

PHÂN TÍCH BIẾN ĐỘNG KHÔNG GIAN - THỜI GIAN VÀ CÁC NGUỒN Ô NHIỄM NƯỚC MẶT TẠI TỈNH AN GIANG BẰNG THỐNG KÊ ĐA BIẾN

Nguyễn Xuân Lan^{1,2}, Nguyễn Thanh Giao³, Huỳnh Vương Thu Minh^{3,*}

¹ Nghiên cứu sinh Môi trường Đất và nước, Đại học Cần Thơ

² Trung tâm Quan trắc Tài nguyên và Môi trường An Giang,
Sở Nông nghiệp và Môi trường tỉnh An Giang

³ Khoa Môi trường và Tài nguyên thiên nhiên, Đại học Cần Thơ

*Email: hvtnminh@ctu.edu.vn

TÓM TẮT

Nghiên cứu này ứng dụng các phương pháp thống kê đa biến, bao gồm: Phân tích cụm (CA) và phân tích thành phần chính (PCA), nhằm đánh giá các quy luật biến động và xác định những nguồn gây ô nhiễm chính ảnh hưởng đến chất lượng nước mặt tại tỉnh An Giang. Phân tích được thực hiện trên bộ dữ liệu gồm 10 thông số chất lượng nước (nhiệt độ, pH, DO, COD, BOD₅, TSS, N-NH₄⁺, N-NO₃⁻, P-PO₄³⁻ và coliform), được thu thập tại 62 vị trí trong giai đoạn 2021 - 2023. Kết quả CA cho thấy, mạng lưới quan trắc có thể được tối ưu hóa hiệu quả bằng cách giảm số vị trí giám sát từ 62 xuống còn 46 vị trí và điều chỉnh tần suất lấy mẫu từ 6 đợt xuống 4 đợt mỗi năm, nhưng vẫn đảm bảo tính đại diện cho toàn khu vực. Trong khi đó, PCA đã xác định các yếu tố chi phối sự suy giảm chất lượng nước chủ yếu là nhóm ô nhiễm hữu cơ, dinh dưỡng và vi sinh vật. Sự biến động của các thông số này có nguồn gốc từ các hoạt động phát triển kinh tế - xã hội như: Sinh hoạt đô thị, thâm canh nông nghiệp, nuôi trồng thủy sản, bên cạnh các quá trình tự nhiên trong lưu vực. Những phát hiện của nghiên cứu cung cấp luận cứ khoa học quan trọng, hỗ trợ các nhà quản lý trong việc tái thiết kế mạng lưới quan trắc và ưu tiên các giải pháp kiểm soát ô nhiễm, góp phần quản lý bền vững tài nguyên nước.

Từ khóa: An Giang, chất lượng nước mặt, phân tích thành phần chính, phân tích cụm, thống kê đa biến.

1. ĐẶT VẤN ĐỀ

Quản lý chất lượng nước mặt đóng vai trò nền tảng trong việc bảo vệ sức khỏe con người, duy trì đa dạng sinh học và đảm bảo sự phát triển kinh tế - xã hội bền vững. Tuy nhiên, dữ liệu quan trắc truyền thống thường rất lớn, đa chiều, khó diễn giải. Trước thách thức đó, phương pháp phân tích thống kê đa biến (multivariate statistical analysis) đặc biệt là PCA và CA đã được sử dụng phổ biến, đã chứng tỏ hiệu quả vượt trội [1 - 3]. PCA cho phép nhận diện các thông số trọng yếu và các nguồn ô nhiễm tiềm tàng [1], [4], trong khi CA hỗ trợ phân nhóm các vị trí quan trắc hoặc các thời điểm lấy mẫu có những đặc điểm chất lượng nước tương đồng vào cùng một cụm [4], từ đó cho phép

tối ưu hóa mạng lưới quan trắc về cả không gian và thời gian, giúp tiết kiệm đáng kể chi phí và nhân lực nhưng vẫn đảm bảo tính đại diện [2], [5].

Tại Việt Nam, đặc biệt là vùng đồng bằng sông Cửu Long (ĐBSCL), chất lượng nước đang suy giảm đáng báo động do chịu tác động cộng hưởng từ đô thị hóa, nông nghiệp thâm canh, nuôi trồng thủy sản quy mô lớn [6], [7], xây dựng các đập thủy điện và biến đổi khí hậu. Một số nghiên cứu đã áp dụng PCA và CA như: Đánh giá chất lượng nước trên sông Hậu [8 - 10], hay tại tỉnh Bến Tre [11], các nghiên cứu này mới dừng ở phạm vi hẹp, chưa có nghiên cứu hệ thống quy mô cấp tỉnh, đặc biệt là tại tỉnh An Giang, với vị trí chiến lược ở thượng nguồn ĐBSCL [12] và mạng lưới sông ngòi

dày đặc, đang đối diện áp lực ô nhiễm cục bộ [13], [14] nhưng chưa có đánh giá toàn diện bằng công cụ hiện đại.

