



EVALUACIÓN DE LA CALIDAD DE AGUA DE LA CIÉNAGA DE LA VIRGEN (CARTAGENA, COLOMBIA) APLICANDO TÉCNICAS ESTADÍSTICAS MULTIVARIADAS

WATER QUALITY ASSESSMENT OF CIÉNAGA DE LA VIRGEN (CARTAGENA, COLOMBIA) APPLYING MULTIVARIATE STATISTICAL TECHNIQUES

Ildefonso Baldiris-Navarro¹

Juan C. García-Ojeda²

José Herrera Franco³

Martha Torres Virviescas⁴

Renzo Herrera-Hoyos⁵

¹ Magíster en Ingeniería, Estudiante de Doctorado en Ingeniería, Universidad de Cartagena.
ORCID: <https://orcid.org/0000-0001-9189-3710>
ibaldirisn@unicartagena.edu.co

² PhD. Information Science and Technology, Universidad de Cartagena.
ORCID: <https://orcid.org/0000-0002-3148-0436>
jcgarciao@unicartagena.edu.co

³ Magíster en Ciencias Ambientales, secretaria Educación de Cartagena.
ORCID: <https://orcid.org/0009-0005-6065-5943>
ambientaltecn@gmail.com

⁴ Magíster en Ciencias del Mar, Universidad del Sinú.
ORCID: <https://orcid.org/0000-0003-1914-0583>
biologiamarina@unisinucartagena.edu.co

⁵ Médico general, Universidad de Cartagena.
ORCID: <https://orcid.org/0009-0005-9239-6632>
rherrerah1@unicartagena.edu.co

Citación: Baldiris-Navarro, I., García-Ojeda, J.C., Herrera, J., Torres, M., y Herrera-Hoyos, R. (2025). Evaluación de la calidad de agua de la Ciénaga de la Virgen (Cartagena, Colombia) aplicando técnicas estadísticas multivariadas. *Revista de Investigación Agraria y Ambiental*, 16(1), 319 - 338. <https://doi.org/10.22490/21456453.8149>

RESUMEN

Contextualización: el agua es uno de los recursos naturales más importantes para la vida y es el que más está siendo afectado diariamente por las comunidades aledañas a los cuerpos de agua.

Vacío de conocimiento: las autoridades ambientales utilizan muy poco el análisis estadístico multivariado para el estudio de la calidad de agua de las regiones, lo cual es una herramienta importante para la toma de decisiones ambientales.

Propósito: determinar la variación de la calidad del agua de la Ciénaga de la Virgen mediante el empleo de métodos estadísticos multivariados (PCA, Clúster y matriz de correlaciones) a datos colectados desde el año 2009 hasta 2017. Con esto se dará una línea base para la toma de decisiones por parte de las autoridades ambientales de la ciudad para el mejor mantenimiento de este ecosistema.

Metodología: las muestras fueron tomadas en la Ciénaga de la Virgen en Cartagena de manera mensual, llevadas al laboratorio y analizadas para amonio (NH_4^+), demanda bioquímica de oxígeno (DBO_5), demanda química de oxígeno (DQO), oxígeno disuel-

to (OD), pH, fósforo total (TP), salinidad, coliformes totales (TC), coliformes fecales (FC), sólidos suspendidos totales (SST). A los datos colectados se le aplicaron análisis multivariados como matriz de correlación, análisis de componentes principales, comparación con diagramas de cajas y bigotes y análisis clúster.

Resultados y conclusiones: el análisis multivariado de los datos obtenidos evidencia diferentes comportamientos de la calidad del agua en la ciénaga y los factores de contaminación más fuertes se da por altos niveles de DBO_5 y coliformes totales y fecales, ambos factores relacionados con aguas residuales y residuos sólidos. La mayor contaminación en la ciénaga es causada por la comunidad y sus residuos sólidos que son arrojados a la ciénaga, por ello es necesario iniciar jornadas de sensibilización de la comunidad cartagenera en temas de cuidado del agua para poder preservar este recurso hídrico tan importante para la ciudad.

Palabras claves: análisis clúster, análisis de componentes principales, calidad de agua, Ciénaga de la Virgen, estadística multivariada

ABSTRACT

Contextualization: Water is one of the most important natural resources for life and is most impacted daily by the communities surrounding the bodies of water.

Knowledge gap: Environmental authorities rarely use multivariate statistical analyzes to examine regional water quality, which is an important tool for environmental decision-making.

Purpose: To determine the variation of the water quality of the Ciénaga de la Virgen using multivariate statistical methods to data collected from the year 2009 to 2017.

Methodology: the samples were taken in the Ciénaga monthly, then in the laboratory the concentrations of Ammonium (NH_4^+), Biochemical Oxygen Demand (BOD_5), Chemical Oxygen Demand (COD), Dissolved Oxygen (DO), pH, Total Phosphorus (TP), Salinity, Total Coliforms (TC), fecal coliforms (FC), total suspended solids (TSS)

were analyzed. Multivariate analyzes such as correlation matrix, principal component analysis, boxplots, and cluster were applied to the collected data.

Results and conclusions: Multivariate data analysis showed different behaviors of water quality around the coastal lagoon and the strongest contamination factors are high levels of BOD_5 and total and fecal coliforms. Both factors related to wastewater and solid waste. Conclusions: the greatest contamination in the swamp is caused by the community and its poorly treated waste, it is necessary to start awareness days in the Cartagena community on issues of water care to preserve this important water resource for the city.

Keywords: Ciénaga de la Virgen, cluster análisis, multivariate statistics, principal component analysis, water quality

