

ĐÁNH GIÁ CHẤT LƯỢNG NƯỚC GIẾNG Ở VÙNG TRŨNG HUYỆN HẢI LĂNG, TỈNH QUẢNG TRỊ DỰA VÀO PHÂN TÍCH THỐNG KÊ

Nguyễn Trọng Hữu¹, Mai Xuân Dũng¹, Nguyễn Trường Khoa², Nguyễn Hữu Nam², Nguyễn Văn Hợp^{3*}

¹ Trung tâm Quan trắc Tài nguyên và Môi trường Quảng Trị, Đông Hà, Quảng Trị, Việt Nam

² Sở Tài nguyên và Môi trường Quảng Trị, 227 Hùng Vương, Đông Hà, Quảng Trị, Việt Nam

³ Khoa Hóa, Trường Đại học Khoa học, Đại học Huế, 77 Nguyễn Huệ, Huế, Việt Nam

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

(Ngày nhận bài: 14-10-2019; Ngày chấp nhận đăng: 04-11-2019)

Tóm tắt. Các số liệu quan trắc chất lượng nước (CLN) giếng ở sáu xã vùng trũng huyện Hải Lăng gồm Hải Thành, Hải Dương, Hải Vĩnh, Hải Ba, Hải Quế và Hải Hoà trong giai đoạn 2015–2017 với tần suất quan trắc một lần/năm (tháng 4–5 hàng năm) được thu thập. Trên cơ sở các số liệu đó và kết hợp với các kết quả lấy mẫu và phân tích CLN giếng ở ba vị trí thuộc xã Hải Ba, Hải Quế và Hải Hoà năm 2018 và 2019 (tần suất quan trắc một lần/năm – tháng 10/2018 và tháng 5/2019), CLN giếng ở vùng khảo sát được đánh giá qua 12 thông số: pH, độ đục (Tur), TDS, COD, Cl⁻, N-NH₄⁺, N-NO₃⁻, N-NO₂⁻, SO₄²⁻, tổng sắt tan (Fe), tổng mangan tan (Mn) và tổng coliform. Phương pháp phân tích phương sai, phân tích thành phần chính (PCA) và phân tích cụm (AHC) cũng được áp dụng để đánh giá biến động CLN theo không gian và thời gian. Vùng khảo sát được phân chia thành 3 tiểu vùng (TV) có đặc trưng chất lượng nước khác nhau: TV1 gồm xã Hải Dương và xã Hải Vĩnh bị nhiễm đáng kể các chất hữu cơ (COD cao hơn); TV2 (xã Hải Thành và xã Hải Hoà) bị nhiễm phen sắt (Fe cao hơn và pH thấp hơn); và TV3 (xã Hải Ba và xã Hải Quế) bị nhiễm đáng kể muối (TDS, Cl⁻, SO₄²⁻), Mn và vi khuẩn có nguồn gốc phân (tổng coliform).

Từ khóa: chất lượng nước giếng, PCA, AHC, huyện Hải Lăng

Statistical analysis-based assessment of well water quality in bottomland area – Hai Lang district, Quang Tri province

Nguyen Trong Huu¹, Mai Xuan Dung¹, Nguyen Truong Khoa², Nguyen Huu Nam², Nguyen Van Hop^{3*}

¹ Quang Tri's Centre for Resources and Environmental Monitoring, Dong Ha City, Quang Tri, Vietnam

² Quang Tri's Resources and Environmental Department, 227 Hung Vuong St., Dong Ha City, Quang Tri, Vietnam

³ Department of Chemistry, University of Sciences, Hue University, 77 Nguyen Hue St., Hue, Vietnam

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

(Received: 14 October 2019; Accepted: 04 November 2019)

Abstract. The data of well water quality monitoring (drilled and dug wells) in six communes of the bottomland area, Hai Lang district, Quang Tri province, including Hai Thanh, Hai Duong, Hai Vinh, Hai Ba, Hai Que, and Hai Hoa from 2015 to 2017 with the frequency of once per year (April or May annually) were collected. On the basis of data and the results obtained from sampling and analysis of

well water quality at three sites in three communes (Hai Ba, Hai Que, and Hai Hoa) between 2018 and 2019 with the frequency of once per year (October 2018 or May 2019), the well water quality in the area was assessed through 12 parameters: pH, turbidity, TDS, COD, Cl^- , N-NH_4^+ , N-NO_3^- , N-NO_2^- , SO_4^{2-} , total dissolved iron (Fe), total dissolved manganese (Mn), and total coliform. Analysis of variance, principal component analysis (PCA), and agglomerate hierarchical clustering (AHC) were also applied to assess the spatial and temporal variation of well water quality. The area was divided into three sub-area with different well water quality: Sub-area I including Hai Duong and Hai Vinh communes considerably polluted with organics (high COD); Sub-area II (Hai Thanh and Hai Hoa communes) contaminated with iron (high Fe and low pH); and Sub-area III (Hai Ba and Hai Que communes) much contaminated with salts (high TDS, Cl^- , SO_4^{2-}), Mn, and total coliform.

Keywords: well water quality, PCA, AHC, Hai Lang

1 Mở đầu

Huyện Hải Lăng, tỉnh Quảng Trị, có 19 xã và 1 thị trấn với tổng diện tích tự nhiên 42.479,7 ha, dân số 81.883 người (thành thị 3.142 người, nông thôn 78.691 người) (số liệu năm 2018). Vùng trung huyện Hải Lăng (gồm 12 xã) có vị trí địa lý $16^\circ 38'$ đến $16^\circ 43'$ vĩ độ Bắc và $107^\circ 17'$ đến $107^\circ 22'$ kinh độ Đông, với tổng diện tích đất tự nhiên 21.356,5 ha. Đây là vùng thuộc hạ lưu của các sông Ô Lâu, Thác Ma, Ô Khê, Tân Vĩnh Định và Cựu Vĩnh Định. Địa hình khu vực có dạng lòng chảo, đáy là khu đồng ruộng rộng lớn có cao độ thấp từ $-1,0$ đến $+0,2$ m; vùng ven có cao độ $+0,6$ đến $+1$ m, xung quanh là các cồn cát và đồi núi với cao độ $+3$ đến $+10$ m. Với địa hình như vậy, khu vực này thường xuyên bị ngập úng vào mùa mưa lũ [1]. Đa số dân cư ở vùng trung huyện Hải Lăng (khoảng 80%) sử dụng nước giếng, chủ yếu là giếng khoan và một số giếng đào (gọi chung là nước giếng) để cấp cho ăn uống và sinh hoạt. Do điều kiện vệ sinh môi trường ở vùng này chưa tốt và khó khăn về nguồn nước, nên lo lắng về rủi ro sức khỏe cộng đồng. Theo Trung tâm y tế huyện Hải Lăng, ở vùng này thường phát sinh nhiều bệnh tật như các bệnh phụ khoa, ngoài da, v.v. Tuy vậy, cho đến nay, thông tin về chất lượng nước (CLN) cấp cho sinh hoạt ở vùng này còn hạn chế.

