

NGHIÊN CỨU LỰA CHỌN PHƯƠNG PHÁP NỘI SUY XÁC ĐỊNH HÀM LƯỢNG MỘT SỐ THÔNG SỐ CHẤT LƯỢNG NƯỚC KHU VỰC SÔNG HỒNG VÀ SÔNG ĐUỐNG

ĐINH THỊ THU HIỀN⁽¹⁾, ĐINH LAM THẮNG⁽²⁾
NGUYỄN THỊ THU NGA⁽³⁾, ĐÀO KHÁNH HOÀI⁽³⁾

⁽¹⁾Trường Đại học Điện lực

⁽²⁾Viện Khoa học Công nghệ Cơ khí, Tự động hóa và Môi trường

⁽³⁾Viện Kỹ thuật Công trình đặc biệt, Học viện Kỹ thuật quân sự, Hà Nội

Tóm tắt:

Các kỹ thuật nội suy đã được sử dụng hiệu quả trong các nghiên cứu cứu về tài nguyên, môi trường trên cơ sở các số liệu quan trắc. Bài báo này trình bày kết quả xác định hàm lượng một số thông số chất lượng nước mặt khu vực sông Hồng và sông Đuống, bao gồm COD, BOD₅ và NH₄⁺ bằng phương pháp nội suy. 56 điểm lấy mẫu chất lượng nước được sử dụng để nội suy hàm lượng các thông số chất lượng nước cho toàn bộ khu vực nghiên cứu, trong đó 43 mẫu được sử dụng để huấn luyện mô hình và 13 mẫu để đánh giá độ chính xác. 04 kỹ thuật nội suy thông dụng, bao gồm Kriging, Inverse Distance Weighted (IDW), Neighbor Natural và Spline được thử nghiệm để lựa chọn phương pháp có độ chính xác cao nhất. Kết quả nhận được cho thấy, thuật toán IDW đạt độ chính xác cao nhất, thể hiện qua so sánh sai số trung phương (RMSE) đối với cả 03 thông số COD, BOD₅ và NH₄⁺. Từ kết quả đạt được, trong nghiên cứu đã xây dựng bản đồ phân bố hàm lượng các thông số COD, BOD₅ và NH₄⁺ khu vực nghiên cứu. Kết quả nhận được trong nghiên cứu có thể sử dụng phục vụ công tác giám sát và đánh giá chất lượng nước mặt ở khu vực sông Hồng và sông Đuống, đoạn chảy qua thành phố Hà Nội.

Từ khóa: ô nhiễm nước, nội suy, các thông số chất lượng nước, sông Hồng, sông Đuống

1. Mở đầu

Ô nhiễm nước mặt là một vấn đề môi trường mang tính toàn cầu, nhất là với các quốc gia có sự phát triển kinh tế - xã hội nhanh chóng như Việt Nam. Áp lực gia tăng dân số và các hoạt động công nghiệp dẫn đến chất lượng nước ở nhiều lưu vực sông, nhất là các đoạn chảy qua các thành phố lớn bị ô nhiễm nghiêm trọng. Nghiên cứu xác định hàm lượng các thông số chất lượng nước mặt là

một vấn đề có tính cấp bách hiện nay, cung cấp thông tin kịp thời giúp các nhà quản lý trong giám sát và ứng phó với ô nhiễm nước.

Nhiều nghiên cứu cho thấy, một số thông số chất lượng nước, đặc biệt là chất lơ lửng hòa tan (TSS), độ đục (turbidity), chất diệp lục (Chl-a) có mối quan hệ chặt chẽ với giá trị phản xạ phổ mặt nước (Nguyễn Thị Thu Hà và nnk, 2016; Đinh Lam Thắng và nnk, 2022; Trinh Le Hung và nnk, 2018; Trinh Le Hung

