

SO SÁNH MỘT SỐ PHƯƠNG PHÁP NỘI SUY KHÔNG GIAN TRONG XÂY DỰNG BẢN ĐỒ CHẤT LƯỢNG NƯỚC SÔNG TIỀN PHỤC VỤ NUÔI CÁ TRA THÂM CANH

Trần Minh Thiện^{1,2*}, Nguyễn Thanh Trúc²

¹Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

²Viện Nghiên cứu Nuôi trồng Thủy sản II

*Email: tranminhthien85@gmail.com

Ngày nhận bài: 23/2/2023; Ngày chấp nhận đăng: 27/3/2023

TÓM TẮT

Ứng dụng GIS trong xây dựng bản đồ chất lượng nước đã được áp dụng rộng rãi trong thời gian qua. Tuy nhiên, chưa có đánh giá cụ thể nào về việc áp dụng thuật toán nội suy phù hợp cho việc xây dựng bản đồ chất lượng nước tại khu vực sông Tiền thuộc Đồng bằng sông Cửu Long. Bài báo này so sánh và đánh giá khả năng sử dụng của 3 phương pháp nội suy không gian thông dụng là IDW (Inverse Distance Weighting), RBF (Radial Basis Functions) và Kriging trên 11 thông số chất lượng nước, bao gồm nhiệt độ, pH, độ dẫn điện (EC – Electrical Conductivity), độ kiềm, oxy hòa tan (DO), tổng chất rắn lơ lửng (TSS), nhu cầu oxy hóa học (COD), nitrite (NO₂-N), TAN (Total Ammonia Nitrogen), phosphate (PO₄³-P) và vi khuẩn *Aeromonas* sp. tại 15 điểm quan trắc với 42 đợt thu mẫu trong thời gian từ 2020 - 2021. Chỉ số Willmott (Willmott's index of agreement) được sử dụng cho việc xem xét đánh giá kết quả. Kết quả cho thấy không có sự khác biệt có ý nghĩa thống kê giữa 3 phương pháp nội suy không gian trên khu vực nghiên cứu ($p > 0,05$). Cả ba phương pháp cho thấy vị trí và thông số chất lượng nước có ảnh hưởng đến độ chính xác của kết quả nội suy. Giá trị trung bình của chỉ số Willmott ứng với các phương pháp IDW, Kriging và RBF lần lượt là $0,72 \pm 0,18$, $0,71 \pm 0,17$ và $0,70 \pm 0,18$. Hay phương pháp IDW có thể được xem là một lựa chọn ưu tiên để nội suy các thông số chất lượng nước trên khu vực nghiên cứu.

Từ khóa: IDW, RBF, Kriging, GIS, bản đồ chất lượng nước.

1. MỞ ĐẦU

Nuôi trồng thủy sản nước ngọt, tiêu biểu là cá tra, là một trong các đối tượng nuôi chủ lực tại Việt Nam. Năm 2020, diện tích nuôi cá tra toàn quốc gần 20 nghìn ha (hơn 6.000 ha nuôi cá tra thâm canh và bán thâm canh, sản lượng 1.494,5 nghìn tấn), trong đó diện tích nuôi cá tra tập trung chủ yếu tại Đồng bằng sông Cửu Long (chiếm 94,9% diện tích của cả nước) [1]. Theo báo cáo của các Chi cục Thủy sản Đồng Tháp, An Giang, Vĩnh Long, Bến Tre, Cần Thơ, 9 tháng đầu năm 2022, tổng diện tích nuôi thuộc 5 tỉnh này là 4.635,7 ha và tổng sản lượng cá tra thu hoạch đạt 1.275 ngàn tấn. Diện tích nuôi giảm 8,1% và sản lượng cá tra tăng 126,1% so với năm 2021. Trong đó, hai tỉnh An Giang và Đồng Tháp dẫn đầu về sản lượng thu hoạch (chiếm 69,9%) và các tỉnh Bến Tre, Cần Thơ, Vĩnh Long chiếm 30,1% về sản lượng thu hoạch của các tỉnh thuộc ĐBSCL [2].

Sông Mekong là nguồn cung cấp nước chính cho các vùng nuôi cá tra trọng điểm của khu vực đang đối mặt với sự suy giảm dòng chảy, xâm nhập mặn từ biển Đông và nguy cơ ô nhiễm từ khu vực thượng nguồn [3]. Theo kết quả quan trắc của Viện Nghiên cứu Nuôi trồng thủy sản II phục vụ công tác nuôi cá tra trọng điểm tại ĐBSCL cũng cho thấy có sự vượt giá trị cho phép theo QCVN 02-20:2014/BNNPTNT và QCVN 08-MT:2015/BTNMT ở một số chỉ tiêu quan trắc, đặc biệt số liệu năm 2021 cho thấy các chỉ tiêu TSS, COD, TAN, NO₂-N, PO₄³-P xuất hiện nhiều lần vượt ngưỡng giá trị [4]. Điều này cho thấy chất lượng nước không đáp ứng yêu cầu ở nhiều nơi và ảnh hưởng đến hoạt động nuôi trồng thủy sản nói chung và cá tra thâm canh nói riêng tại khu vực.

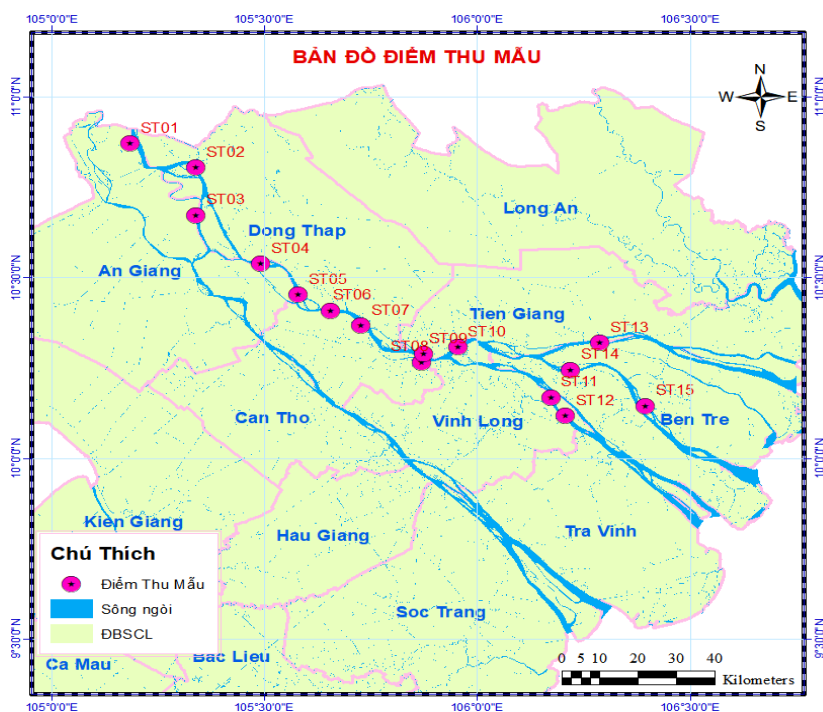
Việc quản lý chất lượng nước lưu vực đang được chính quyền địa phương và cả người nông dân quan tâm. Một hệ thống quan trắc hoàn chỉnh là cần thiết cho khu vực. Việc xác định vùng chất lượng nước phù hợp cho nuôi cá tra thâm canh là rất cần thiết và thiết thực. Tuy nhiên với một lưu vực rộng lớn, nguồn kinh phí trong giới hạn đã gây ra một số khó khăn nhất định trong quản lý và cho người nông dân. Công cụ nội suy không gian đã được đề xuất để dự đoán các giá trị chưa biết trên lưu vực. Tuy nhiên chưa có các nghiên cứu nhằm xác định thuật toán nội suy phù hợp cho lưu vực Đồng bằng sông cửu Long.