RESUMEN GRÁFICO

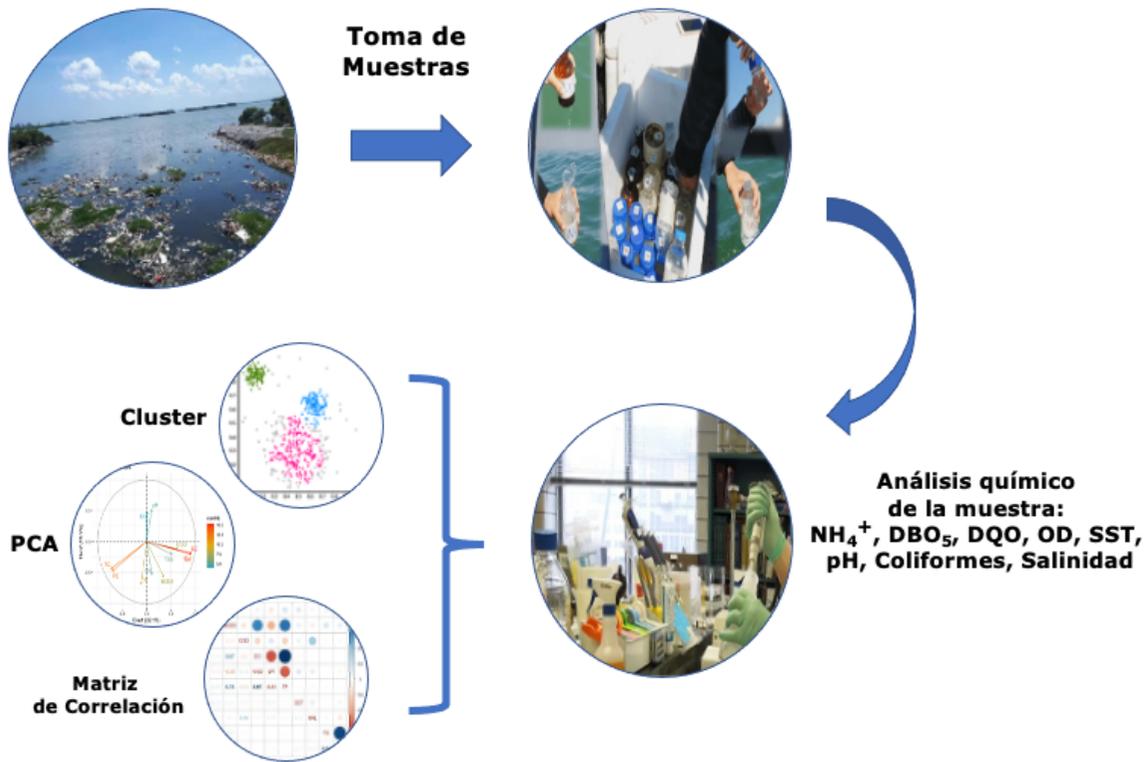


Diagrama de los procesos del desarrollo del estudio

Fuente: autores.

1 INTRODUCCIÓN

El agua es un elemento trascendental para el avance de la sociedad a nivel mundial. La calidad de los cuerpos de agua, hoy día, se ha visto deteriorada debido a la sobrepoblación y la industrialización que han traído consigo una producción elevada de residuos sólidos, líquidos y gaseosos, los cuales afectan las propiedades del agua

de las reservas hídricas (Loi *et al.*, 2022). Uno de los problemas más frecuentes en los cuerpos de agua es el alto contenido de nutrientes como el nitrógeno (nitratos, nitrito y amonio) y el fósforo en forma de fosfatos, los cuales conllevan a un problema llamado eutrofización (Xiong *et al.*, 2022; Zamora-López *et al.*, 2023). La Ciénaga de

La Virgen (CV), que es una laguna costera de Cartagena, presenta altos niveles de contaminación antrópica debido a la pésima disposición final de las aguas servidas de la ciudad, ya que dichos vertimientos llegan a la ciénaga sin tratamiento previo. Aproximadamente el 35% de las aguas servidas eran vertidas en la Bahía de Cartagena y el 65% dispuestas en la ciénaga (Martin y Urrea, 2016).

Para el año 2001, debido a los altos niveles de contaminación, el gobierno nacional y la empresa Holandesa Royal Haskoning lanzaron el proyecto de la bocana estabilizada de la ciénaga (Moor *et al.*, 2002) que pretendía mejorar las propiedades del agua y el entorno. Este programa comprendió la fabricación de una conexión artificial entre la ciénaga y el mar para aumentar el intercambio de agua de una forma más eficiente, con el objetivo de que el agua marina descontaminara la ciénaga al mismo tiempo que optimizara la concentración de oxígeno disuelto y la salinidad, logrando con ello recuperar su capacidad de auto regeneración y mejorando las propiedades del agua (Maldonado *et al.*, 2011).

Posteriormente, se planteó el plan maestro del alcantarillado de Cartagena, cuya finalidad fue la eliminación de los vertimientos a la bahía y a la ciénaga por medio de una descarga al mar llamada Emisario Submarino (Palomino y Serge, 2015). El Proyecto del emisario submarino permitiría en teoría recuperar aproximadamente 15.000 hectáreas de áreas estuarinas en los cuerpos de agua, lo que posibilitaría un ecosistema marino biológicamente más sano.

La alcaldía de Cartagena es la encargada del control y seguimiento de las condiciones

de la ciénaga una vez terminaron los trabajos de construcción del proyecto Bocana. Las variables monitoreadas para evaluar la efectividad del sistema son pH, oxígeno disuelto (OD), demanda bioquímica de oxígeno (DBO_5), demanda química de oxígeno (DQO), sólidos suspendidos totales (SST), amonio, fósforo total (TP), coliformes totales (CT) y fecales (CF).

Varios autores han aplicado diferentes técnicas estadísticas multivariadas con éxito en la valoración del agua de diferentes escenarios, y subrayaron la necesidad y utilidad de estas técnicas estadísticas para la interpretación de complejos conjuntos de datos, identificación de las fuentes de contaminación y la identificación de las variaciones temporales-espaciales en las propiedades del agua para la buena gestión de los recursos naturales (Karangoda y Nanayakkara, 2023; Varol, 2020; Kherifi *et al.*, 2019; Bodrud-Doza *et al.*, 2019; Howladar *et al.*, 2021). Las autoridades ambientales en la ciudad de Cartagena solo aplican técnicas de estadística descriptiva para el estudio de las condiciones de los recursos hídricos, aunque visualmente es poderosa, esta se queda corta para decidir adecuadamente en el manejo administrativo de los cuerpos de agua ya que no explica qué variables o qué fuentes tienen mayor efecto en las propiedades del recurso hídrico. El objetivo de este artículo es el análisis estadístico multivariado de las propiedades de agua de la Ciénaga de La Virgen desde el año 2009 hasta el 2017, con lo cual se dará una línea base para la toma de decisiones por parte de las autoridades ambientales de la ciudad para el mejor mantenimiento de este ecosistema.

2 MATERIALES Y MÉTODOS

Área de estudio

La Ciénaga de La Virgen es considerada un humedal prioritario, está ubicada en Cartagena de Indias en el departamento de Bolívar. La ciénaga se encuentra ubicada en la zona norte de la ciudad, en las coordenadas 10° 26' 45" Norte y 75° 29' 40" Oeste (Figura 1). El estuario abarca 2989 ha, tiene una forma triangular, estrecha al norte y ancha en el sur, con ancho máximo de 4.5 km y una

extensión de unos 7 km. Cuenta con profundidades de hasta 1,2 m y está cercada por la red de drenajes de la ciudad (Grisales *et al.*, 2014). Esta cuenta con un clima tropical semiárido, con una época seca de diciembre a abril y lluviosa de mayo a noviembre (Betancur y Rivero, 2020). El monitoreo en la ciénaga consta de diez estaciones previamente georeferenciadas (Figura 1).