và nnk, 2020). Từ mối quan hệ này, dữ liệu viễn thám quang học được sử dụng hiệu quả trong ước lượng hàm lượng các thông số chất lượng nước này. Tuy nhiên, ảnh quang học như Landsat, Sentinel 2 không khả thi khi xác định hàm lượng các thông số như COD, BOD₅ và NH₄⁺ do hàm lượng các thông số này không có mối quan hệ chặt chẽ với phản xạ phổ mặt nước. Với các thông số này, phương pháp nội suy là một tiếp cận phù hợp để ước lượng hàm lượng cho toàn bộ lưu vực thông qua bộ dữ liệu các điểm quan trắc chất lượng nước. Có thể kể đến các nghiên cứu (Gnauck, 2004; Mielniczuk và nnk; Boumpoulis và nnk, 2023) đã sử dụng nhiều phương pháp nội suy khác nhau để xác định hàm lượng các thông số chất lượng nước ở nhiều khu vực trên thế giới. Kết quả nhận được trong các nghiên cứu này cho thấy, không có phương pháp nội suy chung cho các khu vực có đặc điểm chất lượng nước mặt khác nhau. Do vậy, với mỗi khu vực cụ thể, cần phân tích, thử nghiệm với nhiều phương pháp nội suy khác nhau để lựa chọn phương pháp có độ chính xác cao nhất.

Bài báo này trình bày kết quả xây dựng bản đồ phân bố hàm lượng các thông số chất lượng nước, bao gồm COD, BOD₅ và NH₄⁺ từ dữ liệu các điểm lấy mẫu chất lượng nước bằng phương pháp nội suy. Khu vực thử nghiệm được lựa chọn là sông Hồng và sông Đuống đoạn chảy qua thành phố Hà Nội. 56 điểm lấy mẫu chất lượng nước được chia thành 2 bộ dữ liệu: bộ dữ liệu dùng để xây dựng mô hình nội suy (43 điểm) và bộ dữ liệu để kiểm tra độ chính xác kết quả nội suy (13 điểm). 04 kỹ thuật nội suy thông dụng, bao gồm Kriging, Inverse Distance Weighted (IDW), Neighbor Natural và Spline được thử nghiệm để lựa chọn phương pháp có độ chính xác cao nhất nhằm xây dựng bản đồ phân bố

hàm lượng các thông số COD, BOD₅ và NH₄⁺ tkhu vực nghiên cứu.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Dữ liệu và khu vực nghiên cứu

a) Khu vực nghiên cứu

Khu vực nghiên cứu được lựa chọn trong bài báo là lưu vực sông Hồng và sông Đuống đoạn chảy qua thành phố Hà Nội. Đây là hệ thống sông lớn ở Hà Nội cũng như cả nước, có vai trò hết sức quan trọng trong đời sống sinh hoạt và hoạt động sản xuất. Đây cũng là nguồn cung cấp nước chính cho các nhà máy nước sông Hồng, nhà máy nước Đông Anh và nhà máy nước sông Đuống. Những năm gần đây, do ảnh hưởng của hoạt động sản xuất và sinh hoạt của người dân, ô nhiễm nước khu vực sông Hồng và sông Đuống có những diễn biến phức tạp, ảnh hưởng đến đời sống người dân và các hoạt động sản xuất nông nghiệp.

b) Dữ liệu sử dụng

Để xác định hàm lượng một số thông số chất lượng nước mặt, 56 mẫu nước được thu thập tại khu vực sông Hồng và sông Đuống đoạn chảy qua thành phố Hà Nội, trong đó 43 điểm được sử dụng để xây dựng hàm hồi quy và 13 điểm dùng để đánh giá độ chính xác kết quả xây dựng mô hình. Các mẫu nước này được thu thập trong đợt khảo sát thực địa ngày 21/12/2021 của nhóm nghiên cứu.

Vị trí phân bố các điểm lấy mẫu nước khu vực nghiên cứu được trình bày trong hình 1. Vị trí các điểm lấy mẫu phân bố đều trên khu vực nghiên cứu nhằm đảm bảo thể hiện được đặc trưng phân bố hàm lượng các thông số chất lượng nước. Các mẫu nước được lấy vào chai nhựa màu tối, ướp lạnh và đưa về phân tích trong phòng thí nghiệm trong ngày. Quá trình lấy mẫu, bảo quản và phân tích được thực hiện theo quy định trong Quy chuẩn quốc

gia về chất lượng nước mặt QCVN 08-MT:2015/BTNMT.



