

Đánh giá chất lượng nước sông Lá Buông bằng phương pháp thống kê đa biến theo không gian và thời gian

Trần Đức Dũng^{1*}, Nguyễn Quốc Quân¹, Nguyễn Thị Thanh Huệ¹, Phạm Luân¹

¹ Trung tâm Quản lý Nước và Biến đổi khí hậu, Viện Môi Trường và Tài Nguyên, Đại học Quốc gia TP HCM; dungtranducvn@yahoo.com; quocquannguyen1987@gmail.com; nguyen.tt.hue@gmail.com; hongluanosgeo@gmail.com

*Tác giả liên hệ: dungtranducvn@yahoo.com; Tel.: +84-902007905

Ban Biên tập nhận bài: 26/7/2021; Ngày phản biện xong: 24/8/2021; Ngày đăng bài: 25/11/2021

Tóm Tắt: Sông Lá Buông là chi lưu lớn nằm bên bờ hữu sông Đồng Nai, hoạt động kinh tế của lưu vực này khá đa dạng, từ các hoạt động nông nghiệp, chăn nuôi ở phía thượng lưu cho đến phát triển công nghiệp, dịch vụ ở phía hạ lưu. Tuy nhiên, chính sự phát triển kinh tế xã hội đã gây ra ô nhiễm nước mặt do các hoạt động công nghiệp và nông nghiệp. Mục tiêu của nghiên cứu này nhằm đánh giá chất lượng nước mặt sông Lá Buông, sử dụng phương pháp thống kê đa biến, phân tích các thành phần chính dựa trên các chỉ tiêu chất lượng nước theo không gian và thời gian. Số liệu quan trắc chất lượng nước giai đoạn 2010-2017 được chuẩn hóa và xử lý loại bỏ những giá trị bất thường sử dụng phép kiểm định Shapiro-Wilk và kiểm định phi tham số Wilcoxon signed-rank, trước khi được sử dụng để phân tích. Kết quả chỉ ra rằng nguồn nước sông Lá Buông trong cả mùa khô và mùa mưa bị ô nhiễm cục bộ các chất dinh dưỡng, vi sinh (E. coli, Coliform, N-NH₄) tại khu vực thượng nguồn do hoạt động chăn nuôi; ô nhiễm các chất vô cơ và hữu cơ (BOD₅, COD, Fe, N-NH₄, N-NO₂, TSS, độ đục) từ nước thải sinh hoạt và công nghiệp tại khu vực trung lưu và hạ lưu. Kết quả nghiên cứu giúp đánh giá hiện trạng chất lượng nước sông Lá Buông nhằm đưa ra các giải pháp góp phần bảo vệ nguồn nước mặt của toàn lưu vực hướng đến phát triển bền vững.

Từ khóa: Ô nhiễm; Quan trắc; Thống kê; Phân tích; Thành phần chính.

1. Đặt vấn đề

Đồng Nai là tỉnh có lượng mưa trung bình năm tương đối lớn, dao động trong khoảng 1700–2800 mm/năm [1]; dòng chính sông Đồng Nai có nguồn nước dồi dào, có khả năng đáp ứng các nhu cầu sử dụng nước phục vụ phát triển kinh tế - xã hội của tỉnh. Tuy nhiên, gần một nửa diện tích của tỉnh nằm xa dòng chính sông Đồng Nai, do đó khi nhu cầu sử dụng nước ngày càng lớn do dân số tăng và quá trình nông nghiệp và công nghiệp hóa ngày càng phát triển thì việc khai thác hợp lý nguồn nước mặt từ các sông suối lưu vực và bảo vệ môi trường rất cần thiết [2]. Lưu vực sông Lá Buông nằm trọn trong địa phận tỉnh Đồng Nai, bao gồm một số phường/xã thuộc thành phố Biên Hoà, thị xã Long Khánh và các huyện Long Thành, Trảng Bom, Thống Nhất, Cẩm Mỹ (Hình 1). Lưu vực này không những đóng vai trò lớn trong việc điều tiết, cung cấp nguồn nước cho các thành phố và huyện phụ thuộc, mà còn là một trong những vùng có tiềm năng phát triển kinh tế xã hội lớn của tỉnh về nhiều mặt. Là lưu vực có phân bố dân cư phù hợp để phát triển kinh tế xã hội, tỉnh Đồng Nai đã đầu tư xây dựng các khu công nghiệp tập trung quy mô lớn, và bố trí sản xuất nông nghiệp đạt hiệu quả kinh tế cao kết hợp với phát triển du lịch và dịch vụ [3]. Mặc dù vậy, môi trường nước mặt

trên lưu vực sông Lá Buông đang đứng trước nguy cơ ô nhiễm nghiêm trọng do những hoạt động phát triển trong những năm gần đây, đặc biệt là các đoạn sông chảy qua các vùng công nghiệp và khu dân cư tập trung.

Hiện nay đã có khá nhiều nghiên cứu sử dụng chỉ số đánh giá chất lượng nước (WQI), nhưng rất ít nghiên cứu phân tích đa biến về chất lượng nước sông Lá Buông theo không gian và thời gian để đánh giá tổng thể về hiện trạng môi trường nước trên sông. Việc phân tích theo hệ thống sẽ tạo cơ sở giúp đưa ra các biện pháp quản lý, bảo vệ nguồn nước trên lưu vực phù hợp với điều kiện phát triển bền vững không những cho các huyện vùng thuộc lưu vực, mà còn cho sự phát triển chung của tỉnh Đồng Nai [4]. Nghiên cứu này sử dụng phương pháp phân tích thống kê đa biến mà cụ thể là phân tích thành phần chính (PCA), dựa trên chuỗi số liệu quan trắc giai đoạn 2010–2017 để xác định các thành phần chính tác động đến chất lượng nước từ mười lăm thông số quan trắc, bao gồm nhiệt độ, pH, độ đục, EC, DO, TSS, COD, BOD₅, N-NH₄, N-NO₂, N-NO₃, P-PO₄, Fe, E. Coli, và Coliform. Cùng với kết quả phân tích, nhóm nghiên cứu cũng dựa trên thông tin lược khảo tài liệu và quan sát khảo sát thực tế, để giải thích những nguồn ô nhiễm thực tế, và xác định các thành phần chính ảnh hưởng đến chất lượng nước mặt trên sông.

2. Phương pháp nghiên cứu và thu thập dữ liệu

2.1. Giới thiệu khu vực nghiên cứu

Sông Lá Buông với tổng chiều dài khu vực nghiên cứu khoảng hơn 56 km, được chia làm hai phần, (i) khu vực không có ảnh hưởng triều, dòng chảy 1 chiều, khu vực thượng lưu cầu sông Lá Buông (Quốc lộ 51) và (ii) khu vực có ảnh hưởng triều, khu vực hạ lưu cầu sông Lá Buông, kéo dài đến ngã ba hợp lưu với sông Đồng Nai. Sông Lá Buông có nguồn từ suối Đá Bàn trên cao nguyên An Lộc (Long Khánh) ở độ cao hơn 200 m, và vùng rừng núi thuộc nông trường Ông Quế (Hình 1). Ở thượng nguồn hai suối Gia Dách và suối Cam hợp với nhau thành suối Nhận, xuôi về trung lưu suối Nhận được suối Sáu và suối Cải Hao nhập vào trở thành sông Nhận tại khu vực xã An Viễn. Ở hạ lưu sau khi nhập với rạch Bến Gỗ ở Long Hưng chảy qua ấp Phước Chân nhập vào sông Đồng Nai ở ngã ba Láng Lùn, xã Tam Phước. Đoạn từ An Viễn về đến cửa sông có tên là Lá Buông hoặc được gọi khác là sông Buông. Dòng chính của sông có chiều dài từ nguồn đến cửa sông là 52 km, chảy theo hướng Đông sang Tây, với diện tích lưu vực: khoảng 264 km², trải dài trong miền đón gió mùa Tây Nam, lượng mưa phong phú, trung bình 1800 mm/năm, độ dốc bình quân 5,3‰, hàng năm cung cấp cho sông Đồng Nai lượng nước khoảng $0,23 \times 10^9$ m³. Môđun dòng chảy bình quân năm $M = 28,3$ l/s/km² [5].