Kimleang và cộng sự (2017) chỉ ra rằng phương pháp nội suy Kriging cho kết quả tốt hơn IDW tại hồ Tonle Sap, Cambodia [5]. Nagalakshmi và cộng sự (2016) đã lựa chọn Kriging vì cho rằng phương pháp cho sai số tốt hơn trong việc xây dựng bản đồ chất lượng nước tại Tamil Nadu, Ấn Độ [6]. Trong khi đó, Ahmad và cộng sự (2021) cho rằng IDW và RBF tốt hơn Kriging, đồng thời IDW tốt hơn RBF trong nghiên cứu xây dựng bản đồ nước ngầm tại Qatar [7]. Tại Việt Nam, Nguyễn Thị Thu Hiền và cộng sự (2018) nghiên cứu xây dựng bản đồ chất lượng nước suối Nậm La, tỉnh Sơn La cho rằng không có sự khác biệt lớn giữa phương pháp IDW và phân tích mẫu đối chứng [8]. Ngoài ra, Nông Thị Hải Yến và Nguyễn Hải Hòa (2017) cho rằng hai phương pháp nội suy IDW, Kriging đều có thể sử dụng để xây dựng bản đồ chất lượng nước mặt, trong đó phương pháp IDW cho độ chính xác cao hơn [9]. Trong nghiên cứu này, ba phương pháp nội suy không gian thông dụng là nội suy khoảng cách có trọng số (IDW), Radial Basis Functions (RBF) và Kriging được đánh giá so sánh với dữ liệu nội suy được sử dụng từ 42 đợt thu mẫu trong thời gian 2020 - 2021 tại 15 điểm quan trắc trên khu vực sông Tiền. Các thông số được sử dụng cho đánh giá nội suy bao gồm: nhiệt độ, pH, oxy hòa tan, độ kiềm, TSS, EC, TAN, $\text{NO}_2\text{-N}$, $\text{PO}_4^{3-}\text{-P}$, COD, *Aeromonas* tổng số.

2. SỐ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Thu thập số liệu

Nghiên cứu sử dụng số liệu quan trắc môi trường của Trung tâm Quan trắc Môi trường và Bệnh thủy sản Nam bộ, Viện Nghiên cứu Nuôi trồng Thủy sản II giai đoạn 2020 - 2021 tại 15 điểm quan trắc trên sông Tiền, đoạn chảy qua các tỉnh An Giang, Đồng Tháp, Vĩnh Long và Bến Tre (Hình 1).



Hình 1. Vị trí thu mẫu

Bảng 1. Vị trí thu mẫu

STT	Mã	Huyện	Tỉnh	Kinh độ	Vĩ độ
1	ST01	Tân Châu	An Giang	105.18516	10.87039
2	ST02	Tp. Hồng Ngự	Đồng Tháp	105.34028	10.80389
3	ST03	Thanh Bình	Đồng Tháp	105.34000	10.67250
4	ST04	Chợ Mới	An Giang	105.49197	10.53728
5	ST05	Cao Lãnh	Đồng Tháp	105.58083	10.45167
6	ST06	Lấp Vò	Đồng Tháp	105.65583	10.40806
7	ST07	Sa Đéc	Đồng Tháp	105.72861	10.36667
8	ST08	Châu Thành	Đồng Tháp	105.87082	10.26528
9	ST09	Châu Thành	Đồng Tháp	105.87588	10.28815
10	ST10	Long Hồ	Vĩnh Long	105.95500	10.30694
11	ST11	Mang Thít	Vĩnh Long	106.17549	10.16749
12	ST12	Vũng Liêm	Vĩnh Long	106.20892	10.11615
13	ST13	Châu Thành	Bến Tre	106.28984	10.31930
14	ST14	Châu Thành	Bến Tre	106.22067	10.24369
15	ST15	Giồng Trôm	Bến Tre	106.39701	10.14382

Thời gian quan trắc của dữ liệu được sử dụng trong nghiên cứu này là từ tháng 01/2020 đến 11/2021. Các thông số quan trắc gồm: Nhiệt độ, pH, độ dẫn điện, độ kiềm, DO (Dissolved Oxygen), TAN, NO₂-N, PO₄³⁻-P, TSS (Total suspended solids), COD (Chemical Oxygen Demand), *Aeromonas* tổng số (Bảng 2). Tần suất quan trắc là 2 lần/tháng.

Bảng 2. Các thông số và phương pháp phân tích

STT	Thông số	Phương pháp phân tích
1	pH	TCVN 6492:2011
2	Nhiệt độ	SMEWW 2510 B: 2017
3	DO	TCVN 7325: 2004
4	EC	SMEWW 2510 B: 2017
5	Độ Kiềm	SMEWW 2320: 2017
6	TSS	SMEWW 4500-TSS - 2017
7	COD	Đun hoàn lưu mở, dùng tác nhân KMnO ₄ , acid hóa môi trường bằng H ₂ SO ₄
8	TAN	SMEWW 4500 - NH ₃ - F: 2017
9	PO ₄ ³⁻ -P	SMEWW-4500-P.D:2017
10	NO ₂ -N	SMEWW 4500-NO ₂ B: 2017
11	<i>Aeromonas</i> tổng số	SMEWW 9260 L: 2017

Ghi chú: SMEWW (Standard Methods for the Examination of Water and Wastewater) phương pháp tiêu chuẩn cho kiểm tra nước và nước thải.