Hình 1: Sơ đồ vị trí các điểm lấy mẫu chất lượng nước

2.2. Phương pháp nghiên cứu

Nội suy là phương pháp ước tính giá trị của các điểm dữ liệu chưa biết trong phạm vi của một tập hợp rời rạc chứa một số điểm dữ liệu đã biết. Đây là phương pháp được sử dụng phổ biến trong nghiên cứu ô nhiễm môi trường khi cho phép ước lượng hàm lượng các thông số chất lượng nước, không khí... từ tập dữ liệu quan trắc.

Sơ đồ phương pháp nội suy trong xác định hàm lượng các thông số COD, BOD₅ và NH₄⁺ được thể hiện trên hình 2. Trong phương pháp này, số liệu lấy mẫu chất lượng nước được chia thành 2 tập dữ liệu: tập dữ liệu huấn luyện mô hình nội suy và tập dữ liệu kiểm tra. 04 kỹ thuật nội suy thông dụng bao gồm Inverse Distance Weighted (IDW), Kriging, Neighbor Natural Neighbor và Spline được thử nghiệm để ước lượng hàm lượng các thông số COD, BOD₅ và NH₄⁺ trong nước mặt khu vực sông Hồng và sông Đuống, đoạn chảy qua thành phố Hà Nội.

Nội suy IDW là phương pháp xác định giá trị của các điểm chưa biết bằng cách tính trung bình trọng số khoảng cách các giá trị đã biết trong

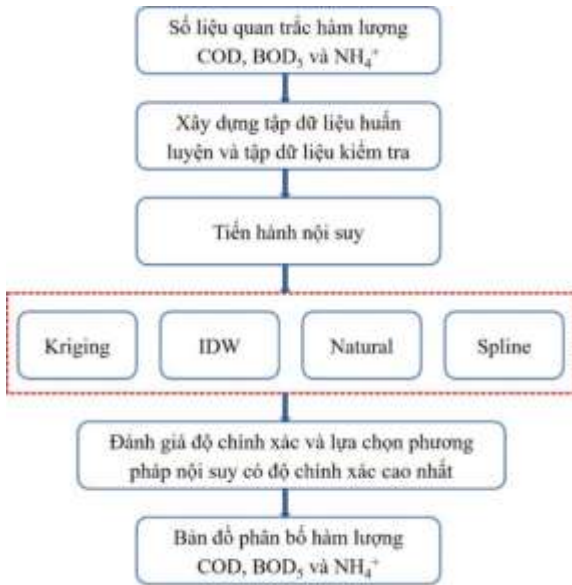
vùng lân cận của mỗi pixel. Trong phương pháp này, các điểm nằm càng xa điểm cần tính, trọng số ảnh hưởng sẽ càng thấp. IDW có hiệu quả trong trường hợp tập dữ liệu dày đặc và phân bố đều trên khu vực nghiên cứu.

Kriging là phương pháp nội suy có nhiều ưu điểm so với các phương pháp khác dựa vào cách tính toán thống kê của bề mặt. Kriging sử dụng khái niệm về sự khác biệt có tính chất vùng với sự khác nhau từ nơi này sang nơi khác song vẫn có sự liên tục. Kriging thuộc nhóm thuật toán bình phương tối thiểu tuyến tính. Ưu điểm của Kriging là giá trị các điểm được nội suy không chỉ phụ thuộc vào khoảng cách mà còn phụ thuộc và sự phân bố không gian của chúng. Đây là một thuật toán nội suy phức tạp, đòi hỏi thời gian tính toán nhiều hơn so với các thuật toán khác.

Neighbor Natural Neighbor là phương pháp nội suy được phát triển bởi Sibson (1981) dựa trên lưới Voronoi của một tập các điểm rời rạc. Thuật toán này cho phép tìm tập con gần nhất của mẫu đầu vào và áp dụng trọng số cho chúng dựa trên các khu vực tương ứng để nội suy các giá trị chưa biết (Phan Quốc Yên, Nguyễn Thị Thu Nga, 2020).

Spline là phương pháp nội suy xây dựng các đường cong trơn đi qua các điểm dữ liệu đã biết, thực tế là đi tìm hàm số $f(x)$ sao cho $f(x_i) = y_i$ với mọi i (Đặng Văn Nam và nnk, 2020). Phương pháp nội suy này tỏ ra hiệu quả trong trường hợp tập dữ liệu không dày, chẳng hạn dữ liệu về khí tượng, thủy văn.