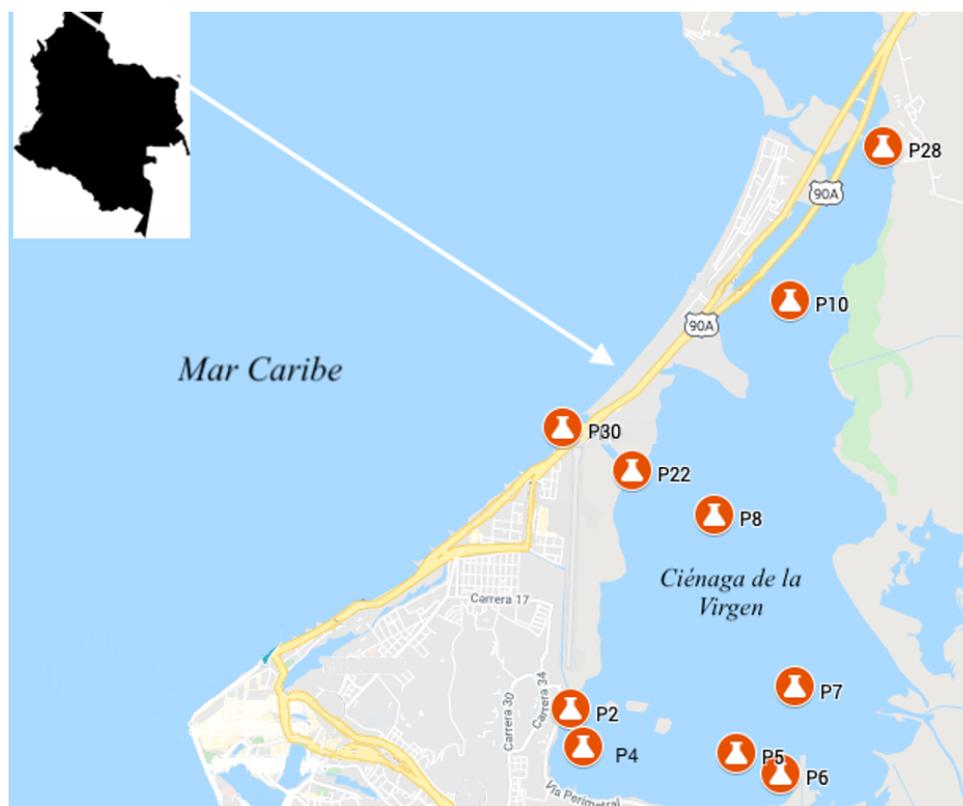


Figura 1. Puntos de toma de muestra en Ciénaga de la Virgen

Fuente: autores.

Las muestras fueron tomadas mensualmente y llevadas al laboratorio donde se evaluaron los parámetros de acuerdo con lo establecido por los métodos estándar de APHA. Se tomaron muestras in situ con la ayuda del equipo multiparámetro, para los ensayos de laboratorio las muestras se preservaron, luego se depositaron en recipientes de color ámbar y refrigeradas para su posterior análisis.

Métodos analíticos

Los parámetros seleccionados fueron: oxígeno (DO), coliformes totales (TC), pH, agua temperatura (T), demanda bioquímica de oxígeno de 5 días (DBO_5), demanda química de oxígeno (DQO), nitrato (NO_3^-), fósforo total (P), sólidos suspendidos totales (TSS), salinidad (Sal) y amonio (NH_4^+). La temperatura, pH, salinidad y concentraciones de OD se midieron in situ mediante el uso del multiparámetro portátil HACH sesión 5465011, respectivamente. La DBO_5 fue determinada por el método Winkler y SST se determinaron gravimétricamente a 105-110 °C. NO_3^- y P, fueron analizados por el método espectrofotométrico de reducción de cadmio y ácido ascórbico usando un espectrofotómetro Varian Cary 100 UV-vis, respectivamente. Para determinar el número de unidades formadoras de colonias, por 100 ml (ufc / 100 mL) de coliformes totales en el agua de muestreo se utilizó el método de filtración por membranas (Baldiris-Navarro *et al.*, 2019; Rice *et al.*, 2017; APHA, 2014).

Análisis estadísticos

Matriz de correlación

Una matriz de correlación es una herramienta que evalúa la relación existente entre dos o más variables. Esta herramienta define la correlación de Pearson entre cada pareja de variables. El coeficiente de correlación varía entre -1 y +1, y mide la relación lineal entre las variables. Cuando el coeficiente de correlación es positivo cercano a +/-1 hay una alta relación entre las variables, si el factor tiende a cero la relación es nula, si el factor es positivo la relación es directamente proporcional, de otra manera es inversamente proporcional (Sun *et al.*, 2019; Kothari *et al.*, 2021).

Análisis de componentes principales

El análisis de componentes principales (PCA) es una técnica de aprendizaje no supervisado para la reducción de variables, no depende de la no-homogeneidad en la muestra de datos, de los valores perdidos y de las tendencias en la información. También, identifica variaciones temporales en las propiedades del agua y el efecto de los diferentes factores sobre ella. La técnica PCA permite extraer los parámetros más relevantes en una estación contaminada (Krishnaraj y Deka, 2020; Ibrahim *et al.*, 2023; Alam *et al.*, 2021).

PCA es un método de reducción que logra disminuir la cantidad de variables originales

consideradas para el estudio, perdiendo la menor información posible. Las nuevas variables se formulan como una combinación lineal de las variables iniciales. El PCA tiene validez una vez se verifica que existe una correlación entre las variables, si no fuera así el análisis carecería de valor. Para escoger el número de componentes se puede usar el criterio del valor propio donde se escogen los componentes con valores mayores a la unidad. Para escoger los parámetros que más influyen en un componente se toman aquellos que tengan cargas mayores a 0.4 (Barakat *et al.*, 2016; Li *et al.*, 2018; Roy *et al.*, 2024).

Análisis clúster

El análisis clúster es una técnica no supervisada cuya intención es agrupar objetos basados en sus características. Este catalo-

ga los objetos para que cada uno sea similar a los demás en el grupo con respecto a un criterio de selección predeterminado. Los grupos resultantes de los objetos deben exhibir una alta homogeneidad dentro del grupo y alta disparidad con los demás grupos. La agrupación jerárquica aglomerativa (HAC) es el método más común, el cual suministra relaciones de similitud sospechada entre cualquier elemento de una muestra y el conjunto total de datos, y se esquematiza con un dendrograma. Este último es una herramienta visual que permite diferenciar fácilmente los grupos dentro del conjunto de datos. En este artículo se utilizó el método de Ward para calcular la distancia entre los grupos y la distancia euclidiana cuadrada para medir la similitud entre los grupos (Liu y You, 2023; Ganguli *et al.*, 2022; Mazari-Hiriart *et al.*, 2024).