2.2. Phương pháp nghiên cứu

2.2.1. Phương pháp nội suy không gian

Nội suy không gian là việc xác định giá trị tại các điểm chưa biết từ tập dữ liệu của các điểm đã biết bằng những thuật toán. Nội suy không gian là công cụ nhằm ước tính dữ liệu không gian, và được phân thành hai loại: phương pháp nội suy xác định (phi địa lý) và nội suy địa lý. Hiện nay, công cụ nội suy không gian đã phát triển nhanh chóng với nhiều thuật toán khác nhau. Một số phương pháp nội suy không

gian được sử dụng phổ biến gồm nghịch đảo khoảng cách (IDW), nội suy không gian Kriging, láng giềng tự nhiên (NN) và hàm cơ sở xuyên tâm (RBF), ... [10]. Trong đó, IDW và RBF là các phương pháp nội suy xác định, sử dụng các mô hình toán học đơn giản để xác định một điểm chưa biết từ các điểm đã biết xung quanh. Trong khi đó, Kriging sử dụng các mô hình địa thống kê phức tạp hơn, bao gồm các mối tương quan thống kê giữa các điểm được lấy mẫu, còn được gọi là tự tương quan không gian [7].

Phương pháp IDW xác định giá trị của các điểm chưa biết dựa trên sự kết hợp tuyến tính của dữ liệu có sẵn. Các điểm gần với vị trí cần ước tính được xác định trọng số lớn hơn, khi khoảng cách tăng lên thì trọng số sẽ giảm [11]. Công thức như sau:

$$Z = \frac{\sum_{i=1}^n w_i z_i}{\sum_{i=1}^n w_i} = \frac{\sum_{i=1}^n \frac{1}{d_i^k} z_i}{\sum_{i=1}^n \frac{1}{d_i^k}} \quad (1)$$

trong đó: Z là giá trị dự đoán; Z_i là giá trị đã biết; n là tổng số điểm đã biết; d là khoảng cách đến điểm i; k là hệ số mũ. Giá trị của k càng cao thì độ ảnh hưởng của các điểm ở xa càng thấp.

Phương pháp RBF được đặt cho một nhóm các bộ nội suy sử dụng một phương trình cơ bản phụ thuộc vào khoảng cách giữa điểm được nội suy và các điểm lấy mẫu. Giá trị dự đoán của RBF có thể được biểu thị dưới dạng tổng của hai thành phần [11].

$$Z = \sum_{i=1}^m a_i f_i(x) + \sum_{j=1}^n b_j \psi(d_j) \quad (2)$$

Trong đó: Z là giá trị dự đoán; $f_i(x)$ là một hàm xu hướng; $\psi(d_i)$ là hàm cơ sở xuyên tâm; d_i là khoảng cách đến điểm i; a_i và b_j được tính bằng cách giải hệ phương trình tuyến tính.

Tương tự như phương pháp IDW, Kriging là sự kết hợp tuyến tính của các giá trị được quan sát với trọng số [11]. Tuy nhiên, trọng số Kriging được xác định bởi khoảng cách và hướng không gian của các điểm hoặc bán phương sai đã biết [7].

$$Z = \sum_{i=1}^n z_i \lambda_i, \quad \text{với} \quad \sum_{i=1}^n \lambda_i = 1 \quad (3)$$

trong đó: Z là giá trị dự đoán; Z_i là giá trị đã biết; n là tổng số điểm đã biết; λ_i là trọng số.

Mỗi phương pháp nội suy đều có những ưu điểm và nhược điểm riêng. Phương pháp IDW có ưu điểm là khá chính xác, ngoại trừ một số trường hợp ngoại lệ. Trong khi đó, Kriging tạo ra các bề mặt nhẵn và tương đối liên tục chứa các thay đổi bất thường trong dữ liệu, nhưng các giá trị nội suy có thể vượt quá điểm đo tối thiểu và tối đa. Ưu điểm của RBF là nó tạo ra bề mặt đều và mịn với mức độ thay đổi dần dần. RBF không phù hợp với những thay đổi đáng kể về giá trị của các điểm mẫu trong khoảng cách ngắn hoặc có những sai số lớn trong tập mẫu [7].

Nói chung, các kỹ thuật nội suy khác nhau sẽ tạo ra các kết quả không giống nhau với cùng một dữ liệu đầu vào, và không có phương pháp nào luôn chính xác hơn các phương pháp khác trong mọi trường hợp. Độ chính xác của phương pháp phụ thuộc vào yếu tố cần nội suy và các đặc điểm của khu vực nghiên cứu [7]. Nghiên cứu này thực hiện nhằm đánh giá độ chính xác của ba phương pháp thông dụng là IDW, RBF và Kriging trong việc nội suy các thông số chất lượng nước khu vực sông Tiền.

2.2.2. Phương pháp đánh giá độ chính xác nội suy không gian

Với 15 điểm quan trắc, nghiên cứu này giữ lại lần lượt 1 điểm trong 5 điểm được lựa chọn làm đối chứng để so sánh với kết quả nội suy. Các điểm được giữ lại là các ST5 đến ST9 (Hình 1). Các điểm này nằm giữa của khu vực nghiên cứu và khoảng cách giữa chúng với các điểm quan trắc còn lại không quá xa.

Để so sánh kết quả nội suy giữa ba phương pháp không gian, nghiên cứu này sử dụng chỉ số Willmott để đánh giá kết quả. Chỉ số này được Willmott (1980) đề xuất như một thước đo tiêu chuẩn hóa về mức độ sai số dự đoán mô hình [12]. Chỉ số này thể hiện tỷ lệ giữa sai số bình phương trung bình và sai số tiềm ẩn. Với chỉ số này, giá trị bằng 1 cho biết sự phù hợp cao nhất và giảm dần khi tiến về 0. Chỉ số này được xác định như sau:

$$d = 1 - \frac{\sum_{i=1}^n (X_{o_i} - X_{s_i})^2}{\sum_{i=1}^n (|X_{s_i} - \bar{X}_o| + |X_{o_i} - \bar{X}_o|)^2} \quad (4)$$

Trong đó X_{o_i} , X_{s_i} tương ứng là giá quan trắc và giá trị mô phỏng, với $i = 1, n$ là độ dài của chuỗi. \bar{X}_o và \bar{X}_s tương ứng là giá trị trung bình của chuỗi quan trắc và chuỗi mô phỏng.

