Tổng cục Môi trường cho giá trị RMSE =

7,3 rất thấp. Do đó, kết quả đánh giá bằng phương pháp GRNN cho chất lượng nước là đáng tin cậy.

Kết quả phân tích hồi quy tương quan giữa hai kết quả đánh giá cũng cho thấy $R^2 = 0,873$ và R^2 hiệu chỉnh = 0,873, $P = 0$ cho thấy phương pháp GRNN cho chất lượng nước là phù hợp với cách đánh giá WQI.

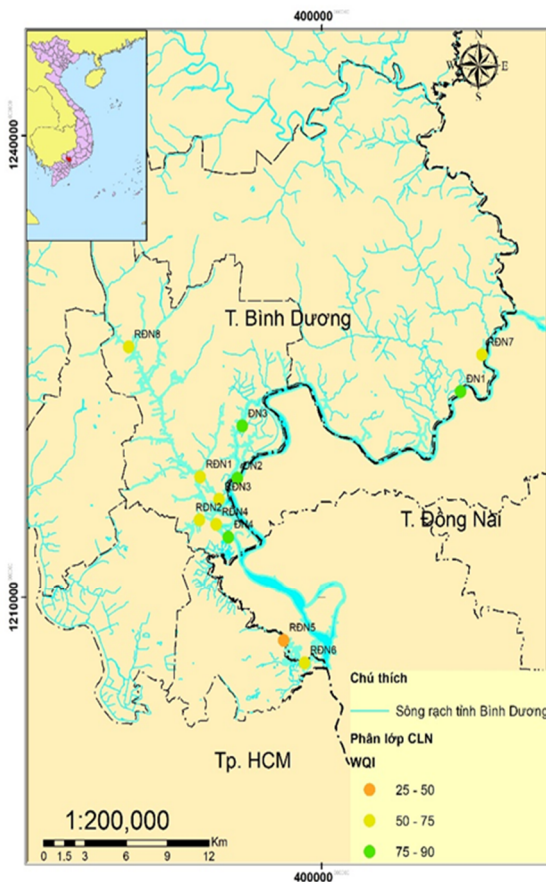


Hình 5: Kết quả ước tính WQI sông Đồng Nai và các chi lưu tỉnh Bình Dương năm 2018 bằng mô hình hồi quy tương quan và GRNN

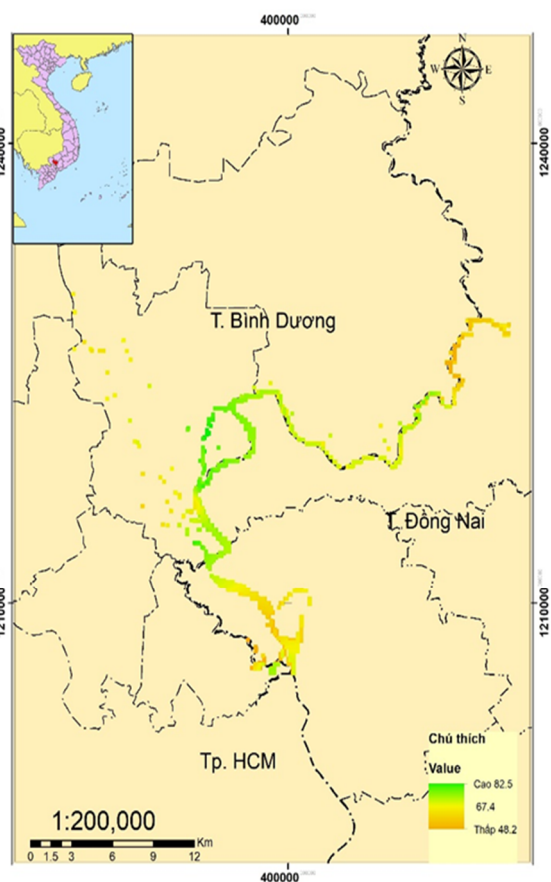
3.3 Phân vùng chất lượng nước sông Đồng Nai và các chi lưu

