

DEPÓSITO LEGAL ZU2020000153

ISSN 0041-8811

E-ISSN 2665-0428

Revista de la Universidad del Zulia

Fundada en 1947
por el Dr. Jesús Enrique Lossada



Ciencias del
Agro,
Ingeniería
y Tecnología

Año 15 N° 42

Enero - Abril 2024

Tercera Época

Maracaibo-Venezuela

Aplicación de la estadística multivariante para el desarrollo de un modelo de calidad del agua del Estero El Macho, Machala-Ecuador

Rafael Gilberto Zhindon-Almeida*
Raúl Alfredo Sánchez-Ancajima**
Walter Javier Castañeda-Guzmán***

RESUMEN

El agua es un elemento indispensable para las personas, así como para el sostenimiento de la vida en el planeta. De esta forma, debido a la contaminación que puede sufrir el agua de tipo químico o biológico, es necesario tener conocimiento sobre su disponibilidad en función de sus características de calidad. En consecuencia, los modelos estadísticos son el medio pertinente para realizar una evaluación y estimaciones futuras sobre la calidad de las aguas superficiales. Por tanto, el presente trabajo va encaminado a formular un modelo estadístico que permita determinar el grado de contaminación de las aguas del estero El Macho, ubicado en la ciudad de Machala, provincia de El Oro. La metodología aplicada fue de tipo básico, ya que se realizó una búsqueda de fundamentos teóricos que permitan profundizar en la modelación estadística para estudios de calidad del agua. El diseño de la investigación es correlacional bajo un modelo estadístico multivariante de regresión lineal múltiple. El modelo de regresión lineal múltiple unificado concluyó que la demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero, de acuerdo al parámetro químico demanda química oxígeno y el parámetro físico sólidos disueltos totales. El coeficiente de determinación indica que estas variables explican el 99,9% de la variación de la variable dependiente.

PALABRAS CLAVE: Modelos estadísticos, calidad de agua, regresión lineal múltiple, Ecuador.

* Docente. Universidad Nacional de Tumbes – Perú. ORCID: <https://orcid.org/0000-0002-3901-1829>. E-mail: rzhindon.itso@gmail.com

** Docente. Universidad Nacional de Tumbes – Perú. ORCID: <https://orcid.org/0000-0003-3341-7382>. E-mail: rsanchez@untumbes.edu.pe

*** Docente. Universidad Nacional de Tumbes – Perú. ORCID: <https://orcid.org/0000-0002-9483-0210>. E-mail: wcasta@untumbes.edu.pe

Recibido: 22/09/2023

Aceptado: 15/11/2023

Application of Multivariate Statistics for the Development of a Water Quality Model of Estuary El Macho, Machala-Ecuador

ABSTRACT

Water is an essential element for people, as well as for sustaining life on the planet. In this way, due to the contamination that water can suffer from a chemical or biological type, it is necessary to have knowledge about its availability based on its quality characteristics. Consequently, statistical models are the relevant means to carry out an evaluation and future estimates about the quality of surface waters. Therefore, the present work is aimed at formulating a statistical model that allows determining the degree of contamination of the waters of the El Macho estuary, located in the city of Machala, province of El Oro. The applied methodology was of a basic type, since a search of theoretical foundations was carried out that allows to deepen in the statistical modeling for studies of water quality. The research design is correlational under a multivariate statistical model of multiple linear regression. The unified multiple linear regression model concluded that the biochemical oxygen demand significantly determines the water quality of the estuary, according to the chemical parameter, chemical oxygen demand and the physical parameter total dissolved solids. The coefficient of determination indicates that these variables explain 99.9% of the variation of the dependent variable.

KEYWORDS: Statistical models, Water quality, Multiple linear regression, Ecuador.

Introducción

El agua es el elemento vital de sustento de la vida en el planeta Tierra. Por la importancia de este recurso es necesario tener conocimiento y control sobre sus características de calidad. La forma tradicional de llevar a cabo un estudio de calidad de agua es midiendo determinados parámetros tales como el oxígeno disuelto, niveles de pH y temperatura, presencia de nitratos, coliformes fecales, sólidos disueltos, así como turbiedad, entre otros (Torres et al., 2009).

No obstante, estos indicadores por sí mismos resultan insuficientes sin un análisis estadístico que ayude a demostrar la variación espacial y temporal de la calidad del agua. En este sentido, los modelos estadísticos son el medio pertinente para realizar una evaluación y estimaciones futuras sobre la calidad de las aguas superficiales. Torres, Hernán y Patiño (2009) señalan que el uso de la estadística en procesos de análisis de los índices de calidad del agua que permiten revelar las variaciones espacio temporal de las propiedades del

recurso hídrico, ayudando además a interpretar una gran cantidad de datos, identificando tendencias sobre la calidad.

