

XÂY DỰNG CHỈ SỐ CHẤT LƯỢNG NƯỚC DƯỚI ĐẤT (GWQI) VÀ ÁP DỤNG CHO HUYỆN HẢI LĂNG, TỈNH QUẢNG TRỊ

Trương Trung Kiên¹, Mai Thị Thanh Tuyền¹, Trương Quý Tùng², Nguyễn Văn Hợp^{3*}

¹ Sở Tài nguyên và Môi trường tỉnh Quảng Trị, Việt Nam

² Khoa Môi trường, Trường Đại học Khoa học Huế, Việt Nam

³ Khoa Hóa, Trường Đại học Khoa học Huế, Việt Nam

* Tác giả liên hệ Nguyễn Văn Hợp <ngvanhopkh@gmail.com>

(Ngày nhận bài: 25-12-2023; Hoàn thành phản biện: 29-02-2024; Ngày chấp nhận đăng: 29-02-2024)

Tóm tắt. Trong nghiên cứu này, Chỉ số chất lượng nước dưới đất (GWQI) được xây dựng và áp dụng cho huyện Hải Lăng, tỉnh Quảng Trị. Mười thông số chất lượng nước (CLN) cơ bản được lựa chọn để xây dựng GWQI bao gồm: pH, TDS, độ cứng (HARD); SO₄, COD, N-NH₄, N-NO₃, Fe, Mn và tổng coliform (TC). Phương pháp phân tích thành phần chính (PCA) được áp dụng cho tập dữ liệu CLN ở huyện Hải Lăng giai đoạn 2015-2018 để xác định trọng số (w_i) của các thông số CLN. Các hàm chỉ số phụ (q_i) tuyến tính được thiết lập dựa trên giới hạn cho phép của các thông số được quy định trong Quy chuẩn kỹ thuật quốc gia về chất lượng nước sử dụng cho mục đích sinh hoạt (QCVN 01-1:2018/BYT) do Bộ Y tế ban hành. Ba (3) phương pháp khác nhau được sử dụng để tính các chỉ số GWQI_M (phương pháp tích), GWQI_A (phương pháp tổng) và GWQI_{Ref} (phương pháp khác) cho tập dữ liệu CLN giai đoạn 2015-2018. Các chỉ số đó đều là hàm số của chỉ số phụ q_i và trọng số w_i . Kết quả cho thấy, chỉ số GWQI_M phù hợp hơn GWQI_A và GWQI_{Ref} trong phản ánh CLN dưới đất. Áp dụng chỉ số GWQI_M cho tập dữ liệu CLN giai đoạn 2019-2021 ở huyện Hải Lăng cho thấy, có 89% giá trị GWQI_M thuộc loại CLN Rất tốt hoặc Tốt, 7% thuộc loại Trung bình, 1% thuộc loại Kém và 3% thuộc loại Rất kém (hay Không phù hợp cho ăn uống).

Từ khoá: GWQI, PCA, huyện Hải Lăng

Establishment of undergroundwater quality index (GWQI) and application to Hai Lang district, Quang Tri province

Trương Trung Kiên¹, Mai Thị Thanh Tuyền¹, Trương Quý Tùng², Nguyễn Văn Hợp^{3*}

¹ Department of Natural Resources and Environment of Quang Tri Province, Vietnam

² Department of Environment, University of Sciences, Hue University, Vietnam

³ Department of Chemistry, University of Sciences, Hue University, Vietnam

* Correspondence to Nguyen Van Hop <ngvanhopkh@gmail.com>

(Received: 25 December 2023; Revised: 29 February 2024; Accepted: 29 February 2024)

Abstract. In this study a Ground Water Quality Index (GWQI) was established and applied to Hai Lang district in Quang Tri province. Ten basic water quality parameters selected for establishment of the GWQI consisted of: pH, TDS, hardness (HARD); SO₄, COD, N-NH₄, N-NO₃, Fe, Mn and Total coliform (TC). Principal component analysis (PCA) was applied to the monitoring data set of the period 2015-2018 in the district to determine the weights (w_i) of the parameters. The linear sub-index

functions (q_i) were set up basing on the allowable limits of the parameters specified in the National Technical Regulation on Domestic Water Quality (QCVN 01-1:2018/BYT) issued by the Ministry of Health. Three (3) different methods were used to calculate index $GWQI_M$ (multiply form), $GWQI_A$ (additive form) and $GWQI_{Ref}$ (reference form) for the water quality data set. The indexes were all functions of sub-index q_i and weights w_i . The obtained results show that the index $GWQI_M$ is more suitable than $GWQI_A$ and $GWQI_{Ref}$ in reflecting the groundwater quality. Applying the $GWQI_M$ to the water quality data set of the period 2019 - 2021 in the district indicated that there are 89% of $GWQI_M$ values classified as Very good or Good water quality, 7% as Medium 1% as Poor and 3% as Very poor (or Unfit for drinking).

Keywords: GWQI, PCA, Hai Lang District

1 Mở đầu

Một trong những nhiệm vụ quan trọng trong quản lý nguồn nước là quan trắc và đánh giá chất lượng nước (CLN). Để đánh giá CLN, có thể dựa vào các thông số CLN riêng biệt, tức là so sánh kết quả quan trắc thông số với giá trị giới hạn được quy định trong quy chuẩn kỹ thuật quốc gia về CLN, hoặc dựa vào Chỉ số CLN. Đánh giá CLN mặt dựa vào Chỉ số chất lượng nước (WQI/water quality index) và đánh giá CLN dưới đất dựa vào Chỉ số chất lượng nước dưới đất (GWQI/groundwater quality index) có nhiều ưu điểm hơn, do chúng cho phép đánh giá CLN tổng quát (tức là dựa vào tổ hợp nhiều thông số CLN), cho phép sánh CLN theo không gian và thời gian, giúp cộng đồng và các nhà hoạch định chính sách dễ hiểu hơn [1], [2]. Chỉ số WQI/GWQI là một giá trị nguyên không có đơn vị, được tính toán theo một công thức toán học xác định từ các kết quả quan trắc được của các thông số CLN lựa chọn [1], [3], [4]. Tổng quát, chỉ số WQI/GWQI được tính toán từ 2 đại lượng: *Chỉ số thông số* hay *Chỉ số phụ* (sub-index) q_i (thể hiện chất lượng của thông số CLN i) và *Trọng số* (weightage) w_i với là thông số CLN lựa chọn ($i = 1 - n$). Chỉ số thông số q_i không có đơn vị, thể hiện chất lượng của thông số CLN i , thường nhận giá trị trong khoảng 0 - 100 và được xác định bằng cách chuyển đổi kết quả quan trắc thông số i (có đơn vị đo khác nhau) thành thang đo chuẩn hóa 0 - 100; q_i càng gần 100, chất lượng của thông số càng tốt và ngược lại. Trọng số w_i thể hiện tầm quan trọng tương đối