2.2. Dữ liệu sử dụng

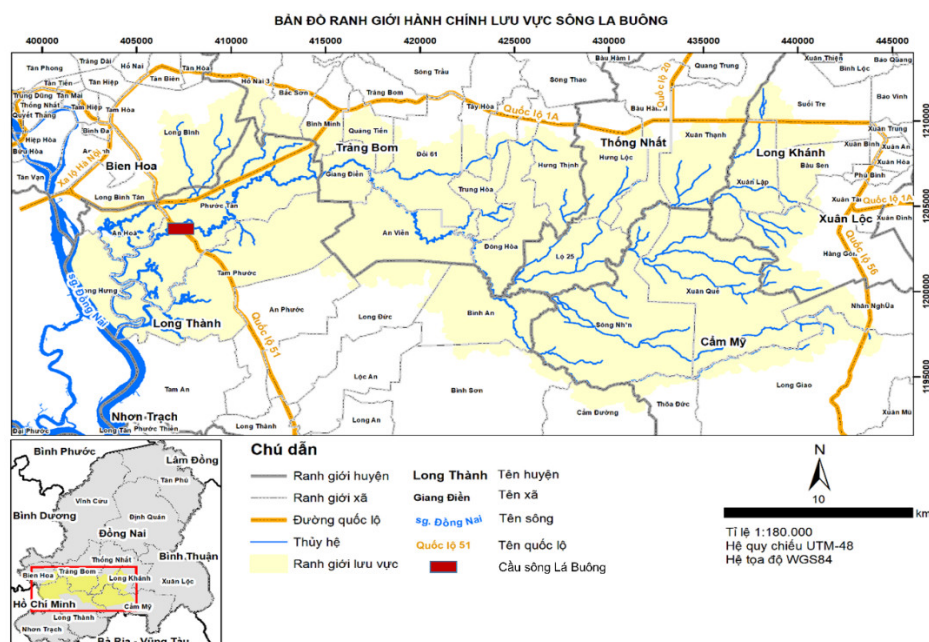
Thu thập dữ liệu quan trắc chất lượng nước mặt giai đoạn 2010–2017 tại 7 trạm quan trắc tự động dọc theo lưu vực sông Lá Buông (Hình 2). Cụ thể trình bày trong Bảng 1.

Bảng 1. Vị trí và thời gian lấy mẫu chất lượng nước

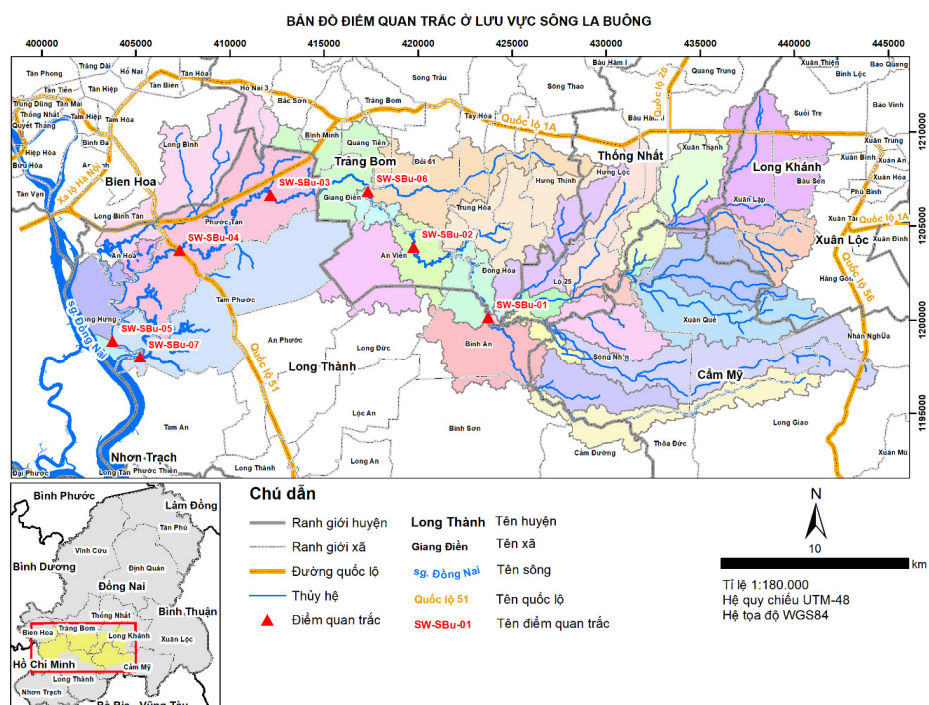
STT	Vị trí lấy mẫu	Thời gian lấy mẫu (năm)
1	SW-SBu-01	2010–2017
2	SW-SBu-02	2010–2017
3	SW-SBu-03	2010–2013
4	SW-SBu-04	2010–2017
5	SW-SBu-05	2012–2013
6	SW-SBu-06	2012–2017
7	SW-SBu-07	2016–2017

Trong đó:

- Năm 2010: lấy mẫu 6 đợt (tháng 2, 4, 6, 7, 12);
- Năm 2011, 2012, 2013: lấy mẫu 6 đợt (tháng 2, 4, 6, 8, 10, 12);
- Năm 2014: lấy mẫu 5 đợt (tháng 2, 4, 6, 8, 10);
- Năm 2015: lấy mẫu 6 đợt (tháng 1, 2, 4, 6, 8, 10);
- Năm 2016: lấy mẫu 6 đợt (tháng 2, 4, 6, 8, 10, 12);
- Năm 2017: lấy mẫu 2 đợt (tháng 2, 4). Thông số giám sát: pH, độ đục, EC, TSS, DO, BOD₅, COD, N-NH₄⁺, Độ mặn, N-NO₂⁻, N-NO₃⁻, P-PO₄³⁻, As, Pb, Zn, Fe, Tổng dầu mỡ, Phenol, Aldrin, Endosulfan, Heptachlor, Paration, E. Coli và Coliform.



Hình 1. Lưu vực sông La Bung.



Hình 2. Vị trí các trạm quan trắc.

2.3. Phương pháp phân tích

Hiện nay có rất nhiều phương pháp sử dụng để phân tích thành phần chính (PCA) [6–9]. Trong đó, để thực hiện PCA thì cần nhiều phương pháp thống kê xử lý số liệu. Nghiên cứu này sử dụng các phương pháp khác nhau trong phân tích số liệu, có thể phân thành 04 giai đoạn:

Giai đoạn 1: Phân tích thống kê mô tả với các tính toán giá trị trung bình và độ lệch chuẩn của hàm lượng các thông số chất lượng nước trong 2 mùa mưa (các tháng 6, 7, 8, 10) và khô (các tháng 12, 1, 2, 4) được thực hiện sau khi đã loại bỏ các giá trị ngoại lai/dị biệt (outliers) trong dữ liệu của từng thông số sử dụng phương pháp Khoảng tứ phân vị (IQR) [10–11]. Giai đoạn này cũng thực hiện chuẩn hóa dữ liệu để chuẩn bị cho giai đoạn 2.

Giai đoạn 2: Để lựa chọn phép kiểm định phù hợp cho việc đánh giá ý nghĩa thống kê của sự khác biệt chất lượng nước theo từng thông số giữa hai mùa, trước hết cần kiểm định phân phối chuẩn của dữ liệu chất lượng nước trong 2 mùa bằng cách sử dụng phép kiểm định Shapiro–Wilk. Sau đó, phép kiểm định phi tham số Wilcoxon signed–rank được sử dụng để so sánh và đánh giá ý nghĩa thống kê của sự khác biệt chất lượng nước theo từng thông số giữa hai mùa [12–13].

Giai đoạn 3: Phân tích tương quan Pearson, ma trận tương quan và các tính toán trị số p để đánh giá ý nghĩa thống kê của các tương quan được thực hiện để xác định sự tương quan giữa các thông số chất lượng nước trong mùa khô và mùa mưa [14].

Giai đoạn 4: Phân tích đa biến được sử dụng để xác định mối quan hệ giữa dữ liệu chất lượng nước tại 7 vị trí quan trắc. Nghiên cứu này áp dụng phân tích thành phần chính (PCA) nhằm xác định mối quan hệ không gian và thời gian của các biến chất lượng nước. Phần này sẽ được trình bày chi tiết trong nội dung bài báo.

Phân tích thành phần chính là một kỹ thuật phân tích biến đổi được nhiều nghiên cứu ứng dụng. Thuật toán này lần đầu tiên được giới thiệu bởi Pearson [15]. Ý tưởng chính của thuật toán phân tích thành phần chính là giảm chiều của tập dữ liệu có sự tương quan lớn giữa các biến. Bản chất của phương pháp phân tích thành phần chính là một thuật toán thống kê toán học nhằm biến đổi tập dữ liệu đa biến tương quan vào trong một tập dữ liệu đa biến không tương quan – còn được gọi là các thành phần chính.

Phân tích PCA là một kỹ thuật giảm biến, được sử dụng để hiểu các mối quan hệ cơ bản giữa các biến. Việc giảm biến sẽ giúp trích xuất thông tin về các thông số có ý nghĩa nhất mô tả toàn bộ tập dữ liệu mà không làm mất thông tin gốc. Kết quả phân tích PCA cung cấp các thông tin mà các phân tích thống kê mô tả đơn giản không thể cung cấp, trong đó quan trọng nhất là các xu hướng chính hiện diện trong ma trận dữ liệu. Kỹ thuật PCA liên quan đến việc biến đổi các biến tương quan ban đầu trong dữ liệu gốc thành các thành phần chính trực giao không tương quan (PC), trong đó các thành phần chính (PC) là sự kết hợp tuyến tính của các biến tương quan ban đầu như được trình bày trong công thức bên dưới:

$$PC_i = a_{1i}X_1 + a_{2i}X_2 + \dots + a_{ni}X_n \quad (1)$$

Trong đó PC_i là thành phần chính i và a_{ji} là tải thành phần (hệ số tương quan) của biến ban đầu x .

Tổng số PC được trích xuất từ PCA cho biết tổng số các nguồn biến đổi có thể có trong tập dữ liệu. PC đầu tiên có giá trị riêng cao nhất và đại diện cho nguồn biến đổi quan trọng nhất trong bộ dữ liệu. PC cuối cùng ít biến động nhất trong bộ dữ liệu. Dữ liệu đầu vào của PCA được chuẩn hóa về thang 0–1 để đồng bộ giữa biến (thông số chất lượng nước).