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Đánh giá các thông số chất lượng nước trên khu vực nghiên cứu

Từ kết quả thống kê cho thấy các thông số nhiệt độ, pH, COD, $\text{NO}_2\text{-N}$, $\text{PO}_4^{3-}\text{-P}$, TAN trong các đợt khảo sát hầu hết đều đạt dưới ngưỡng cho phép theo QCVN 08-MT:2015/BTNMT (cột A1, giới hạn các thông số COD, $\text{NO}_2\text{-N}$, TAN, $\text{PO}_4^{3-}\text{-P}$ lần lượt là 10 mg/L, 0,05 mg/L, 0,3 mg/L và 0,1 mg/L) [13]. Trong đó, kết quả khảo sát có 5,2% số liệu COD lớn hơn 10 mg/L, 16,8% số liệu $\text{NO}_2\text{-N}$ lớn hơn 0,05 mg/L, 11,0% số liệu thông số TAN lớn hơn 0,3 mg/L và 3,5% số liệu $\text{PO}_4^{3-}\text{-P}$ lớn hơn 0,1 mg/L. Thông số DO có hơn 75% số liệu khảo sát vượt ngưỡng ($\leq 6,0$ mg/L), kết quả khảo sát DO dao động từ 2,0 - 7,0 mg/L, trung bình $4,3 \pm 0,8$ mg/L. Nguyên nhân hàm lượng oxy thấp là do nguồn nước bị ô nhiễm hữu cơ cục bộ từ các nguồn nông nghiệp, đô thị. Thông số TSS ghi nhận trung bình 39 ± 34 mg/L, có 67,3% kết quả khảo sát TSS vượt ngưỡng theo QCVN 08-MT:2015/BTNMT.

Bảng 3. Kết quả thống kê các thông số quan trắc

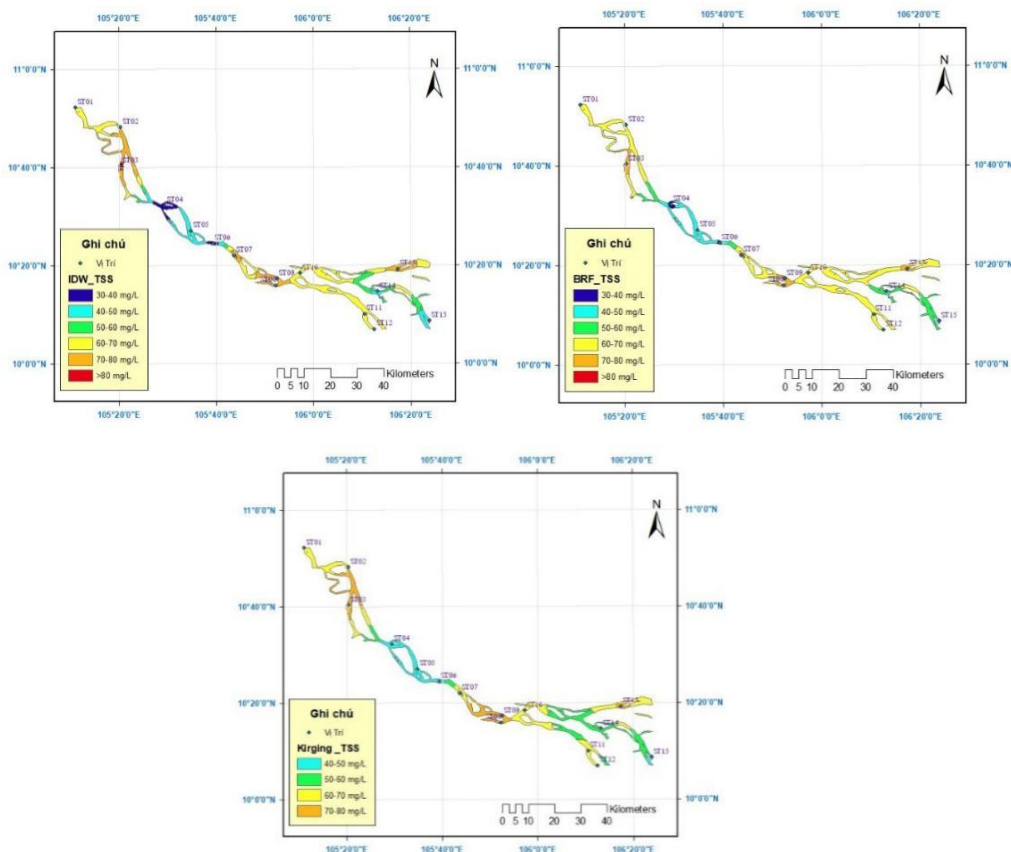
	T	pH	DO	EC	Alk	$\text{NO}_2\text{-N}$	TAN	$\text{PO}_4^{3-}\text{-P}$	TSS	COD	<i>Aeromonas</i>
Count	630	630	630	630	630	630	630	630	630	630	630
Mean	29,4	7,2	4,3	989	57	0,03	0,16	0,04	39	4,6	1.495
SD	1,5	0,4	0,8	3.380	14	0,04	0,32	0,07	34	3,5	3.638
Min	25,0	6,0	2,0	83	31	0,00	0,00	0,00	1	0,3	0
25%	28,4	7,0	4,0	172	47	0,01	0,03	0,02	18	2,7	150
50%	29,0	7,0	4,5	233	59	0,01	0,05	0,03P	29	3,8	420
75%	30,0	7,5	5,0	273	66	0,03	0,13	0,04	48	5,2	1.188
Max	36,0	8,5	7,0	29.500	181	0,28	3,11	1,04	309	23,7	30.000
CV (%)	5	5	19	342	25	133	200	175	88	75	243