Trong nhiều năm qua, đã có một số nghiên cứu đánh giá CLN giếng ở tỉnh Quảng Trị như CLN giếng ở xã Cam Thành và Cam Nghĩa, huyện Cam Lộ năm 2007 [2]; Quy hoạch quản lý, khai thác, sử dụng và bảo vệ tài nguyên nước dưới đất miền đồng bằng tỉnh Quảng Trị năm 2008 [3]; Bộ chỉ số theo dõi – đánh giá nước sạch và vệ sinh môi trường nông thôn năm 2010 [4]. Mặt khác, trong chương trình quan trắc hàng năm do Trung tâm Quan trắc Tài nguyên và Môi trường (TNMT) thuộc Sở TNMT Quảng trị thực hiện, CLN giếng cũng được quan trắc định kỳ tại 25 điểm với 1–2 đợt/năm từ 2015 đến nay, trong đó có quan trắc nước giếng ở vùng trung huyện Hải Lăng [5]. Tuy nhiên, các nghiên cứu trên chủ yếu đánh giá CLN qua so sánh kết quả thu được với giá trị được quy định trong các Quy chuẩn Việt Nam (QCVN) hiện hành. Những nghiên cứu chi tiết để nhận biết sự biến động theo không gian và thời gian, tương quan giữa các thông số CLN, phân vùng dựa vào đặc điểm CLN hầu như chưa được thực hiện. Những điều này có thể thực hiện được bằng cách áp dụng phương pháp phân tích thống kê đa biến như phân tích thành phần chính (PCA – principal component analysis), phân tích cụm (AHC – agglomerate hierarchical clustering).

Bài báo này cung cấp thông tin về CLN giếng ở sáu xã thuộc vùng trung huyện Hải Lăng, tỉnh Quảng Trị (Hải Thành, Hải Dương, Hải Vĩnh,

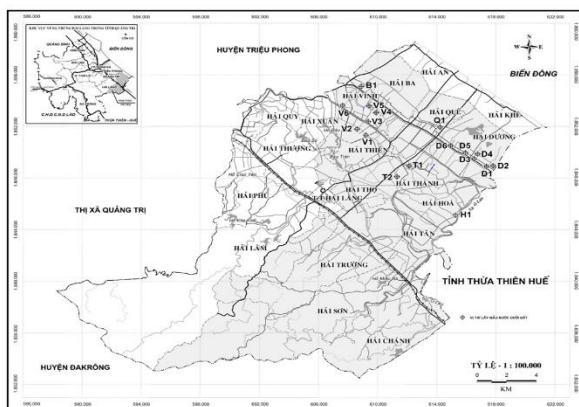
Hải Ba, Hải Quế và Hải Hoà) trên cơ sở phân tích thống kê các dữ liệu CLN giếng trong giai đoạn 2015–2019.

2 Phương pháp

Thông tin về các vị trí quan trắc và chuẩn bị mẫu

Thông tin về vị trí quan trắc (gọi tắt là *vị trí*) ở vùng trung huyện Hải Lăng trong ba năm (2015–2017): Các vị trí quan trắc là các giếng đang được sử dụng cho sinh hoạt gia đình; tần suất quan trắc: một đợt/năm (vào mùa khô) – tháng 4 ở ba xã Hải Thành, Hải Dương, Hải Vĩnh và tháng 5 ở ba xã Hải Ba, Hải Quế, Hải Hòa) [4, 5]:

– *Xã Hải Thành* (ký hiệu tên xã là T): hai vị trí (một giếng đào ở thôn Trung Đơn, một giếng khoan ở thôn Kim Sanh); *Xã Hải Dương* (D): sáu vị trí (hai giếng khoan/thôn ở ba thôn Xuân Viên, Đông Dương và Diên Khánh); *Xã Hải Vĩnh* (V): sáu vị trí (một giếng đào và một giếng khoan ở thôn Thi Ông, hai giếng khoan/thôn ở thôn Thuận Nhon và thôn Lam Thủy). Tổng cộng ở ba xã này là 14 vị trí. Các giếng đào có độ sâu trung bình là 4 m; các giếng khoan có độ sâu 8–20 m. Do ở mỗi vị trí, chỉ lấy mẫu và quan trắc một lần trong năm, nên CLN có thể thay đổi từ năm này sang năm khác và do vậy, để cho tiện, các mẫu được gọi là các *giếng*. Theo đó, tổng số giếng được quan trắc trong ba năm (2015–2017) là $n_1 = 42$.



Hình 1. Các vị trí quan trắc ở 6 xã vùng trung huyện Hải Lăng trong giai đoạn 2015–2019

– *Xã Hải Ba* (B) – thôn Phương Lan, *xã Hải Quế* (Q) – thôn Kim Long và *xã Hải Hòa* (H) – thôn An Thơ: ba vị trí (mỗi xã có một vị trí được lấy mẫu và quan trắc – là giếng khoan có độ sâu 8–15 m). Theo cách tương tự trên, tổng số giếng được quan trắc trong ba năm là $n_2 = 9$.

Chuẩn bị mẫu. Trong năm 2018 và 2019, tiến hành lấy mẫu và phân tích nước giếng tại ba vị trí ở ba xã Hải Ba (ký hiệu tên xã là B), Hải Quế (Q) và Hải Hòa (H); ba vị trí này trùng với ba vị trí được quan trắc trong giai đoạn 2015–2017. Tần suất quan trắc: một đợt/năm (tháng 10/2018 thuộc mùa mưa và tháng 5/2019 thuộc mùa khô). Tổng số giếng được quan trắc trong năm 2018 và 2019 là $n_3 = 6$. Quy cách lấy mẫu và bảo quản mẫu tuân thủ quy định của TCVN 6663-11:2011 – Hướng dẫn lấy mẫu nước dưới đất và TCVN 6663-3:2016 – Hướng dẫn bảo quản và xử lý mẫu nước. Các mẫu được bảo quản trong chai nhựa PET (polyetyleneterephthalat) sạch và nước đá lạnh, vận chuyển về và phân tích tại phòng thí nghiệm của Trung tâm Quan trắc TNMT, Sở TNMT tỉnh Quảng Trị.