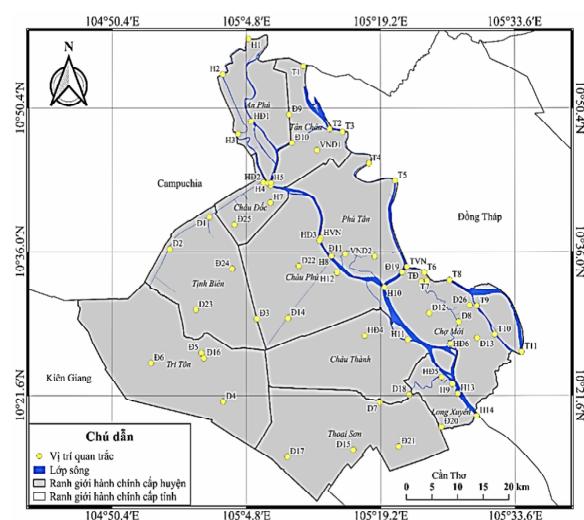
Từ thực tiễn đó, nghiên cứu này được thực hiện nhằm ứng dụng các phương pháp phân tích đa biến để làm rõ mối tương quan giữa các chỉ số chất lượng nước, xác định các nguồn ô nhiễm chính và tối ưu hóa mạng lưới quan trắc môi trường nước mặt tại tỉnh An Giang. Để đạt được mục tiêu tổng quát, nghiên cứu tập trung vào ba mục tiêu cụ thể: (1) Phân loại các khu vực nước mặt theo mức độ tương đồng về chất lượng nước nhằm xác định các vùng ô nhiễm trọng điểm; (2) Xác định các thông số quan trắc chính yếu và nhận dạng các nguồn ô nhiễm chủ đạo; (3) đề xuất một mạng lưới quan trắc chất lượng nước mặt được tối ưu hóa về không gian (vị trí) và thời gian (tần suất), giúp nâng cao hiệu quả và giảm thiểu chi phí cho công tác giám sát môi trường của tỉnh. Nghiên cứu sử dụng bộ dữ liệu quan trắc chất lượng nước mặt của tỉnh An Giang trong ba năm liên tiếp (2021 – 2023), bao gồm các hệ thống sông chính và kênh rạch nội đồng. Kết quả của nghiên cứu được kỳ vọng sẽ cung cấp những luận cứ khoa học xác đáng, giúp các nhà quản lý môi trường có cái nhìn toàn diện về hiện trạng và các yếu tố tác động đến chất lượng nước, từ đó xây dựng các chiến lược và chính sách quản lý, bảo vệ tài nguyên nước một cách hiệu quả, phù hợp với định hướng phát triển kinh tế - xã hội bền vững của tỉnh An Giang trong giai đoạn tới.

2. PHƯƠNG PHÁP NGHIÊN CỨU

Nghiên cứu sử dụng bộ dữ liệu thứ cấp về chất lượng nước mặt tỉnh An Giang trong giai đoạn 2021 - 2023, được thu thập từ các báo cáo của Sở Tài nguyên và Môi trường (nay là Sở Nông nghiệp và Môi trường tỉnh An Giang) [15 - 17]. Dữ liệu bao gồm kết quả quan trắc của 10 thông số: Nhiệt độ, pH, oxy hòa tan (DO), nhu cầu oxy hóa học (COD), nhu cầu oxy sinh hóa (BOD_5), tổng chất rắn lơ lửng (TSS), amoni ($N-NH_4^+$), nitrate ($N-NO_3^-$), orthophosphat ($P-PO_4^{3-}$) và coliform tại 62 vị trí trên sông Tiền, sông Hậu và hệ thống kênh rạch nội đồng.

Toàn bộ số liệu được tổng hợp, xử lý sơ bộ bằng phần mềm Microsoft Excel, sau đó được

phân tích chuyên sâu bằng phần mềm Primer. Phương pháp tiếp cận chính của nghiên cứu là áp dụng các kỹ thuật thống kê đa biến, trong đó CA và PCA là hai phương pháp được sử dụng phổ biến nhất trong các nghiên cứu đánh giá chất lượng nước [12], [18]. Các phương pháp này được dùng để đánh giá các quy luật biến động theo không gian, thời gian và xác định các nguồn gây ô nhiễm tiềm tàng. CA được thực hiện nhằm mục tiêu phân nhóm các vị trí và thời điểm quan trắc dựa trên sự tương đồng về đặc điểm chất lượng nước, từ đó tối ưu hóa mạng lưới giám sát [2]. Kỹ thuật phân nhóm theo phương pháp liên kết của Ward (Ward's method) và thước đo khoảng cách Euclide bình phương (squared Euclidean distance) đã được sử dụng để xác định các cụm có tính đồng nhất nội nhóm cao và khác biệt ngoại nhóm lớn [19], [20]. Kết quả phân cụm được trực quan hóa bằng biểu đồ hình cây (dendrogram), làm cơ sở cho việc đề xuất điều chỉnh mạng lưới quan trắc. PCA được áp dụng để giảm chiều dữ liệu và xác định các yếu tố hoặc nguồn chính gây ra sự thay đổi chất lượng nước [4], [2]. Các thành phần chính (PC) có giá trị riêng (Eigenvalue) lớn hơn 1 được xem là có ý nghĩa và được giữ lại để diễn giải [21]. Mức độ đóng góp và tầm quan trọng của từng thông số ban đầu vào mỗi PC được đánh giá thông qua hệ số tải (factor loadings), từ đó giúp nhận dạng các nguồn ô nhiễm tiềm tàng như từ hoạt động nông nghiệp, sinh hoạt hay các quá trình tự nhiên [11].



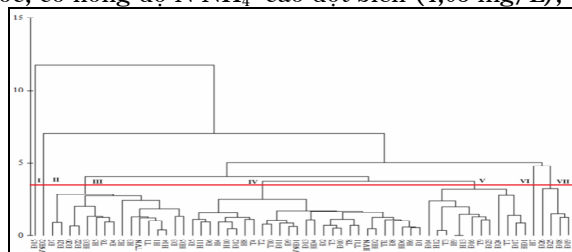
Hình 1. Vị trí thu mẫu tỉnh An Giang

3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Phân tích biến động chất lượng nước theo không gian và hiện trạng ô nhiễm

Phân tích biến động chất lượng nước theo không gian là bước đi nền tảng nhằm mục tiêu "vẽ bản đồ" các vùng có đặc tính ô nhiễm tương đồng, từ đó xác định các điểm nóng (hotspots) cần được ưu tiên quản lý. Nghiên cứu này đã tiến hành CA trên bộ dữ liệu của từng năm riêng lẻ từ năm 2021 đến năm 2023 để theo dõi sự thay đổi trong cấu trúc ô nhiễm. Kết quả cho thấy, một quy luật phân hóa không gian rõ rệt và tương đối ổn định qua các năm.