Các thông số nhiệt độ, DO và độ kiềm có độ biến thiên ở mức thấp ($< 25\%$), thông số EC có độ biến thiên ở mức cao nhất 342% (Bảng 3). Nguyên nhân giá trị EC có CV lớn là do khu vực khảo sát chịu ảnh hưởng của xâm nhập mặn tại các điểm hạ lưu. Các thông số COD, TSS, $\text{NO}_2\text{-N}$, $\text{PO}_4^{3-}\text{-P}$, TAN, và *Aeromonas* sp. có độ biến thiên lớn, lần lượt từ 75%, 88%, 133%, 175%, 200% và 243%, hàm lượng các chất này thường bị ảnh hưởng của chế độ triều, dòng chảy, mùa và mức độ xả thải. Các vị trí từ ST7 đến ST9 thường xuyên ghi nhận giá trị $\text{NO}_2\text{-N}$ vượt ngưỡng, trong đó ST7 có số lần vượt ngưỡng cao nhất 33,3%, kế tiếp ST8 là 26,2% và ST9 là 23,8%. Ngoài ra, khu vực này cũng thường xuyên ghi nhận số lần vượt ngưỡng của các thông số TAN, TSS và *Aeromonas* sp. Cụ thể, thông số TAN ghi nhận tại các vị trí ST7, ST8 có số lần vượt ngưỡng tương đương nhau là 26,2%, trong khi ST9 có số lần vượt ngưỡng là 31,0%. Thông số TSS ghi nhận số lần vượt ngưỡng lần lượt tại các vị trí ST7, ST8, ST9 là 69%, 73,8% và 69%. Thông số *Aeromonas* sp. ghi nhận số lần vượt ngưỡng lần lượt là 57,1%, 66,7% và 59,5% tại các vị trí ST7, ST8, ST9. Khu vực này thường xuyên ghi nhận các thông số về ô nhiễm vượt ngưỡng giới hạn là do vị trí điểm ST8, ST9 thuộc huyện Châu Thành tỉnh Đồng Tháp, vị trí thu mẫu này gần với nhánh sông dẫn vào khu nuôi cá tra thâm canh lớn của huyện, thường xuyên tiếp nhận lượng nước thải ra từ các hoạt động nuôi trồng thủy sản khu vực. Ngoài ra, khu vực khảo sát từ ST7 đến ST9 đi qua khu vực dân cư và đô thị, tiếp nhận các nguồn hữu cơ từ hoạt động công nghiệp, sinh hoạt tại khu vực thành phố Sa Đéc, tỉnh Đồng Tháp. Các vị trí ST13 và ST15 ghi nhận thông số TSS có số lần vượt ngưỡng cao, lần lượt là 81,1% và 90,5%. Nguyên nhân khiến thông số TSS tại các điểm quan trắc này cao hơn các điểm còn lại phần lớn là do dòng chảy. Khu vực ST13 và ST15 nằm ở khu vực có nhiều cồn, cù lao trên sông.

Xem xét phân bố các chất ô nhiễm theo thời gian, các thông số NO_2^- -N, PO_4^{3-} -P, TAN số lần vượt ngưỡng trong mùa khô cao hơn mùa mưa. Cụ thể, thông số NO_2^- -N vượt ngưỡng 17,3% vào mùa khô cao hơn so với 14,5% trong mùa mưa. Tương tự, thông số PO_4^{3-} -P và TAN ghi nhận số lần vượt ngưỡng trong mùa khô lần lượt là 5,6% và 14,6% so với số lần vượt ngưỡng hai thông số này trong mùa mưa là 1,5% và 7,5%. Nguyên nhân là do vào mùa khô lưu lượng dòng chảy từ thượng nguồn giảm. Ngược lại, thông số TSS thường xuyên ghi nhận số lần vượt ngưỡng vào mùa mưa hơn so với mùa khô. Cụ thể, thông số TSS ghi nhận số lần vượt ngưỡng vào mùa mưa là 79,8% so với 55,2% vào mùa khô. Nguyên nhân là do mùa mưa lượng nước nhiều, tăng cường dòng chảy làm cho thông số TSS nhiều lần ghi nhận vượt ngưỡng. Trong khi đó, thông số *Aeromonas* sp. giữa mùa mưa và mùa khô không có nhiều sự khác biệt trên khu vực nghiên cứu.

Thông số TSS được lựa chọn cho việc xây dựng bản đồ do thường xuyên ghi nhận vượt ngưỡng giá trị cho phép (67,3%), chỉ số CV ở mức cao (88%) đồng thời chỉ số Willmott được xác định theo 3 phương pháp đều được đánh giá kết quả mức cao (lần lượt IDW, RBF, Kriging là $0,78 \pm 0,13$; $0,79 \pm 0,13$; $0,77 \pm 0,14$). Ngoài ra, tháng 10 được lựa chọn cho thời gian nội suy, vì đây là thời điểm Nam bộ chuyển mùa, từ mùa mưa sang mùa khô và khu vực nghiên cứu đang trải qua mùa nước nổi thường kỳ mỗi năm tại Đồng bằng sông cửu Long.

Kết quả nội suy theo không gian cho thấy có hai khu vực thường có nồng độ ô nhiễm cao là khu vực từ ST2 đến ST3 và khu vực từ ST7 đến ST9. Ô nhiễm tại các khu vực này có thể bắt nguồn từ các khu đô thị thuộc thành phố Hồng Ngự và thành phố Sa Đéc, tỉnh Đồng Tháp. Sự tiếp nhận nguồn thải hữu cơ từ các khu vực đô thị, công nghiệp ven sông Tiền đã có những ảnh hưởng nhất định đến chất lượng nước mặt trong suốt thời gian nghiên cứu. Ngoài ra, các khu vực ST3, ST8, ST9 là khu vực nuôi cá tra tập trung của huyện Thanh Bình, và huyện Châu Thành, tỉnh Đồng Tháp, thường xuyên tiếp nhận các nguồn ô nhiễm từ các hoạt động nuôi trồng thủy sản tại khu vực.



Hình 2. Kết quả nội suy TSS tháng 10 theo phương pháp nội suy IDW (a); RBF (b) và Kriging (c)

So sánh các phương pháp nội suy theo không gian cho thấy sự khác biệt về phân bố các chất ô nhiễm. Hình 2 minh họa sự khác biệt này dựa trên thông số TSS. Theo trật tự IDW, Kriging, RBF thì

khoảng giá trị của các thông số nội suy thường giảm dần, hay nói cách khác là mức độ làm trơn của chúng theo thứ tự này tăng dần. Sự khác biệt này cho thấy ba phương pháp có độ chính xác khác nhau và cần được đánh giá kỹ lưỡng để chọn lựa.

3.2. Đánh giá độ chính xác trung bình của các phương pháp nội suy

Trong mục này, độ chính xác của phương pháp nội suy được đánh giá trung bình dựa trên 5 trạm quan trắc được giữ lại, từ ST5 đến ST9, và 11 thông số môi trường. Việc xác định giá trị của d trong phương trình (4) được thực hiện cho từng thông số. Với 42 đợt khảo sát, ta có giá trị của n trong phương trình (4) là 42.

Từ kết quả tính toán chỉ số d , tiến hành phân tích phương sai Anova để đánh giá quan hệ chỉ số d giữa các phương pháp nội suy, với giả thuyết $H_0: d_{IDW} = d_{RBF} = d_{Kriging}$ và H_a : có ít nhất 1 chỉ số d khác biệt với các chỉ số còn lại. Kết quả phân tích thống kê cho kết quả p -value = 0,88, chấp nhận giả thuyết H_0 , điều này đồng nghĩa không có sự khác biệt có ý nghĩa thống kê giữa 3 phương pháp nội suy trên khu vực nghiên cứu.