của thông số i và thông thường tổng các trọng số $\sum w_i = 1$ [1], [2]. Đến nay, đã có nhiều nghiên cứu xây dựng GWQI trên thế giới [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15]. Trong đó, để xác định các trọng số w_i , một số tác giả dựa vào ý kiến cá nhân của các chuyên gia. Nhiều tác giả cho rằng, để xác định một cách khách quan w_i , cần áp dụng phương pháp phân tích thống kê đa biến như phương pháp phân tích thành phần chính (PCA/principal component analysis) và một dạng tương tự là phân tích yếu tố (FA/factor analysis) như Narsimha A., Hui Qian [9], Iuliana Paun [16], N. R. Ekere [11], José [14], H. M. Joung [17], M. Tripathi and S. K. Singal [18], M. Vega [19], Y. Qian [20], C. Coletti [21]. Ở Việt Nam, Bộ Tài nguyên và Môi trường đã ban hành chỉ số WQI để áp dụng đánh giá chất lượng nước của các nguồn nước mặt (từ năm 2011 và điều chỉnh năm 2019), nhưng chưa nghiên cứu xây dựng chỉ số GWQI. Nguyễn Hải Âu [22] đã áp dụng chỉ số GWQI do N. Subba Rao đề xuất năm 2019 để đánh giá CLN dưới đất tầng Pleistocen tại xã Phú Mỹ, tỉnh Bà Rịa Vũng Tàu: lựa chọn 10 thông số CLN, xác định trọng số entropy. Nói chung, các nghiên cứu về GWQI ở nước ta còn rất hạn chế.

Bài báo này đề cập đến các kết quả nghiên cứu xây dựng chỉ số GWQI và áp dụng để đánh giá CLN ở huyện Hải Lăng, tỉnh Quảng Trị, trong đó áp dụng phương pháp PCA để xác định trọng số w_i của các thông số CLN lựa chọn.

2 Phương pháp nghiên cứu

2.1 Khu vực nghiên cứu

Hải Lăng là huyện ven biển phía Nam của tỉnh Quảng Trị, có tọa độ địa lý từ 16°33'40" đến 16°48'00" độ vĩ Bắc và 107°04'10" đến 107°23'30" độ kinh Đông; tổng diện tích đất tự nhiên 21.356,5 ha; bao gồm 16 đơn vị hành chính (15 xã và 01 thị trấn – thị trấn Diên Sanh).

Ở huyện Hải Lăng, tầng chứa nước Holocen phân bố rộng khắp, chỉ vắng mặt ở khu vực phía Tây nơi diện lộ của tầng chứa nước Pleistocen, các đá trầm tích biển chất và xâm nhập. Chúng bao gồm các thành tạo trầm tích đa nguồn gốc (a, am, amb, m, mv) Q_2^3 , aQ_2^{2-3} , (am, m) Q_2^2 , (a, am) Q_2^{1-2} . Thành phần thạch học của chúng khá đa dạng phụ thuộc vào nguồn gốc với thành phần từ hạt thô đến hạt mịn bao gồm cuội, sỏi, cát, bột, sét, vật chất hữu cơ. Tổng chiều dày chung của tầng của khu vực biến đổi từ 9,64m - 51m, trung bình 20m.- 30m. Trên cơ sở kết quả hút nước thí nghiệm trong tầng chứa nước Holocen tại các lỗ khoan trên khắp đồng bằng ven biển tỉnh Quảng Trị cho thấy, mức độ chứa nước thay đổi từ nghèo đến giàu. Về tính chất thủy lực nước thuộc loại không áp. Mực nước tĩnh thay đổi từ 0,0m – 1,85m, phụ thuộc vào bề mặt địa hình và điều kiện thế nằm. Động thái của tầng chứa nước biến đổi theo mùa, nhưng biên độ dao động không lớn. Mùa mưa nước trong các cồn cát hầu như tràn trên mặt đất, mùa khô mực nước nằm cách mặt đất khoảng từ 0,5m đến 1,85m [23].

Trong chương trình quan trắc môi trường hàng năm do Trung tâm Tài nguyên và Môi trường (TNMT) Quảng Trị thực hiện, từ năm 2015 đến nay, người ta quan trắc nước dưới đất ở 16 đến 36 vị trí ở huyện Hải Lăng với tần suất 2 đợt/năm (01 đợt mùa khô, tháng 1-8 và 01 đợt mùa mưa, tháng 9-12). Số vị trí quan trắc (để trong ngoặc đơn) ở các xã, thị trấn như sau: Diên Sanh (8), Hải Ba (5), Hải Quế (4), Hải Phong (6),

Hải Chánh (4), Hải Sơn (5), Hải Khê (2), Hải Định (2).

Tập dữ liệu 1 là dữ liệu CLN tại 16 vị trí quan trắc (hay 16 giếng khoan hoặc đào) trong giai đoạn 2015-2018 với 62 mẫu. Tập dữ liệu 2 là dữ liệu CLN tại 36 vị trí quan trắc trong giai đoạn 2019-2021 với 101 mẫu. Từ 02 tập dữ liệu đó, 10 thông số được lựa chọn để xây dựng chỉ số GWQI và đánh giá CLN dưới đất gồm: pH, tổng chất rắn hòa tan (TDS) (độ cứng (HARD), nhu cầu oxy hóa học (COD), sunfat (SO_4), amoni tính theo nitơ (N-NH₄), nitrat tính theo nitơ (N-NO₃), tổng sắt tan (Fe), tổng mangan tan (Mn) và tổng coliform (TC).