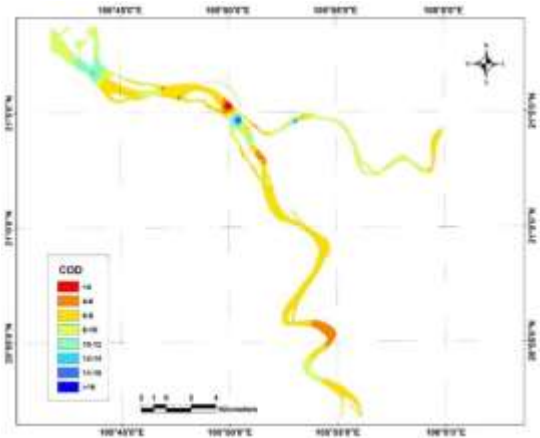
Kết quả nhận được sẽ được đánh giá độ chính xác thông qua so sánh sai số trung phương (RMSE), sau đó lựa chọn kỹ thuật nội suy có giá trị RMSE nhỏ nhất để xây dựng bản đồ phân bố hàm lượng COD, BOD₅ và NH₄⁺ cho toàn bộ khu vực nghiên cứu.



Hình 2: Sơ đồ phương pháp nội suy trong xác định hàm lượng các thông số chất lượng nước

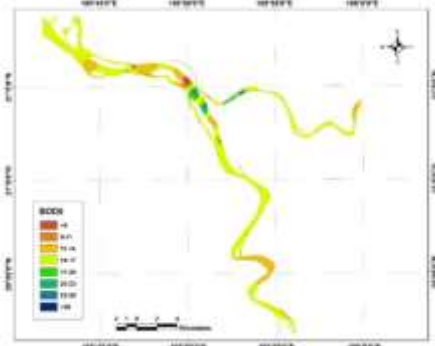
3. Kết quả nghiên cứu và thảo luận

Từ số liệu quan trắc, lấy mẫu chất lượng nước ở 56 điểm, trong nghiên cứu tiến hành ước lượng hàm lượng COD khu vực sông Hồng và sông Đuống đoạn chảy qua thành phố Hà Nội bằng phương pháp nội suy. Kết quả nhận được cho thấy, từ 10 điểm kiểm tra, giá trị RMSE khi nội suy bằng 04 kỹ thuật Kriging, IDW, Neighbor Natural và Spline đạt lần lượt là 1,93 (mg/l), 0,084 (mg/l), 0,269 (mg/l) và 0,201 (mg/l). Như vậy, kỹ thuật nội suy IDW có độ chính xác cao nhất đối với bộ dữ liệu khu vực nghiên cứu. Trên hình 3, hàm lượng COD được chia thành các khoảng: < 4, 4 - 6, 6 - 8, 8 - 10, 10 - 12, 12 - 14, 14 - 16 và > 16 (mg/l). Phân tích kết quả nhận được cho thấy, hàm lượng COD khu vực sông Hồng và sông Đuống đoạn chảy qua thành phố Hà Nội chủ yếu đạt dưới 10 (mg/l). Một số vị trí có hàm lượng COD cao cục bộ (trên 16 mg/l) được thể hiện bởi màu xanh trên hình 3.