Giá trị trung bình của chỉ số Willmott ứng với các phương pháp IDW, Kriging và RBF lần lượt là $0,72 \pm 0,18$, $0,71 \pm 0,17$ và $0,70 \pm 0,18$. Dựa trên các giá trị này có thể xem IDW là một lựa chọn ưu tiên để nội suy các thông số chất lượng nước trên khu vực nghiên cứu. Theo Liu và cộng sự (2013) và Emmanuel Eze (2020) chỉ số $d > 0,9$ kết quả đạt “xuất sắc”; từ $0,8 \leq d < 0,9$: kết quả “Tốt”, từ $(0,7 \leq d < 0,8)$: kết quả “trung bình” và $(d < 0,7)$ kết quả “kém” [13, 14]. Như vậy, cả 3 phương pháp nội suy đều cho kết quả ở mức trung bình từ $(0,7 \leq d < 0,8)$.

Cả ba phương pháp đều có thể sử dụng để nội suy không gian cho khu vực nghiên cứu ở mức trung bình. Kết quả nghiên cứu tương đồng với báo cáo của Nông Thị Hải Yến, Nguyễn Hải Hòa (2017) trong nghiên cứu ứng dụng công nghệ GIS kết hợp thuật toán nội suy không gian xây dựng bản đồ chất lượng nước tại khu vực khai thác khoáng sản huyện Hoà Bình, tỉnh Quảng Ninh. Nhóm tác giả so sánh số liệu nội suy không gian của 2 phương pháp IDW và Kriging với kết quả phân tích tại phòng thí nghiệm trên mẫu quan trắc được thu thập tại khu vực nghiên cứu. Nhóm tác giả đã chỉ ra cả hai phương pháp nội suy không gian IDW và Kriging đều cho kết quả sai số thấp và cả hai phương pháp đều có thể áp dụng tại khu vực nghiên cứu [9].

3.3. Đánh giá độ chính xác của các phương pháp nội suy theo vị trí quan trắc

Kiểm định giả thuyết cho từng vị trí có ảnh hưởng đến kết quả nội suy của 3 phương pháp nghiên cứu, phân tích phương sai Anova được sử dụng để đánh giá quan hệ chỉ số d giữa các vị trí từ ST5 đến ST9. Sau đó, tiến hành phân tích Post-hoc với kiểm định Turkey HSD kiểm định sự khác biệt chỉ số d giữa các vị trí.

Từ kết quả thống kê (Bảng 4), đối với phương pháp IDW và RBF giữa các vị trí không có sự khác biệt ý nghĩa thống kê ($p > 0,05$; $n = 11$). Trong khi đó, phương pháp Kriging cho thấy, vị trí ST5, ST6, ST7 không có sự khác biệt có ý nghĩa thống kê ($p > 0,05$; $n = 11$) và hai vị trí này có khác biệt ý nghĩa thống kê đối với vị trí ST8 và ST9 (Bảng 4).

Tại các vị trí từ ST5 đến ST7, kết quả nội suy ở mức kém ($< 0,7$) theo mức đánh giá chỉ số d của Liu và cộng sự (2013). Trong khi đó, chỉ số d tại vị trí ST8 của phương pháp IDW và RBF đạt mức trung bình (từ 0,7 - 0,8) và phương pháp Kriging nội suy đạt mức tốt (0,82). Đối với vị trí ST9, ngoại trừ RBF có mức đánh giá nội suy đạt mức trung bình (0,77) còn lại phương pháp IDW và Kriging được đánh giá đạt mức nội suy tốt (0,83 đối với IDW, 0,82 đối với Kriging).

Bảng 4. Kết quả thống kê chỉ số d theo vị trí

Vị trí	IDW	RBF	Kriging
ST5	0,65±0,20 ^a (0,24-0,95)	0,65±0,18 ^a (0,24-0,95)	0,63±0,18 ^a (0,29-0,96)
ST6	0,65±0,19 ^a (0,39-0,97)	0,64±0,18 ^a (0,40-0,97)	0,64±0,17 ^a (0,41-0,97)
ST7	0,68±0,15 ^a (0,46-0,96)	0,69±0,15 ^a (0,46-0,97)	0,63±0,15 ^a (0,47-0,97)
ST8	0,78±0,17 ^a (0,39-0,95)	0,73±0,21 ^a (0,20-0,94)	0,82±0,11 ^b (0,55-0,96)
ST9	0,83±0,12 ^a (0,57-0,95)	0,77±0,12 ^a (0,59-0,93)	0,82±0,11 ^b (0,56-0,96)

Ghi chú: Các giá trị trong cùng 1 cột có các chữ cái giống nhau thì không có khác biệt có ý nghĩa thống kê ($p > 0,05$). Các giá trị thể hiện là số trung bình và độ lệch chuẩn, giá trị thấp nhất và cao nhất.

Từ đó, có thể kết luận rằng khoảng cách giữa các vị trí được lựa chọn có ảnh hưởng đến kết quả nội suy. Cụ thể, khoảng cách từ ST4 đến ST5 là 13 km, từ ST5 đến ST6 và từ ST6 đến ST7 đều là 10 km, do đó khi lựa chọn 1 vị trí từ ST5 đến ST7 để nội suy thì vị trí đến trạm quan trắc gần nhất là 10 km đến 13 km. Trong khi đó, khi chọn nội suy tại ST8 hoặc ST9, khoảng cách từ điểm lựa chọn đến trạm gần nhất chỉ là 4 km. Do đó, khi lựa chọn vị trí ST8 hoặc ST9 cho phép nội suy thì vị trí từ điểm lựa chọn đến trạm gần nhất tương ứng 4 km. Kết quả này tương ứng với việc nội suy tại các vị trí từ ST5 đến ST7 cho kết quả kém, trong khi tại ST8 và ST9, kết quả nội suy đạt từ trung bình đến tốt.

3.4. Đánh giá độ chính xác của các phương pháp nội suy theo các thông số quan trắc

Xét theo các thông số quan trắc, hầu hết các thông số không có sự khác biệt có ý nghĩa thống kê trên khu vực nghiên cứu ($p > 0,05$; $n = 5$) ngoại trừ thông số độ kiềm có kết quả khác biệt và cao nhất trong cả 3 phương pháp, trong khi thông số pH có kết quả thấp nhất trong cả 3 phương pháp được lựa chọn.