Như vậy, trong giai đoạn 2015–2019, tổng số vị trí quan trắc là $q = 17$ và tổng số giếng được quan trắc là $n = n_1 + n_2 + n_3 = 57$. Các giếng được ký hiệu là X_{ij} , trong đó, X là ký tự chỉ xã, $i = 1 \div q$ là thứ tự vị trí quan trắc tại xã và $j = 5 \div 9$ ứng với số cuối của năm quan trắc 2015–2019. Chẳng hạn, giếng T15 là giếng ở xã Hải Thành, vị trí 1 và được quan trắc năm 2015... Các vị trí quan trắc CLN giếng được chỉ ra ở Hình 1.

2.1 Phương pháp đo/phân tích các thông số chất lượng nước

Các phương pháp đo/phân tích các thông số CLN là các phương pháp tiêu chuẩn của Việt Nam và các phương pháp chuẩn phân tích nước và nước thải (SMEWW, APHA, USA) [6]. Số thông số CLN được đo/phân tích là 12, bao gồm: pH, độ đục (Tur), tổng muối tan (TDS), nhu cầu oxy hóa học (COD) – phép đo pemanganat, clorua (Cl⁻), N-amoni (viết tắt là NH₄), N-nitrat (NO₃), N-nitrit (NO₂), sunfat (SO₄), tổng sắt tan (Fe), tổng mangan

tan (Mn) và tổng coliform (TC). Đây là những thông số bắt buộc phải phân tích đối với nước dưới đất trong các QCVN hiện hành. Một số kết quả quan trắc các kim loại độc (Hg, Cd, As, Pb, Cr, Ni, Cu và Zn) trong nước giếng ở vùng này (giai đoạn 2015–2017) rất thấp (cỡ <math><5+10\ \mu\text{g/L}</math>, riêng Hg <math><1\ \mu\text{g/L}</math>) và nhiều giá trị nhỏ hơn giới hạn phát hiện của phương pháp phân tích, nên không được đề cập trong nghiên cứu này. Mặt khác, do nước giếng vùng này có độ cứng rất thấp (<math><50+100\ \text{mg/L}</math>), nên độ cứng cũng không được đề cập ở đây.

2.2 Phương pháp đánh giá chất lượng nước

Chất lượng nước được đánh giá qua từng thông số bằng cách so sánh kết quả thu được ($m \pm \varepsilon$ với m là trung bình số học và ε là biên giới tin cậy 95%; $\varepsilon = \pm t_{(p,f)} \cdot S$; S là độ lệch chuẩn; t là giá trị t của phân bố student ở độ tin cậy $p = 95\%$ và bậc tự do $f = n - 1$; n là số kết quả quan trắc) với các giá trị giới hạn được quy định trong QCVN 09-MT:2015/BTNMT (viết tắt là QCVN09) về CLN dưới đất, QCVN 01:2009/BYT (viết tắt là QCVN01) về CLN ăn uống và QCVN 02:2009/BYT (viết tắt là QCVN02) về CLN sinh hoạt. Hệ số biến động (CV – coefficient of variation) đối với mỗi thông số CLN cũng được tính toán: $CV = S \cdot 100/m$.

2.3 Phương pháp xử lý số liệu và phân tích đa biến – phương pháp PCA và AHC

Sử dụng phần mềm Microsoft-Excel 2013 với công cụ Data Analysis để xử lý thống kê các số liệu thu được: Tính toán các đại lượng thống kê cơ bản – trung bình số học và độ lệch chuẩn; trung vị và độ lệch tuyệt đối trung vị /MAD – medium absolute deviation (đối với trường hợp số liệu dao động trong khoảng rộng, dẫn đến $S > m$); biên giới tin cậy 95%; phân tích phương sai và phân tích tương quan.

Phân tích thành phần chính: Phân tích thành phần chính và AHC là hai phương pháp điển hình thuộc nhóm các phương pháp phân tích đa biến. Phân tích thành phần chính cho phép giảm dữ liệu từ không gian p chiều của tập dữ liệu gốc

(là một ma trận gồm m dòng ứng với các *đối tượng*/individual như các vị trí quan trắc, các mẫu/giếng... và n cột ứng với các *biến*/variable như các thông số pH, COD, TDS...) thành không gian ít chiều hơn (k chiều và <math>k < p</math>). Mỗi chiều mới được chiết rút ra từ tập dữ liệu gốc được gọi là một thành phần chính/principal component (PC). Các PC không tương quan với nhau, mà trực giao với nhau và mỗi PC là tổ hợp tuyến tính của các biến gốc x_{ij} (với $i = 1+m$ và $j = 1+n$) và chúng giải thích được đa số các biến động (hay phương sai) của tập dữ liệu gốc. Mô hình PCA cho phép trực quan và khám phá dữ liệu dễ hơn – dễ nhận ra hơn quan hệ/tương quan giữa các quan sát (các biến và các đối tượng), các biến có ảnh hưởng mạnh/yếu đến các đối tượng và các cấu trúc (pattern) tồn tại trong tập dữ liệu gốc. Từ các kết quả PCA, khi phát hiện có các cấu trúc tồn tại trong dữ liệu gốc, chẳng hạn, có sự tách ra các cụm (hay nhóm) các đối tượng, cần thực hiện phân tích sâu hơn theo phương pháp AHC để phân tách các đối tượng khảo sát thành các cụm (cluster). Việc gộp các đối tượng thành các cụm là dựa vào độ tương tự (similarity)/bất tương tự (dissimilarity) giữa chúng. Độ tương tự/bất tương tự được thể hiện qua hàm khoảng cách (distance function) giữa các đối tượng hoặc cụm đối tượng. Trong nghiên cứu này, sử dụng khoảng cách Oclit (Euclidean distance) làm thước đo độ tương tự/bất tương tự giữa các đối tượng/cụm đối tượng [7].