2.2 Quy trình phát triển chỉ số GWQI

Tổng quát, quy trình xây dựng chỉ số WQI hoặc GWQI gồm 4 bước: (i) Lựa chọn các thông số CLN; (ii) Tính *Chỉ số phụ* thông qua việc chuyển nồng độ quan trắc được của mỗi thông số thành giá trị không có đơn vị theo thang chất lượng chuẩn hóa; (iii) Xác định *Trọng số* của mỗi thông số; (iv) Tổ hợp các chỉ số phụ và trọng số thành giá trị WQI hoặc GWQI [1], [2], [14], [24]. Cuối cùng, đánh giá chất lượng nước dựa vào thang điểm phân loại WQI hoặc GWQI.

Lựa chọn thông số

Cho đến nay chưa có một phương pháp chung nào được thừa nhận rộng rãi để lựa chọn các thông số trong xây dựng chỉ số GWQI [1], [2], [14]. Số thông số lựa chọn (n) để xây dựng chỉ số GWQI thường khác nhau, tùy thuộc vào đặc điểm nguồn nước, khoảng 9 – 13 thông số và trong đó chủ yếu là các thông số CLN cơ bản [1], [5], [7], [8], [9], [14]. Trong nghiên cứu này, lựa chọn 10 thông số như nêu ở mục 2.1. Trong đó, các thông số pH, TDS, SO_4 thể hiện đặc trưng của nguồn nước; thông số COD thể hiện mức ô nhiễm các chất hữu cơ; thông số N-NH₄ và N-NO₃ thể hiện mức ô nhiễm các chất dinh dưỡng; Fe và Mn thể hiện mức nhiễm kim loại nặng; TC thể hiện mức

nhiễm các vi khuẩn có nguồn gốc từ phân người và động vật. Các thông số này cũng được quy định trong QCVN 09-MT:2015/BTNMT về CLN dưới đất cho đa mục đích sử dụng và QCVN 01-1:2018/BYT về CLN sạch sử dụng cho mục đích sinh hoạt. Các kim loại nặng và độc (Hg, Cd, Pb, As, Ni, Cr, Cu, Zn) cũng như các hóa chất bảo vệ thực vật nhóm clo (nhóm DDT, HCH) đều có hàm lượng rất nhỏ, hoặc nhỏ hơn giới hạn phát hiện của phương pháp phân tích hoặc nhỏ hơn nhiều so với giới hạn cho phép trong QCVN 09-MT:2015/BTNMT (quy định về CLN dưới đất), nên chúng không được đưa vào để xây dựng chỉ số GWQI.

Phương pháp xác định chỉ số phụ

Bước này nhằm mục đích chuyển đổi nồng độ của thông số CLN thành thang đo chuẩn hóa không có đơn vị, được gọi là *Chỉ số phụ* q_i . Trong các nghiên cứu phát triển chỉ số GWQI, nhiều tác giả xác định giá trị q_i theo công thức (1) [1], [8], [25], [26], [27]:

$$q_i = \frac{C_i}{C_{oi}} \times 100 \quad (1)$$

trong đó, C_i và C_{oi} tương ứng là nồng độ của mỗi thông số trong mẫu (mg/L) và giới hạn được quy định trong hướng dẫn CLN của quốc gia hoặc WHO đối với mỗi thông số. Theo công thức (1), đối với mỗi thông số, q_i là hàm tuyến tính của C_i ; q_i càng tăng, chất lượng của thông số càng giảm và ngược lại; khi $C_i > C_{oi}$, giá trị của $q_i > 100$.

Việc sử dụng hướng dẫn CLN để hỗ trợ chia nhỏ các giá trị chỉ số phụ q_i sẽ cung cấp thêm thông tin và thuận tiện cho người dùng [28]. Mặt khác, nếu sử dụng thang đo q_i trong khoảng 0 – 100 sẽ dễ hiểu hơn so với thang đo q_i theo công thức (1). Hiện nay ở nước ta người dân thường sử dụng nước giếng khoan, giếng đào cho mục đích sinh hoạt và do vậy, cần xây dựng chỉ số GWQI để đánh giá CLN cấp cho sinh hoạt, tức là phải đáp ứng QCVN 01-1:2018/BYT (viết tắt là

QCVN01-1) [29]. Trong nghiên cứu này, từ QCVN01-1, thiết lập các hàm tuyến tính để tính q_i (có giá trị 1 – 100, q_i càng gần 100, chất lượng thông số càng tốt và ngược lại). Phương trình tuyến tính có dạng $y = a + bx$, trong đó, $y = q_i$ được tính toán từ nồng độ quan trắc được của thông số CLN (x); a và b là hai hệ số được suy ra từ hai phương trình tuyến tính (2) và (3):

$$100 = a + b \times (\text{giới hạn cho phép của QCVN 01-1 đối với thông số}) \quad (2)$$

trong đó, $y = q_i = 100$ ứng với chất lượng tốt của thông số; Khi kết quả quan trắc thông số (x) \leq giới hạn cho phép của QCVN 01-1 [29], giá trị q_i của nó = 100;

$$1 = a + b \times (\text{giới hạn cho phép của QCVN 01-1 đối với thông số} + 3 \times S) \quad (3)$$

trong đó, $y = q_i = 1$ ứng với chất lượng kém của thông số; Khi kết quả quan trắc của thông số (x) \geq giới hạn cho phép + $3 \times S$, giá trị q_i của nó = 1. Điều này được hiểu là: Giới hạn cho phép đối với mỗi thông số là trung bình của một phân bố chuẩn (hay một quần thể); Chỉ những giá trị x nằm ngoài khoảng tin cậy 99,87% của giá trị trung bình đó mới được xem là phân bố khác (hay quần thể khác), tức là lớn hơn giới hạn cho phép; Số 3 là giá trị của biến chuẩn hóa (u) ứng với xác suất tin cậy $P = 0,9987$; S là độ lệch chuẩn của quần thể [30].

Độ lệch chuẩn S có thể tính toán từ Độ lệch chuẩn tương đối (RSD, %):

$$S = \frac{\text{RSD} \times \text{Trung bình số học}}{100} \quad (4)$$

trong đó, Trung bình số học chính là giới hạn cho phép của thông số CLN trong QCVN 01-1; RSD được xác định dựa vào phương trình Horwitz [31]:

$$\text{RSD} = 2^{1 - 0,5 \lg C}, \quad (5)$$

với C là nồng độ được biểu diễn bằng phân số.