Sử dụng GIS để lập bản đồ chất lượng nước sông Đồng Nai và các chi lưu, kết quả bản đồ chất lượng nước (Hình 6) cho thấy chất lượng nước sông Đồng Nai và các chi lưu dao động từ mức màu da cam đến màu xanh nước biển. Trong đó, có 4 điểm quan trắc tại sông chính Đồng Nai đều có chất lượng nước tốt có thể sử dụng cho sinh hoạt thông qua biện pháp xử lý. Bản đồ nội suy không gian chất lượng nước tại Hình 7, chất lượng nước thấp về phía hạ nguồn sông

Đồng Nai nơi tiếp nhận nguồn nước thải từ hoạt động đô thị và công nghiệp của các tỉnh thuộc lưu vực. Chất lượng nước trên sông Đồng Nai hiện nay còn khá tốt trong phạm vi khu vực nghiên cứu. Bên cạnh đó, có thể thấy chất lượng nước tại các khu vực giao giữa chi lưu sông và sông Đồng Nai chất lượng nước thấp hơn ở khu vực khác của nhánh sông chính. Điều này cho thấy các kênh rạch hiện nay là nguồn dẫn đến suy giảm chất lượng nước sông Đồng Nai, đặc biệt là các chi lưu có chất lượng nước thấp như RĐN5 và RĐN7.



Hình 6: Bản đồ chất lượng nước sông Đồng Nai và các chi lưu thuộc tỉnh Bình Dương bằng mô hình GRNN



Hình 7: Bản đồ lan truyền chất lượng nước sông Đồng Nai và các chi lưu thuộc tỉnh Bình Dương bằng mô hình GRNN

3.4 Lợi ích kinh tế từ mô hình đánh giá chất lượng GRNN

Mô hình đánh giá chất lượng nước bằng mạng GRNN với số lượng các biến tham gia đánh giá bao gồm 4 thông số quan trắc môi trường: pH, TSS, độ đục, EC và 2 thông số khí tượng độ ẩm và nhiệt độ không khí trung bình. Trong đó, dữ liệu khí tượng được kế thừa từ dữ liệu khí tượng thủy văn. Do đó,

chi phí cho mô hình đánh giá GRNN được tính toán cho 4 thông số pH, TSS, độ đục, EC. Căn cứ Quyết định số 22/2018/QĐ-UBND ngày 20 tháng 8 năm 2018 của Ủy ban nhân dân tỉnh Bình Dương ban hành quy định về đơn giá hoạt động quan trắc và phân tích môi trường trên địa bàn tỉnh Bình Dương, lượng chi phí đánh giá chất lượng nước từ mô hình GRNN và WQI được tính toán theo Quyết định số 879 của Tổng cục Môi trường như sau:

Bảng 4: Khái toán chi phí đánh giá chất lượng nước cho một mẫu nước

Mô hình	Thông số đầu vào	Chi phí đánh giá	Thành tiền
GRNN	pH, TSS, độ đục, EC nhiệt độ không khí trung bình tháng	Chi phí = pH + TSS + độ đục + EC	Chi phí = 56.400 + 188.300 + 69.600 + 67.200 = 381.500 VNĐ
WQI 879	Nhiệt độ nước, pH, DO, COD, BOD ₅ , TSS, NH ₄ ⁺ , PO ₄ ³⁻ , Độ đục, Coliform.	Chi phí = Nhiệt độ nước + pH + DO + COD + BOD ₅ + TSS + NH ₄ ⁺ + PO ₄ ³⁻ + Độ đục + Coliform.	Chi phí = 56.400 + 56.400 + 82.800 + 235.400 + 184.700 + 188.300 + 198.900 + 210.100 + 69.600 + 470.500 = 1.753.100 VNĐ

Theo kết quả tính toán chi phí đánh giá chất lượng nước Bảng 4, chi phí tính toán chất lượng

nước theo mô hình GRNN là 381.500 VNĐ/mẫu và mô hình tính toán của Tổng cục Môi trường là

1.753.100 VNĐ/mẫu. Điều này cho thấy mô hình tính toán chất lượng nước bằng mạng GRNN tiết kiệm được 1.371.600 VNĐ/mẫu. Mô hình GRNN sẽ tiết kiệm chi phí đáng kể cho kinh phí đánh giá chất lượng nước. Bên cạnh đó, mô hình đánh giá chất lượng nước này sẽ tiết kiệm được thời gian đánh giá vì số lượng các biến phân tích trong phòng thí nghiệm chỉ có thông số TSS. Các thông số pH, độ đục và EC được đo đạc nhanh ngoài hiện trường.