Cụ thể, vào năm 2021, 62 vị trí quan trắc đã được phân thành 7 cụm riêng biệt (Hình 2). Nổi bật nhất là cụm 1, chỉ bao gồm một vị trí duy nhất (Đ15), thể hiện mức độ ô nhiễm hữu cơ và vi sinh cực kỳ nghiêm trọng, với nồng độ COD 150,6 mg/L và BOD₅ 98,0 mg/L, vượt xa các giới hạn cho phép của QCVN 08:2023/BTNMT [22], kể cả mức C dành cho giao thông thủy. Mật độ coliform tại đây cũng ở mức báo động (8.041.667 MPN/100 mL). Vị trí này nằm trên kênh nội đồng, được xác định chịu tác động trực tiếp từ nước thải sinh hoạt của khu đô thị tập trung, không qua xử lý [15]. Các cụm 3 (16 vị trí) và cụm 4 (27 vị trí) bao gồm phần lớn các điểm quan trắc, có mức độ ô nhiễm hữu cơ và dinh dưỡng thấp hơn nhưng vẫn ở mức đáng quan ngại, cho thấy sự ảnh hưởng lan tỏa từ các hoạt động nông nghiệp và dân sinh dọc các tuyến sông. Đáng chú ý, cụm 5 (12 vị trí) có chỉ số DO trung bình rất thấp (3,95 mg/L), dưới ngưỡng cho phép tại mức C (≥ 4 mg/L), báo hiệu tình trạng thiếu hụt oxy nghiêm trọng có thể gây hại cho hệ thủy sinh. Cụm 6, đại diện bởi vị trí gần khu xử lý nước thải tập trung của thành phố Châu Đốc, có nồng độ N-NH₄⁺ cao đột biến (4,05 mg/L),

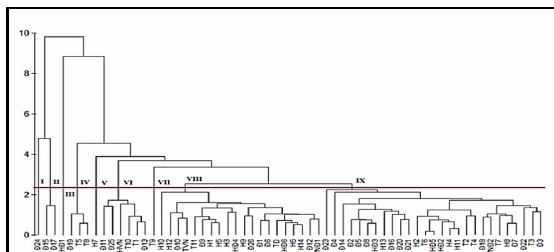


Hình 2. Diễn biến chất lượng nước mặt theo không gian tại tỉnh An Giang năm 2021

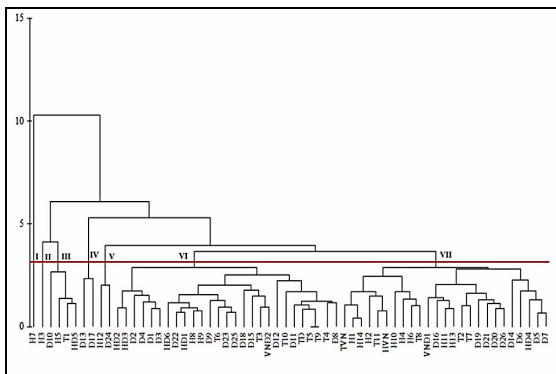
chỉ rõ đây là một nguồn điểm gây ô nhiễm amoni cần được kiểm soát chặt chẽ.

Bước sang năm 2022, cấu trúc phân cụm có sự thay đổi nhẹ, phân thành 9 cụm (Hình 3), tuy nhiên các quy luật cơ bản vẫn được duy trì. Kết quả cho thấy, các vị trí nội đồng như Đ24 (cụm 1), Đ15 và Đ17 (cụm 2) tiếp tục là những điểm ô nhiễm nặng nhất. Vị trí giao giữa sông Hậu và kênh nội đồng (HĐ1) thuộc cụm 3 cũng cho thấy hàm lượng hữu cơ cao, khẳng định vai trò của các kênh rạch trong việc gom chất ô nhiễm từ các khu dân cư ra sông chính [16]. So với năm 2021, chất lượng nước năm 2022 có xu hướng giảm nhẹ ở một số chỉ tiêu hữu cơ tại các cụm lớn, nhưng vấn đề ô nhiễm amoni và vi sinh vẫn còn rất phổ biến. Phát hiện này tương đồng với kết quả các nghiên cứu khác tại ĐBSCL, vốn xác định ô nhiễm hữu cơ và vi sinh là hai thách thức môi trường hàng đầu của vùng [5].

Phân tích của năm 2023 (Hình 4) tiếp tục củng cố các nhận định trên. Chất lượng nước được phân thành 7 cụm, với các vị trí nội đồng và các khu vực chịu ảnh hưởng từ điểm xả thải tập trung vẫn là những nơi có chất lượng nước kém nhất. Một điểm đáng chú ý trong năm 2023 là nồng độ TSS và coliform tại cụm 4 và 5 tương đối cao, cho thấy sự gia tăng ảnh hưởng từ xói mòn và cuốn trôi bề mặt, có thể liên quan đến các hoạt động xây dựng hoặc thay đổi sử dụng đất ven bờ, bên cạnh các nguồn thải từ chăn nuôi [5]. Nhìn chung, phân tích riêng lẻ từng năm khẳng định một thực trạng đáng lo ngại: Các kênh rạch nội đồng tại tỉnh An Giang đang hoạt động như những "ống dẫn" tập trung chất ô nhiễm từ các hoạt động dân sinh, kinh tế và đổ ra các dòng sông chính, trong khi khả năng tự làm sạch của chúng rất hạn chế do mật cát hẹp và lưu lượng thấp.



Hình 3. Diễn biến chất lượng nước mặt theo không gian tại tỉnh An Giang năm 2022

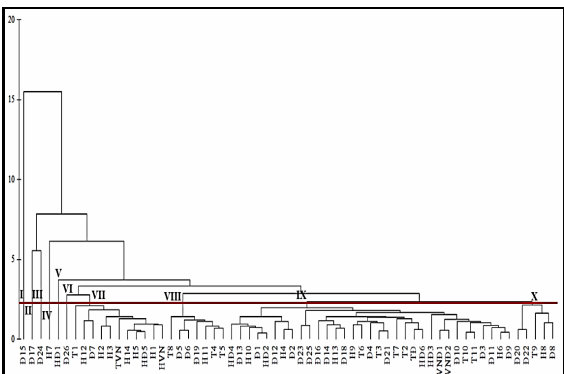


Hình 4. Diễn biến chất lượng nước mặt theo không gian tại tỉnh An Giang năm 2023

3.2. Nghiên cứu mạng lưới quan trắc không gian dựa trên CA

Việc phân tích dữ liệu riêng lẻ từng năm cung cấp cái nhìn về sự biến động, tuy nhiên để đưa ra các đề xuất mang tính chiến lược và dài hạn cho mạng lưới quan trắc, cần dựa trên những quy luật ổn định và bền vững nhất. Do đó, nghiên cứu đã tiến hành CA trên bộ dữ liệu trung bình của cả giai đoạn 2021 - 2023. Phương pháp này giúp loại bỏ các biến động ngẫu nhiên hoặc bất thường trong một năm cụ thể, từ đó làm nổi bật lên cấu trúc không gian cốt lõi của chất lượng nước. Kết quả được trình bày ở hình 5 là cơ sở khoa học vững chắc nhất để tái cấu trúc mạng lưới quan trắc.