Phương pháp xác định trọng số

Trọng số (w_i) thể hiện tầm quan trọng tương đối của thông số CLN và chúng có thể bằng nhau hoặc khác nhau [1], [2]. Một số chỉ số GWQI chấp nhận trọng số của các thông số CLN bằng nhau ($w_i = 1$) [5], [7]. Nhiều chỉ số GWQI sử dụng trọng số w_i khác nhau và được tính theo công thức (6) và (7) [1], [6], [8], [9], [11], [25], [27]:

$$w_{i^*} = \frac{K}{C_{oi}} \quad (6)$$

trong đó, hệ số $K = 1/\sum(1/C_{oi})$; C_{oi} là giới hạn của thông số được quy định trong hướng dẫn CLN của quốc gia hoặc WHO; w_{i^*} là trọng số của mỗi thông số CLN được xác định theo ý kiến chuyên gia, tùy thuộc vào tầm quan trọng của thông số trong đánh giá CLN; các giá trị $w_{i^*} = 1 - 5$ và w_{i^*} càng lớn, thông số càng quan trọng và ngược lại. Từ các giá trị w_{i^*} xác định được, tính toán trọng số w_i theo công thức (7):

$$w_i = \frac{w_{i^*}}{\sum_{i=1}^n w_{i^*}} \quad (7)$$

Trong nghiên cứu này, phương pháp PCA được sử dụng để xác định trọng số của các thông số CLN. Phương pháp PCA là phương pháp chiếu, cho phép giảm dữ liệu từ không gian m chiều (mỗi chiều ứng với một biến/thông số CLN) thành không gian ít chiều hơn: k chiều ($k < m$); mỗi chiều được gọi là một thành phần chính (PC/principal component) – là tổ hợp tuyến tính của các biến gốc (x_{ij}) và các PC không tương quan với nhau, mà trực giao nhau. Thông thường, chỉ cần giữ lại một vài PC đầu tiên, chẳng hạn PC1, PC2 vì chúng vẫn giải thích được đa số biến động hay phương sai của tập dữ liệu gốc. Những PC được giữ lại là những PC có *Giá trị riêng* lớn hơn 1 (giá trị riêng là lượng phương sai của các biến gốc được giải thích bởi PC tương ứng) và phương sai tích lũy của chúng lớn hơn 60%. Mặt khác, từ kết quả tính toán theo PCA, sẽ xác định được *Tải lượng bình phương* (squared loading) của mỗi biến trên mỗi PC. Tải lượng bình phương của mỗi biến chính là lượng phương sai của nó được giải thích

bởi mỗi PC [32], [33]. Từ *Tổng tải lượng bình phương* của mỗi biến trên các PC giữ lại (communality) và *Tổng tải lượng bình phương* của tất cả các biến trên các PC giữ lại (được gọi là *Tổng tích lũy*), sẽ tính được Trọng số (hay tầm quan trọng tương đối) của mỗi biến theo công thức (8). Để thực hiện các tính toán theo phương pháp PCA, nghiên cứu này sử dụng phần mềm R – phần mềm được áp dụng phổ biến trên thế giới: R version 4.0.3/64-bit (10-10-2020) với module R-Studio, dùng gói (package) Factoextra (version 1.0.7).

$$\text{Trọng số } (w_i) = \frac{\text{Tổng tải lượng bình phương của thông số } i \text{ trên các PC giữ lại}}{\text{Tổng tải lượng bình phương của tất cả các thông số trên các PC giữ lại}} \quad (8)$$

Phương pháp tính chỉ số GWQI và thang điểm phân loại

Để chọn công thức tính phù hợp, tiến hành tính chỉ số GWQI cuối cùng cho Tập dữ liệu 1 theo phương pháp tích – công thức (9) và phương pháp tổng – công thức (10):

$$GWQI_M = \prod_{i=1}^n q_i^{w_i} \quad (9)$$

$$GWQI_A = \sum_{i=1}^n q_i \times w_i \quad (10)$$

Để thuận lợi cho người dùng chỉ số GWQI, nghiên cứu này sử dụng thang điểm phân loại CLN dưới đất dựa vào GWQI tương tự như thang điểm phân loại CLN mặt do Tổng cục Môi trường quy định [34]: GWQI càng lớn, CLN dưới đất càng tốt và ngược lại (Bảng 1).

Bảng 1. Phân loại chất lượng nước dưới đất theo chỉ số GWQI

Loại/Mức	Chỉ số GWQI	Đánh giá chất lượng
I	91 – 100	Rất tốt (RT)
II	76 – 90	Tốt (T)
III	51 – 75	Trung bình (TB)
IV	26 – 50	Kém (K)
V	1 - 25	Rất kém (RK/ Không phù hợp cho ăn uống)

Mặt khác, để so sánh, ở đây cũng tính toán giá trị GWQI cuối cùng theo phương pháp khác - công thức (11) - đã được một số tác giả đã đề xuất như B. A. Elubid [6], M. Prasad [8], N. Adimalla and H. Qian [9], N. R. Ekere [11], M. Vasanthavigar [25], G. S. Solangi [27]:

$$GWQI_{Ref} = \sum_{i=1}^n q_i \times \frac{W_i^*}{\sum_{i=1}^n W_i^*} \quad (11)$$

trong đó, K và wi như ở (6) và (7), Chỉ số phụ qi được tính theo công thức (12):

$$q_i = \frac{100 \times (C_i - C_o)}{S_i - C_o} \quad (12)$$

Ở đây, Ci và Co tương ứng là nồng độ của thông số CLN trong mẫu và nồng độ lý tưởng của thông số trong nước tinh khiết = 0 (ngoại trừ pH có Co = 7,0).

Theo công thức (11), GWQI được phân chia thành 5 mức CLN: < 50 (Rất tốt); 51 – 100 (Tốt); 101 – 200 (Kém); 201 – 300 (Rất kém) và > 300 (Không phù hợp cho ăn uống).