Khi đánh giá chỉ số d theo từng thông số chất lượng nước, thông số độ kiềm cho kết quả nội suy cao nhất ở mức xuất sắc (d lớn hơn 0,9), thông số COD cho kết quả nội suy ở mức tốt (0,82 đối với IDW và 0,81 đối với Kriging), thông số nhiệt độ ($d_{IDW} = d_{Kriging} = 0,7$; $d_{RBF} = 0,71$), DO ($d_{IDW} = 0,73$; $d_{RBF} = 0,68$; $d_{Kriging} = 0,71$), EC ($d_{IDW} = 0,73$, $d_{RBF} = 0,66$; $d_{Kriging} = 0,68$), TSS ($d_{IDW} = 0,78$, $d_{RBF} = 0,79$; $d_{Kriging} = 0,77$) có kết quả nội suy ở mức trung bình còn các thông số còn lại như pH, TAN, NO_2^- -N, PO_4^{3-} -P và *Aeromonas* sp. có mức nội suy kém trong vùng nghiên cứu (Bảng 5). Nhìn chung, các thông số độ kiềm, nhiệt độ, DO có CV nhỏ hơn 30% cho kết quả nội suy ở mức tốt đến xuất sắc; thông số COD, TSS có CV lần lượt là 75% và 88%, với kết quả nội suy tốt và trung bình. Các thông số TAN, NO_2^- -N, PO_4^{3-} -P và *Aeromonas* sp. có kết quả nội suy kém và có CV cao hơn 100%. Thông số có hệ số biến thiên thấp hơn thường được nội suy chính xác hơn so với các thông số có hệ số biến thiên cao hơn, tương đồng với báo cáo của Ahmad và ctv (2021) [11]. Riêng thông số EC có CV ở mức 342%, với kết quả nội suy trung bình, do đặc điểm thủy văn của khu vực nghiên cứu, các vị trí ST13 đến ST15 chịu ảnh hưởng của xâm nhập mặn theo mùa.

Ngoài ra, các thông số ô nhiễm như TAN, NO_2^- -N, PO_4^{3-} -P, *Aeromonas* sp. có nguy cơ ô nhiễm cục bộ tại một số vị trí lấy mẫu trong thời gian khảo sát nhất định. Các phương pháp nội suy dựa trên ước lượng từ kết quả của các khu vực lân cận, do đó, kết quả nội suy không chính xác đối với các thông số ô nhiễm. Các thông số mang tính chất vật lý hoặc dễ lan truyền theo dòng chảy như nhiệt độ, TSS, độ kiềm, và EC có kết quả nội suy tốt hơn các thông số ô nhiễm với phạm vi lan truyền thấp hơn, chịu ảnh hưởng bởi khả năng hấp thụ của thực vật và khả năng tự làm sạch của sông.

Bảng 5. Kết quả thống kê chỉ số d theo thông số quan trắc

Thông số	IDW	RBF	Kriging
Alk	0,94±0,03 ^a (0,90-0,97)	0,95±0,02 ^a (0,92-0,97)	0,94±0,03 ^a (0,90-0,97)
COD	0,82±0,16 ^{ab} (0,52-0,95)	0,78±0,12 ^{ab} (0,55-0,79)	0,82±0,15 ^{ab} (0,56-0,96)
TSS	0,78±0,13 ^{ab} (0,63-0,94)	0,79±0,13 ^{bc} (0,64-0,94)	0,77±0,14 ^{ab} (0,63-0,94)
T	0,70±0,12 ^{ab} (0,55-0,85)	0,71±0,10 ^{bc} (0,59-0,83)	0,70±0,12 ^{ab} (0,55-0,83)
DO	0,73±0,04 ^{ab} (0,69-0,79)	0,68±0,05 ^{bc} (0,62-0,76)	0,71±0,03 ^{ab} (0,65-0,75)
EC	0,73±0,19 ^{ab} (0,39-0,95)	0,66±0,24 ^{bc} (0,20-0,87)	0,68±0,15 ^a (0,48-0,86)
NO_2^- -N	0,67±0,18 ^{ab} (0,42-0,86)	0,69±0,19 ^{bc} (0,41-0,86)	0,67±0,18 ^{ab} (0,40-0,85)
TAN	0,64±0,24 ^{ab} (0,24-0,88)	0,62±0,20 ^{bc} (0,29-0,83)	0,63±0,23 ^{ab} (0,29-0,88)
PO_4^{3-} -P	0,68±0,13 ^{ab} (0,50-0,81)	0,67±0,15 ^{bc} (0,46-0,83)	0,64±0,13 ^{ab} (0,48-0,79)
<i>Aeromonas</i>	0,63±0,20 ^{ab} (0,39-0,88)	0,62±0,14 ^{bc} (0,42-0,81)	0,66±0,18 ^{ab} (0,46-0,88)
pH	0,55±0,12 ^b (0,41-0,70)	0,51±0,07 ^c (0,40-0,59)	0,56±0,13 ^b (0,41-0,72)

Các giá trị trong cùng 1 cột có các chữ cái giống nhau thì không có khác biệt có ý nghĩa thống kê ($P > 0,05$). Các giá trị thể hiện là số trung bình và độ lệch chuẩn, giá trị thấp nhất và cao nhất.

4. KẾT LUẬN

Từ kết quả khảo sát trên vùng nghiên cứu sông Tiền, việc sử dụng các phương pháp nội suy IDW, RBF và Kriging cho 11 thông số quan trắc phục vụ nuôi cá tra thâm canh cho thấy không có sự khác biệt có ý nghĩa giữa ba phương pháp nội suy không gian. Giá trị trung bình của chỉ số Willmott theo các phương pháp IDW, Kriging, và RBF lần lượt là $0,72 \pm 0,18$; $0,71 \pm 0,17$, và $0,70 \pm 0,18$. Hay phương pháp IDW có thể được xem là một lựa chọn ưu tiên để nội suy các thông số chất lượng nước trên khu vực nghiên cứu.

Các thông số có hệ số biến thiên thấp hơn thì kết quả nội suy chính xác hơn. Thông số EC có hệ số biến thiên lớn và kết quả nội suy ở mức trung bình do ảnh hưởng của thủy triều. Ngoài ra, các thông số ô nhiễm như TAN, $\text{NO}_2\text{-N}$, $\text{PO}_4^{3-}\text{-P}$, và *Aeromonas* sp. có kết quả nội suy thấp hơn các thông số còn lại, do mức độ ảnh hưởng cục bộ của vị trí thu mẫu, phạm vi lan truyền và khả năng tự làm sạch của sông.