Phân tích tổng hợp đã nhóm 62 vị trí thành 10 cụm, phản ánh một cách chi tiết hơn các mức độ ô



Hình 5. Diễn biến chất lượng nước mặt trung bình năm 2021 - 2023 theo không gian

nhễm khác nhau. Các cụm từ 1 - 6, mỗi cụm chỉ gồm một vị trí, đại diện cho những điểm dị biệt và ô nhiễm nặng nhất, bao gồm: Đ15, Đ17, Đ24, H7, HĐ1 và Đ26. Các vị trí này cần được xem là các "điểm nóng" và phải được giám sát chặt chẽ, thậm chí với tần suất cao hơn. Các cụm còn lại có quy mô lớn hơn: Cụm 7 (11 vị trí), cụm 8 (7 vị trí), cụm 9 (33 vị trí) và cụm 10 (5 vị trí). Sự hình thành các cụm lớn này, đặc biệt là cụm 9 cho thấy một vùng không gian rộng lớn bao gồm các vị trí trên sông chính và các kênh lớn có đặc tính chất lượng nước tương đối đồng nhất. Đây chính là cơ hội để tối ưu hóa mạng lưới. Việc duy trì quan trắc tại tất cả các vị trí trong cùng một cụm lớn có thể dẫn đến sự dư thừa thông tin và lãng phí nguồn lực.

Bảng 1. Các vị trí quan trắc cụm 9 có đặc tính giống nhau

Ký hiệu vị trí	Năm	Thông số quan trắc									
		pH	Nhiệt độ (°C)	DO (mg/L)	COD (mg/L)	BOD (mg/L)	TSS (mg/L)	N-NO ₃ ⁻ (mg/L)	N-NH ₄ ⁺ (mg/L)	P-PO ₄ ³⁻ (mg/L)	Coliform (MPN/100 mL)
HĐ2	2021	7,17	30,07	4,93	21,33	14,00	54,33	0,29	0,85	0,06	21.100
	2022	7,11	29,40	5,49	16,67	10,33	50,00	0,09	0,43	0,04	18.017
	2023	7,15	29,98	5,04	20,83	13,33	55,67	0,47	0,35	0,05	16.150
H4	2021	7,17	29,17	4,68	16,67	10,67	49,00	0,14	0,34	0,05	30.600
	2022	7,03	28,73	4,64	16,33	10,50	50,17	0,10	0,30	0,04	16.067
	2023	7,13	29,62	5,13	18,17	11,50	55,17	0,28	0,35	0,04	14.517
Đ23	2021	7,23	30,43	0,00	18,67	12,33	48,67	0,05	0,44	0,00	14.200
	2022	7,15	28,80	3,96	16,83	11,00	45,17	0,03	0,34	0,01	13.717
	2023	7,14	29,12	4,56	23,83	15,33	60,67	0,25	0,61	0,03	16.800
H13	2021	7,13	30,33	0,00	21,33	13,67	46,33	0,22	0,58	0,00	14.500
	2022	7,16	29,15	4,27	32,50	21,00	81,17	0,12	0,81	0,18	32.500
	2023	7,19	29,63	4,54	20,00	13,17	56,67	0,30	0,58	0,04	13.650

Dựa trên nguyên tắc này, nghiên cứu đưa ra đề xuất cốt lõi: Giảm số lượng vị trí quan trắc từ 62 xuống còn 46 vị trí, tức là giảm 16 điểm. Các vị trí được đề xuất loại bỏ bao gồm: H3, H4, H5, H8, H9, T8, Đ5, Đ7, Đ11, Đ19, Đ21, Đ22, Đ23, TĐ, HVN và HD2. Việc lựa chọn các vị trí này dựa trên sự tương đồng ổn định của chúng với các vị trí khác được giữ lại trong cùng một cụm qua nhiều năm. Ví dụ, trong cụm 9, các vị trí HD2, H4, Đ23, H13 có đặc tính rất giống nhau; do đó, việc giữ lại một hoặc hai vị trí đại diện là đủ để theo dõi xu hướng chung của tiểu vùng. Đề xuất này không chỉ giúp tiết kiệm chi phí vận hành, lấy mẫu và phân tích, mà quan trọng hơn, nó cho phép tái phân bổ nguồn lực để thiết lập các điểm quan trắc mới tại những khu vực đang phát triển công nghiệp hoặc đô thị hóa nhanh, nơi có nguy cơ phát sinh các nguồn ô nhiễm mới. Cách tiếp cận này sử dụng CA để hợp lý hóa mạng lưới giám sát, đã được cộng đồng khoa học quốc tế công nhận và áp dụng thành công tại nhiều quốc gia, chứng tỏ tính hiệu quả và khoa học của nó trong công tác quản lý môi trường hiện đại [3]. Do đó, việc áp dụng đề xuất này sẽ giúp chương trình quan trắc của tỉnh An Giang trở nên hiệu quả, linh hoạt và khoa học hơn.

3.3. Phân tích biến động chất lượng nước theo thời gian và giảm tần suất quan trắc

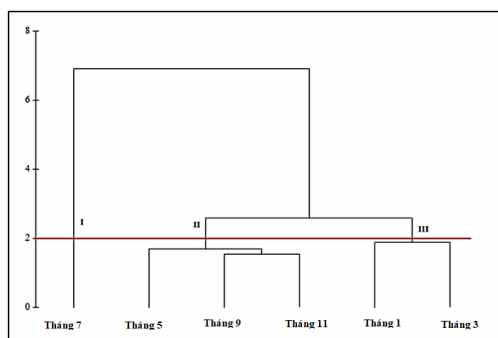
Bên cạnh biến động không gian, chất lượng nước mặt còn chịu sự chi phối mạnh mẽ của yếu tố thời gian, đặc biệt là quy luật mùa trong năm. CA theo thời gian đã được thực hiện để làm rõ quy luật này và đề xuất tối ưu hóa tần suất lấy mẫu. Kết quả phân tích cho cả năm 2022 và 2023 đều cho thấy một sự phân tách rõ ràng giữa các tháng mùa mưa và các tháng mùa khô.