3 Kết quả và thảo luận

3.1 Chỉ số phụ của các thông số chất lượng nước

Bảng 2. Tổng hợp các đại lượng thống kê mô tả cho 10 thông số được chọn của Tập dữ liệu 1 (n=62)

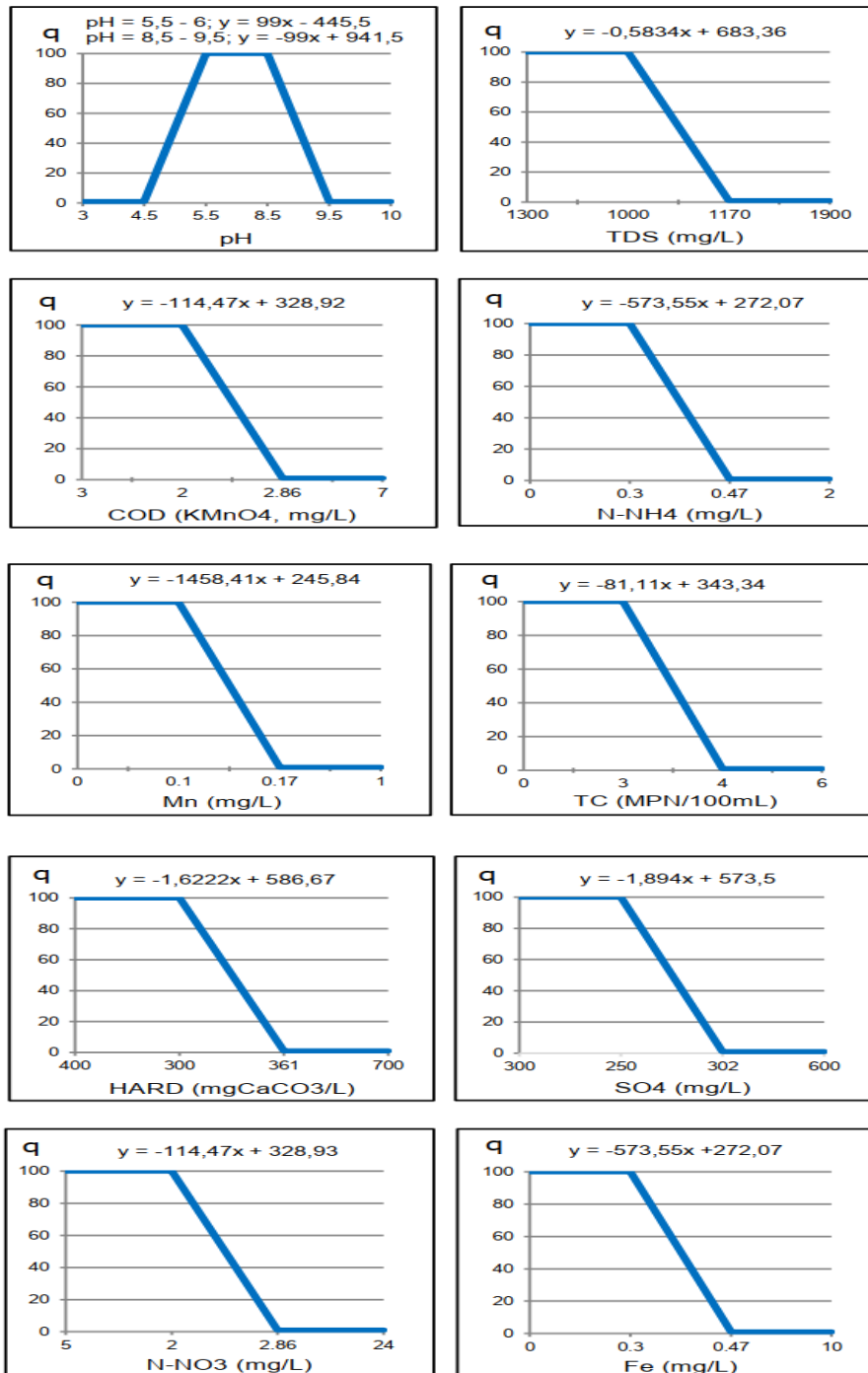
Đại lượng thống kê (*)	pH	TDS (mg/L)	HARD (mg/L)	SO4 (mg/L)	COD (mg/L)	N-NH4 (mg/L)	N-NO3 (mg/L)	Fe (mg/L)	Mn (mg/L)	TC (MPN/100mL)
min	5.1	43	10	9.0	0.5	0.02	0.03	0.02	0.01	3
Mean	6.1	348	115	67	1.7	0.10	0.30	0.30	0.20	49
max	7.3	2674	590	628	6.0	1.30	6.10	3.50	0.90	1100
± S	0.5	489	91	115	1.1	0.20	0.80	0.70	0.20	156
Median	6.1	220	99	33	1.5	0.10	0.10	0.10	0.10	3.0
MAD	0.4	81	35	16	0.5	0.10	0.10	0.10	0.10	3
CV(%)	100	63	86	50	88	52	33	32	52	6
CL95%	0.1	66	29	13	0.41	0.04	0.06	0.06	0.05	3
QCVN 01-1	6.0-8.5	1000	300	250	2	0.3	2	0.3	0.1	<3

(*) min và max: giá trị nhỏ nhất và giá trị lớn nhất; Mean: giá trị trung bình; MAD: độ lệch tuyệt đối trung bình; CV: hệ số biến thiên; CL95%: giới hạn tin cậy 95%; QCVN 01-1: viết tắt của QCVN 01-1:2018/BYT – Quy chuẩn kỹ thuật quốc gia về chất lượng nước dùng cho mục đích sinh hoạt; RSD (Horwitz) và ± S: tương ứng là độ lệch chuẩn tương đối và độ lệch chuẩn được tính từ phương trình Horwitz với Co là phân số (Co là mức giới hạn đối với thông số được quy định trong QCVN 01-1).

Số liệu thống kê mô tả, được xử lý từ Microsoft-Excel bằng công cụ Real Statistics, được mô tả trong Bảng 2. Quy chuẩn kỹ thuật quốc gia về chất lượng nước sạch dùng cho mục đích sinh hoạt do Bộ Y tế ban hành năm 2018 (QCVN 01-1:2018/BYT) [29] cũng được đưa vào trong Bảng 2

để chỉ ra các giới hạn cho phép của các tham số được sử dụng để thiết lập các hàm chỉ số con tuyến tính.