2.4. Phần mềm hỗ trợ phân tích số liệu

Dữ liệu ban đầu được tổng hợp bằng phần mềm Excel (Microsoft Office). Phần mềm R phiên bản 4.0.4 được sử dụng để phân tích dữ liệu theo các hình ảnh phân tích đa biến và

thành phần chính, kết hợp bảng biểu. Phần mềm ArcGIS 10.3 được sử dụng phân tích dữ liệu không gian và xây dựng các bản đồ.

3. Kết quả và thảo luận

3.1. Thống kê mô tả các thông số chất lượng nước trong mùa khô và mùa mưa

Do các thông số As, Pb, Zn, Dầu mỡ, Phenol, Heptachlor có các giá trị quan trắc thấp hơn rất nhiều lần so với quy chuẩn chất lượng nước mặt (QCVN 08-MT:2015/BTNMT) cột A2 (dùng cho mục đích cấp nước sinh hoạt nhưng phải dùng công nghệ xử lý phù hợp hoặc dùng cho các mục đích khác với yêu cầu chất lượng nước thấp) các thông số Endrin, Endosulfan, Paration có các giá trị quan trắc rất thấp và không được quy định giá trị ngưỡng trong quy chuẩn, thông số Độ mặn có các giá trị quan trắc dao động trong khoảng hàm lượng bình thường đối với nước mặt (0,5‰ hoặc thấp hơn [5]) nên các thông số này không được phân tích thêm trong phần này. Kết quả tính toán được trình bày trong Bảng 2.

Bảng 2. Giá trị trung bình và độ lệch chuẩn của hàm lượng các thông số chất lượng nước trong hai mùa.

STT	Thông số	Đơn vị	Giá trị trung bình \pm độ lệch chuẩn	
			Mùa khô	Mùa mưa
1	Nhiệt độ	°C	28,431 \pm 1,641	28,925 \pm 1,327
2	pH	-	7,055 \pm 0,453	6,96 \pm 0,381
3	Độ đục	NTU	14,994 \pm 12,913	59,378 \pm 43,791
4	EC	μ S/cm	155,845 \pm 48,66	166,496 \pm 41,59
5	DO	mg/l	5,684 \pm 1,69	6,198 \pm 1,118
6	TSS	mg/l	19,365 \pm 18,879	67,101 \pm 50,702
7	COD	mg/l	13,423 \pm 7,432	18,471 \pm 8,12
8	BOD ₅	mg/l	4,936 \pm 2,21	5,848 \pm 2,486
9	N-NH ₄	mg/l	0,543 \pm 0,569	0,435 \pm 0,386
10	N-NO ₂	mg/l	0,141 \pm 0,15	0,231 \pm 0,236
11	N-NO ₃	mg/l	1,44 \pm 0,822	1,937 \pm 1,441
12	P-PO ₄	mg/l	0,451 \pm 0,285	0,471 \pm 0,222
13	Fe	mg/l	1,458 \pm 0,805	6,665 \pm 5,074
14	E. coli	MPN/100 ml	1622,936 \pm 2490,5	1943,341 \pm 2137,025
15	Coliform	MPN/100 ml	11496,54 \pm 14533,5	14485,22 \pm 11467,58

Kết quả trong Bảng 2 cho thấy một số thông số có giá trị trung bình không chênh lệch nhiều giữa hai mùa trong khi đó một số thông số khác lại có các giá trị trung bình biến động mạnh giữa hai mùa (Độ đục, TSS, COD, Fe, E. Coli, Coliform). Để xác định liệu những khác biệt này là do các yếu tố ngẫu nhiên của số liệu hay thật sự là do sự khác biệt trong chất lượng nước giữa hai mùa (do ảnh hưởng của nước mưa, các dòng chảy tràn, sự rửa trôi đất đá, các đặc trưng dòng chảy theo mùa, đặc tính nước thải theo mùa...) cần thực hiện đánh giá ý nghĩa thống kê của sự khác biệt chất lượng nước theo từng thông số giữa hai mùa như được trình bày sau đây.

3.2. Đánh giá ý nghĩa thống kê của sự khác biệt chất lượng nước theo từng thông số giữa mùa khô và mùa mưa

3.2.1. Kiểm định phân phối chuẩn

Kết quả kiểm định phân phối chuẩn của dữ liệu chất lượng nước theo từng thông số vào mùa khô và mùa mưa được thể hiện trong Bảng 3.

Bảng 3. Kết quả kiểm định phân phối chuẩn của dữ liệu chất lượng nước theo từng thông số trong 2 mùa.

STT	Thông số	Trị số p Mùa khô	Trị số p Mùa mưa
1	t°C	0,002025	0,565665
2	pH	0,200272	0,003076
3	Độ đục	5,48E-15	1,29E-06
4	EC	1,34E-05	0,003193
5	DO	0,000198	0,000156
6	TSS	1,62E-14	1,87E-07
7	COD	1,73E-05	0,000215
8	BOD ₅	1,11E-05	3,82E-08
9	N-NH ₄	2,64E-10	2,84E-12
10	N-NO ₂	1,06E-07	1,94E-09
11	N-NO ₃	1,78E-10	0,00011
12	P-PO ₄	3,5E-05	0,03442
13	Fe	2,88E-14	3,73E-06
14	E. coli	2,63E-14	5,84E-10
15	Coliform	1,25E-12	8,79E-08

Kết quả trong Bảng 3 cho thấy:

- Trong mùa khô chỉ có thông số pH có phân phối chuẩn với trị số $p > 0,05$; tất cả các thông số còn lại không có phân phối chuẩn do có trị số $p < 0,05$.
- Trong mùa mưa chỉ có thông số to có phân phối chuẩn với trị số $p > 0,05$; tất cả các thông số còn lại không có phân phối chuẩn do có trị số $p < 0,05$.

3.2.2. Đánh giá ý nghĩa thống kê của sự khác biệt chất lượng nước theo từng thông số giữa hai mùa

Vì hầu hết các thông số không có phân phối chuẩn trong cả hai mùa nên Phép kiểm định phi tham số Wilcoxon signed-rank được sử dụng để so sánh và đánh giá ý nghĩa thống kê của sự khác biệt chất lượng nước theo từng thông số giữa hai mùa. Kết quả đánh giá được thể hiện trong Bảng 4.

Bảng 4. Kết quả đánh giá ý nghĩa thống kê của sự khác biệt chất lượng nước theo từng thông số giữa hai mùa.

STT	Thông số	Số mẫu mùa khô	Số mẫu mùa mưa	Trị số p	Có ý nghĩa thống kê *
1	T °C	101	96	0,282223	
2	pH	101	96	0,162502	
3	Độ đục	100	83	1,4E-17	X

STT	Thông số	Số mẫu mùa khô	Số mẫu mùa mưa	Trị số p	Có ý nghĩa thống kê *
4	EC	97	96	0,860874	
5	DO	101	96	0,051756	
6	TSS	100	87	3,46E-14	X
7	COD	100	94	9,34E-06	X
8	BOD ₅	99	93	0,040997	X
9	N-NH ₄	88	92	0,319901	
10	N-NO ₂	99	95	0,016243	X
11	N-NO ₃	100	91	0,026428	X
12	P-PO ₄	99	95	0,26143	
13	Fe	101	87	2,43E-19	X
14	E. coli	93	91	0,000295	X
15	Coliform	93	95	4,14E-07	X

* Sau khi so sánh với giá trị $\alpha = 0,05$

X: giá trị có ý nghĩa thống kê

Theo kết quả kiểm tra trong Bảng 4, các thông số chất lượng nước có khác biệt có ý nghĩa thống kê giữa 2 mùa gồm có: Độ đục, (TSS), COD, BOD₅, Nitrit (N-NO₂), Nitrat (N-NO₃), Fe, E. Coli và Coliform (9 thông số). Đáng quan tâm, tất cả 9 thông số này đều có hàm lượng trong mùa mưa cao hơn mùa khô. Đặc biệt các thông số Độ đục, TSS và Fe mùa mưa tăng gấp 3,5 – 4,5 lần so với mùa khô. Như vậy, nhìn chung nước sông trong mùa mưa có chất lượng thấp hơn so với mùa khô.

Nguyên nhân của hiện tượng này có thể được giải thích là do đất đá trong lưu vực và các chất ô nhiễm phát sinh từ các hoạt động nông nghiệp tồn lưu trong đất bị nước mưa rửa trôi và cuốn theo chảy vào nước sông. Ngoài ra, thực tế cho thấy có nhiều nhà máy, cơ sở sản xuất công nghiệp và nông nghiệp hoạt động gần bờ các con sông lợi dụng lưu lượng dòng chảy lớn vào mùa mưa để xả trộm nước thải (chưa đạt chuẩn chất lượng xả thải) xuống sông làm cho nước sông vào mùa mưa ô nhiễm hơn mùa khô ([16–18]).