Cụ thể, kết quả phân tích năm 2022 (Hình 6) đã nhóm 6 đợt quan trắc thành 3 cụm. Cụm 3 bao gồm tháng 1 và 3, đại diện cho mùa khô. Cụm 1 (tháng 7) và cụm 2 (tháng 5, 9, 11) đại diện cho mùa mưa. Đặc điểm chất lượng nước của hai mùa này rất khác biệt. Vào mùa mưa, đặc biệt là đầu mùa, nồng độ TSS thường tăng cao (58,2 mg/L tại cụm 1) do sự xói mòn và rửa trôi các chất từ bề mặt

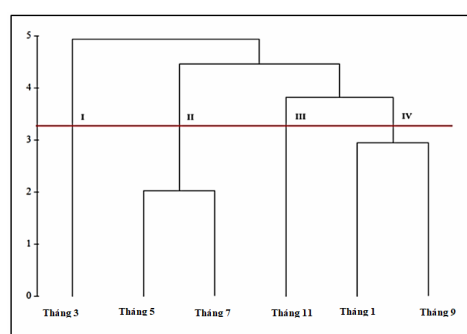
lưu vực [23]. Ngược lại, vào mùa khô (cụm 3), nồng độ các chất ô nhiễm hữu cơ như COD (20,8 mg/L) và BOD₅ (13,4 mg/L) có xu hướng thấp hơn do dòng chảy ổn định và ít bị xáo trộn. Tuy nhiên, nồng độ N-NH₄⁺ tại cụm 3 (0,35 mg/L) lại cao hơn cụm 2 (0,24 mg/L), có thể do lưu lượng nước thấp làm giảm khả năng pha loãng chất ô nhiễm từ các nguồn thải.

Quy luật này tiếp tục được khẳng định trong năm 2023 (Hình 7). Bốn cụm được hình thành, nhưng vẫn thể hiện sự khác biệt mùa rõ rệt. Cụm 2 (tháng 5, 7) đại diện cho giai đoạn đầu và giữa mùa mưa, có nồng độ các chất ô nhiễm cao nhất. Điều này có thể được giải thích bởi hiện tượng "cuốn trôi đầu mùa" (first flush), khi lượng mưa đầu mùa mang theo một lượng lớn chất ô nhiễm tích tụ từ mùa khô vào sông [24]. Cụm 3 (tháng 11), đại diện cho cuối mùa mưa, lại có chất lượng nước tốt nhất, với chỉ số COD (19,1 mg/L) đã nằm trong giới hạn cho phép của QCVN 08:2023/BTNMT [22] tại mức C. Điều này có thể do các chất ô nhiễm đã được rửa trôi ở các đợt mưa trước đó, đồng thời lượng mưa lớn làm tăng khả năng pha loãng và tái tạo oxy trong nước [8]. Cụm 4 (tháng 1, 9) đại diện cho giai đoạn chuyển tiếp giữa hai mùa, có đặc tính chất lượng nước trung gian.

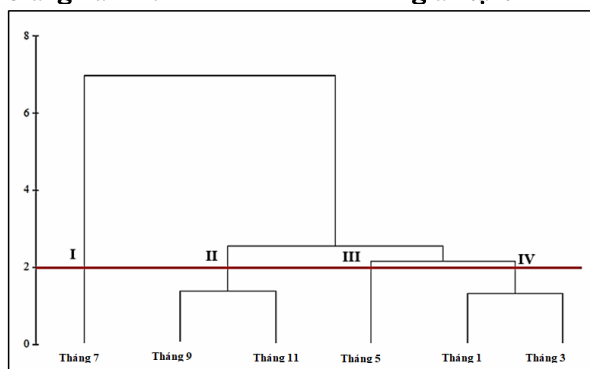
Khi tổng hợp dữ liệu trung bình năm 2022 và 2023 (Hình 8), quy luật phân tách mùa mưa - mùa khô càng trở nên rõ nét. Sự tương đồng giữa các tháng trong cùng một mùa (tháng 1 và 3 trong cụm 4; tháng 9 và 11 trong cụm 2) cho thấy, việc lấy mẫu ở tất cả các tháng này có thể không cần thiết. Do đó, nghiên cứu đề xuất có thể giảm tần suất quan trắc từ 6 đợt xuống còn 4 đợt mỗi năm. Việc lựa chọn 4 đợt này cần đảm bảo tính đại diện cho các giai đoạn đặc trưng: Đầu mùa khô, cuối mùa khô - đầu mùa mưa, đỉnh mùa mưa và cuối mùa mưa - đầu mùa khô. Đề xuất này không chỉ giúp tối ưu hóa nguồn lực mà còn đảm bảo chương trình quan trắc vẫn nắm bắt được đầy đủ các xu hướng biến động chất lượng nước quan trọng nhất trong năm, phù hợp với các nghiên cứu về ảnh hưởng của mùa và thủy triều đến chất lượng nước trong khu vực [9].



Hình 6. Diễn biến chất lượng nước mặt theo thời gian tại tỉnh An Giang năm 2022



Hình 7. Diễn biến chất lượng nước mặt theo thời gian tại tỉnh An Giang năm 2023



Hình 8. Diễn biến chất lượng nước mặt trung bình năm 2022 - 2023 theo thời gian

3.4. Nhận dạng các nguồn ô nhiễm tiềm tàng bằng PCA

Phân tích PCA theo không gian qua các năm (Bảng 2 - 4) luôn cho thấy sự nổi bật của một vài PC đầu tiên, giải thích phần lớn sự biến thiên của dữ liệu. PC thứ nhất (PC1) hầu như luôn có hệ số tải (loading) cao với các chỉ tiêu COD, BOD₅, N-NH₄⁺ và đôi khi là coliform. Điều này cho thấy, nhóm các chỉ tiêu này luôn biến động cùng nhau. Do đó, PC1 có thể được định danh là "yếu tố ô nhiễm nhân tạo", đại diện cho sự tác động tổng hợp của các nguồn thải giàu chất hữu cơ và dinh

dưỡng. Nguồn gốc của yếu tố này chủ yếu đến từ nước thải sinh hoạt chưa qua xử lý từ các khu dân cư, nước thải từ các hoạt động chăn nuôi và các làng nghề, dư lượng phân bón từ hoạt động nông nghiệp và chất thải từ các khu nuôi trồng thủy sản [25]. Thành phần chính thứ hai (PC2) thường có hệ số tải cao với các thông số vật lý như: Nhiệt độ, pH, DO. PC2 có thể được xem là "yếu tố thủy văn - tự nhiên", phản ánh các quá trình tự nhiên như sự thay đổi nhiệt độ theo mùa, mối quan hệ nghịch đảo giữa nhiệt độ và khả năng hòa tan oxy và các quá trình quang hợp/hô hấp của thủy sinh vật.