Các hàm tuyến tính (biểu diễn đồ thị và phương trình) xác định chỉ số phụ qi đối với thông số CLN i được nêu ở Hình 1.



Hình 1. Đồ thị biểu diễn sự phụ thuộc giữa chỉ số phụ (q) và nồng độ/hàm lượng của thông số chất lượng nước (Các phương trình để tính q_i của các thông số được đưa ra trên đồ thị)

3.2 Trọng số của các thông số chất lượng nước từ tính toán theo PCA

Chuyển dạng các biến gốc x_{ij} (10 thông số CLN) thành biến chuẩn hóa z_{ij} : $z_{ij} = (x_{ij} - \text{Độ lệch}$

chuẩn)/Trung bình số học; biến z_{ij} có giá trị trung bình quần thể bằng 0 và phương sai bằng 1 [32], [33]. Thực hiện phương pháp PCA trên ma trận tương quan của các biến z_{ij} , thu được các giá trị

riêng, phần phương sai và phương sai tích lũy của các thành phần chính (PC) ở Bảng 3. Các kết quả cho thấy: Chỉ cần giữ lại 04 thành phần chính (PC1 – PC4) vì chúng đều có giá trị riêng ≥ 1 và giải thích được 74,4 % phương sai của tập dữ liệu gốc; Các thành phần chính còn lại (PC5 – PC10) đều có giá trị riêng < 1 và chỉ giải thích 25,6 % phương sai, được xem là biến động nền hay “nhiều nền”.

Bảng 3. Giá trị riêng, phần phương sai và phương sai tích lũy của 04 thành phần chính đầu tiên

Thành phần chính	PC1	PC2	PC3	PC4
Giá trị riêng	3,6	1,6	1,2	1,0
Phần phương sai (%)	36,2	15,7	12,3	10,2
Phương sai tích lũy (%)	36,2	51,9	64,1	74,4

Bảng 4. Trọng số w của thông số chất lượng nước^(*)

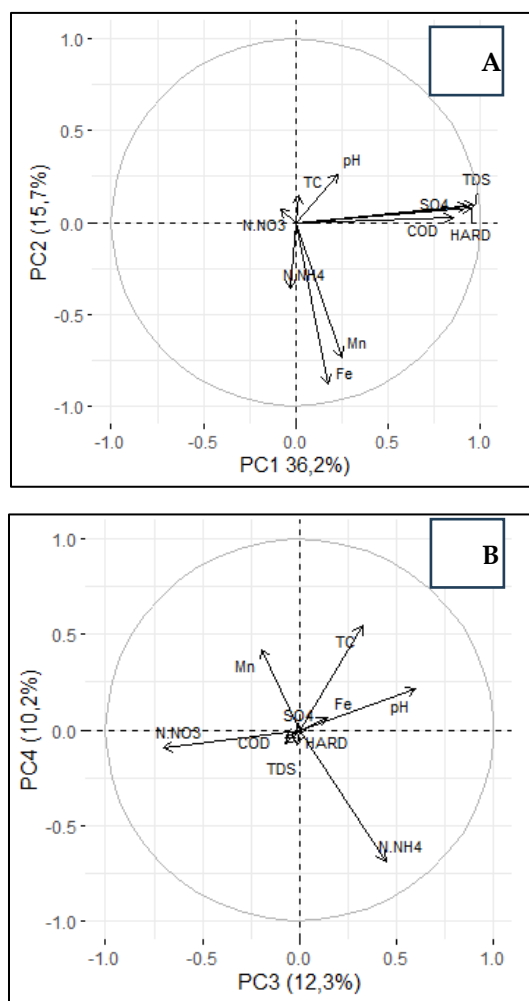
Thông số	Tổng tích lũy Trên 4 PC (1-4)	Trọng số (w)
pH	0,701	0,10
TDS	0,950	0,14
Hard	0,839	0,12
SO4	0,941	0,14
COD	0,468	0,07
N-NH4	0,575	0,08
N-NO3	0,581	0,08
Fe	0,689	0,10
Mn	0,626	0,09
TC	0,500	0,08

^(*) Ngoại trừ các giá trị w , số con số sau dấu phẩy của các số liệu trong bảng được giữ lại chỉ để phục vụ tính toán. Tổng tích lũy của 10 biến trên 4 PC giữ lại bằng tổng giá riêng của chúng.

Biểu diễn các vector biến trong không gian các thành phần chính PC1-PC2 và PC3-PC4, thu được đồ thị tải lượng yếu tố (factor loading plot)

ở Hình 2. Hình 2 cho thấy: PC1 giải thích tốt các biến TDS, HARD, SO4 và COD; PC2 giải thích tốt biến Fe, Mn; PC3 chủ yếu giải thích các biến N-NO3, pH; còn PC4 chủ yếu giải thích biến N-NH4 và TC (Tải lượng mỗi biến trên mỗi PC là hình chiếu vuông góc của nó trên PC đó).

Từ tải lượng của các biến ở Hình 2 tính được *Tải lượng bình phương* của mỗi biến và *Tổng tải lượng bình phương* của nó trên các PC giữ lại (communality). Từ đó tính được Tổng tải lượng bình phương của 10 biến trên các PC giữ lại (*Tổng tích lũy*) và Trọng số của mỗi biến theo công thức (8). Trọng số w_i của 10 thông số CLN được nêu trong Bảng 4.



Hình 2. Đồ thị tải lượng yếu tố A. PC1 × PC2 và B. PC3 × PC4

3.3 Kiểm tra sự phù hợp của chỉ số GWQI xây dựng được

Để kiểm tra sự phù hợp của các công thức tính GWQI khi phản ánh CLN thực tế ở vùng khảo sát, tiến hành so sánh kết quả tính toán và phân loại CLN theo 03 (ba) công thức khác nhau chỉ số GWQI được tính theo công thức tích (9), tổng (10) và khác (11) được ký hiệu tương ứng là $GWQI_M$, $GWQI_A$ và $GWQI_{Ref}$.