Các nghiên cứu [6, 19], cho thấy đối với hầu hết các thông số chất lượng nước có khác biệt có ý nghĩa thống kê giữa 2 mùa (mùa khô/mùa đông và mùa mưa/mùa hè) thì nước sông trong mùa mưa có hàm lượng các chất ô nhiễm cao hơn đáng kể so với mùa khô. Nghiên cứu [20] đã kết luận rằng TSS, N-NO₃, P-PO₄ và COD có tương quan thuận từ yếu đến trung bình với lượng mưa, trong khi đó DO, pH và EC có mối tương quan nghịch từ khá đến mạnh với tất cả các thông số thủy văn. Bên cạnh đó, một số nghiên cứu đã chỉ ra rằng lượng mưa lớn có thể làm tăng lượng vi sinh vật chỉ thị ô nhiễm vi sinh và mầm bệnh tiềm ẩn trong nguồn nước [21] và tác động của lượng mưa đối với sự gia tăng mức độ vi khuẩn chỉ thị là đáng kể trong tất cả các mùa [22].

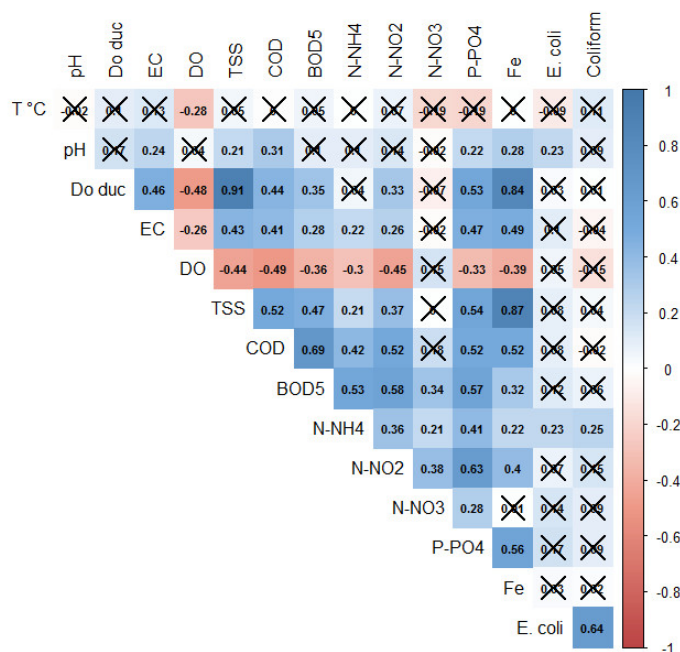
Các thông số không có khác biệt có ý nghĩa thống kê giữa 2 mùa: nhiệt độ (t°), pH, độ dẫn điện (EC), oxy hòa tan (DO), amoni (N-NH₄) và phosphat (P-PO₄) (6 thông số).

3.3 Phân tích tương quan

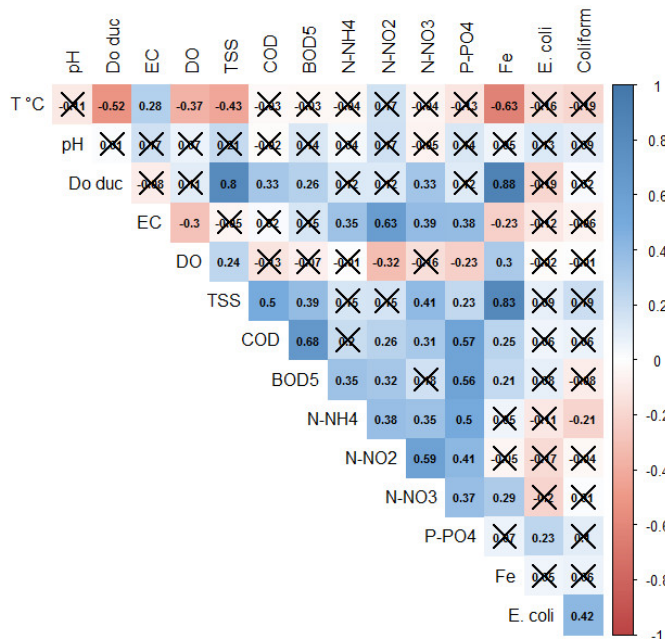
Trong phần này, phân tích tương quan Pearson, ma trận tương quan và các tính toán trị số p để đánh giá ý nghĩa thống kê của các tương quan được thực hiện để xác định sự tương quan giữa các thông số chất lượng nước trong mùa khô và mùa mưa. Kết quả phân tích được trình bày trong Hình 3 và Hình 4.

So sánh các hệ số tương quan giữa các thông số chất lượng nước trong hai mùa cho thấy trong mùa khô các hệ số tương quan có giá trị tuyệt đối lớn hơn và có nhiều tương quan có ý

nghĩa thống kê hơn so với mùa mưa. Từ đó có thể thấy các thông số chất lượng nước trong mùa khô có tương quan với nhau mạnh mẽ hơn trong mùa mưa.



Hình 3. Kết quả phân tích tương quan giữa các thông số chất lượng nước mùa khô. Ghi chú: Các tương quan không có ý nghĩa thống kê (trị số $p \geq 0,05$) được đánh dấu X.



Hình 4. Kết quả phân tích tương quan giữa các thông số chất lượng nước mùa mưa. Ghi chú: Các tương quan không có ý nghĩa thống kê ($p \geq 0,05$) được đánh dấu X.

Trong mùa khô:

- DO có tương quan nghịch có ý nghĩa thống kê với phần lớn các thông số chất lượng nước còn lại (TSS, Fe, P-PO₄, COD, Độ đục, BOD₅, N-NO₂, N-NH₄, t°), trong đó hệ số tương quan nghịch cao nhất là với COD (-0,49). Các phát hiện tương tự cũng được báo cáo bởi nghiên cứu [23] trong đó DO có tương quan nghịch có ý nghĩa thống kê với BOD₅, EC, N-NO₃ và P-PO₄ do nhiều quá trình phân hủy và chuyển hóa các chất ô nhiễm sử dụng và làm tiêu hao lượng oxy hòa tan trong nước.

- Phần lớn các thông số đại diện cho hàm lượng các chất ô nhiễm trong nước có tương quan thuận có ý nghĩa thống kê với nhau. Trong đó, các thông số có tương quan thuận cao nhất là TSS - Độ đục (0,91), TSS - Fe (0,87), Fe - Độ đục (0,84), COD - BOD₅ (0,69) và E. Coli - Coliform (0,64).

- Hai thông số chỉ thị cho khả năng ô nhiễm vi sinh trong nước mặt là E. Coli và Coliform có tương quan thuận có ý nghĩa thống kê với thông số ô nhiễm dinh dưỡng N-NH₄ với các hệ số tương quan lần lượt là 0,23 và 0,25. E. coli cũng có tương quan thuận có ý nghĩa thống kê với pH (0,23).

Trong mùa mưa:

- DO có tương quan nghịch có ý nghĩa thống kê với t^o, N-NO₂, EC, P-PO₄ và có tương quan thuận có ý nghĩa thống kê với Fe và TSS.

- Tương tự như trong mùa khô, các thông số chất lượng nước có tương quan thuận cao nhất là Fe-Độ đục (0,88), Fe-TSS (0,83), TSS-Độ đục (0,8), COD-BOD₅ (0,68), nhưng phần lớn giá trị các hệ số tương quan giảm so với mùa khô. Hệ số tương quan giữa E. Coli và Coliform cũng giảm mạnh xuống còn 0,42.

- E. Coli có tương quan thuận có ý nghĩa thống kê với P-PO₄ và Coliform có tương quan nghịch có ý nghĩa thống kê với N-NH₄.

3.4. Phân tích thành phần chính (Principle Component Analysis - PCA)

3.4.1. Mùa khô

a) Các thành phần chính

Sử dụng lược đồ Scree plot để xác định số thành phần chính của tập dữ liệu (Hình PL1 trong Phụ lục). Tổng phương sai của 4 thành phần chính đầu tiên (PC1, PC2, PC3, và PC4) là 63,51% của bộ dữ liệu được sử dụng để phân tích sâu hơn, từ thành phần chính thứ 5 trở đi có tỷ lệ phương sai nhỏ và thay đổi rất ít (dưới 10%). Giá trị riêng và % phương sai của từng thành phần chính và lược đồ Scree plot mùa khô được trình bày cụ thể trong bảng PL1 trong phần Phụ lục.

Các thông số chất lượng nước có ảnh hưởng chính trong thành phần chính thứ nhất (PC1) là: TSS, Fe, COD, P-PO₄, Độ đục, N-NO₂, BOD₅, N-NH₄, DO. Nhóm thông số chỉ thị cho sự ô nhiễm nước có nguồn gốc tự nhiên hoặc do con người tạo ra (nước thải sinh hoạt, nông nghiệp và công nghiệp) với sự có mặt của các chất ô nhiễm vô cơ và hữu cơ.

Các thông số E. Coli, Coliform, N-NH₄ đại diện cho ô nhiễm dinh dưỡng, và vi sinh đóng góp chủ yếu cho thành phần chính thứ 2 (PC2). Nước thải từ các hoạt động chăn nuôi gia súc gia cầm có chứa hàm lượng lớn phân động vật chưa qua hệ thống xử lý bị xả thải trực tiếp ra sông gây ô nhiễm phân, tạo điều kiện thuận lợi cho các loại vi sinh gây hại phát triển.