Khoảng cách giữa 2 vị trí có ảnh hưởng đến kết quả nội suy. Ngoài ra, độ chính xác nội suy phụ thuộc vào từng thông số chất lượng nước. Độ kiểm có kết quả nội suy tốt nhất trong khi các chỉ tiêu pH, TAN, $\text{NO}_2\text{-N}$, $\text{PO}_4^{3-}\text{-P}$ và *Aeromonas* sp. có kết quả nội suy kém hơn các thông số còn lại.

Từ kết quả nghiên cứu, có thể áp dụng thuật toán nội suy IDW cho việc xây dựng bản đồ chất lượng nước khu vực sông Tiền phục vụ công tác quản lý và phát triển cá tra thâm canh. Hạn chế của nghiên cứu này là chưa xác định được khoảng cách tối ưu cho các phương pháp nội suy, cần có hướng nghiên cứu tiếp theo trong việc xác định khoảng cách thích hợp giữa 2 điểm quan trắc để cho kết quả nội suy được tối ưu đồng thời đảm bảo thực tế về kinh phí và nguồn lực có hạn.

TÀI LIỆU THAM KHẢO

1. Tổng cục Thống kê - Các báo cáo phân tích và dự báo thống kê năm 2020, báo cáo Phân tích nhân tố ảnh hưởng đến kết quả sản xuất cá tra, tôm sú, tôm thẻ chân trắng và dự báo sản xuất đến năm 2025, Hà Nội (2021) 393-398.
2. Lê Hồng Phước - Quan trắc, cảnh báo và giám sát môi trường vùng nuôi trồng thủy sản năm 2022, Báo cáo tổng kết đề tài, Viện Nghiên cứu Nuôi trồng Thủy sản II, TP. HCM (2023).
3. Bộ Tài nguyên và Môi trường - Báo cáo hiện trạng môi trường quốc gia năm 2018, chuyên đề: Môi trường nước các lưu vực sông, Hà Nội (2018), tr. 66.
4. Lê Hồng Phước - Quan trắc, cảnh báo và giám sát môi trường vùng nuôi trồng thủy sản năm 2021, Báo cáo tổng kết nhiệm vụ, Viện Nghiên cứu Nuôi trồng Thủy sản II, TP. Hồ Chí Minh (2022).
5. Kimleang C., Ty S., Ratboren C., Bunmanut H., Sokly S., Heejun Y., Chihiro Y., Layheang S., Chantha O. - Assessment of spatial interpolation methods to map water quality in Tonle Sap Lake, The 2nd International Symposium on Conservation and Management of Tropical Lakes (2017).
6. Nagalakshmi., Prasanna K., and Prakash C.S. - Water quality analysis using GIS interpolation method in serthalaikadu Lagoon, east coast of India, Rasayan Journal of Chemistry **9** (2016) 634-640.
7. Ahmad A.Y., Saleh I.A, Balakrishnan P., Al-Ghouti, M.A. - Comparison GIS-based interpolation methods for mapping groundwater quality in the state of Qatar, Groundwater for Sustainable Development **13** (2021) 100573. <https://doi.org/10.1016/j.gsd.2021.100573>
8. Nguyễn Thị Thu Hiền, Nguyễn Hải Nam, Nguyễn Hải Hòa, Nguyễn Thị Khanh - 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 **189** (2018) 39-43.
9. Nông Thị Hải Yến, Nguyễn Hải Hòa - 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 Hoàn Bồ, tỉnh Quảng Ninh, Tạp chí khoa học & Công nghệ Đại học Thái Nguyên **169** (09) (2017) 75-80.

10. Phan Quoc Yen, Nguyen Thi Thu Nga, and Tong Thi Hanh - Research and experimental comparison of topographic modeling methods, *VNU Journal of Science: Earth and Environmental Sciences* **35** (2019). <https://doi.org/10.25073/2588-1094/vnuées.4445>
11. Xie Y., Chen T.B., Lei M., Yang J., Guo Q.J., Song B., Zhou X.Y. - Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: Accuracy and uncertainty analysis, *Chemosphere* **82** (2011) 468-476. <https://doi.org/10.1016/j.chemosphere.2010.09.053>
12. Willmott C. J., Wicks D. E. - An empirical method for the spatial interpolation of monthly precipitation within california, *Physical Geography* **1** (1980) 59-73. <https://doi.org/10.1080/02723646.1980.10642189>
13. Bộ Tài nguyên và Môi trường - QCVN 08-MT:2015/BTNMT, Quy chuẩn kỹ thuật quốc gia về chất lượng nước mặt (2015).
14. Liu S., Yang J. Y., Zhang X. Y., Drury C. F., Reynolds W. D., Hoogenboom G. - Modelling crop yield, soil water content and soil temperature for a soybean-maize rotation under conventional and conservation tillage systems in Northeast China, *Agricultural Water Management* **123** (2013) 32-44. <https://doi.org/10.1016/j.agwat.2013.03.001>
15. Eze E., Girma A., Zenebe A., Kourouma J. M., and Zenebe G. - Exploring the possibilities of remote yield estimation using crop water requirements for area yield index insurance in a data-scarce dryland, *Journal of Arid Environments* **183** (2020) 104261. <https://doi.org/10.1016/j.jaridenv.2020.104261>.

ABSTRACT

COMPARISON OF SPATIAL INTERPOLATION METHODS IN CONSTRUCTING WATER QUALITY MAPS OF TIEN RIVER FOR INTENSIVE CATFISH CULTURE

Tran Minh Thien^{1,2*}, Nguyen Thanh Truc²

¹*Industrial University of Ho Chi Minh City*

²*Research Institute for Aquaculture No.2*

*Email: tranminhthien85@gmail.com

The application of GIS in water quality mapping has been commonly applied in recent years. However, there is no specific assessment for the application of the appropriate interpolation algorithm for the construction of water quality maps in Tien River belonging to the Mekong Delta. This paper compared and evaluated the usability of three common spatial interpolation methods, namely IDW, RBF and Kriging, on 11 water quality parameters, including temperature, pH, conductivity, alkalinity, dissolved oxygen, total suspended solid, COD, nitrite, TAN, phosphate and total *Aeromonas* bacteria at 15 monitoring points with 42 sampling periods from 2020 to 2021. Willmott's index of agreement was used for the assessment of results. The results show that there was no statistically significant difference between the two spatial interpolation methods on the study area ($p > 0.05$). The location, water quality parameters affected the results of three spatial interpolation methods. The average results of three spatial interpolation methods (IDW, RBF and Kriging) were 0.72 ± 0.18 , 0.70 ± 0.18 and 0.71 ± 0.17 , respectively. IDW can be considered as a preferred choice to interpolate water quality parameters on the study area.

Keywords: IDW, RBF, Kriging, GIS, Water quality map.