En este sentido, el presente estudio se sitúa en la aplicación de la estadística multivariante para el desarrollo de un modelo predictivo de la calidad del agua del Estero El Macho, de la ciudad de Machala en la provincia El Oro, mismo que recibe la descarga de aguas residuales de aproximadamente el 48% de los habitantes de la ciudad (Serie Palas, 2020), además sirve de drenaje a 800 ha cultivo de banano.

Bajo este contexto, se plantea la siguiente interrogante al problema: ¿El desarrollo de un modelo estadístico permitirá determinar en forma significativa la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro? En consecuencia, la problemática se orienta a la utilización de la estadística multivariante para analizar simultáneamente una serie de datos provenientes de varias variables, medidas para cada objeto de estudio, es contribuir con una mejor comprensión de las múltiples correlaciones de dicho fenómeno.

Respecto a investigaciones sobre estudios de calidad del agua, Samboni, Reyes y Carvajal (2011) consideran importante la realización de estudios de indicadores de calidad del agua, para los ámbitos domésticos, industriales o agrícolas. Bajo estos argumentos, un adecuado procedimiento de medición consiste en integrar parámetros de tipo fisicoquímicas y biológicas.

Respecto a la aplicación de técnicas de estadística multivariante Sotomayor (2016) considera que la temática es de suma importancia como consecuencia del avance de las ciudades e incremento de las actividades industriales y agrícolas que incrementa el uso de recursos hídricos que a la vez deterioran su funcionalidad. En este sentido, señala que se puede aplicar el análisis de componentes principales (PCA) de datos multivariados para garantizar un nivel alto de rigurosidad científica en la evaluación de la calidad de agua

Por su parte Gómez y Peñuela (2017) señalan que dentro de los métodos estadísticos apropiados para interpretar matrices sobre la calidad del agua se encuentran los siguientes tipos de análisis: clúster, de componentes principales, factorial y discriminante. Mediante la aplicación de estos métodos es posible desarrollar modelos estadísticos para estudios de calidad de agua, tales como: a) correlación canónica; b) modelo jerárquico o de efectos mixtos; c) modelo de ecuaciones estructurales: estudia de forma simultánea las relaciones de las variables endógenas y exógenas. Al respecto, Gil Marín (2020) estos métodos y

modelos contribuyen a obtener e interpretar información valiosa para realizar diagnósticos y predicciones importantes para la toma de decisiones respecto a los procesos de manejo de calidad de agua.

En función de este conjunto de investigaciones se puede concluir que los principales métodos estadísticos para evaluar la calidad del agua es el análisis de regresión lineal múltiple, método de análisis de componentes principales, análisis clúster y análisis factorial. Estos métodos permiten del desarrollo de modelos que contribuyen a establecer estimaciones y predicciones sobre la calidad del agua, y por ende, una mejor gestión de los recursos hídricos ya sea para el consumo, la industria o la agricultura (Quiñones Huatangari, 2021).

Bajo estos fundamentos teóricos el presente estudio tiene el objetivo de desarrollar un modelo estadístico para determinar la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro.

1. Metodología

1.1. Tipo y diseño de la investigación

Tomando como referencia la investigación de Sokolov y Black (1996), sobre la modelación de la evolución temporal de parámetros de calidad del agua, y de Shen y otros (2022), sobre proyectos de recuperación de agua de un estero, la presente investigación se plantea de tipo aplicada. En este sentido, en el estudio se pone en práctica los aspectos conceptuales relacionados con la modelación estadística para estudios de calidad del agua, a fin de describir y predecir la evolución de los parámetros de calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, en Ecuador.

Según una revisión crítica de bibliografía realizada por Syeed y otros (2023) para orientar la investigación a la caracterización y definición de las propiedades del agua mediante métodos estadísticos, en nivel de profundización de la investigación debe tener al menos un alcance descriptivo. El diseño de la investigación es no experimental, llegando al nivel correlacional de análisis con un modelo estadístico multivariante de regresión lineal múltiple. Según Morantes-Quintana y otros (2019), el modelo de regresión lineal múltiple con q variables responde a la ecuación: $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_q X_{iq} + \varepsilon_i$

Donde:

Y = variable dependiente, valor observado

X = variable independiente, variable predictora

β = magnitud de las observaciones

ε = error por factores no controlados

1.2. Muestreo

El objeto de estudio son las aguas del estero El Macho de la ciudad de Machala, provincia de El Oro. Teniendo en cuenta la metodología empleada por Morantes-Quintana y otros (2019), en el que se utilizaron las muestras de material particulado recolectadas durante un periodo de 72 horas reservadas para validar el modelo, se seleccionaron 4 puntos de recolección de muestras que fueron sometidas a los estudios de laboratorio.

1.3. Hipótesis general

La demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, conforme a los parámetros físico, químicos y microbiológicos.