Nhiệt độ (t^o) và hàm lượng oxy hòa tan (DO) có mối tương quan nghịch với nhau là các thông số đóng góp chính cho PC3. Nghiên cứu đã cho thấy sự tăng nhiệt độ làm giảm sự hòa tan của oxy trong môi trường xung quanh vào nước sông [24].

Coliform là thông số chính đóng góp cho PC4 đại diện cho ô nhiễm vi sinh.

Bảng 5. Đóng góp theo % của các thông số chất lượng nước vào PC1, PC2, PC3 và PC4 trong mùa khô.

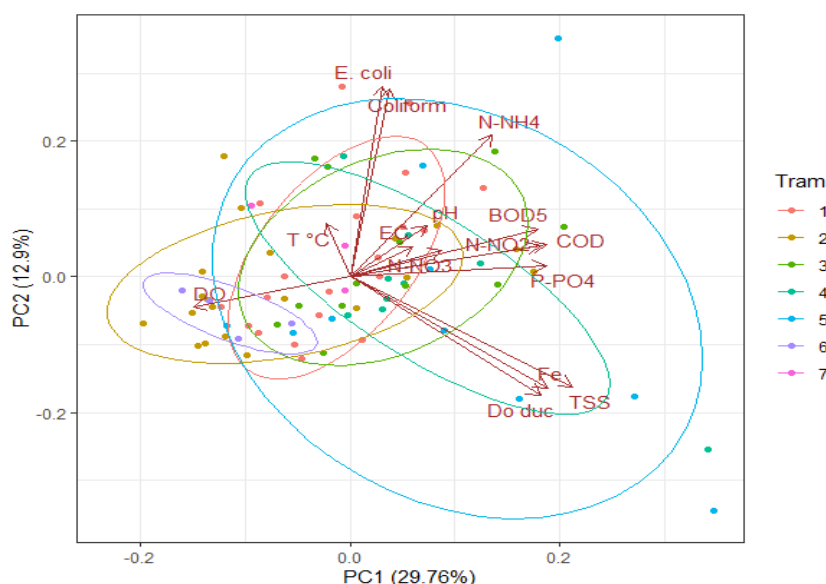
Thông số	PC1	PC2	PC3	PC4
t ^o	0,16	1,98	29,52	0,03
pH	1,74	1,81	12,31	3,03
Độ đục	10,70	9,84	2,10	10,38
EC	1,17	0,65	3,01	0,16
DO	7,27	0,66	23,45	0,40

Thông số	PC1	PC2	PC3	PC4
TSS	14,58	8,49	0,43	7,09
COD	11,42	0,71	0,06	8,25
BOD ₅	10,30	1,62	2,17	10,84
N-NH ₄	5,96	14,13	0,71	1,04
N-NO ₂	10,70	0,67	0,05	5,35
N-NO ₃	2,48	0,50	12,72	14,37
P-PO ₄	11,30	0,09	7,39	0,00
Fe	11,46	8,66	1,08	8,51
E. coli	0,31	25,53	3,63	13,41
Coliform	0,47	24,68	1,37	17,15

b) Đặc trưng ô nhiễm theo trạm quan trắc

Ứng với PC1 và PC2:

Đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC1 (các thông số đóng góp chính: TSS, Fe, COD, P-PO₄, Độ đục, N-NO₂, BOD₅, N-NH₄, DO) và PC2 (E. coli, Coliform, N-NH₄) trong mùa khô được thể hiện trong Hình 5.



Hình 5. Đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC1 và PC2 trong mùa khô.

Từ tỷ lệ đóng góp của các thông số chất lượng nước vào PC1, PC2, PC3, PC4 (Bảng 4) và đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC1 và PC2 (Hình 5) có thể nhận thấy chất lượng nước sông tại các trạm quan trắc có đặc trưng ô nhiễm không khác nhau nhiều. Tuy nhiên vẫn có những điểm cá biệt rõ ràng:

- Tại trạm SW-SBu-01 chất lượng nước đặc trưng bởi giá trị cao của các thông số E. coli, Coliform và N-NH₄ hơn các trạm khác. Trên thực tế khảo sát cho thấy gần các vị trí quan trắc chất lượng nước ở khu vực thượng lưu (trạm SW-SBu-01) có các hoạt động chăn nuôi quy mô hộ gia đình và quy mô công nghiệp, nước thải từ các hoạt động này đổ vào nguồn nước gây ô nhiễm dinh dưỡng và phát triển vi sinh có hại.

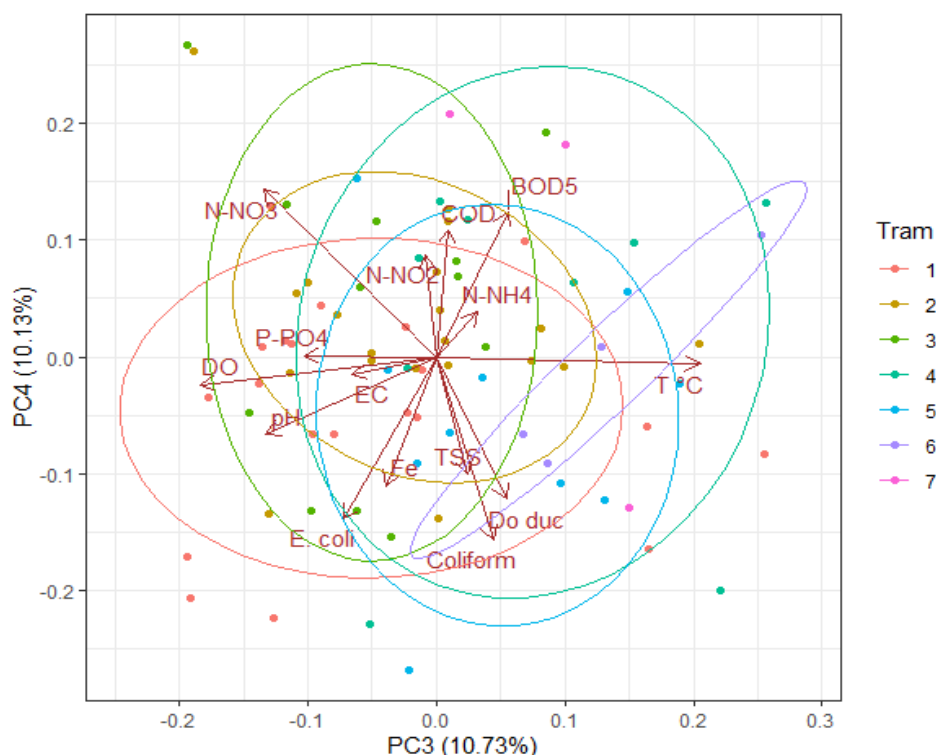
- Trạm SW-SBu-02 và SW-SBu-06 đặc trưng bởi DO. Khu vực xung quanh hai trạm này ít bị tác động bởi các hoạt động xả thải của con người. Tại trạm SW-SBu-02 chủ yếu trồng cao su, và tại trạm SW-SBu-06 là khu đất trồng giải tỏa. Điều này lý giải cho chất lượng nước tại đây ít bị ô nhiễm hơn các vị trí quan trắc khác.

- Trạm SW-SBu-03 và SW-SBu-04 đặc trưng bởi hầu hết các thông số chính trong PC1. Riêng trạm SW-SBu-05 đặc trưng bởi tất cả thông số trong PC1 và PC1. Trên thực tế khảo sát, chất lượng nước tại ba trạm này chịu tác động nhiều nhất do nằm trong khu vực có nhiều dân cư và hoạt động sản xuất công nghiệp, riêng khu vực trạm SW-SBu-05 còn có hoạt động khai thác đá, lý giải cho thông số độ đục và TSS cao tại đây.

- Trạm SW-SBu-07 không đủ dữ liệu nên không thể hiện được các phân tích chi tiết.

Ứng với PC3 và PC4:

Đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC3 (Nhiệt độ, DO) và PC4 (Coliform) trong mùa khô được thể hiện trong Hình 6. Có thể thấy chất lượng nước ở các trạm có ít khác biệt và không có đặc trưng rõ ràng vì PC3 và PC4 giải thích các tỷ lệ phần trăm rất nhỏ của tổng phương sai và các thông số chính của 2 thành phần này đã có mặt trong số các thông số chính của PC1 và PC2.



Hình 6. Đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC3 và PC4 trong mùa khô.

3.4.2. Mùa mưa

a) Các thành phần chính

Sử dụng biểu đồ Scree để xác định số thành phần chính của tập dữ liệu (Hình PL2 phần Phụ lục). Tổng phương sai của 4 thành phần chính đầu tiên (PC1, PC2, PC3, và PC4) là 70,21% của bộ dữ liệu được sử dụng để phân tích sâu hơn, từ thành phần chính thứ 5 trở đi có tỷ lệ phương sai nhỏ và thay đổi rất ít. Giá trị riêng và % phương sai của từng thành phần chính mùa khô được đính kèm trong Bảng PL2 phần Phụ lục.