3 RESULTADOS Y DISCUSIÓN

Matriz de correlación

Con esta herramienta se evaluó la relación existente entre las variables monitoreadas en la ciénaga. El p-valor es el del test de significancia estadística de la correlación estimada entre dos variables. Los p-valores por debajo de 0.05 indica una correlación significativa diferente de cero con un nivel de confianza

del 95%. Los siguientes pares de variables tuvieron p-valores por debajo de 0.05: NH_4^+ -DBO; DBO-DO; DBO-pH; DBO-TP; DQO-SAL; DO-pH; DO-TP; pH-TP y CT-CF. La Tabla 1 evidencia que los puntos en negrillas y de mayor valor son correlaciones fuertes entre las variables, como por ejemplo pH con fósforo total (TP).

Tabla 1.

Matriz de Correlación de variables estudiadas

	NH₄⁺	DBO₅	DQO	OD	pH	PT	SST	SAL	CT	CF
NH₄⁺	1.000									
DBO₅	0.27	1.000								
DQO	0.00	-0.14	1.000							
OD	0.00	0.67	-0.2	1.000						
pH	-0.13	-0.36	0.15	-0.63	1.000					
PT	0.17	0.74	-0.22	0.97	-0.61	1.000				
SST	0.15	0.11	0.09	0.04	0.14	0.05	1.000			
SAL	0.07	-0.1	0.33	0.03	0.11	0.02	0.12	1.000		
CT	0.08	0.17	0.02	0.01	0.01	0.01	0.01	0.09	1.000	
CF	0.05	0.12	0.05	0.02	0.05	0.02	0.07	-0.12	0.91	1.000

Fuente: autores.

Las relaciones entre algunas variables pueden evidenciarse mejor en la Figura 2. El fósforo y la DBO₅ presentan relación directa fuerte debido al hecho de que la mayor fuente de fósforo son las aguas residuales que conllevan al aumento de la DBO₅. La demanda química de oxígeno se ve relacionada directamente de manera moderada por el aumento o disminución de la salinidad, esto va en relación con la época del año (lluvioso o seco). La relación positiva fuerte entre el fósforo y el oxígeno disuelto es causada por el aumento de los nutrientes en aguas residuales, las cuales durante el día incremen-

tan el contenido de microalgas y estas por el proceso de la fotosíntesis aumentan los niveles de oxígeno disuelto. La DBO₅ y el OD, aunque su relación debe ser inversa puede mostrar este valor por la presencia de algas y microalgas que aumentan el valor del oxígeno disuelto por efecto de la fotosíntesis. De manera similar, la relación fuerte e inversa del oxígeno disuelto con el pH puede ser causada por el ingreso de aguas residuales que bajan el pH del agua y a su vez aumentan el contenido de microalgas lo que ocasiona un aumento de la concentración del oxígeno durante el día.

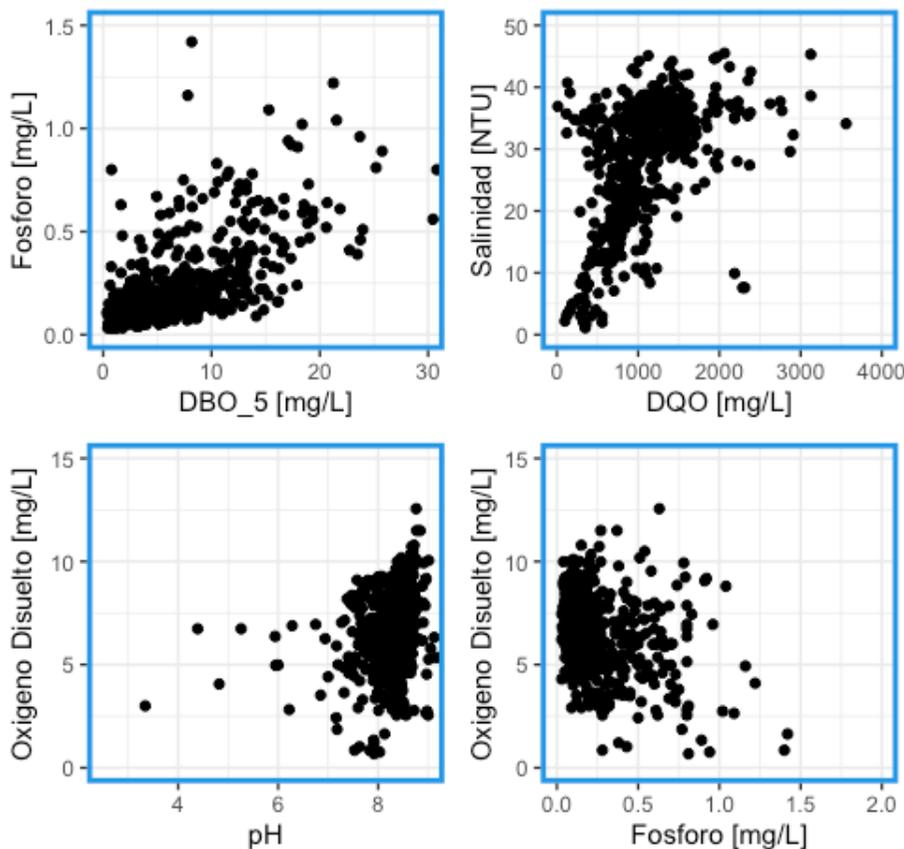


Figura 2. Correlaciones de variables

Fuente: autores.

Las variables coliformes totales y fecales varían directamente ya que ambas son generadas por las mismas bacterias presentes en aguas residuales. Con estos resultados de correlaciones entre las diferentes variables se evidencia la pertinencia del uso del análisis de componentes principales al dataset del cuerpo de agua.

Análisis de componentes principales

El estadístico de Kaiser–Meyer–Olkin

(KMO) tuvo un valor de 0.65 (>0.5), lo que indica que es viable aplicar el método de componentes principales al set de datos. Además, el test de esfericidad de Bartlett con un nivel de significancia menor a 0.05 indica que el análisis PCA puede ser utilizado en este caso, como lo expresa Tripathi y Singal (2019) en el trabajo realizado para analizar los índices de calidad de agua en el río Ganga, India. La Figura 3 muestra la gráfica de sedimentación que indica el número de componentes principales de la matriz de

datos estudiada. La línea punteada indica un valor de eigenvalor igual a 1. Los componentes con eigenvalor mayor a 1 se escogen, los que tengan eigenvalores menores a 1 se

desechan. La Figura 3 indica que del set de datos se pueden obtener tres componentes principales con eigenvalores mayores a 1.