4 KẾT LUẬN

Đánh giá chất lượng nước bằng mạng nơ ron hồi quy tổng quát cho kết quả tốt và có độ chính xác cao. Kết quả nghiên cứu ghi nhận chất lượng nước sông Đồng Nai năm 2018 cải thiện 11/12 điểm so với năm 2012. Cấu trúc mô hình GRNN I_6HO_1 thu được hoạt động rất tốt với $R^2 = 0,948$, $RMSE = 0,052$ và chỉ số Nash–Sutcliffe = 0,945 trong mô hình huấn luyện và $R^2 = 0,916$, $RMSE = 0,061$ và chỉ số Nash–Sutcliffe = 0,913 kiểm tra mô hình. Kết quả ước tính WQI năm 2018 từ mô hình cũng thu được kết quả rất tốt, giá trị tính toán từ mô hình GRNN có sự liên hệ chặt với chỉ số WQI theo Quyết định 879 ($R^2 = 0,938$, $RMSE = 0,055$).

Mô hình đánh giá chất lượng nước GRNN với các biến tham gia đánh giá hầu hết có thể thu thập nhanh ngoài hiện trường và dữ liệu sẵn có. Từ đó, chi phí đánh giá chất lượng nước chỉ khoảng 381.500 VNĐ/mẫu, tiết kiệm 1.371.600 VNĐ/mẫu so với phương pháp đánh giá hiện nay.

Kết quả nghiên cứu cho thấy mô hình mạng nơ ron hồi quy tổng quát cho kết quả đánh giá tốt về chất lượng nước. Mạng nơ ron GRNN cần được áp dụng vào lĩnh vực quản lý và quan trắc môi trường như là công cụ mạnh mẽ hỗ trợ quản lý chất lượng nước để phân nào giảm được thời gian đánh giá và tiết kiệm ngân sách.

LỜI CẢM ƠN

Nhóm nghiên cứu xin trân trọng cảm ơn Trường Đại học Thủ Dầu Một đã hỗ trợ kinh phí thực hiện nghiên cứu này và các phản biện đã góp ý chỉnh sửa bản thảo để bài báo được hoàn thiện hơn.

TÀI LIỆU THAM KHẢO

Abyaneh, H.Z., 2014. Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. *Journal of Environmental Health Science & Engineering*. 12(1): 1-8.

Alte, P. D., 2015. Water quality prediction by using ANN. *International Journal of Advance Foundation And Research in Science & Engineering*. 1: 178-285.

Banejad, H. and Olyaie, E., 2011. Application of an artificial neural network model to rivers water

quality indexes prediction – a case study. *Journal of American Science*. 7(1): 60-65.

Csábrági, A., Molnár, S., Tanos, P. and Kovács, J., 2017. Application of artificial neural networks to the forecasting of dissolved oxygen content in the Hungarian section of the river Danube. *Ecological Engineering*. 100: 63–72.

Yan, H., Zou, Z. and Wang, H., 2010. Adaptive neuro fuzzy inference system for classification of water quality status. *Journal of Environmental Sciences*. 22(10): 1891–1896.

Hanrahan, G., 2011. Model design and selection considerations. *In: Lochmüller, C.H. (Eds.). Artificial neural networks in biological and environmental analysis*. CRC Press. New York, pp. 40.

Heddam, S., 2016. Secchi disk depth estimation from water quality parameters: artificial neural network versus multiple linear regression models. *Environmental Process*. DOI 10.1007/s40710-016-0144-4.

Huỳnh Thị Như Quỳnh. 2011. Đánh giá chất lượng môi trường nước bằng chỉ số tổng hợp (WQI) và đề xuất giải pháp bảo vệ môi trường nước mặt trên địa bàn tỉnh Bình Dương. Luận văn cao học. Viện Môi trường và Tài Nguyên. Thành phố Hồ Chí Minh.