Bảng 2. Thành phần ảnh hưởng đến chất lượng nước mặt năm 2021 theo không gian

Thông số	PC1	PC2	PC3	PC4	PC5
pH	-0,117	-0,394	0,45	-0,041	-0,491
Nhiệt độ (°C)	0,132	-0,331	0,342	-0,701	-0,014
DO (mg/L)	-0,258	0,376	0,415	-0,249	0,309
COD (mg/L)	0,505	0,008	0,09	0,084	0,051
BOD ₅ (mg/L)	0,505	0,01	0,092	0,085	0,054
TSS (mg/L)	0,339	0,284	0,057	-0,125	-0,172
N-NO ₃ ⁻ (mg/L)	-0,105	0,254	0,58	0,511	-0,286
N-NH ₄ ⁺ (mg/L)	0,358	0,25	-0,164	-0,134	-0,53
P-PO ₄ ³⁻ (mg/L)	-0,063	0,613	0,071	-0,328	-0,019

KHOA HỌC & CÔNG NGHỆ

Thông số	PC1	PC2	PC3	PC4	PC5
Coliform (MPN/100 mL)	0,367	-0,096	0,346	0,167	0,515
Giá trị riêng	3,76	1,85	1,18	0,89	0,79
% phương sai	37,6	18,5	11,8	8,9	7,9
% tích lũy	37,6	56,1	67,9	76,8	84,7

Bảng 3. Thành phần ảnh hưởng đến chất lượng nước mặt năm 2022 theo không gian

Thông số	PC1	PC2	PC3	PC4	PC5
pH	0,076	-0,629	-0,255	0,034	-0,194
Nhiệt độ (°C)	0,032	-0,518	-0,144	0,448	0,534
DO (mg/L)	0,335	0,073	-0,041	0,27	0,268
COD (mg/L)	-0,475	-0,008	-0,114	0,13	-0,068
BOD ₅ (mg/L)	-0,474	-0,008	-0,126	0,128	-0,07
TSS (mg/L)	-0,185	-0,228	-0,286	-0,731	0,356
N-NO ₃ ⁻ (mg/L)	0,046	0,49	-0,462	-0,013	0,496
N-NH ₄ ⁺ (mg/L)	-0,437	-0,071	0,096	-0,075	0,062
P-PO ₄ ³⁻ (mg/L)	-0,386	0,181	-0,26	0,388	-0,055
Coliform (MPN/100 mL)	-0,233	-0,028	0,715	0,012	0,465
Giá trị riêng	4,15	1,76	1,07	0,93	0,8
% phương sai	41,5	17,6	10,7	9,3	8
% tích lũy	41,5	59,1	69,7	79	87

Bảng 4. Thành phần ảnh hưởng đến chất lượng nước mặt năm 2023 theo không gian

Thông số	PC1	PC2	PC3	PC4	PC5
pH	-0,099	-0,465	0,367	-0,168	0,365
Nhiệt độ (°C)	-0,117	-0,429	0,229	-0,545	-0,289
DO (mg/L)	-0,297	-0,393	-0,043	-0,065	0,175
COD (mg/L)	0,486	-0,155	-0,072	-0,04	0,173
BOD ₅ (mg/L)	0,486	-0,155	-0,069	-0,038	0,176
TSS (mg/L)	0,161	0,341	0,466	-0,132	0,607
N-NO ₃ ⁻ (mg/L)	0,015	0,217	-0,583	-0,719	0,195
N-NH ₄ ⁺ (mg/L)	0,431	-0,147	0,09	-0,063	-0,399
P-PO ₄ ³⁻ (mg/L)	0,422	-0,275	-0,144	0,122	0,032
Coliform (MPN/100 mL)	0,161	0,369	0,464	-0,338	-0,355
Giá trị riêng	3,71	2,21	1,14	0,88	0,8
% phương sai	37,1	22,1	11,4	8,8	8
% tích lũy	37,1	59,2	70,6	79,5	87,5

PCA theo thời gian (Bảng 5 - 6) tiếp tục củng cố mô hình này. Các kết quả cho thấy, sự biến động theo mùa của chất lượng nước là sự tương tác giữa hai yếu tố chính trên. Vào mùa khô, do lưu lượng thấp, ảnh hưởng của "yếu tố ô nhiễm nhân tạo" (PC1) trở nên rõ rệt hơn do khả năng pha loãng kém. Ngược lại, vào mùa mưa, "yếu tố thủy văn - tự nhiên" (PC2) trở nên mạnh mẽ hơn, với sự gia tăng của TSS do xói mòn và sự thay đổi của DO do xáo trộn dòng chảy.

Một điểm thảo luận quan trọng là sự liên kết chặt chẽ giữa kết quả của PCA và CA. Các vị trí được CA xếp vào nhóm "ô nhiễm nặng" (Đ15, Đ24) chính là những vị trí có điểm số (factor scores) cao nhất trên "yếu tố ô nhiễm nhân tạo" (PC1) của PCA. Sự tương thích này không chỉ làm tăng độ tin cậy của cả hai phương pháp mà còn khẳng định một cách mạnh mẽ rằng, các điểm nóng ô nhiễm không gian tại tỉnh An Giang bị chi phối chủ yếu bởi các nguồn thải nhân tạo. Tóm lại,

PCA đã cung cấp một bức tranh toàn cảnh về các "động lực" đằng sau sự thay đổi chất lượng nước, chỉ rõ rằng, để cải thiện môi trường nước, các nỗ

lực quản lý cần phải tập trung vào việc kiểm soát và giảm thiểu các nguồn thải từ hoạt động dân sinh, nông nghiệp và thủy sản.