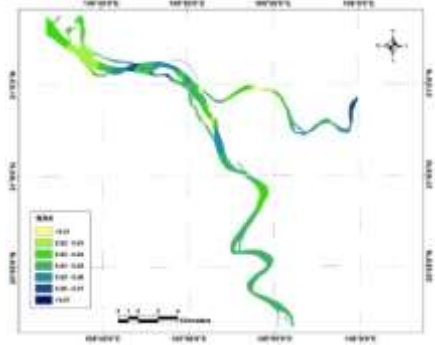


Hình 3: Kết quả xác định hàm lượng COD khu vực nghiên cứu

Tương tự như với thông số COD, trong nghiên cứu cũng tiến hành xác định phân bố các thông số BOD₅ và NH₄⁺ từ tập dữ liệu quan trắc bằng 4 kỹ thuật nội suy khác nhau. Kết quả nhận được cũng cho thấy, phương pháp nội suy IDW có độ chính xác cao hơn so với các phương pháp Kriging, Neighbor Natural và Spline, thể hiện qua giá trị RMSE đạt 0,092 (mg/l) đối với thông số BOD₅ và 0,0009 (mg/l) đối với thông số NH₄⁺. Trong khi đó, khi nội suy bằng các kỹ thuật Kriging, Neighbor Natural và Spline, giá trị RMSE đối với thông số BOD₅ đạt tương ứng là 3,28; 0,553 và 0,264 (mg/l), với thông số NH₄⁺ đạt 0,014; 0,0025 và 0,002 (mg/l). Từ kết quả này, trong nghiên cứu đã sử dụng phương pháp nội suy IDW để tiến hành xây dựng bản đồ phân bố hàm lượng BOD₅ và NH₄⁺ cho khu vực sông Hồng và sông Đuống, đoạn chảy qua thành phố Hà Nội. Kết quả xây dựng bản đồ phân bố hàm lượng BOD₅ và NH₄⁺ cho khu vực thực nghiệm được thể hiện trên các hình 4 và 5.



Hình 4: Kết quả xác định hàm lượng BOD₅ khu vực nghiên cứu



Hình 5: Kết quả xác định hàm lượng NH₄⁺ khu vực nghiên cứu

Để trực quan, hàm lượng BOD₅ và NH₄⁺ cũng được phân vùng thành các giá trị khác nhau trên hình 4 và 5. Từ kết quả nhận được cho thấy, khu vực giao nhau giữa sông Hồng và sông Đuống có hàm lượng BOD₅ và NH₄⁺ cao hơn so với các khu vực khác.

Như vậy, có thể khẳng định, phương pháp nội suy cho phép đánh giá được phân bố hàm lượng các thông số chất lượng nước từ tập dữ liệu quan trắc với độ chính xác khoảng 80% (thể hiện qua giá trị sai số trung phương RMSE). Tuy nhiên, phương pháp này cũng có nhược điểm khi kết quả không có tính liên tục, chi tiết đến từng pixel như đối với phương pháp hồi quy dựa trên mối quan hệ với phản xạ phổ mặt nước. Do vậy, nội suy chỉ được sử dụng với các thông số chất lượng nước không thể xác định từ dữ liệu viễn thám.

4. Kết luận

Trong nghiên cứu này, 4 kỹ thuật nội suy thông dụng, bao gồm Kriging, IDW, Neighbor

Natural Neighbor và Spline được thử nghiệm để ước lượng hàm lượng các thông số chất lượng nước (COD, BOD₅ và NH₄⁺) khu vực sông Hồng và sông Đuống đoạn chảy qua thành phố Hà Nội. 56 điểm lấy mẫu chất lượng nước, bao gồm 43 điểm để nội suy và 13 điểm để đánh giá độ chính xác được sử dụng để ước lượng hàm lượng các thông số trên cho toàn bộ khu vực nghiên cứu. Kết quả nhận được cho thấy, phương pháp nội suy IDW có độ chính xác cao nhất trong ước lượng hàm lượng các thông số COD, BOD₅ và NH₄⁺, thể hiện qua giá trị sai số trung phương (RMSE) đạt thấp nhất so với 3 phương pháp nội suy còn lại. Từ kết quả này, trong nghiên cứu đã lựa chọn phương pháp IDW để xây dựng bản đồ phân bố hàm lượng COD, BOD₅ và NH₄⁺ cho toàn bộ khu vực nghiên cứu. Kết quả nhận được cho thấy, hàm lượng các thông số COD, BOD₅ và NH₄⁺ đều đạt cao nhất ở khu vực giao nhau giữa sông Hồng và sông Đuống. ○

Tài liệu tham khảo

[1]. Nguyễn Thị Thu Hà, Bùi Đình Cảnh, Nguyễn Thiên Phương Thảo, Bùi Thị Nhị (2016). Thử nghiệm mô hình hóa sự phân bố không gian của hàm lượng chlorophyll-a và chỉ số trạng thái phú dưỡng nước Hồ Tây sử dụng ảnh Sentinel-2A, Tạp chí Khoa học ĐHQGHN: Các Khoa học Trái đất và Môi trường 32, 2S, trang 121-130.