1.4. Hipótesis específicas

- La demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, conforme a los parámetros físicos.
- La demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, conforme a los parámetros químicos.
- La demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, conforme a los parámetros microbiológicos.

1.5. Plan de procesamiento y análisis de datos

Teniendo en cuenta lo sugerido por Álvarez y otros (2012), Sotomayor (2016) y Bluhm Gutiérrez (2008), el plan para el procesamiento y análisis de datos consisten en:

- Selección de 4 puntos de recolección de muestras.
- Obtención de los análisis de laboratorio de los parámetros fisicoquímicos y biológicos de las muestras del agua.
- Organización y clasificación de los datos recabados.
- Determinación del grado de incidencia de los parámetros estadístico para la calidad del agua del estero El Macho mediante el modelo estadístico multivariante de regresión lineal múltiple. Siguiendo un método similar al empleado por Nel y otros (2022).
- Obtención de la ecuación y gráficas para la modelación, estudio y control de la calidad del agua del estero mediante la aplicación de los softwares estadísticos: programa estadístico SPSS. Siguiendo un método similar al empleado por Bravo Moreno (2022).
- Validar el modelo estadístico mediante los métodos de linealidad, normalidad, homocedasticidad e independencia.
- Interpretar los resultados obtenidos, resaltando la importancia de la demanda bioquímica de oxígeno como indicador clave de la calidad del agua en el estero El Macho.
- Discutir sobre las implicaciones y posibles acciones para mejorar la calidad del agua en función de los hallazgos.

2. Resultados

La base de datos fue tomada del monitoreo realizado por la Empresa Pública Aguas Machala en el año 2017, en cuatro puntos o sectores:

- A-1 Drenaje Macho El Cambio
- A-2 La Primavera 1
- A-3 La Primavera 2
- A-4 Estero Salinas (ver anexo 5)

Tabla 1. Resultados de monitoreo de agua

Parámetros		Sector			
		A-1	A-2	A-3	A-4
Parámetro biológico	Demanda Bioquímica Oxígeno	13,59	4,75	10,71	37,07
	Sólidos Disueltos Totales	1280,00	638,00	1046,00	22784,00
Parámetros físicos	Temperatura	25,00	24,80	25,00	21,10
	Turbidez	9,50	20,70	17,80	15,30
	Aceites Grasas IR	0,2	0,2	0,2	0,2
Parámetros químicos	Potencial Hidrógeno	7,87	7,44	7,17	7,38
	Demanda Química Oxígeno	30,80	10,00	22,00	75,30
	Oxígeno Disuelto	97,55	64,83	14,93	33,26
	Fosfatos	5,54	1,23	1,74	1,56
	Nitritos	0,01	0,032	0,01	0,012
	Nitratos	1	1	1	1
	Sulfatos	151,20	102,50	103,50	1582,00
	Arsénico	0,008	0,005	0,006	0,003
	Cadmio	0,0004	0,0004	0,0008	0,0004
	Cobre	0,002	0,0024	0,002	0,001
	Hierro	0,54	0,8	0,9	0,56
Parámetros microbiológicos	Coliformes Fecales	540,00	13,00	68,00	24,00
	Coliformes Totales	920,00	27,00	330,00	40,00

Fuente: Informe de Rendición de Cuentas 2017 (Aguas Machala EP, 2022).

-Regresión lineal múltiple

Según Mendoza y otros (2023), un modelo de regresión lineal múltiple es una herramienta estadística flexible que permite analizar las relaciones entre una variable objetivo continua y varios predictores preestablecidos. Siguiendo la lógica del análisis, en seguida, se presentan las comprobaciones.

- Comprobación de la hipótesis I: el primer modelo de correlación es entre la variable dependiente (*Demanda bioquímica de oxígeno*) y las cuatro variables independientes

(*Demanda Bioquímica Oxígeno, Sólidos Disueltos Totales, Temperatura, Turbidez, Aceites y Grasas IR*) se empleó el método de “Stepwise” o “Por pasos”, con la finalidad de que el programa SPSS evalúe y determine un modelo o modelos que presenten mayores correlaciones.

Tabla 2. Correlaciones

		Demanda Bioquímica Oxígeno	Sólidos Disueltos Totales	Temperatura	Turbidez	Aceites y Grasas IR
Correlación de Pearson	Demanda Bioquímica Oxígeno	1,000	,972	-,952	-,302	.
	Sólidos Disueltos Totales	,972	1,000	-,997	-,096	.
	Temperatura	-,952	-,997	1,000	,039	.
	Turbidez	-,302	-,096	,039	1,000	.
	Aceites y Grasas IR	1,000

La variable Sólidos Disueltos Totales tiene una correlación de Pearson de ,972.