So sánh với mùa khô cho thấy các thông số EC, BOD₅, N-NH₄, N-NO₃ có % đóng góp vào PC1 tăng lên nhiều vào mùa mưa, đặc biệt là N-NO₃. Nghiên cứu cho thấy tương đối ít Nitrat được tìm thấy trong nước tự nhiên có nguồn gốc từ khoáng chất mà hầu hết đến từ các nguồn hữu cơ như chất thải động thực vật và vô cơ chủ yếu là phân bón nhân tạo. Ở các con sông, hàm lượng Nitrat cao có nhiều khả năng là chỉ thị cho sự cuốn trôi đáng kể nitrat trong phân bón từ các khu vực đất nông nghiệp do nước mưa chảy tràn gây ra [25].

Các thông số đóng góp chính cho PC2 là Fe, Độ đục, TSS, EC, N-NO₂, nhiệt độ. Như vậy các thông số Fe, Độ đục, TSS, EC, N-NO₂ đóng góp chính cho cả 2 thành phần PC1 và

PC2 cho thấy sự ảnh hưởng mạnh mẽ của các yếu tố liên quan các chất rắn lơ lửng trong nước sông gây ra do các dòng chảy tràn cuốn trôi vật chất trong lưu vực và lưu lượng dòng chảy lớn trong mùa mưa. TSS cũng có thể là chỉ thị cho việc xả nước rửa từ các hố cát, mỏ đá hoặc hầm mỏ đã làm giảm sự xâm nhập của ánh sáng trong vùng nước mặt và cản trở đời sống thực vật thủy sinh. Những hoạt động này phá hoại nghiêm trọng vùng nước đánh bắt và có thể ảnh hưởng đến đời sống của cá, có thể hình thành các chất lắng đọng trên lòng sông và làm phát sinh các tình trạng tự hoại, và liên quan sự hiện diện của các dòng nước thải không đạt yêu cầu [25].

PC3 được đóng góp chủ yếu bởi các thông số E. coli và Coliform đặc trưng cho ô nhiễm vi sinh, trong khi đó DO và Coliform là các thông số chính đóng góp cho PC4.

Bảng 6. Đóng góp theo % của các thông số chất lượng nước vào PC1, PC2, PC3 và PC4 trong mùa mưa.

Thông số	PC1	PC2	PC3	PC4
t°	1,35	20,03	2,24	1,38
pH	0,4	0,49	17,03	8,36
Độ đục	8,65	13,08	3,46	1,96
EC	6,15	10,29	1,71	4
DO	0,01	6,7	0,79	34,95
TSS	11,86	10,16	0	0,59
COD	11,53	1,37	0,32	3,4
BOD ₅	12,1	2,23	0	1,96
N-NH ₄	7,94	2,33	0,38	6,82
N-NO ₂	9,26	9,15	0,59	1,44
N-NO ₃	10,58	0,27	0,51	9,4
P-PO ₄	11,93	2,23	6,88	2,6
Fe	7,1	19,92	1,41	0,38
E. coli	1,02	1,24	35,31	1,99
Coliform	0,13	0,5	29,38	20,76

b) Đặc trưng ô nhiễm theo trạm quan trắc

Ứng với PC1 và PC2:

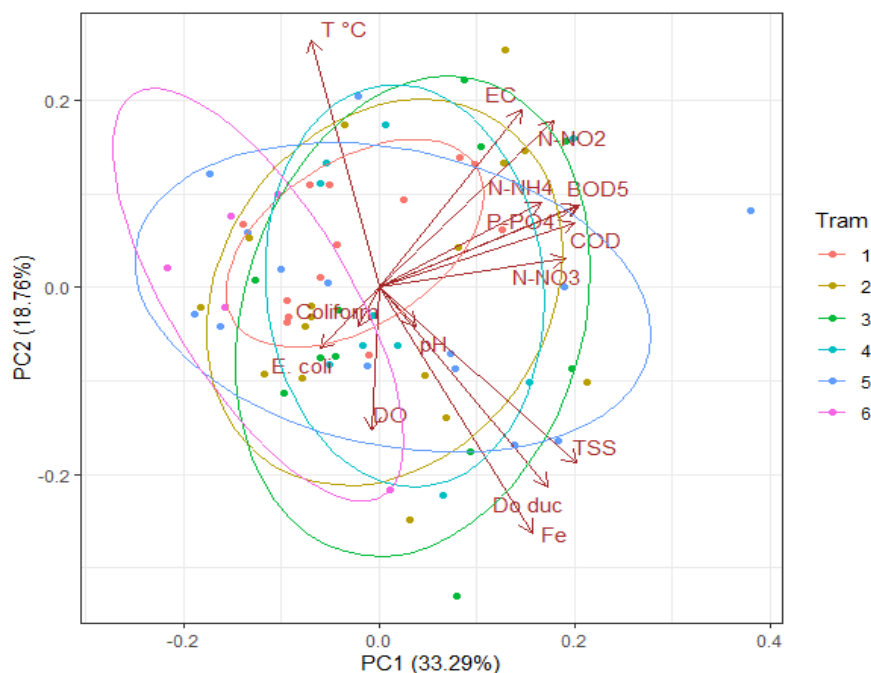
Đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC1 (các thông số đóng góp chính: BOD₅, TSS, P-PO₄, COD, N-NO₃, N-NO₂, Độ đục, Fe, N-NH₄) và PC2 (Nhiệt độ, Fe, Độ đục, TSS, EC, N-NO₂) trong mùa mưa được thể hiện trong Hình 7.

Từ tỷ lệ đóng góp của các thông số chất lượng nước vào PC1, PC2, PC3, PC4 (Bảng 5) và đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC1 và PC2 (Hình 7) có thể nhận thấy chất lượng nước ở các trạm có ít khác biệt và đặc trưng riêng.

Tuy nhiên, trạm SW-SBu-01 bị ô nhiễm BOD₅, P-PO₄, COD, N-NO₃, N-NH₄ (PC1) cũng như Fe, Độ đục, TSS (PC2) ít hơn so với các trạm khác.

Các trạm SW-SBu-02, SW-SBu-03, SW-SBu-04 khá giống nhau và bị ô nhiễm bởi các thông số ở cả 2 thành phần PC1 và PC2, tuy nhiên các thông số chính trong PC2 có nhiều ảnh hưởng hơn, và trạm SW-SBu-03 có phạm vi giá trị các thông số ô nhiễm cao hơn 2 trạm còn lại.

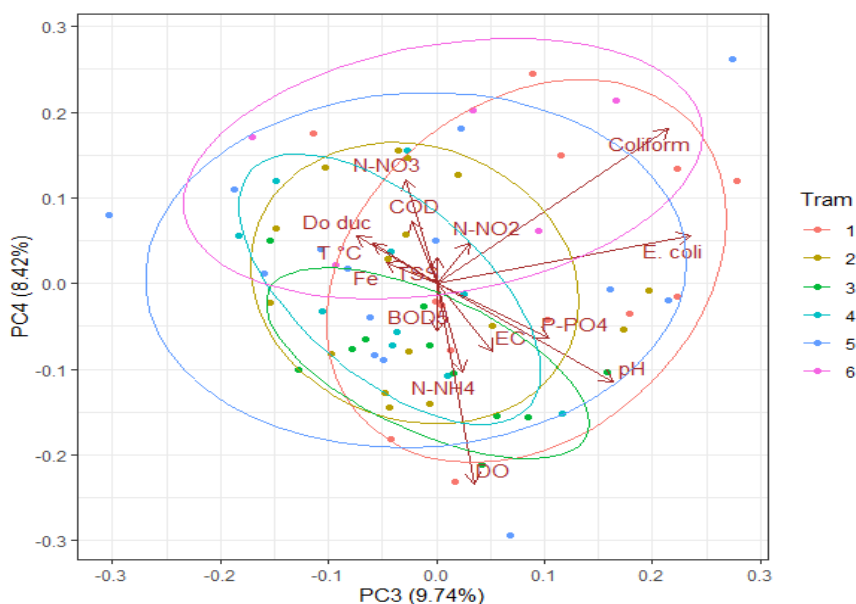
Chất lượng nước tại trạm SW-SBu-05 cũng bị ô nhiễm bởi các thông số ở cả 2 thành phần PC1 và PC2, tuy nhiên PC1 (trục hoành) có ảnh hưởng nhiều hơn. Như vậy, trong mùa mưa có nhiều trạm bị ảnh hưởng bởi Fe, Độ đục, TSS hơn và đặc tính ô nhiễm ở các trạm có nhiều điểm tương đồng hơn so với mùa khô.



Hình 7. Đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC1 và PC2 trong mùa mưa.

Ứng với PC3 và PC4:

Đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC3 (thông số đóng góp chính: E. coli, Coliform) và PC4 (DO, Coliform) trong mùa mưa được thể hiện trong Hình 8. Trạm SW-SBu-01 có nhiều mẫu nước bị ô nhiễm Coliform và E. coli hơn các trạm còn lại, tiếp theo là trạm SW-SBu-05 và các trạm SW-SBu-02, SW-SBu-02, và SW-SBu-04.



Hình 8. Đặc điểm chất lượng nước tại các trạm quan trắc ứng với PC3 và PC4 trong mùa mưa.