[2]. Đặng Văn Nam, Hoàng Quý Nhân, Ngô Văn Mạnh, Nguyễn Thị Hiền (2020). Các phương pháp chuẩn hóa dữ liệu thủy văn áp dụng cho trạm 74129 - Yên Bái, Tạp chí Khí tượng Thủy văn, số 6, 18 - 29.

[3]. Đinh Lam Thắng, Đinh Thị Thu Hiền, Chu Việt Thức, Trịnh Lê Hùng, Nguyễn Thị Thu Nga (2022). Xây dựng quy trình xác định một số thông số chất lượng nước từ ảnh viễn thám quang học Sentinel 2 và số liệu quan trắc, Tạp chí Khoa học Đo đạc và Bản đồ, số 53, trang 57 - 64.

[4]. Phan Quốc Yên, Nguyễn Thị Thu Nga (2020). Thực nghiệm, đánh giá các phương pháp nội suy bề mặt địa hình cho các dạng địa hình khác nhau, Tạp chí Khoa học Kỹ thuật Mô Địa cát, số 61(2), 116 - 125.

[5]. Boumpoulis V., Michalopoulou M., Depountis N. (2023). Comparison between different spatial interpolation methods for the development of sediment distribution maps in coastal areas, Earth Science Informatics, 16, 2069 - 2087.

[6]. Gnauck A. (2004). Interpolation and approximation of water quality time series and process identification, Analytical and Bioanalytical Chemistry, 380, 484 - 492.

[7]. Mielniczuk U., Mielniczuk J., Obrosiak R., Przystupa W. (2019). A comparison of some interpolation techniques for determining spatial distribution of

nitrogen compounds in groundwater, International Journal of Environmental Research, 13, 679 - 687.

[8]. Trinh Le Hung, Zablotskii V.R., Nguyen T.T.N. (2020). Determining the concentration of suspended sediment in the lower Day River (Northern Vietnam) using MSI Sentinel 2 high spatial resolution data, Izvestiya, Atmospheric and Oceanic Physics, Vol. 56(12), 1645 - 1652.

[9]. Trinh Le Hung, Zablotskii V., Thi Giang Le, Thi Trinh Le, Thi Tham Trinh, Thi Thu Hien Dinh, Thi Thu Nga Nguyen (2018). Suspended sediment concentration monitoring using VNREDSat - 1A multispectral data, a case study in Red River, Hanoi, Vietnam, Geography, Environment, Sustainability, Vol.11(3), 49 - 60.○

Summary

Estimation the concentration of water quality parameters in the Red River and Duong River using the spatial interpolation method

Dinh Thi Thu Hien, Electric Power University, Hanoi, Vietnam

Dinh Lam Thang, Institute for Mechanical, Automation and Environmental Science and Technologies

Nguyen Thi Thu Nga, Dao Khanh Hoai, Institute of Techniques for Special Engineering, Le Quy Don Technical University, Hanoi, Vietnam

Spatial interpolation technique have been used effectively in research on Neighbor Natural resources and the environment based on in situ data. This article presents the results of estimating the concentration of some surface water quality parameters in the Red River and Duong River, including COD, BOD₅ and NH₄⁺ using the interpolation method. 56 waters quality samples were used to estimated the water quality parameters concentration, in which 43 samples were used to train the interpolation model and 13 samples were used to evaluate accuracy. Four common interpolation techniques, including Kriging, Inverse Distance Weighted (IDW), Neighbor Natural and Spline methods were tested to select the method with the highest accuracy. The results showed that the IDW interpolation method achieved the highest accuracy, shown by comparing the root mean square error (RMSE) for all three parameters COD, BOD₅ and NH₄⁺. From the obtained results, in the study, the spatial distribution maps of COD, BOD₅ and NH₄⁺ parameters in the study area was built. The results received in the study can be used to monitor and evaluate surface water quality in the Red River and Duong River - the section flowing through Hanoi city.○

Keywords: water pollution, spatial interpolation, water quality parameters, Red River, Duong River.