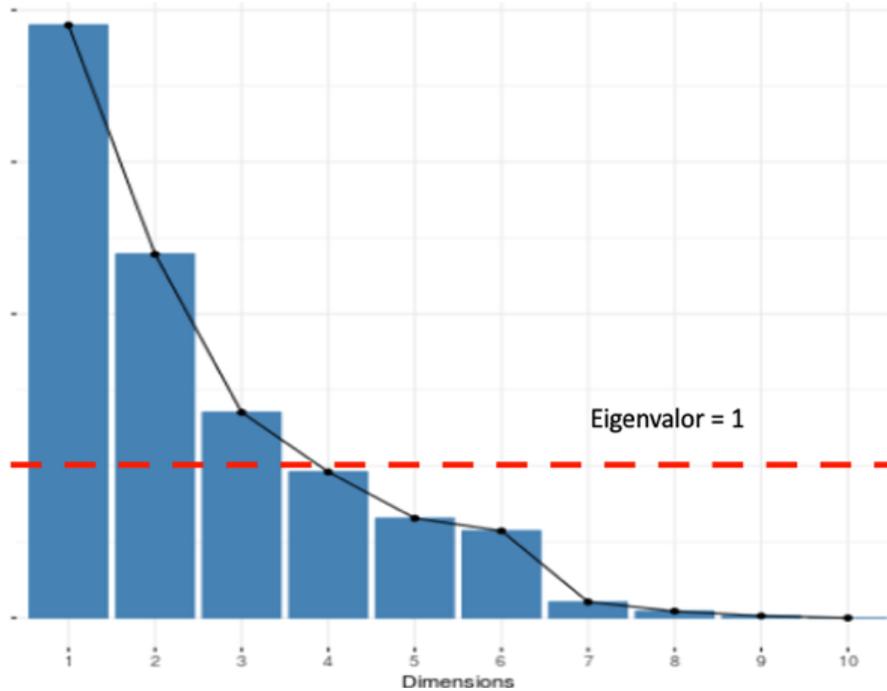


Figura 3. Gráfica de sedimentación análisis PCA

Fuente: autores.

Analíticamente, se escogen los componentes principales que tengan eigenvalores mayores a la unidad. Es así como solo se

escogieron los tres primeros componentes para el análisis de la información (Tabla 2).

Tabla 2.

Eigenvalores de los datos experimentales

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Eigenvalores	1.975	1.547	1.163	0.980	0.809	0.756	0.323
Aporte a varianza	0.390	0.239	0.135	0.096	0.065	0.057	0.010
Acumulativo (%)	39	62.9	76.4	86	92.6	98.3	99.4

Fuente: autores.

En la Tabla 2 puede observarse que el PC1 explica el 39% de la variación en la información, el PC2 explica el 23.9% y el PC3 explica el 13.5% del total de la varianza, estos tres componentes explican el 76.4% de la variación total de los datos.

En la Tabla 3 se aprecian los pesos de cada variable en cada uno de los componentes principales. Se tomaron las variables con pe-

sos mayores a 0.4 los cuales fueron resaltados en la Tabla 3. Para el componente uno (PC1) que explica el 39% de la varianza puede evidenciarse que tiene una carga positiva moderada por el contenido de amonio, los coliformes totales y los coliformes fecales, este factor representa la contaminación por aguas residuales con alto contenido de heces fecales que son vertidas por las comunidades cercanas a la ciénaga.

Tabla 3.

Pesos de las variables medidas en los componentes

	PC1	PC2	PC3
Amonio	0.476	-0.039	0.117
DBO ₅	0.319	-0.018	-0.439
DQO	-0.273	-0.258	-0.263
Oxígeno disuelto	-0.098	-0.320	-0.558
pH	0.116	0.535	-0.241
Fósforo total	0.387	-0.006	-0.538
SST	0.048	-0.425	-0.036
Salinidad	-0.273	-0.435	0.018
Coliformes totales	0.436	-0.231	0.208
Coliformes fecales	0.439	-0.342	0.241

Fuente: autores.

En la Tabla 3 se puede observar en negrillas que el componente principal dos (PC2) tiene una contribución positiva alta por el pH y una contribución moderada negativa por los sólidos suspendidos totales y por la salinidad, este factor puede ser atribuido a la entrada de minerales a la ciénaga por efecto de las mareas. El componente principal tres (PC3) tiene alta carga negativa por oxígeno disuelto y fósforo total, además tiene una carga moderada negativa por DBO₅, estos factores orgánicos indican principalmente las fuentes puntuales de contaminación de las aguas residuales do-

mésticas y residuos sólidos de alto contenido orgánico que son arrojados al cuerpo de agua (Ali Abed *et al.*, 2019).

Análisis clúster

El objetivo del análisis de conglomerados (CA) es ensamblar objetos (en este caso, puntos de toma de muestra) en agregaciones basadas en sus similitudes y variables o características interdependientes. El análisis arrojó los clústeres de puntos graficados en la Figura 4.

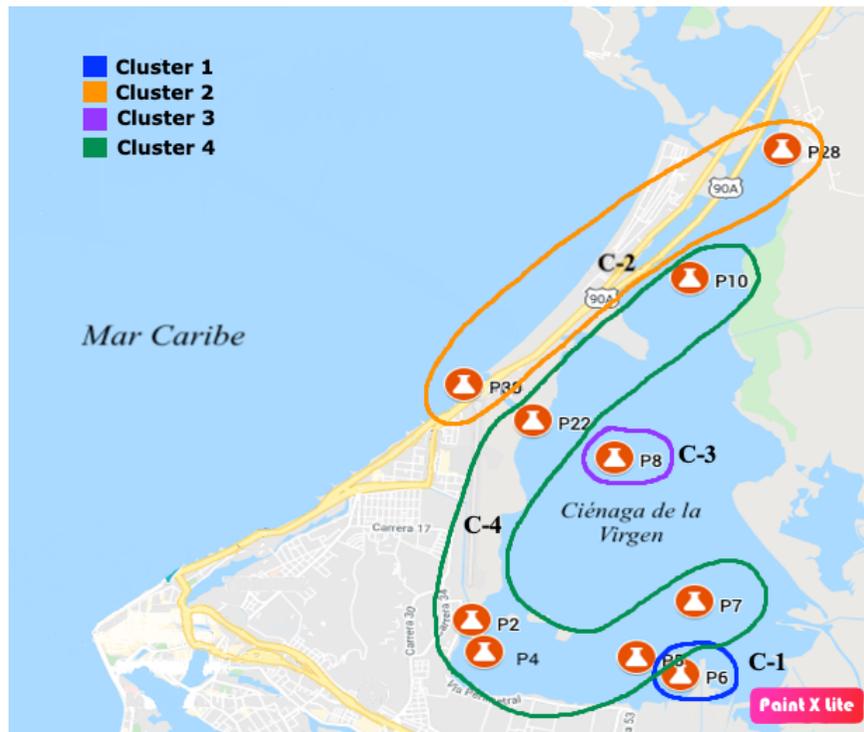


Figura 4. Resultado del análisis clúster

Fuente: autores.