Bảng 5. Thành phần ảnh hưởng đến chất lượng nước mặt năm 2022 theo thời gian

Thông số	PC1	PC2	PC3	PC4	PC5
pH	-0,117	-0,502	-0,021	0,038	-0,011
Nhiệt độ (°C)	-0,266	-0,063	0,677	0,518	0,262
DO (mg/L)	-0,231	-0,335	-0,3	0,556	-0,369
COD (mg/L)	-0,436	0,196	0,06	-0,162	0,155
BOD ₅ (mg/L)	-0,438	0,188	0,059	-0,171	0,116
TSS (mg/L)	0,398	0,185	-0,12	0,454	0,016
N-NO ₃ ⁻ (mg/L)	0,146	0,461	-0,284	0,188	0,235
N-NH ₄ ⁺ (mg/L)	-0,295	0,354	-0,229	0,35	0,195
P-PO ₄ ³⁻ (mg/L)	-0,392	0,268	-0,132	-0,025	-0,627
Coliform (MPN/100 mL)	0,246	0,339	0,528	0,04	-0,52
Giá trị riêng	4,38	3,73	0,99	0,79	0,11
% phương sai	43,8	37,3	9,9	7,9	1,1
% tích lũy	43,8	81,1	91	98,9	100

Bảng 6. Thành phần ảnh hưởng đến chất lượng nước mặt năm 2023 theo thời gian

Thông số	PC1	PC2	PC3	PC4	PC5
pH	-0,363	-0,247	-0,082	0,487	-0,533
Nhiệt độ (°C)	-0,313	-0,229	-0,472	-0,447	0,142
DO (mg/L)	-0,136	-0,476	0,038	-0,558	-0,106
COD (mg/L)	-0,445	0,042	-0,095	0,055	0,102
BOD ₅ (mg/L)	-0,445	0,047	-0,093	0,073	0,053
TSS (mg/L)	0,14	-0,475	0,314	-0,049	-0,483
N-NO ₃ (mg/L)	0,25	-0,41	-0,265	0,276	0,269
N-NH ₄ ⁺ (mg/L)	0,257	-0,433	0,087	0,152	0,413
P-PO ₄ ³⁻ (mg/L)	-0,272	0,023	0,744	-0,152	0,234
Coliform (MPN/100 mL)	-0,364	-0,274	0,148	0,344	0,375
Giá trị riêng	4,95	3,4	1,11	0,43	0,11
% phương sai	49,5	34	11,1	4,3	1,1
% tích lũy	49,5	83,5	94,6	98,9	100

4. KẾT LUẬN

Bằng cách áp dụng các phương pháp thống kê đa biến, nghiên cứu đã giải mã thành công các quy luật phức tạp chi phối chất lượng nước mặt trên địa bàn tỉnh An Giang. Kết quả khẳng định, hiện trạng môi trường nước chịu sự tác động của một hệ thống kép: Các nguồn ô nhiễm nhân tạo và các yếu tố thủy văn theo mùa. Về mặt khoa học, nghiên cứu đã xác định ô nhiễm hữu cơ (BOD₅, COD), dinh dưỡng (N-NH₄⁺) và vi sinh vật (coliform) là những yếu tố chính gây suy giảm chất lượng nước, trong đó các kênh rạch nội đồng đóng vai trò là

những "điểm nóng" tích tụ và lan truyền ô nhiễm. Từ những phát hiện này, nghiên cứu cung cấp hai ý nghĩa ứng dụng quan trọng cho công tác quản lý. Thứ nhất, đưa ra luận cứ khoa học vững chắc để tái cấu trúc mạng lưới quan trắc, đề xuất giảm số vị trí giám sát từ 62 xuống 46 và tần suất lấy mẫu từ 6 xuống 4 đợt mỗi năm. Việc tối ưu hóa này giúp tiết kiệm đáng kể chi phí mà vẫn đảm bảo hiệu quả giám sát toàn diện. Thứ hai, chỉ rõ các nhà quản lý cần ưu tiên nguồn lực để kiểm soát chặt chẽ các nguồn thải từ hoạt động sinh hoạt đô thị và thâm canh nông nghiệp, vốn là nguyên nhân chính gây

ra tình trạng ô nhiễm tại các điểm nóng. Nghiên cứu không chỉ đóng góp tri thức mới vào việc tìm hiểu môi trường nước tại ĐBSCL mà còn cung cấp một công cụ hỗ trợ ra quyết định dựa trên bằng chứng, hướng tới việc quản lý tài nguyên nước tại tỉnh An Giang một cách bền vững và hiệu quả hơn.