4. Kết luận

Kết quả phân tích thống kê cho thấy nước sông Lá Buông trong mùa mưa có chất lượng thấp hơn so với mùa khô. Các thông số chất lượng nước có khác biệt về ý nghĩa thống kê giữa 2 mùa, bao gồm 9 thông số độ đục, tổng chất rắn lơ lửng (TSS), COD, BOD₅, Nitrit (N-NO₂), Nitrat (N-NO₃), Fe, E. coli và Coliform. Đáng chú ý, các thông số độ đục, TSS và Fe

mùa mưa cao gấp 3,5 – 4,5 lần so với mùa khô. Kết quả tính toán khá tin cậy nhờ sử dụng các phương pháp tổng hợp và phân tích theo hệ thống mặc dù nguồn dữ liệu đo đạc chưa liên tục và đồng bộ ở một số thời điểm giữa các năm giai đoạn 2010-2017.

Kết quả phân tích thành phần chính cho thấy các thông số chất lượng nước có ảnh hưởng chính trong Thành phần chính 1 (PC1) khá tương đồng với nhau giữa 2 mùa, bao gồm độ đục, EC, TSS, COD, BOD₅, N-NH₄, N-NO₂, N-NO₃, P-PO₄, Fe, DO mặc dù thứ tự độ lớn % đóng góp vào PC1 của các thông số là khác nhau. Ngoài ra, các kết quả tính toán và phân tích cho thấy ảnh hưởng đáng kể của hoạt động canh tác nông nghiệp trồng trọt đến chất lượng nước sông trong mùa mưa dưới tác động của các dòng chảy tràn (được chỉ thị chính bởi thông số N-NO₃ với sự tăng mạnh % đóng góp vào PC1 trong mùa mưa so với mùa khô). Trong mùa khô, các thông số đại diện cho ô nhiễm dinh dưỡng, và vi sinh (E. coli, Coliform, N-NH₄) đóng góp chủ yếu cho PC2. Nhiệt độ và DO là các thông số đóng góp chính cho PC3 và Coliform là thông số chính đóng góp cho PC4. Trong mùa mưa, đặc điểm PC2 (được đóng góp chính bởi Fe, Độ đục, TSS, EC, N-NO₂, Nhiệt độ) cho thấy sự ảnh hưởng mạnh mẽ của các yếu tố liên quan các chất rắn lơ lửng trong nước sông gây ra do các dòng chảy tràn cuốn trôi vật chất trong lưu vực và lưu lượng dòng chảy lớn trong mùa mưa. PC3 được đóng góp chủ yếu bởi các thông số E. coli và Coliform. DO và Coliform là các thông số chính đóng góp cho PC4. Nghiên cứu có thể cần được bổ sung áp dụng nếu có đủ các nguồn dữ liệu liên tục được đo đạc trên sông Lá Buông trong những năm sắp tới, đồng thời có thể tham khảo phương pháp thực hiện cho các nghiên cứu ở những lưu vực sông tương tự.

Nghiên cứu áp dụng phương pháp PCA tĩnh cho bộ số liệu chuỗi quan trắc chất lượng nước theo thời gian còn hạn chế và rời rạc giữa các trạm, dựa trên giả thiết giữa các điểm dữ liệu không có mối tương quan theo thời gian. Cách tiếp cận này do đó đã bỏ qua thông tin quan trọng để tính toán và đánh giá hiệp phương sai theo thời gian. Nhóm nghiên cứu đề nghị áp dụng các phương pháp đa biến động như “Functional PCA” hoặc “Dynamic Factor Analysis” trong các nghiên cứu tương lai, với số liệu chất lượng nước đầy đủ hơn, đồng đều và liên tục theo thời gian.

Đóng góp của tác giả: Trần Đức Dũng và Nguyễn Quốc Quân viết chính bài báo, Nguyễn Thị Thanh Huệ và Phạm Luân phân tích số liệu và diễn giải kết quả.

Lời cảm ơn: Nghiên cứu được Đại học Quốc Gia TP.HCM cấp kinh phí thực hiện theo đề tài C2020-24-09/HĐ – KHCN.

Lời cam đoan: Các tác giả tuyên bố rằng không có xung đột lợi ích.

Tài liệu tham khảo

1. Sở Kế Hoạch và Đầu Tư tỉnh Đồng Nai, ‘Địa hình, đất đai, khí hậu, dân số tỉnh Đồng Nai, 2020. Available: =<http://dpidongnai.gov.vn/Pages/gioithieu.aspx?CatID=24>.
2. Than, N.H.; Ly, C.D.; Tat, P.V. The performance of classification and forecasting Dong Nai River water quality for sustainable water resources management using neural network techniques. *J. Hydrol.* **2021**, *596*, 126099.
3. Khoi, D.N.; Nguyen, V.T.; Sam, T.T.; Nhi, P.T.T. Evaluation on effects of climate and land-use changes on streamflow and water quality in the La Buong River Basin, Southern Vietnam. *Sustain.* **2019**, *11*, 24.
4. Nguyen, H.D.; Quan, N.H.; Quang, N.X.; Hieu, N.D.; Thang, L.V. Spatio-temporal pattern of water quality in the Saigon-Dong Nai river system due to waste water pollution sources. *Int. J. River Basin Manag.* **2021**, *19*(2), 221–243.
5. Longley, W. Freshwater Inflows - Ecological Relationships and Methods for Determination of Needs, 1994.
6. Sharma, M.; Kansal, A.; Jain, S.; Sharma, P. Application of multivariate statistical

- techniques in determining the spatial temporal water quality variation of Ganga and Yamuna rivers present in Uttarakhand state, India. *Water Qual. Exposure Health* **2015**, 7, 567–581.
7. 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), 579–599.
8. Tripathi, M.; Singal, S.L. Use of Principal Component Analysis for parameter selection for development of a novel Water Quality Index: A case study of river Ganga India. *Ecol. Indic.* **2019**, 96, 430–436.
9. Zeinalzadeh, K.; Rezaei, E. Determining spatial and temporal changes of surface water quality using principal component analysis. *J. Hydrol. Reg. Stud.* **2017**, 13, 1–10.
10. Kessing, L.V. et al., Lithium in drinking water and the incidence of bipolar disorder: A nation-wide population-based study. *Bipolar Disord.* **2017**, 19(7), 563–567.
11. Okeowo, M.A.; Lee, H.; Hossain, F.; Getirana, A. Automated Generation of Lakes and Reservoirs Water Elevation Changes From Satellite Radar Altimetry. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, 10(8), 3465–3481.
12. Sultan, A.A.; Mashrei, M.A.; Washer, G.A. Utilization of Wilcoxon-Mann-Whitney statistics in assessing the reliability of nondestructive evaluation technologies. *Structures* **2020**, 27, 780–787.
13. Montocchio, D.; Chow-Fraser, P. Influence of water-level disturbances on the performance of ecological indices for assessing human disturbance: A case study of Georgian Bay coastal wetlands. *Ecol. Indic.* **2021**, 127, 107716.
14. Sudhakaran, S.; Mahadevan, H.; Arun, V.; Krishnakumar, A. P.; Krishnan, K.A. A multivariate statistical approach in assessing the quality of potable and irrigation water environs of the Netravati River basin (India). *Groundw. Sustain. Dev.* **2020**, 11, 100462.
15. Shlens, J. A Tutorial on Principal Component Analysis, February, 2014.
16. Nam, K. Nhiều cơ sở sản xuất xả nước thải trực tiếp ra sông Dương Đông, Báo tuổi trẻ, 2019.
17. Chi, T. Xử lý nghiêm vụ xả nước thải ra sông Trà Khúc của Nhà máy cồn - rượu Quảng Ngãi, Báo Nhân dân, 2010.
18. Nam, P. Vedan Bình Thuận gây ô nhiễm sông Phan. Tin tức pháp luật, 2008.
19. Shrestha, S.; Kazama, F.; Nakamura, T. Use of principal component analysis, factor analysis and discriminant analysis to evaluate spatial and temporal variations in water quality of the Mekong River. *J. Hydroinf.* **2008**, 10(1), 43–56.
20. Prathumratana, L.; Sthiannopkao, S.; Kim, K.W. The relationship of climatic and hydrological parameters to surface water quality in the lower Mekong River. *Environ. Int.* **2008**, 34(6), 860–866.
21. Tryland, I.; Robertson, L.; Blankenberg, A.G.B.; Lindholm, M.; Rohrlack, T.; Liltved, H. Impact of rainfall on microbial contamination of surface water. *Int. J.*

Clim. Chang. Strateg. Manag. **2011**, 3(4), 361–373.

22. Tornevi, A.; Bergstedt, O.; Forsberg, B. Precipitation effects on microbial pollution in a river: Lag structures and seasonal effect modification. *PLoS One*, **2014**, 9(5), e98546.
23. Khatoon, N. Correlation Study For the Assessment of Water Quality and Its Parameters of Ganga River, Kanpur, Uttar Pradesh, India. *IOSR J. Appl. Chem.* **2013**, 579(3), 2278–5736.
24. Kumari, M.; Tripathi, S.; Pathak, V.; Tripathi, B.D. Chemometric characterization of river water quality. *Environ. Monit. Assess.* **2013**, 185(4), 3081–3092.
25. Robert Ongom, M.A.; Lukubye, B. Physico-Chemical Quality of Lake Kyoga at Selected Landing Sites and Anthropogenic Activities. *J. Water Resour. Prot.* **2017**, 9(11), 1225–1243.