Trước khi tính toán theo PCA và AHC, thường phải chuyển dạng dữ liệu để thỏa mãn điều kiện – các biến tuân theo phân bố chuẩn và đồng thời, tránh ảnh hưởng của các thang đo khác nhau của các biến. Ở đây, toàn bộ dữ liệu gốc về các thông số CLN đều được chuyển dạng về dạng chuẩn hóa (standardized data) như biến z của hàm phân bố chuẩn chuẩn hóa (standardized normal distribution): $Z_{ij} = (x_{ij} - m)/S$; trong đó m và S tương ứng là trung bình số học và độ lệch chuẩn của biến x_{ij}. Các tính toán và biểu diễn đồ thị theo mô hình PCA và AHC được thực hiện trên phần mềm R –

một trong những phần mềm được sử dụng phổ biến hiện nay, được phép truy cập và sử dụng miễn phí [8]. Trong nghiên cứu này, sử dụng module R-studio (R version 3.6.0, 26–4–2019, 64-bit) với package Factoextra (version 1.0.5).

3 Kết quả và thảo luận

3.1 Hiện trạng chất lượng nước giếng ở vùng trung huyện Hải Lăng

Tổng hợp các kết quả quan trắc CLN giếng ở ba xã T, D, V (2015–2017) từ tài liệu [4] và ba xã B, Q, H (2015–2017) từ tài liệu [5] và các kết quả của nghiên cứu này trong năm 2018 và 2019 (ở ba xã B, Q và H) cho thấy (các số liệu chi tiết không đưa ra ở đây):

(i) Đối với ba xã Hải Thành (T), Hải Dương (D) và Hải Vĩnh (V) trong giai đoạn 2015–2017 (Bảng 1): Nhiều giếng khảo sát có thông số pH, Tur, COD, Fe và TC không đạt yêu cầu của QCVN01 và QCVN02. Các thông số CLN còn lại đều đạt yêu cầu của QCVN 09, QCVN01 và QCVN02. Biến động mạnh nhất (xếp theo thứ tự giảm dần của hệ số biến động CV) là các thông số TC (196%) > Tur > Fe > NO₃ > Cl⁻ > SO₄ > NH₄ > TDS (77%). Độ đục (Tur) lớn nhất (>5 NTU) là ở các giếng T25–T27, V15–V17 và V45–V47 và không đạt yêu cầu của QCVN01 và QCVN02. Sỏi dĩ độ đục cao ở các giếng

này là do nồng độ Fe cao (2–3 mg/L ở giếng T25–T27; 0,3–0,5 mg/L ở các giếng V15–V17 và V45–V47); Fe bị thủy phân tạo thành các kết tủa ở dạng hydroxit và dạng keo dẫn đến làm tăng độ đục và giảm pH nước. Mật độ tổng coliform (TC) khá cao trong nhiều giếng (12/42 giếng, chiếm 29%) và không đạt yêu cầu của QCVN09, QCVN01 và QCVN02. Biến động thấp hơn (theo thứ tự tăng dần của CV) là các thông số pH (9%) < Mn < COD < NO₂ (53%). Tuy pH dao động không nhiều (khoảng 4,5–6,8), nhưng phần lớn các giếng ở xã Hải Dương có giá trị pH thấp (pH < 5,5) và không đạt QCVN01 và QCVN02.

(ii) Đối với ba xã Hải Ba (B), Hải Quế (Q) và Hải Hòa (H) trong giai đoạn 2015–2019 (Bảng 1): Nhiều giếng khảo sát có thông số pH, Tur, Fe và TC không đạt yêu cầu của QCVN01 và QCVN02. Tuy nhiên, mức nhiễm Fe và các vi khuẩn có nguồn gốc phân (TC) ở ba xã này thấp hơn so với ba xã trên: Mật độ TC trong tất cả các giếng khảo sát đều thấp hơn 240 MPN/100 mL. Biến động mạnh nhất (theo thứ tự giảm dần của CV) là thông số TC (240%) > Tur > Fe > NH₄ > NO₂ > Cl⁻ > Mn (84%). Độ đục lớn nhất (24 NTU) ở giếng H19, không đạt yêu cầu của QCVN01 và QCVN02. Biến động thấp hơn (theo thứ tự tăng dần của CV) là các thông số pH (7%) < COD < NO₃ < SO₄ < TDS (66%).

Bảng 1. Kết quả quan trắc CLN giếng ở ba xã T, D, V (2015–2017, n = 42) và ba xã B, Q, H (2015–2019, n = 15)^(*)

Thông số (đơn vị đo)	Min	Max	m	S	ε	Trung vị	MAD	CV(%)	QCVN0 1	QCVN0 2	QCVN0 9
pH	4,5	6,8	5,8	0,5	0,2	–	–	9	6,5–8,5	6,0–8,5	5,5–8,5
	5,3	6,6	5,9	0,4	0,2	–	–	7			
Tur (NTU)	0,1	29	3,6	6,1	1,9	1,1	0,8	170	2	5	–
	0,1	23,6	2,6	6,1	3,4	0,4	0,3	233			
TDS (mg/L)	34	382	132	102	32	–	–	77	1.000	–	1.500
	104	750	253	166	92	–	–	66			
COD (mg/L)	0,5	4,5	2,1	1,1	0,3	–	–	52	2	4	–
	0,5	2,2	1,5	0,6	0,3	–	–	38			
Cl ⁻ (mg/L)	3	129	12,9	4,2	7,8	–	–	105	250	300	250
	9	244	59	64	36	34	23	109			
NH ₄ (mg/L)	0,02	0,31	0,11	0,09	0,03	–	–	85	3	3	1
	0,04	1,25	0,19	0,30	0,20	0,11	0,07	157			
NO ₂ (mg/L)	0,015	0,077	0,034	0,018	0,006	–	–	53	3	–	1
	0,004	0,070	0,013	0,017	0,010	0,004	0,001	138			

Thông số (đơn vị đo)	Min	Max	<i>m</i>	<i>S</i>	ϵ	Trung vị	MAD	CV(%)	QCVN0 1	QCVN0 2	QCVN0 9
NO ₃ (mg/L)	0,06 0,04	3,25 0,14	0,48 0,08	0,55 0,04	0,17 0,02	0,32 –	0,21 –	113 50	50	–	15
SO ₄ (mg/L)	2,7 6,0	44,6 27,0	11,7 14,3	10,6 9,1	3,3 5,0	– –	– –	90 64	250	–	400
Fe (mg/L)	0,02 0,09	3,21 2,04	0,38 0,3	0,63 0,5	0,20 0,3	0,19 0,10	0,12 0,03	166 168	0,3	0,5	5
Mn (mg/L)	0,02 0,06	0,11 0,92	0,04 0,27	0,02 0,23	0,01 0,12	– –	– –	40 84	0,3	–	0,5
TC (MPN/100 mL)	3 3	1.100 240	133 29	259 70	81 39	16 3	13 0	196 240	0	150	3

^(*) Đối với mỗi thông số, dòng 1 và dòng 2 tương ứng là số liệu đối với ba xã T, D, V và ba xã B, Q, H. Đối với QCVN01 và QCVN02, amoni, nitrit và nitrat được tính theo NH₄, NO₂ và NO₃; Đối với các thông số có $S > m$, việc tính *m* và *S* chỉ để tính CV; Dấu (–) ở các QCVN là không quy định.