Các kết quả ở Bảng 5 cho thấy: các chỉ số $GWQI_M$, $GWQI_A$ và $GWQI_{Ref}$ phản ánh CLN khác nhau. Đối với các trường hợp đó, mặc dù có một

hoặc một vài thông số không đạt yêu cầu (so với quy định trong QCVN 01-1:2018/BYT), nhưng chỉ số $GWQI_A$ và $GWQI_{Ref}$ vẫn phản ánh CLN loại TỐT hoặc RẤT TỐT là không phù hợp. Nói cách khác, trong những trường hợp đó, chỉ số $GWQI_A$ và $GWQI_{Ref}$ bị mắc tính “mập mờ” hoặc “che khuất” (eclipsing) trong phản ánh CLN [1]. Trong khi đó, chỉ số $GWQI_M$ phản ánh CLN phù hợp với thực tế hơn. Rõ ràng, chỉ số $GWQI_M$ phù hợp hơn trong phản ánh CLN dưới đất cho mục đích ăn uống ở vùng khảo sát. Từ đây, để cho gọn, ký hiệu chỉ số $GWQI_M$ là GWQI.

Bảng 5. So sánh kết quả tính các chỉ số $GWQI_M$, $GWQI_A$ và $GWQI_{Ref}$, và phân loại CLN đối với một số trường hợp của Tập dữ liệu 1

Ký hiệu mẫu	TDS (mg/L)	SO ₄ (mg/L)	Fe (mg/L)	GWQI _A - phân loại	GWQI _M - phân loại	GWQI _{Ref} - phân loại	TC (MPN/100mL)
DS18	82	10	0.02	92 - RT	69 - TB	12 - RT	43 (*)
HB17-P	1357 (*)	446 (*)	0.37 (*)	79 - T	64 - TB	44 - T	3
HB18-P	2674 (*)	546 (*)	0.22	72 - TB	28 - K	43 - T	3
HQ18	372	38	0.02	92 - RT	69 - TB	69 - TB	23 (*)
HK16-T	99	9	0.14	92 - RT	69 - TB	82 - K	23 (*)
HP17	1013 (*)	155	0.02	91 - RT	66 - TB	111 - RK	3
HC17-B	223	12	3.50 (*)	91 - RT	66 - TB	92 - K	3
HD15	371	45	0.06	92 - RT	69 - TB	3087 - RK	1100 (*)

(*): Không đạt yêu cầu so với quy định của QCVN 01-1:2018/BYT

3.4 Áp dụng chỉ số GWQI cho huyện Hải Lăng

Áp dụng công thức dạng tích - công thức (10) cho Tập dữ liệu 2 (dữ liệu CLN dưới đất ở huyện Hải Lăng giai đoạn 2019 – 2021) để tính các giá trị GWQI – là tích của các chỉ số phụ q_i được lũy thừa các giá trị trọng số w_i tương ứng. Kết quả cho thấy:

- Trong giai đoạn 2019 – 2021, CLN dưới đất huyện Hải Lăng hầu hết đều đạt loại Rất tốt với các giá trị GWQI = 91 – 100 (89/101 trường hợp, chiếm 88%) và Tốt (GWQI = 88, 1/101 trường hợp, chiếm 1%); Chỉ 7% (7/101 trường hợp) các

giá trị GWQI = 63 – 69, phản ánh CLN thuộc loại Trung bình; 1% trường hợp ứng với GWQI = 44, phản ánh CLN thuộc loại Kém và 3% (3/101 trường hợp) các giá trị GWQI = 11 – 18 phản ánh CLN thuộc loại Rất kém.

- Có 7,6 % số liệu quan trắc các thông số CLN vượt quá mức cho phép (hay không đạt yêu cầu) của QCVN 01-1. Phần trăm này khá phù hợp với % các giá trị GWQI (11 %) phản ánh CLN thuộc loại Trung bình, Kém và Rất kém.

4 Kết luận

Trên cơ sở áp dụng phương pháp PCA để xác định trọng số (w_i) của các thông số CLN một cách khách quan, đã xây dựng được chỉ số GWQI áp dụng phù hợp trong đánh giá CLN dưới đất ở huyện Hải Lăng, tỉnh Quảng Trị. Các giá trị chỉ số phụ (q_i) của các thông số CLN được xác định dựa trên tương quan tuyến tính giữa q_i và nồng độ/hàm lượng của thông số CLN, nên thuận tiện cho người dùng chỉ số GWQI. Mặt khác, trên cơ sở so sánh các công thức tính GWQI khác nhau, đã khẳng định rằng, công thức dạng tích ($GWQI_M$) phù hợp hơn công thức dạng tổng ($GWQI_A$) và công thức khác ($GWQI_{Ref}$) trong phản ánh CLN. Cuối cùng, quy trình xây dựng chỉ số GWQI trong nghiên cứu này có thể áp dụng nhân rộng để đánh giá CLN dưới đất ở địa phương hoặc khu vực khác ở Việt Nam.