Assessment of water quality on La Buong River using spatio-temporal principal component analysis

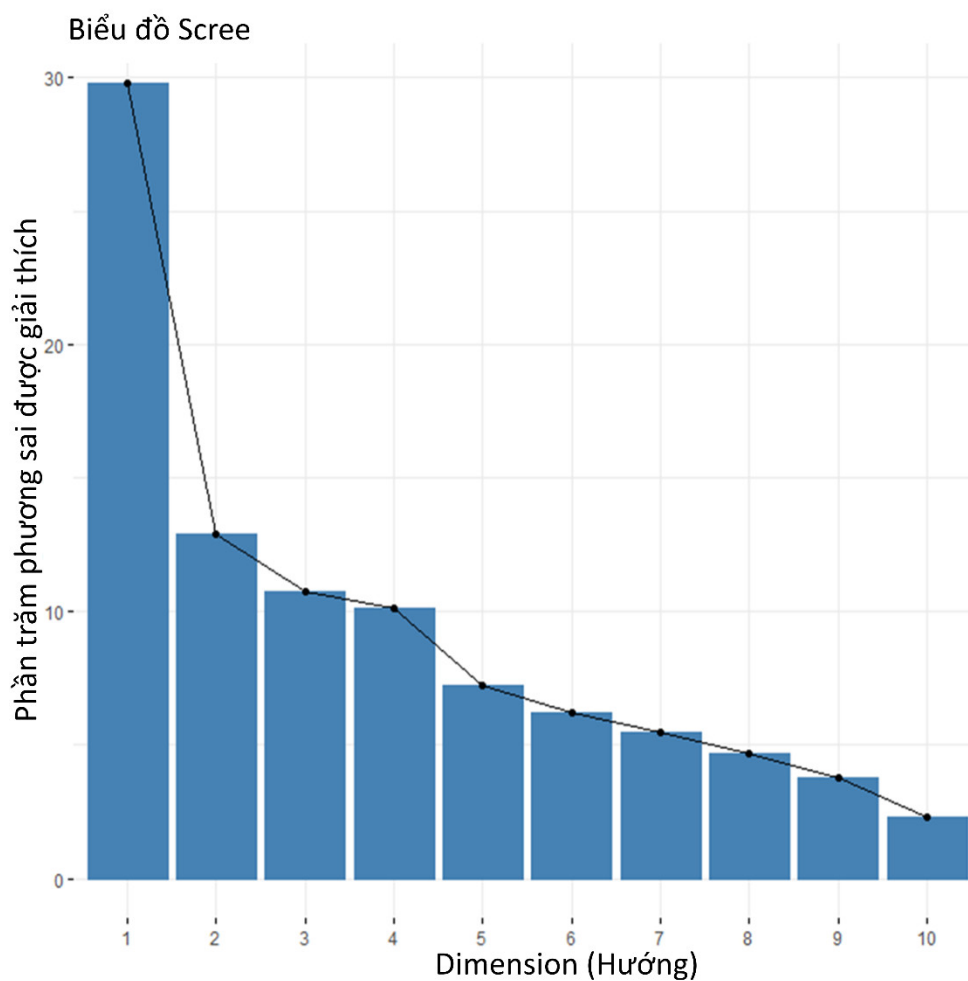
Dung Duc Tran^{1*}, Nguyen Quoc Quan¹, Nguyen Thi Thanh Hue¹, Pham Luan¹

¹ Center of Water Management and Climate Change, Institute for Environment and Resources, Vietnam National University – Ho Chinh city (VNU – HCM), Ho Chi Minh, Vietnam; dungtranducvn@yahoo.com; quocquannguyen1987@gmail.com; nguyen.tt.hue@gmail.com; hongluanosgeo@gmail.com

Abstract: La Buong River is a large tributary located on the right bank of Dong Nai River. The economic activities of this basin are relatively diverse, from agricultural and livestock activities in the upstream to industrial and service development in the downstream area. However, socio-economic development has caused surface water pollution due to industrial and agricultural activities. This study aims to assess the surface water quality of La Buong River, using multivariate statistical methods, analyzing the main components based on water quality indicators in space and time. Monitoring data for the period 2010-2017 is processed to remove outliers and standardize data using Shapiro-Wilk test và non-parametric test of Wilcoxon signed-rank, before being used for analysis. The results show that the water of La Buong River in both the dry season and the rainy season is locally contaminated with nutrients (N-NH₄) and bacteria (E. coli, Coliform) in the upstream area due to livestock activities. In addition, we found pollution of inorganic and organic substances (BOD₅, COD, Fe, N-NH₄, N-NO₂, TSS, turbidity) from domestic and industrial wastewater in the middle and downstream areas. The research results help assess the current status of La Buong river's water quality to propose solutions to contribute to the environmental protection of surface water of the whole basin towards a sustainable development.

Keywords: Polluted; Monitoring; Statistical; Analysis; Principal Component Analysis.

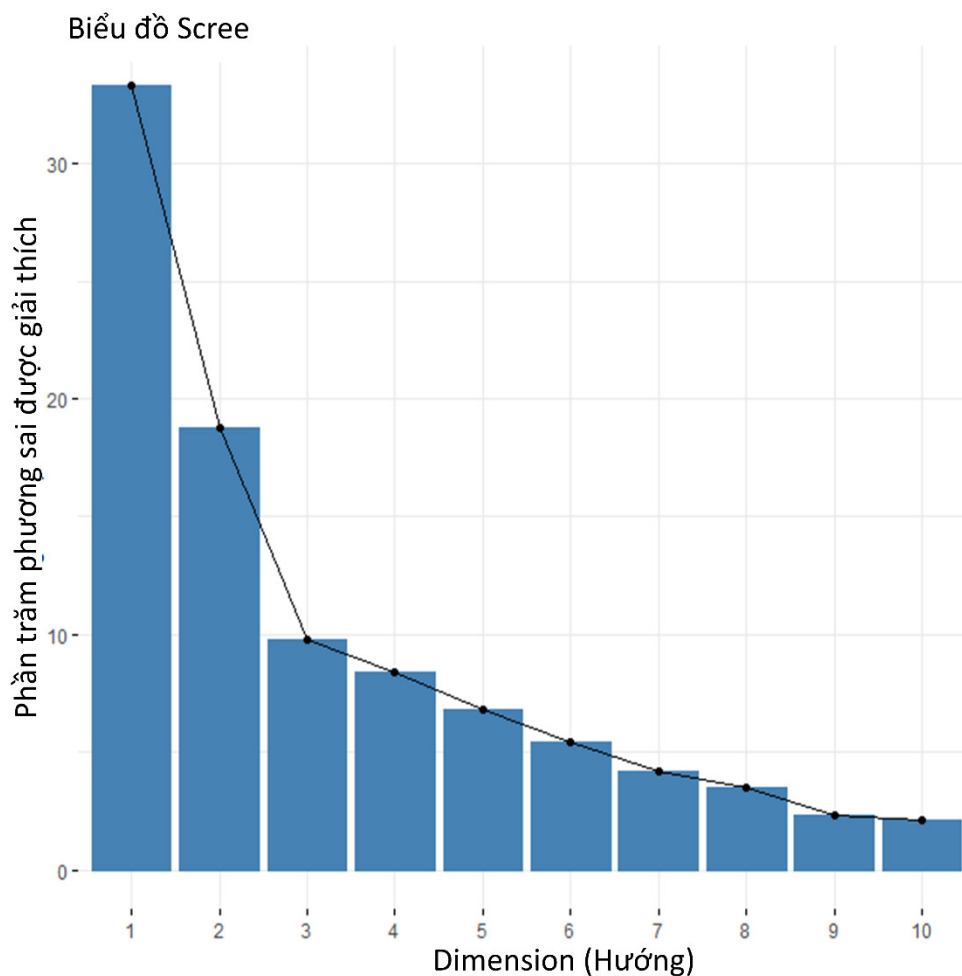
Phụ lục



Hình PL1. Lược đồ Scree plot của phân tích PCA cho dữ liệu chất lượng nước mùa khô.

Bảng PL1. Giá trị riêng và % phương sai mùa khô.

Giá trị riêng	% Phương sai	% Phương sai tích lũy
4,46	29,76	29,76
1,93	12,90	42,66
1,61	10,73	53,38
1,52	10,13	63,51
1,08	7,23	70,74
0,93	6,18	76,92
0,82	5,48	82,40
0,70	4,68	87,07
0,57	3,78	90,86
0,35	2,33	93,18
0,31	2,09	95,28
0,27	1,80	97,08
0,23	1,52	98,60
0,16	1,05	99,65



Hình PL2. Lược đồ Scree plot của phân tích PCA cho dữ liệu chất lượng nước mùa mưa.

Bảng PL2. Giá trị riêng và % phương sai mùa mưa.

Giá trị riêng	% Phương sai	% Phương sai tích lũy
4,99	33,29	33,29
2,81	18,76	52,06
1,46	9,74	61,79
1,26	8,42	70,21
1,02	6,79	77
0,81	5,41	82,41
0,63	4,22	86,63
0,53	3,5	90,13
0,35	2,34	92,47
0,32	2,12	94,59
0,25	1,69	96,28
0,24	1,58	97,86
0,16	1,09	98,95
0,11	0,75	99,7
0,04	0,3	100