El análisis clúster diferenció cuatro agrupaciones para el cuerpo de agua de acuerdo con la base de datos. El clúster 1 de muy alta contaminación está conformado solo por el punto 6; el clúster 2 que es de contaminación baja está conformado por el punto 28 y 30; el clúster 3 de alta polución está conformado por el punto 8; y el clúster 4 de contaminación moderada a alta está formado por los puntos 2, 4, 5, 22, 7 y 10. Se procedió a comparar por medio de boxplots las variables DBO_5 , coliformes totales (CT) y fósforo total (PT) en los clústeres, tomando para el clúster 1 el punto

6, para el clúster 2 el punto 30, para el clúster 3 el punto 8 y para el clúster 4 el punto 2.

El clúster 1 fue el que mostró valores más altos de DBO_5 , lo que indica mayores valores de contaminación por materia orgánica, lo mismo que para coliformes totales y fecales (Figura 5). Ambas variables están relacionadas con altas descargas de aguas residuales domésticas e industriales no tratadas como es lo expuesto por Fraga *et al.* (2020). El punto 6 (clúster 1) es el más afectado por las aguas residuales y residuos sólidos transportados por caños que descargan en la Ciénaga de La Virgen.

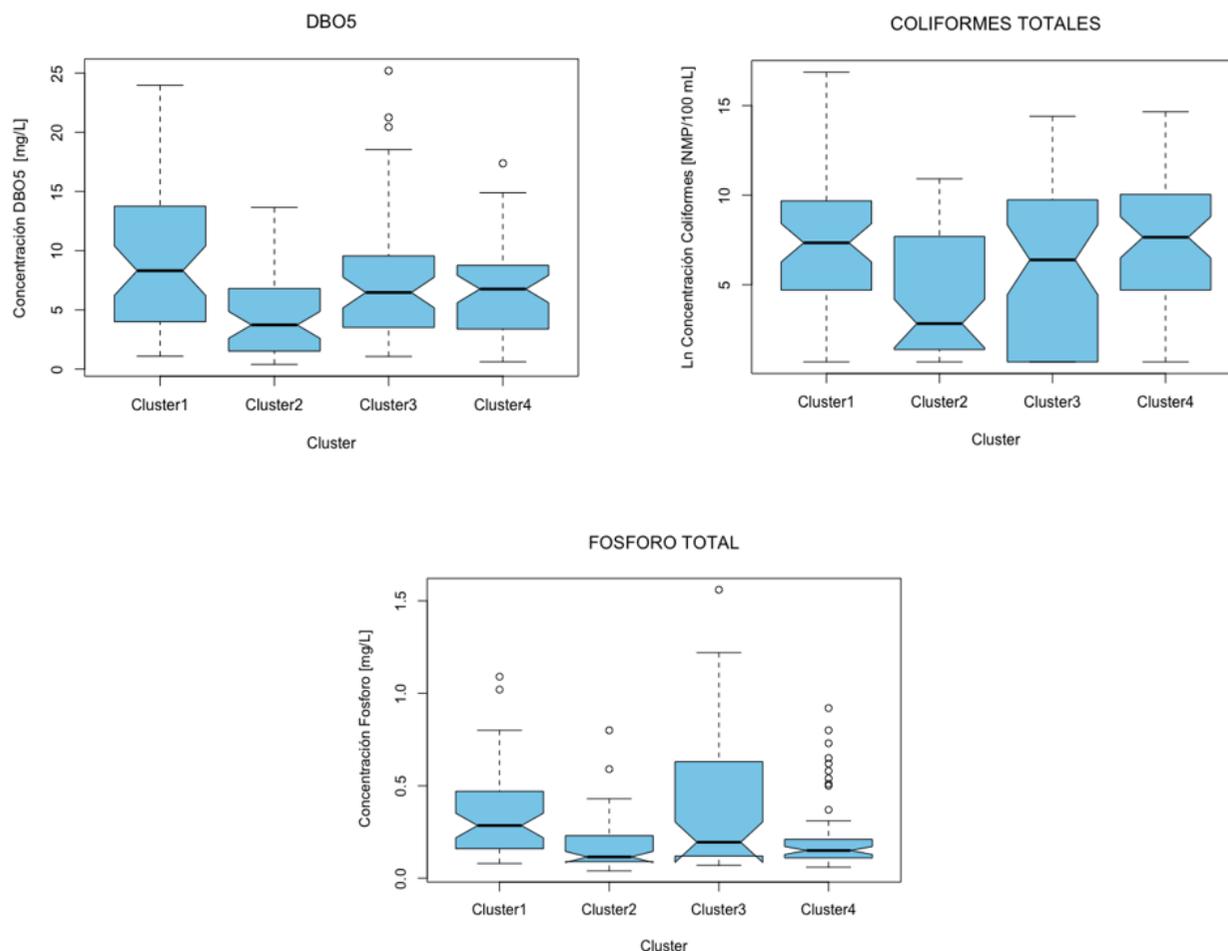


Figura 5. Comportamiento de DBO5, coliformes totales y fósforo por cluster

Fuente: autores.

El clúster 2 se caracteriza por estar alejado de la fuente antropogénica, ambos puntos tienen poco contacto con las poblaciones y sus vertimientos. Este clúster es el que presenta valores más bajos para DBO₅, CT y nutrientes. Los comportamientos son similares a los reportados por el trabajo de Angello *et al.* (2020) en el río Little Akaki, Etiopía.

El clúster 1 y 4 con mayores valores de coliformes (Figura 5), son aquellos puntos de toma de muestra que se encuentran en desembocaduras de grandes canales de re-

colección de aguas residuales domésticas e industriales. El clúster 3 representa geográficamente el centro de la ciénaga donde se dan valores moderados de DBO5 y coliformes, también evidencia altos valores de nutrientes, lo cual puede ser causado por acumulación de estos contaminantes que están próximos a salir de la ciénaga debido a los cambios de marea. Todos los puntos de muestreo tienen valores altos de coliformes que pueden representar una amenaza para la salud humana (Figura 5). La media de los coliformes estuvo alrededor de los 1800 NMP/100 mL,

datos similares a estos fueron obtenidos por Fulke et al (2024), sugiriendo que es urgente tomar medidas de contención sobre la contaminación para proteger la salud pública y al ecosistema.

El clúster 4, es el que mayor cantidad de puntos agrupa, teniendo como características moderados niveles de DBO_5 además de altos niveles de sólidos y coliformes. Este clúster también se ve afectado en menor proporción por caños pequeños que vierten aguas residuales a la ciénaga.

En cuanto al comportamiento del fósforo, este presenta valores de un estado eutrófico y hasta hipertrófico en el cuerpo de agua, lo

que según lo expuesto por los autores en Lencha *et al.* (2021) y Yilma *et al.* (2019), puede llevar a altos niveles de DBO_5 , sobrepoblación de algas y niveles muy bajos de oxígeno en el agua; el clúster 3 es el que presenta los resultados más elevados de este parámetro. Con base en los datos, se puede evidenciar que el clúster 4 y 1 son los que presentan mayor carga de contaminantes orgánicos y aguas residuales, esto debido a los altos valores de DBO_5 y coliformes, causados probablemente por su cercanía a la orilla de la ciénaga y a la comunidad aledaña. El clúster 2 es el que presenta valores más moderados en estas variables debido a su lejanía de la zona poblada de la ciénaga.