TÀI LIỆU THAM KHẢO

1. Feher, I. C., Moldovan, Z. & Oprean, I. (2016). Spatial and seasonal variation of organic pollutants in surface water using multivariate statistical techniques. *Water Science and Technology*, 74(7), 1726 - 1735.
2. Giao, N. T. & Minh, V. Q. (2022). The categorized of surface water quality variation using multivariate statistical approaches: A case study of Ben Tre province, Vietnam. *Trends in Science*, 19, 1 - 12.
3. Giao, N. T. & Nhien, H. T. H. (2022). Assessment of surface water quality and monitoring in Southern Vietnam using multicriteria statistical approaches. *Sustainable Environment Research*, 32(1), 1 - 12.
4. Chounlamany, V., Tanchuling, M. A. & Inoue, T. (2017). Spatial and temporal variation of water quality of a segment of Marikina river using multivariate statistical methods. *Water Science and Technology*, 76(6), 1510 - 1522.
5. Hong, T. T. K. & Giao, N. T. (2022). Assessment of surface water quality Duyen hai district, Tra Vinh province, VietNam. *Journal of Energy Technology and Environment*, 4(4), 1 - 11.
6. Jung, S. Y. & Kim, I. K. (2017). Analysis of water quality factor and correlation between water quality and Chl-a in middle and downstream weir section of Nakdong river. *Journal of Korean Society of Environmental Engineers*, 39, 89 - 96.
7. Karroum, L. A., Baghdadi, M. E., Barakat, A., Meddah, R., Aadraoui, M., Oumenskou, H. & Ennaji, W. (2019). Assessment of surface water quality using multivariate statistical techniques: El Abid river, Middle Atlas, Morocco as a case study. *Desalination and water treatment*, 143, 118 - 125.
8. Kong, Z., Song, Y., Shao, Z. & Chai, H. (2021). Biochar-pyrite bi-layer bioretention system for dissolved nutrient treatment and by-product generation control under various stormwater conditions. *Water Research*, 206, 117737.
9. Lee, K., Kang, T., Ryu, H., Hwang, S. & Kim, K. (2020). Analysis of spatiotemporal variation in river water quality using clustering techniques: A case study in the Yeongsan river, Republic of Korea. *Environmental Science and Pollution Research*, 27, 29327 - 29340.
10. Ni, D. V., Viet, L. H. & Ly, N. H. T. (2022). Seasonal and tidal influence on surface water quality - A case study in the Hau river segment, Vietnamese Mekong Delta province. *Journal of Ecological Engineering*, 23(8), 290 - 298.
11. Liu, C. W., Lin, K. H. & Kuo, Y. M. (2003). Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Science of the Total Environment*, 313(1 - 3), 77 - 89.
12. Mutea, F. G., Nelson, H. K., Au, H. V., Huynh, T. G. & Vu, U. N. (2021). Assessment of water quality for aquaculture in Hau river, Mekong Delta, Vietnam using multivariate statistical analysis. *Water*, 13(22), 3307.
13. Nguyễn Thanh Giao (2020). Chất lượng nước mặt tại các nhánh sông tiếp giáp sông Hậu thành phố Cần Thơ. *Tạp chí Nông nghiệp và Phát triển nông thôn*, 15, 79 - 86.
14. Wehrheim, C., Lübken, M., Stolpe, H. & Wichern, M. (2023). Identifying key influences on surface water quality in freshwater areas of the Vietnamese Mekong Delta from 2018 to 2020. *Water*, 15(7), 1295.
15. Sở Tài nguyên và Môi trường tỉnh An Giang (2021). *Báo cáo tổng hợp kết quả quan trắc chất lượng môi trường tỉnh An Giang năm 2021*.
16. Sở Tài nguyên và Môi trường tỉnh An Giang (2022). *Báo cáo tổng hợp kết quả quan trắc chất lượng môi trường tỉnh An Giang năm 2022*.
17. Sở Tài nguyên và Môi trường tỉnh An Giang (2023). *Báo cáo tổng hợp kết quả quan trắc chất lượng môi trường tỉnh An Giang năm 2023*.
18. Nguyễn Thanh Giao, Phùng Thị Hằng, Dương Văn Ni, Lê Thị Diễm Mi & Huỳnh Bá Lộc (2021). Đánh giá chất lượng nước mặt huyện Kế Sách, tỉnh Sóc Trăng sử dụng phương pháp phân tích đa biến. *Tạp chí Khoa học Tài nguyên và Môi*

trường, Trường Đại học Tài nguyên và Môi trường Hà Nội, (35), 68 - 79.

19. Ward, J. H. (1963). Ward's method. *Journal of American Statistical Association*, 58, 236 – 246.

20. Salah, E. A., Turki, A. M. & Al-Othman, E. M. (2012). Assessment of water quality of Euphrates river using cluster analysis. *Journal of Environmental Protection*, 3(12), 1629.

21. Zhang, Q., Li, Z., Zeng, G., Li, J., Fang, Y., Yuan, Q., Wang, Y. & Ye, F. (2009). Assessment of surface water quality using multivariate statistical techniques in red soil hilly region: A case study of Xiangjiang watershed, China. *Environmental Monitoring and Assessment*, 152, 123 - 131.

22. Quy chuẩn kỹ thuật Quốc gia QCVN 08:2023/BTNMT về Chất lượng nước mặt.

23. Võ Thị Ngọc Giàu, Phan Thị Bích Tuyền & Nguyễn Hiếu Trung (2019). Đánh giá biến động chất lượng nước mặt sông Cần Thơ giai đoạn 2010 - 2014 bằng phương pháp tính toán chỉ số chất lượng nước (WQI). *Tạp chí Khoa học Trường Đại học Cần Thơ*, 55(2), 105 - 113.

24. Yuan, H., Yin, H., Yang, Z., Yu, J., Liu, E., Li, Q., Tai, Z. & Ca, Y. (2020). Diffusion kinetic process of heavy metals in lacustrine sediment assessed under different redox conditions by DGT and DIFS model. *Science of The Total Environment*, 741, 140418.

25. Wang, X., Cuthbertson, A., Gualtieri, C. & Shao, D. (2020). A review on mariculture effluent: Characterization and management tools. *Water*, 12(11), 2991.

SPATIOTEMPORAL VARIATION AND POLLUTION SOURCE APPORTIONMENT OF SURFACE WATER IN AN GIANG PROVINCE USING MULTIVARIATE STATISTICS

Nguyen Xuan Lan^{1,2}, Nguyen Thanh Giao³, Huynh Vuong Thu Minh³

¹PhD student in Soil and Water Environment, College of Environment and Natural Resources, Can Tho University

²Center of Monitoring Environment and Resources An Giang, Department of Agriculture and Environment of An Giang province

³College of Environment and Natural Resources, Can Tho University

Abstract

This study applies multivariate statistical methods, including Cluster Analysis (CA) and Principal Component Analysis (PCA), to assess variation patterns and identify the main pollution sources affecting surface water quality in An Giang province. The analysis was performed on a dataset of 10 water quality parameters (temperature, pH, DO, COD, BOD₅, TSS, N-NH₄⁺, N-NO₃⁻, P-PO₄³⁻ and coliform), collected from 62 locations from 2021 to 2023. The Cluster Analysis results indicated that the monitoring network could be effectively optimized by reducing the number of monitoring sites from 62 to 46 and adjusting the sampling frequency from six to four times per year, while still ensuring representativeness for the entire area. Meanwhile, Principal Component Analysis identified that the governing factors of water quality degradation are primarily organic, nutrient and microbiological pollution. The variation in these parameters originates from socioeconomic activities such as urban wastewater, intensive agriculture and aquaculture, in addition to natural processes within the catchment area. The findings of this study provide a crucial scientific basis to support managers in redesigning the monitoring network and prioritizing pollution control measures, thus contributing to the sustainable management of water resources.

Keywords: Surface water quality, multivariate statistics, Principal Component Analysis, Cluster Analysis, An Giang.

Ngày nhận bài: 25/6/2025

Ngày chuyển phản biện: 3/7/2025

Ngày thông qua phản biện: 25/7/2025

Ngày duyệt đăng: 19/9/2025