Lee, W. and Resdi, T.A.T., 2016. Simultaneous hydrological prediction at multiple gauging stations using the NARX network for Kemaman catchment, Terengganu, Malaysia. *Hydrological Sciences Journal*. DOI: 10.1080/02626667.2016.1174333.

Nasir, M.F.M., Zali, M.A., Juahir, H., Hussain, H., Zain, S.M. and Ramli, N., 2012. Application of receptor models on water quality data in source apportionment in Kuantan River Basin. *Iranian journal of environmental health science & engineering*. 9(1): 18.

Nguyễn Phạm Huyền Linh và Bùi Tá Long, 2013. Xây dựng bản đồ phân vùng chất lượng nước mặt trường hợp tỉnh Bình Dương. Kỷ yếu hội nghị khoa học và công nghệ lần thứ 13, ngày 01/11/2013, Thành phố Hồ Chí Minh.

Nguyễn Thị Thu Hiền, Phạm Hải Nam, Nguyễn Hải Hòa và Nguyễn Thị Khanh, 2018. Ứng dụng GIS và thuật toán nội suy không gian xây dựng bản đồ chất lượng nước suối Nậm La chảy qua Thành Phố Sơn La. *Tạp chí Khoa học và Công nghệ-Đại học Thái Nguyên*. 189(3): 39-43.

Nông Thị Hải Yến và Nguyễn Hải Hòa, 2017. Ứng dụng gis và thuật toán nội suy không gian xây dựng bản đồ chất lượng nước mặt khu vực khai thác khoáng sản tại huyện Hoành Bồ, Tỉnh Quảng Ninh. *Tạp chí Khoa học và Công nghệ-Đại học Thái Nguyên*. 169: 75-80.

Palani, S., Liang, S.Y., and Tkalic, P., 2008. An ANN application for water quality forecasting. *Marine Pollution Bulletin*. 56: 1586–1597.

- OECD (Organization for Economic Co-operation and Development), 2008. Joint Research Centre-European Commission. Handbook on constructing composite indicators: methodology and user guide. OECD publishing, France, pp. 89 – 91.
- Ostad-Ali-Askari, K., Shayannejad, M. and Ghorbanizadeh-Kharazi, H., 2017. Artificial neural network for modeling nitrate pollution of groundwater in Marginal Area of Zayandeh-rood River, Isfahan, Iran. *KSCE Journal of Civil Engineering*, 21(1): 134-140.
- Quang, N. M., Mau, T. X., Nhung, N. T. A., An, T. N. M. and Van Tat, P., 2019. Novel QSPR modeling of stability constants of metal-thiosemicarbazone complexes by hybrid multivariate technique: GA-MLR, GA-SVR and GA-ANN. *Journal of Molecular Structure*. 1195: 95-109.
- Than, N. H. and Ly, C. Đ., 2014. Assessment and prediction of water quality in Binh Duong province. *Tạp chí Khoa học và Công nghệ*. 52(5C): 19-26.
- Tổng Cục Môi Trường, 2011. Quyết định số 879/QĐ-TCMT, ngày 01/07/2011 về việc ban hành số tay hướng dẫn tính toán chỉ số chất lượng nước, ngày truy cập 18/7/2019. Địa chỉ: http://tnmtquangnam.gov.vn/index.php?option=com_content&view=article&id=2371:quyt-nh-879q-tcmt&catid=102:tai-nguyen-nc&Itemid=517.
- Xing, Z., Fu, Q. and Liu, D., 2011. Water quality evaluation by the fuzzy comprehensive evaluation based on EW method. Eighth International Conference on Fuzzy Systems and Knowledge Discovery, 26-28 July 2011, Shanghai, China. IEEE, 476-479.
- Zulin, H. and Benlin, D., 2012. Water quality health assessment of offshore area in North Branch of the Yangtze River Estuary. 2010 4th International Conference on Bioinformatics and Biomedical Engineering, 18-20 June 2010, Chengdu, China. IEEE, 1-3.