4 CONCLUSIONES

En este estudio con el análisis de correlación de variables, se pudo evidenciar una fuerte relación entre los valores de las variables oxígeno disuelto (OD), demanda bioquímica de oxígeno (DBO_5), el fósforo total (PT), y relaciones moderadas entre las demás variables, razón por la que se procedió a realizar el análisis de PCA. Luego, el análisis de componentes principales mostró que la fuente de polución más significativa en la ciénaga es la entrada de aguas residuales cargadas de materia orgánica y nutrientes. Por su parte, el análisis clúster dividió el cenegal en cuatro grupos diferentes, siendo el clúster 1 y 4 los que presentaron mayores valores negativos respecto a la calidad del agua, estos se encuentran ubicados en las zonas más cercanas a la población.

Lo anterior indica que se deben hacer jornadas de sensibilización y cuidado del medio ambiente a estas poblaciones para alcanzar unos niveles aceptables en las variables DBO_5 , coliformes y fósforo en la ciénaga. De igual manera, es importante llevar a cabo un estricto programa de control y vigilancia en sitios puntuales donde se detecten fugas o vertimientos no permitidos sobre la ciénaga. Así, estos resultados basados en ciencia de datos pueden ayudar a las autoridades distritales a tomar decisiones que mejoren de las condiciones ambientales de la Ciénaga de La Virgen, el cuerpo de agua más importante de la ciudad de Cartagena.

AGRADECIMIENTOS

Los autores agradecen a la Universidad de Cartagena por su colaboración para el desarrollo de este trabajo.

CONTRIBUCIÓN DE LA AUTORÍA

Ildelfonso Baldiris-Navarro: metodología, investigación, análisis de datos, escritura, borrador original. **Juan Carlos García Ojeda:** metodología, conceptualización, revisión. José Herrera Franco: análisis de

datos, revisión y edición. **Martha Torres Virviescas:** metodología, investigación y edición. **Renzo Herrera-Hoyos:** análisis y visualización de los datos.

FINANCIAMIENTO

No aplica.

LITERATURA CITADA

Ali Abed, S., Hussein, S., & Al-Ansari, N. (2019). Evaluation of Water quality in the Tigris River within Baghdad, Iraq using Multivariate Statistical Techniques. *Journal of Physics: Conference Series*, 1294(7), 072025. <https://doi.org/10.1088/1742-6596/1294/7/072025>

Alam, R., Ahmed, Z., Seefat, S. M., & Nahin, K. T. K. (2021). Assessment of surface water quality around a landfill using multivariate statistical method, Sylhet, Ban-

gladesh. *Environmental Nanotechnology, Monitoring & Management*, 15, 100422.

Angello, Z., Tränckner, J., & Behailu, B. (2020). Spatio-Temporal Evaluation and Quantification of Pollutant Source Contribution in Little Akaki River, Ethiopia: Conjunctive Application of Factor Analysis and Multivariate Receptor Model. *Polish Journal of Environmental Studies*, 30(1), 23-34. <https://doi.org/10.15244/pjoes/119098>

- APHA. (2014). *Standard Methods for the Examination of Water and Wastewater* (23rd ed.). American Works; Public Health Association/American Water Federation; Association/Water Environment.
- Baldiris-Navarro, I., Acosta-Jimenez, J. C., González-Delgado, A. D., Realpe-Jiménez, A., & Fajardo-Cuadro, J. G. (2019). Multivariate Statistical Analysis Applied to Water Quality of a Tropical Coastal Lagoon, Cartagena, Colombian Caribbean. *Indonesian Journal of Chemistry*, 20(1), 141. <https://doi.org/10.22146/ijc.43035>
- Barakat, A., El Baghdadi, M., Rais, J., Aghez-zaf, B., & Slassi, M. (2016). Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques. *International Soil and Water Conservation Research*, 4(4), 284-292. <https://doi.org/10.1016/j.iswcr.2016.11.002>
- Betancur, S. P., y Rivero, J. P. (2020). Variabilidad de la pendiente espectral de la materia orgánica disuelta cromofórica (cdom) en la estación de monitoreo multianual antares, Cartagena. *Boletín Científico CIOH*, 39(2). <https://doi.org/10.26640/22159045.2020.506>
- Bodrud-Doza, Md., Bhuiyan, M. A. H., Islam, S. M. D.-U., Rahman, M. S., Haque, Md. M., Fatema, K. J., Ahmed, N., Rakib, M. A., & Rahman, Md. A. (2019). Hydrogeochemical investigation of groundwater in Dhaka City of Bangladesh using GIS and multivariate statistical techniques. *Groundwater for Sustainable Development*, 8, 226-244. <https://doi.org/10.1016/j.gsd.2018.11.008>
- Fraga, M. D. S., Reis, G. B., Da Silva, D. D., Guedes, H. A. S., & Elesbon, A. A. A. (2020). Use of multivariate statistical methods to analyze the monitoring of surface water quality in the Doce River basin, Minas Gerais, Brazil. *Environmental Science and Pollution Research*, 27(28), 35303-35318. <https://doi.org/10.1007/s11356-020-09783-0>
- Fulke, A. B., Panigrahi, J., Eranezhath, S., Karthi, J., & Dora, G. U. (2024). Environmental variables and its association with faecal coliform at Madh Island beaches of megacity Mumbai, India. *Environmental Pollution*, 341, 122885.
- Ganguli, S., Hosen Rifat, Md. A., Howlader, S., Hasan, Md. A., Islam, S., Alam, Md. N. E., & Islam, Md. N. (2022). Assessment of Bhatiari Lake water quality: Pollution indices, hydrochemical signatures and hydro-statistical analysis. *Journal of the Indian Chemical Society*, 99(8), 100585. <https://doi.org/10.1016/j.jics.2022.100585>
- Grisales, C., Salgado, J., y Molaes, R. (2014). Proceso de intercambio de masas de agua de la bahía de Cartagena (Caribe colombiano) basado en la medición de parámetros oceanográficos. *Boletín Científico CIOH*, 32, 47-70.
- Howladar, M. F., Chakma, E., Jahan Koley, N., Islam, S., Numanbakth, M. A. A., Ahmed, Z., Chowdhury, T. R., & Akter, S. (2021). The water quality and pollution sources assessment of Surma river, Ban-