Tài liệu tham khảo

- Li P, Abbasi T and Abbasi SA: Water Quality Indices. *Environ Earth Sci.* 2014;71(10):4625-8.
- Noori R, Berndtsson R, Hosseinzadeh M, Adamowski JF, Abyaneh MR. A critical review on the application of the National Sanitation Foundation Water Quality Index. *Environmental Pollution.* 2019;244:575-87.
- Bordalo AA, Teixeira R, Wiebe WJ. A water quality index applied to an international shared river basin: The case of the Douro River. *Environ Manage.* 2006;38(6):910-20.
- Sánchez E, Colmenarejo MF, Vicente J, Rubio A, García MG, Travieso L, et al. Use of the water quality index and dissolved oxygen deficit as simple indicators of watersheds pollution. *Ecol Indic.* 2007;7(2):315-28.
- Backman B, Bodiš D, Lahermo P, Rapant S, Tarvainen T. Application of a groundwater contamination index in Finland and Slovakia. *Environmental Geology.* 1998;36(1-2):55-64.
- Elubid BA, Huag T, Ahmed EH, Zhao J, Elhag KM, Abbass W, et al. Geospatial distributions of groundwater quality in Gedaref state using geographic information system (GIS) and drinking water quality index (DWQI). *Int J Environ Res Public Health.* 2019;16(5).
- Soltan ME. Evaluation of ground water quality in Dakhla Oasis (Egyptian Western Desert). *Environ Monit Assess.* 1999;57(2):157-68.
- Prasad M, Sunitha V, Reddy YS, Suvarna B, Reddy BM, Reddy MR. Data on water quality index development for groundwater quality assessment from Obulavaripalli Mandal, YSR district, A.P India. *Data Brief.* 2019;24.
- Adimalla N, Qian H. Groundwater quality evaluation using water quality index (WQI) for drinking purposes and human health risk (HHR) assessment in an agricultural region of Nanganur, south India. *Ecotoxicol Environ Saf.* 2019 Jul 30;176:153-61.
- Kayemah N, Al-Ruzouq R, Shanableh A, Gokhan Yilmaz A. Evaluation of Groundwater Quality Using Groundwater Quality Index (GWQI) in Sharjah, UAE. In: *E3S Web of Conferences.* EDP Sciences; 2021.
- Ekere NR, Agbazue VE, Ngang BU, Ihedioha JN. Hydrochemistry and Water Quality Index of groundwater resources in Enugu north district, Enugu, Nigeria. *Environ Monit Assess.* 2019;191(3).
- Ahmed S, Khurshidali S, Yunus P, Koli S. Hydrochemical Appraisal of ground water quality and its water quality index: A case study in Mathura district, India. *Int J Adv Res (Indore).* 2018 Jun 30;6(6):1130-45.
- Abdelaziz S, Gad MI, El Tahan AHMH. Groundwater quality index based on PCA: Wadi El-Natron, Egypt. *Journal of African Earth Sciences.* 2020;172(October 2019):103964.
- Barbosa Filho J, de Oliveira IB. Development of a groundwater quality index: GWQI, for the aquifers of the state of Bahia, Brazil using multivariable analyses. *Sci Rep.* 2021;11(1).
- Štambuk-Giljanović N. Water quality evaluation by index in Dalmatia. *Water Res.* 1999;33(16):3423-40.
- Paun I, Cruceru LV, Chiriac FL, Niculescu M, Vasile GG, Marin NM. Water quality indices - Method for evaluating the quality of drinking water. In: *SIMI 2016. National Research and Development Institute for Industrial Ecology;* 2016. p. 395-402.
- Joung HM, Miller WW, Mahannah CN, Guitjens JC. A Generalized Water Quality Index Based on

- Multivariate Factor Analysis. *J Environ Qual.* 1979;8(1):95-100.
18. Tripathi M, Singal SK. Allocation of weights using factor analysis for development of a novel water quality index. *Ecotoxicol Environ Saf.* 2019;183(August):109510.
 19. Vega M, Pardo R, Barrado E, Debán L. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res.* 1998;32(12):3581-92.
 20. Qian Y, Migliaccio KW, Wan Y, Li Y. Surface water quality evaluation using multivariate methods and a new water quality index in the Indian River Lagoon, Florida. *Water Resour Res.* 2007;43(8):1-10.
 21. Coletti C, Testezlaf R, Ribeiro TAP, de Souza RTG, Pereira D de A. Water quality index using multivariate factorial analysis. *Revista Brasileira de Engenharia Agrícola e Ambiental.* 2010;14(5):517-22.
 22. Âu NH, Bảo TM, Nhi PTT, Vy THM, Hiền TT, Hiệp TN, et al. Ứng dụng chỉ số chất lượng nước dưới đất (GWQI) và GIS phân vùng chất lượng nước dưới đất tầng Pleistocen tại thị xã Phú Mỹ, tỉnh Bà Rịa - Vũng Tàu. *Tạp chí Phát triển Khoa học và Công nghệ - Khoa học Trái đất và Môi trường.* 2020;4(1):149-61.
 23. Báo cáo đề tài cấp tỉnh Quảng Trị. Xây dựng CSDL tài nguyên khoáng sản và nước dưới đất làm luận cứ khoa học cho quy hoạch khai thác sử dụng hợp lý tài nguyên bảo vệ môi trường trong chiến lược phát triển kinh tế xã hội tỉnh Quảng Trị; 2013.
 24. Sutadian AD, Muttill N, Yilmaz AG, Perera BJC. Development of river water quality indices—a review. *Environ Monit Assess.* 2016;188(1):1-29.
 25. Vasanthavigar M, Srinivasamoorthy K, Vijayaragavan K, Rajiv Ganthi R, Chidambaram S, Anandhan P, et al. Application of water quality index for groundwater quality assessment: Thirumanimuttar sub-basin, Tamilnadu, India. *Environ Monit Assess.* 2010;171(1-4):595-609.
 26. Subba Rao N, Sunitha B, Adimalla N, Chaudhary M. Quality criteria for groundwater use from a rural part of Wanaparthy District, Telangana State, India, through ionic spatial distribution (ISD), entropy water quality index (EWQI) and principal component analysis (PCA). *Environ Geochem Health.* 2020;42(2):579-99.
 27. Solangi GS, Siyal AA, Babar MM, Siyal P. Application of water quality index, synthetic pollution index, and geospatial tools for the assessment of drinking water quality in the Indus Delta, Pakistan. *Environ Monit Assess.* 2019;191(12).
 28. House MA. A Water Quality Index for River Management. *Water and Environment Journal.* 1989;3(4):336-44.
 29. MOH Q. National technical regulation on Domestic Water Quality; 2018.
 30. Miller JN, Miller JC. The quality of analytical measurements. *Statistics and Chemometrics for Analytical Chemistry.* 2010. 74-151 p.
 31. Rivera C, Rodriguez R. Horwitz equation as quality benchmark in ISO-IEC 17025 Horwitz Ratio (HorRat). *IIE Annual Conference.* 2010;5:1-6.
 32. Denis DJ. Univariate, bivariate, and multivariate statistics using R: Quantitative tools for data analysis and data science. *Univariate, Bivariate, and Multivariate Statistics Using R: Quantitative Tools for Data Analysis and Data Science.* 2020. 1-366 p.
 33. Afifi A, May S, Donatello RA, Clark VA. Practical Multivariate Analysis. *Practical Multivariate Analysis.* 2019. 1-418 p.
 34. MORE. Hướng dẫn kỹ thuật tính toán và công bố chỉ số chất lượng nước Việt Nam (VN_WQI). 2019;1-10.