3.2 Biến động chất lượng nước theo không gian và thời gian

Do TDS bao hàm cả Cl⁻ và SO₄ (các muối), NH₄ liên quan đến NO₂ và NO₃ (các hợp chất chứa nito), còn Fe liên quan đến độ đục và pH, nên ở đây chỉ đánh giá biến động TDS, NH₄ và Fe trong nước giếng theo không gian (vị trí quan trắc) và thời gian (năm). Kết quả áp dụng phương pháp ANOVA hai yếu tố không lặp lại cho thấy (i) Về TDS: Ở ba xã T, D và V (2015–2017, 14 vị trí), nước giếng có TDS khác nhau theo không gian ($p = 0,0001$); nhưng không khác nhau theo thời gian ($p = 0,76$); Ở ba xã B, Q và H (2015–2019, ba vị trí) nước giếng có TDS không khác nhau theo không gian ($p = 0,44$) và thời gian ($p = 0,50$); (ii) Về NH₄: Ở ba xã T, D và V, nước giếng có NH₄ không khác nhau theo không gian ($p = 0,79$) nhưng khác nhau theo thời gian ($p < 0,001$); Đối với ba xã B, Q và H, nồng độ NH₄ trong nước giếng không khác nhau theo không gian ($p = 0,32$) và thời gian ($p = 0,32$); (iii) Về Fe: Nồng độ Fe trong nước giếng ở ba xã T, D và V khác nhau theo cả không gian ($p < 0,0001$) và thời gian ($p = 0,006$); đối với ba xã B, Q và H, không khác nhau theo không gian ($p = 0,57$) và thời gian ($p = 0,60$). Nói chung, trong giai đoạn 2015–2019, nước giếng có TDS, NH₄ và Fe biến động đáng kể (tức là khác nhau) theo không gian và thời gian chỉ ở ba xã T, D và V nhưng lại biến động không đáng kể ở ba xã B, Q và H.

3.3 Phân cụm và phân vùng chất lượng nước giếng dựa vào phương pháp PCA và AHC

Để xác định quan hệ giữa các thông số CLN và phân cụm/phân vùng dựa vào đặc điểm CLN của các giếng, áp dụng phương pháp PCA và AHC cho dữ liệu “đầu vào” là tập dữ liệu gốc về 12 thông số CLN giếng (hay 12 biến) ở 57 giếng (hay 57 đối tượng) ở 6 xã vùng khảo sát (2015–2019). Để thuận lợi trong biểu diễn các đồ thị, các đối tượng được gán các số thứ tự lần lượt 1, 2, 3, ..., 42 ứng với các giếng T15, T16, T17, ..., V67 ở ba xã T, D, V ($n = 42$) và 43, 44, 45, ..., 57 ứng với các giếng B15, B16, B17, ..., H19 ở ba xã B, Q, H ($n = 15$).

Phân tích các biến (các thông số chất lượng nước)

(i) Các thành phần chính (PC): Kết quả tính toán giá trị riêng (eigen vlue), phần phương sai và phương sai tích lũy theo phương pháp PCA ở Bảng 2 cho thấy: Chỉ nên giữ lại 4 thành phần chính (PC1–PC4) vì chúng đều có giá trị riêng lớn hơn 1 và giải thích được 72,1% biến động (hay phương sai) của tập dữ liệu gốc (x_{ij}); các thành phần chính còn lại (PC5–PC12) đều có giá trị riêng nhỏ hơn 1 và chỉ giải thích 27,9% biến động của tập dữ liệu. Phần phương sai này được coi là biến động “nền” của tập dữ liệu.

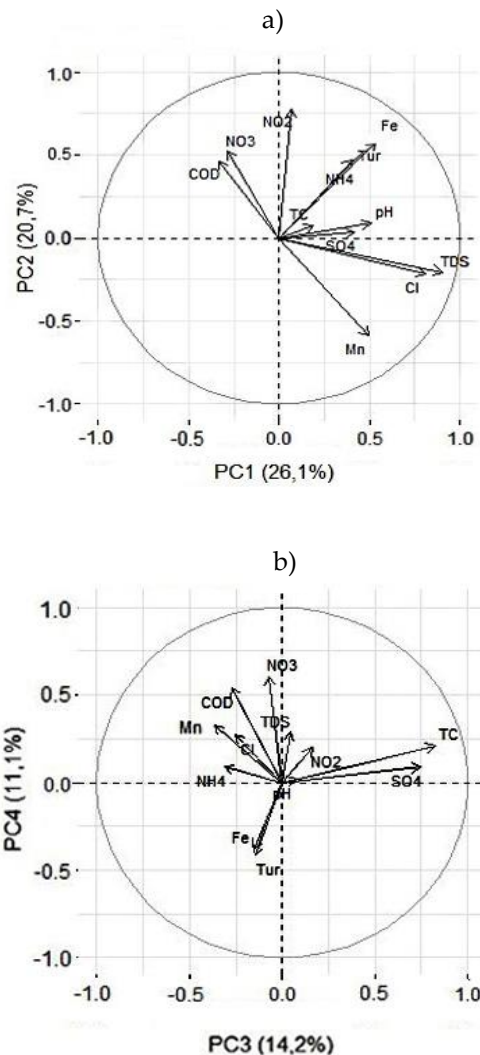
Bảng 2. Các giá trị riêng, phần phương sai và phương sai tích lũy của năm thành phần chính đầu tiên

Đại lượng thống kê	PC1	PC2	PC3	PC4	PC5
Giá trị riêng	3,135	2,480	1,702	1,337	0,841
Phần phương sai (%)	26,1	20,7	14,2	11,1	7,0
Phương sai tích lũy (%)	26,1	46,8	61,0	72,1	77,1