- gladesh using, hydrochemical, multivariate statistical and water quality index methods. *Groundwater for Sustainable Development*, 12, 100523. <https://doi.org/10.1016/j.gsd.2020.100523>
- Ibrahim, A., Ismail, A., Juahir, H., Iliyasu, A. B., Wailare, B. T., Mukhtar, M., & Aminu, H. (2023). Water quality modelling using principal component analysis and artificial neural network. *Marine Pollution Bulletin*, 187, 114493.
- Karangoda, R. C., & Nanayakkara, K. G. N. (2023). Use of the water quality index and multivariate analysis to assess groundwater quality for drinking purpose in Ratnapura district, Sri Lanka. *Groundwater for Sustainable Development*, 21, 100910. <https://doi.org/10.1016/j.gsd.2023.100910>
- Kherifi, W., Hecini, L., Bekiri, F., & Kherici-Bousnoubra, H. (2019). Faecal contamination of water in the Lake Mellah and its catchment area, north-eastern Algeria. *Journal of Water and Land Development*, 42(1), 110-116. <https://doi.org/10.2478/jwld-2019-0051>
- Kothari, V., Vij, S., Sharma, S., & Gupta, N. (2021). Correlation of various water quality parameters and water quality index of districts of Uttarakhand. *Environmental and Sustainability Indicators*, 9, 100093. <https://doi.org/10.1016/j.in-dic.2020.100093>
- Krishnaraj, A., & Deka, P. C. (2020). Spatial and temporal variations in river water quality of the Middle Ganga Basin using unsupervised machine learning techniques. *Environmental Monitoring and Assessment*, 192(12), 744. <https://doi.org/10.1007/s10661-020-08624-4>
- Lencha, S. M., Tränckner, J., & Dananto, M. (2021). Assessing the Water Quality of Lake Hawassa Ethiopia—Trophic State and Suitability for Anthropogenic Uses—Applying Common Water Quality Indices. *International Journal of Environmental Research and Public Health*, 18(17), 8904. <https://doi.org/10.3390/ijerph18178904>
- Li, T., Li, S., Liang, C., Bush, R. T., Xiong, L., & Jiang, Y. (2018). A comparative assessment of Australia's Lower Lakes water quality under extreme drought and post-drought conditions using multivariate statistical techniques. *Journal of Cleaner Production*, 190, 1-11. <https://doi.org/10.1016/j.jclepro.2018.04.121>
- Liu, L., & You, X. (2023). Water quality assessment and contribution rates of main pollution sources in Baiyangdian Lake, northern China. *Environmental Impact Assessment Review*, 98, 106965. <https://doi.org/10.1016/j.eiar.2022.106965>
- Loi, J. X., Chua, A. S. M., Rabuni, M. F., Tan, C. K., Lai, S. H., Takemura, Y., & Syutsu-bo, K. (2022). Water quality assessment and pollution threat to safe water supply for three river basins in Malaysia. *Science of The Total Environment*, 832, 155067. <https://doi.org/10.1016/j.scitotenv.2022.155067>
- Maldonado, W., Baldiris, I., y Díaz, J. (2011). Evaluación de la calidad del agua en la Ciénaga de la Virgen (Cartagena, Colom-

- bia) durante el período 2006-2010. *Revista Científica Guillermo de Ockham*, 9(2), 79-87.
- Martin, I., y Urrea, M. (2016). *Estudio para la optimización del sistema de tratamiento de aguas residuales de Cartagena para su descarga mediante emisario submarino en el mar caribe*. La Salle. https://ciencia.lasalle.edu.co/ing_ambiental_sanitaria/54
- Mazari-Hiriart, M., Fernández-Reyes A., Alvarado-Velázquez J., Gradilla-Hernández, M., Díaz-Vázquez D. (2024). Water quality management in a tropical karstic system influenced by land use in Chiapas, Mexico. *Environmental Challenges*, 100981. <https://doi.org/10.1016/j.envc.2024.100981>
- Moor, R., van Maren, M., & van Laarhoven, C. (2002). A Controlled Stable Tidal Inlet at Cartagena de Indias, Colombia. *Terra et Aqua*, 88, 3-14.
- Palomino, D., y Serge, N. (2015). *Variación de la calidad del agua de la ciénaga de la virgen producto de la implementación del emisario submarino* [Tesis de grado]. Universidad de Cartagena. <https://repositorio.unicartagena.edu.co/handle/11227/1381?show=full&locale-attribute=en>
- Rice, E., Baird, R., & American Public Health Association. (2017). *Standard methods for the examination of water and wastewater* (23.^a ed.). American public health association.
- Roy, B. N., Roy, H., Rahman, K. S., Mahmud, F., Bhuiyan, M. M. K., Hasan, M., ... & Islam, M. S. (2024). Principal component analysis incorporated water quality index modeling for Dhaka-based rivers. *City and Environment Interactions*, 23, 100150.
- Sun, X., Zhang, H., Zhong, M., Wang, Z., Liang, X., Huang, T., & Huang, H. (2019). Analyses on the Temporal and Spatial Characteristics of Water Quality in a Sea-going River Using Multivariate Statistical Techniques: A Case Study in the Duliujian River, China. *International Journal of Environmental Research and Public Health*, 16(6), 1020. <https://doi.org/10.3390/ijerph16061020>
- Tripathi, M., & Singal, S. K. (2019). Use of Principal Component Analysis for parameter selection for development of a novel Water Quality Index: A case study of river Ganga India. *Ecological Indicators*, 96, 430-436. <https://doi.org/10.1016/j.ecolind.2018.09.025>
- Varol, M. (2020). Spatio-temporal changes in surface water quality and sediment phosphorus content of a large reservoir in Turkey. *Environmental Pollution*, 259, 113860. <https://doi.org/10.1016/j.envpol.2019.113860>
- Xiong, J., Lin, C., Cao, Z., Hu, M., Xue, K., Chen, X., & Ma, R. (2022). Development of remote sensing algorithm for total phosphorus concentration in eutrophic lakes: Conventional or machine learning? *Water Research*, 215, 118213. <https://doi.org/10.1016/j.watres.2022.118213>

Yilma, M., Kiflie, Z., Windsperger, A., & Gessese, N. (2019). Assessment and interpretation of river water quality in Little Akaki River using multivariate statistical techniques. *International Journal of Environmental Science and Technology*, 16(7), 3707-3720. <https://doi.org/10.1007/s13762-018-2000-8>

Zamora-López, A., Guerrero-Gómez, A., Torralva, M., Zamora-Marín, J. M., Guillén-Beltrán, A., & Oliva-Paterna, F. J. (2023). Shallow waters as critical habitats for fish assemblages under eutrophication-mediated events in a coastal lagoon. *Estuarine, Coastal and Shelf Science*, 291, 108447. <https://doi.org/10.1016/j.ecss.2023.108447>

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.



Licencia de Creative Commons

Revista de Investigación Agraria y Ambiental is licensed under a Creative Commons Reconocimiento - No Comercial - Compartir Igual 4.0 Internacional License.