(ii) Đồ thị tải lượng hay đồ thị các biến (loading plot):

– Đối với PC1 và PC2 (Hình 2a): PC1 giải thích tốt các biến TDS, Cl⁻ và ở mức độ thấp hơn là Fe, Tur và pH; theo hướng PC1, các biến có tương quan chặt với nhau là TDS – Cl⁻ (với hệ số tương quan $R = 0,87$); Fe – Tur ($R = 0,84$), TDS – Mn ($R = 0,63$) và tương quan ở mức độ thấp hơn là Fe – NH₄ ($R = 0,36$). PC2 giải thích tốt các biến NO₂ và ở mức độ thấp hơn là NO₃ và COD; theo hướng PC2, các biến có tương quan với nhau là COD – NO₃ ($R = 0,55$) và ở mức độ thấp hơn là NO₃ – NO₂ ($R = 0,38$), NH₄ – NO₂ ($R = 0,39$) và chúng tương quan nghịch với Mn ở mức độ thấp: NO₂ – Mn ($R = -0,42$); NO₃ – Mn ($R = -0,20$) và COD – Mn ($R = -0,16$). Như vậy, khi TDS và Cl⁻ cao thì sự hòa tan Fe, Mn vào nước tăng lên, dẫn đến làm tăng nồng độ Fe và Mn; nồng độ Fe cao sẽ làm tăng độ đục của nước. Mặt khác, tương quan giữa các biến COD – NH₄ – NO₂ – NO₃ phản ánh các quá trình tự nhiên xảy ra trong nước: Khi COD và NH₄ cao thì nồng độ NO₂ và NO₃ trong nước tăng (do quá trình nitrat hóa). Tương quan nghịch giữa Mn và NO₂, NO₃, COD có thể là do nồng độ Mn cao sẽ thúc đẩy quá trình oxy hóa các chất hữu cơ và quá trình khử nitrat, dẫn đến làm giảm phần nào COD và NO₃, NO₂.

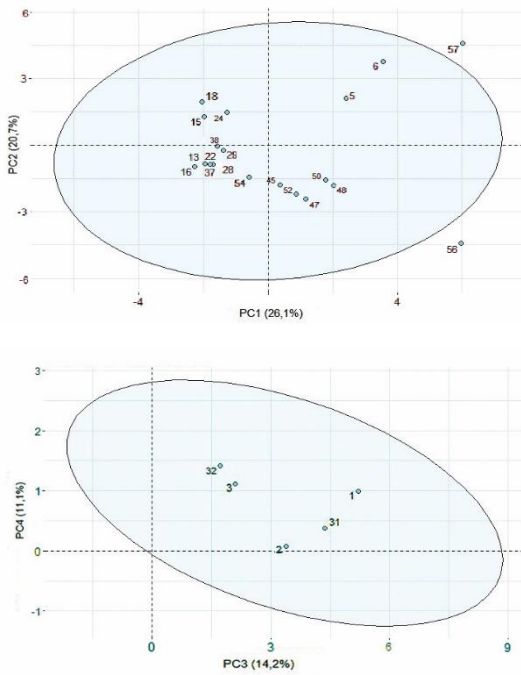
– Đối với PC3 và PC4 (Hình 2b): PC3 giải thích tốt các biến TC và SO₄, trong khi PC4 (cũng gần tương tự PC2 nhưng ở mức độ thấp hơn) giải thích được các biến NO₃ và COD.



Hình 2. Đồ thị các biến trong không gian 2 thành phần chính PC1 và PC2 (a); PC3 và PC4 (b)

Phân tích các đối tượng (các giống)

Do số đối tượng khá lớn (57 đối tượng), nên khó thể hiện cả 57 đối tượng trên đồ thị các đối tượng (score plot) trong không gian các PC vì chúng chồng lấp lên nhau. Ở đây chỉ biểu diễn lên đồ thị không gian 2 chiều (PC1 – PC2 và PC3 – PC4) những đối tượng có hệ số xác định lớn (R^2 hay $\cos^2 \geq 0,5$, nghĩa là trên 50% phương sai của các biến ở mỗi đối tượng được giải thích bởi 2 thành phần chính tương ứng) (Hình 3).



Hình 3. Đồ thị các đối tượng trong không gian PC1 – PC2 (20 đối tượng) (phía trên) và PC3 – PC4 (5 đối tượng) (phía dưới)

Hình 3 cho thấy:

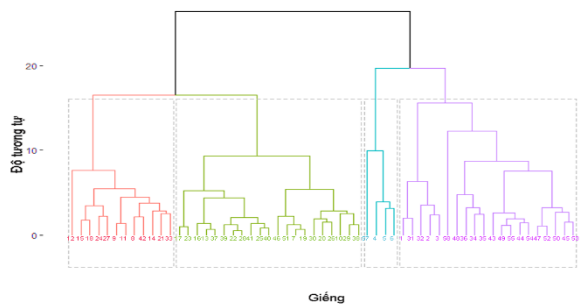
– Đối với PC1 và PC2, 2 đối tượng (56 và 57 ứng với giếng H18 và H19) nằm ngoài vùng tin cậy 95%, nên 2 đối tượng này được xem là bất thường (outlier): Giếng H19 có nồng độ Fe (2,04 mg/L) và NH₄ (1,25 mg/L) cao bất thường, còn giếng H18 có TDS (750 mg/L) và Cl⁻ (244 mg/L) cao bất thường so với các đối tượng khác. Tuy vậy, do chỉ có 2/57 đối tượng (chiếm 3,5%), nên không cần thiết phải loại bỏ chúng trước khi thực hiện PCA [7].

– Các đối tượng ở góc phần tư thứ I (góc trên bên phải hệ tọa độ PC1 – PC2) là 4, 5, 6, 57 (ứng với các giếng T25, T26, T27 và H19) có nồng độ Fe, Tur, NH₄ và NO₂ cao. Ngược lại, các đối tượng ở góc phần tư thứ III (góc dưới bên trái), có nồng độ Fe,

Tur và NH₄, NO₂ thấp (các đối tượng 13, 16, 22, 28, 29, 37... ứng với các giếng D35, D45, D65 và V25, V26, V55...). Các đối tượng ở góc phần tư thứ IV (góc dưới bên phải) 45, 47, 48, 50, 52, 56 ứng với các giếng B17, B19, Q15, Q17, Q19 và H18 có nồng độ TDS, Cl⁻ và Mn cao hơn so với các đối tượng khác nhưng có nồng độ thấp của NO₃, NO₂ và COD. Ngược lại, các đối tượng ở góc phần tư thứ II (góc trên bên trái) – 15, 18, 24 ứng với các giếng D37, D47 và D67, có nồng độ TDS, Cl⁻ và Mn thấp nhưng lại có nồng độ cao của NO₃, NO₂ và COD.

– Đối với PC3 và PC4, các đối tượng phân bố theo 2 hướng này (1, 2, 3, 31, 32 ứng với các giếng T15, T16, T17 và V35, V36) có nồng độ cao của TC và SO₄.

Phân nhóm (hay cụm) các đối tượng và phân vùng chất lượng nước giếng. Các kết quả phân tích PCA chưa cho phép phân cụm các đối tượng dựa vào đặc điểm về CLN của chúng. Áp dụng phương pháp AHC và tổ hợp các đối tượng theo phương pháp Ward (Ward method linkage), thu được đồ thị phân cụm đối tượng (dendrogram) ở Hình 4 và Bảng 3.



Hình 4. Giản đồ phân cụm (4 cụm) các đối tượng ở vùng khảo sát

Bảng 3. Bốn cụm đối tượng (giếng) được phân chia theo đặc điểm CLN (*)

Cụm (số giếng/số vị trí)	Thôn (giếng)/xã	Thông số CLN(*)
Cụm 1 (12 giếng / 9 vị trí)	– Thôn Xuân Viên (D16, D17, D26, D27), Đông Dương (D36, D37, D47), Diên Khánh (D57, D67) / xã Hải Dương	– COD, NO ₂ và NO ₃ ↑
	– Thôn Thi Ông (V17), Thuận Nhon (V37), Lam Thủy (V67) / xã Hải Vĩnh	– Mn, TDS, Cl ⁻ ↓

Cụm (số giếng/số vị trí)	Thôn (giếng)/xã	Thông số CLN ^(*)
Cụm 2 (21 giếng / 13 vị trí)	– Thôn Xuân Viên (D15, D25), Đông Dương (D35, D45, D46), Diên Khánh (D55, D56, D65, D66) / xã Hải Dương – Thôn Thi Ông (V15, V16, V25, V26, V27), Thuận Nhơn (V47), Lam Thủy (V55, V57, V65, V66) / xã Hải Vĩnh – Thôn Phương Lang (B18) / xã Hải Ba – Thôn Kim Long (Q18) / xã Hải Quế	– COD, NO ₃ ↑ – Fe, Tur, NH ₄ ↓ – pH↓ (ở nhiều giếng: D25, D35, D45, D46, D65, D66, V15, V25, V26, V55, V57, V65)
Cụm 3 (4 giếng/2 vị trí)	– Thôn Kim Sanh (T25, T26, T27) / xã Hải Thành – Thôn An Thọ (H19) / xã Hải Hoà	Fe, Tur↑↑, NH ₄ ↑ (Fe, NH ₄ cao nhất ở giếng H19) – TC, SO ₄ ↑ (khá cao ở giếng T15, T16, T17, V35, V36)
Cụm 4 (20 giếng / 6 vị trí)	– Thôn Trung Đon (T15, T16, T17) / xã Hải Thành – Thôn Thuận Nhơn (V35, V36, V45, V46, V47) / xã Hải Vĩnh – Thôn Phương Lang (B15, B16, B17, B19) / xã Hải Ba – Thôn Kim Long (Q15, Q16, Q17, Q19) / xã Hải Quế – Thôn An Thọ (H15, H16, H17, H18) / xã Hải Hoà	– TDS, Cl ⁻ ↑ (cao nhất ở H18) – Mn↑ (khá cao ở giếng B15, B17, B19, Q15, Q17, H16) – Fe, COD, NH ₄ , NO ₃ ↓

^(*) Số vị trí ở cột 1 thuộc 4 cụm bị trùng nhau, nên tổng số vị trí ở đây lớn hơn tổng số vị trí thực tế là 17; Các mũi tên thể hiện cao (↑), rất cao (↑↑) và thấp (↓) so với giá trị trung bình số học của thông số ($n = 57$): pH = 5,8; Tur = 3,3 NTU; TDS = 164 mg/L; COD = 1,9 mg/L; Cl⁻ = 33,2 mg/L; NH₄ = 0,13 mg/L; NO₂ = 0,03 mg/L; NO₃ = 0,38 mg/L; SO₄ = 12,4 mg/L; Fe = 0,36 mg/L; Mn = 0,1 mg/L; TC = 105 MPN/100 mL.

Kết quả ở Bảng 3 cho thấy đặc điểm CLN ở các cụm giếng không phân tách ra theo vị trí địa lý (xã), mà nhiều trường hợp, các giếng ở cùng một xã (trong những năm quan trắc khác nhau) lại thuộc các cụm khác nhau. Có thể cho rằng đặc điểm địa chất thủy văn tầng nông (giếng đào và giếng khoan) ở vùng khảo sát không như nhau. Từ kết quả ở Bảng 3, về mặt địa lý, có thể phân chia vùng này thành ba tiểu vùng CLN:

– **Tiểu vùng I** (gồm cụm 1 và 2): Tiểu vùng này, gồm 2 xã ở xa nhau – Hải Dương (6 vị trí, D1–D6) và Hải Vĩnh (6 vị trí, V1–V6), có đặc điểm CLN giếng gần tương tự nhau: Nồng độ các chất hữu cơ (COD), NO₃, NO₂ và TC cao hơn, nhưng pH, nồng độ muối (TDS, Cl⁻), Fe, Tur và NH₄ thấp hơn tiểu vùng II và III. Tuy cao hơn, nhưng nồng độ NO₃, NO₂ vẫn đạt yêu cầu của QCVN01 và QCVN09. Riêng nồng độ COD trong nhiều giếng ở xã Hải Dương (2–4 mg/L) và xã Hải Vĩnh (0,5–3 mg/L) không đạt QCVN01 (quy định ≤2 mg/L). Đáng chú

ý là pH khá thấp ở nhiều vị trí thuộc xã Hải Dương (pH = 4,5÷6,2 ở vị trí 2, 3, 4) và Hải Vĩnh (pH = 5,1÷6,3 ở cả 6 vị trí), không đạt yêu cầu của QCVN01, QCVN02 và nhiều trường hợp không đạt cả QCVN09.

– **Tiểu vùng II** (cụm 3): Tiểu vùng này gồm 2 xã gần nhau – Hải Thành (2 vị trí, T1 và T2) và Hải Hòa (1 vị trí là H1). Riêng vị trí T2 và H1 có Fe, độ đục và NH₄ cao hơn so với tiểu vùng I và III: Vị trí T2 có Tur (17–29 NTU) và Fe (1,88–3,21 mg/L) không đạt QCVN01 và QCVN02; Vị trí H1 (thôn An Thọ) có bất thường năm 2019 là Fe, Tur và NH₄ khá cao: Nồng độ Fe (2,04 mg/L) và Tur (24 NTU) không đạt QCVN01 và QCVN02; Nồng độ NH₄ (1,25 mg/L) không đạt QCVN09.

– **Tiểu vùng III** (cụm 4) gồm 2 xã gần nhau – xã Hải Ba và Hải Quế (2 vị trí, B1 và Q1). Ở tiểu vùng này, mức nhiễm muối (TDS, Cl⁻, SO₄) và Mn cao hơn, nhưng COD, NO₃ và TC thấp hơn so với tiểu vùng I và II. Tuy mức nhiễm muối cao, nhưng nồng độ TDS,

Cl⁻ và SO₄ vẫn đạt yêu cầu của QCVN01, QCVN02 và QCVN09; riêng nồng độ Mn (năm 2015, 2017 và 2019) ở xã Hải Ba (0,3–0,4 mg/L) và xã Hải Quế (0,3–0,5 mg/L) không đạt yêu cầu của QCVN01.

Đáng lưu ý là mức nhiễm các vi khuẩn có nguồn gốc phân trong hầu hết các giếng ở cả ba tiểu vùng khá cao (TC = 3±1.100 MPN/100 mL), không đạt yêu cầu của QCVN01, QCVN02 và QCVN09.

Cuối cùng, cần thấy rằng, do mỗi năm chỉ quan trắc một lần ở một vị trí, nên các kết quả thu được chưa cho phép nhận ra xu thế diễn biến CLN theo thời gian. Tuy nhiên, việc nhận biết các cụm giếng và các tiểu vùng có đặc điểm CLN như nhau ở trên sẽ tạo điều kiện thuận lợi cho việc lựa chọn các vị trí quan trắc và giám sát diễn biến CLN sau này. Mặt khác, để tìm hiểu nguyên nhân gây ra hiện tượng nhiễm phen (Fe), chất hữu cơ (COD), các hợp chất chứa nitơ (NH₄, NO₂, NO₃), nhiễm muối (TDS, Cl⁻, SO₄) và vi khuẩn có nguồn gốc phân (TC), cần phải điều tra, khảo sát chi tiết hơn về đặc điểm địa chất thủy văn và nguồn phát thải từ các hoạt động dân sinh, sản xuất công nghiệp, nông nghiệp và dịch vụ ở vùng khảo sát.

4 Kết luận

Hiện trạng CLN giếng và phân cụm/tiểu vùng theo đặc điểm CLN ở 6 xã vùng trung huyện Hải Lăng trong giai đoạn 2015–2019 đã được làm sáng tỏ bằng cách áp dụng hiệu quả các phương pháp phân tích thống kê, đặc biệt là phương pháp PCA và AHC. Ở xã Hải Dương và xã Hải Vĩnh, các giếng có mức nhiễm các chất hữu cơ (COD) cao hơn và pH thấp hơn so với ở các xã khác. Các giếng ở xã Hải Thành và xã Hải Hòa có mức nhiễm Fe, độ đục và NH₄ cao hơn, còn các giếng ở xã Hải Ba và xã Hải Quế bị nhiễm Mn và nhiễm muối (TDS,

Cl⁻, SO₄) cao hơn so với các xã khác. Hầu hết các giếng ở cả 6 xã khảo sát đều bị nhiễm các vi khuẩn có nguồn gốc phân và không đạt yêu cầu về CLN cấp cho ăn uống và sinh hoạt. Để đánh giá đầy đủ hơn về CLN giếng và các tác động ở vùng trung huyện Hải Lăng, cần thực hiện các nghiên cứu tiếp theo về CLN giếng trong toàn vùng trung (12 xã) và tiến hành điều tra, xác định các nguồn ô nhiễm (tự nhiên và nhân tạo) tác động đến CLN giếng trong vùng.

Tài liệu tham khảo

1. Hợp NV. Báo cáo đánh giá tác động môi trường Dự án “Chống lũ cho vùng trung Hải Lăng, tỉnh Quảng Trị”. Quảng Trị: Sở TNMT Quảng Trị; 2009.
2. Hợp NV. Hoạt động quản lý nguồn nước (thuộc Dự án “Thí điểm phục hồi đất lâm nghiệp bị ảnh hưởng chất độc da cam ở tỉnh Quảng Trị”. Quảng Trị: UBND tỉnh Quảng Trị; 2007.
3. Sơn NT. Quy hoạch quản lý, khai thác, sử dụng và bảo vệ tài nguyên nước dưới đất miền đồng bằng tỉnh Quảng Trị (báo cáo tổng kết đề tài). Quảng Trị: Sở TNMT Quảng Trị; 2008.
4. Trung tâm Nước sạch và Vệ sinh Môi trường Nông thôn. Bộ chỉ số theo dõi – đánh giá nước sạch và vệ sinh môi trường nông thôn năm 2015–2017 (báo cáo tổng kết nhiệm vụ). Quảng Trị: Sở NN&PTNT Quảng Trị; 2017.
5. Sở TNMT Quảng Trị – Trung tâm Quan trắc TNMT. Báo cáo tổng hợp kết quả quan trắc tài nguyên và môi trường tỉnh Quảng Trị năm 2018. Quảng Trị: Sở TNMT Quảng Trị; 2018.
6. Clesceri LS, Greenberg AE, Eaton AD. Standard methods for the examination of water and wastewater. 20th ed. USA: American Public Health Association; 1998.
7. Esbensen K, Midtgaard T, Schönkopf S. Multivariate Analysis in Practice. Camo AS. Trondheim; 1994.
8. Tuấn NV. Phân tích dữ liệu với R – Hỏi và đáp. Hồ Chí Minh: Nhà xuất bản Tổng hợp Thành phố Hồ Chí Minh; 2018.