Por tanto, se comprobó que la variable dependiente tiene una correlación significativa con: Sólidos disueltos totales. En consecuencia, se descartaron o eliminaron las variables de temperatura y turbidez, además, se excluyó la variable Aceites y Grasas IR por mantener valores constantes.

Ecuación de regresión 1

Tabla 3. Coeficientes ^a

		Totales	
Modelo		a. Variable dependiente: Demanda Bioquímica Oxígeno	
		B	Error estándar
1	(Constante)	8,394	2,481
	Sólidos Disueltos	,001	,000

$$y = B_0 + B_1 \cdot x_1 + B_2 \cdot x_2 + B_n \cdot x_n$$

$$y = 8,394 + 0,001x \quad (1)$$

Validación

Tabla 4. R Cuadrado e Independencia

Modelo ^b	R	R cuadrado	R cuadrado ajustado	Durbin-Watson ^c
1	,972 ^a	,944	,916	2,837

a. Predictores: (Constante), Sólidos Disueltos Totales

b. Variable dependiente: Demanda Bioquímica Oxígeno

c. Criterio: Si el valor está entre 1 y 3 aceptamos el supuesto de independencia de errores.

a) **Coefficiente de determinación – R cuadrado:** la variable independiente Sólidos disueltos totales, se explica el 94,4% de la varianza de la variable dependiente (R cuadrado: ,944)

b) **Prueba de Durbin-Watson – Independencia:** El puntaje indica que hay independencia de errores (2,837)

b. Criterio: Ningún valor por encima de 10 y en conjunto todos los valores cercanos a 1.

Tabla 5. Colinealidad

Modelo ^a	Estadísticas de colinealidad ^b	
	Tolerancia	VIF
1	Sólidos Disueltos Totales	1,000
		1,000

a. Variable dependiente: Demanda Bioquímica Oxígeno

c) **Colinealidad:** El factor de varianza inflada (FIV) indica que se cumple el supuesto de no colinealidad (FIV = 1,000).

Tabla 6. Prueba de normalidad

	Shapiro-Wilk		
	Estadístico	gl	Sig.
Standardized Residual	,968	4	,828

d) **Prueba de normalidad:** La prueba de normalidad SW indica que se cumple el supuesto de normalidad (Estadístico = ,968; gl = 4; Sig. ,828)

La demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, conforme al parámetro físico sólidos disueltos totales

- **Comprobación de la hipótesis 2:** el segundo modelo de correlación entre la variable dependiente Demanda bioquímica de oxígeno y las once variables independientes, se empleó el método de “Stepwise” o “Por pasos”, con la finalidad de que el programa SPSS evalúe y determine un modelo o modelos que presenten mayores correlaciones.

Al efectuar el análisis de correlación de Pearson, se comprobó una correlación significativa con la variable demanda química oxígeno de 0,999.

Por tanto, se descartaron o eliminaron las variables de potencial hidrógeno, oxígeno disuelto, fosfatos, nitritos, sulfatos, arsénico, cadmio, cobre y hierro, al no presentar correlaciones significativas con la variable dependiente. Asimismo, se excluyó la variable nitratos, por presentar valores constantes.

- **Comprobación de la hipótesis 3:** Se aplicó el método de regresión múltiple por pasos, el cual no generó modelo de regresión, debido a que no hay correlaciones significativas entre las variables Demanda Bioquímica Oxígeno con las variables Coliformes Fecales y Coliformes Totales.

Modelo unificado

El propósito de este proceso es verificar si existe un modelo que se puede configurar con las variables predictoras que presentaron mayor correlación respecto a la variable dependiente demanda bioquímica de oxígeno (Montero Granados, 2016). En este sentido se utilizó en el programa SPSS el método de regresión múltiple aplicado fue el jerárquico, dando los siguientes resultados:

Tabla 7. Correlaciones

	Demanda Bioquímica Oxígeno	Potencial Hidrógeno	Demanda Química Oxígeno	Oxígeno Disuelto	Fosfatos	Nitritos	Nitratos	Sulfatos	Arsénico	Cadmio	Cobre	Hierro
Demanda Bioquímica Oxígeno	1,000	-,074	,999	-,294	-,103	-,479	.	,971	-,629	-,274	-,994	-,626
Potencial Hidrógeno	-,074	1,000	-,027	,957	,878	-,076	.	-,165	,594	-,669	,142	-,729
Demanda Química Oxígeno	,999	-,027	1,000	-,252	-,052	-,500	.	,961	-,592	-,293	-,991	-,661
Oxígeno Disuelto	-,294	,957	-,252	1,000	,760	,196	.	-,332	,598	-,692	,373	-,556
Fosfatos	-,103	,878	-,052	,760	1,000	-,465	.	-,286	,807	-,256	,122	-,585
Nitritos	-,479	-,076	-,500	,196	-,465	1,000	.	-,264	-,239	-,374	,542	,336
Nitratos	1,000
Sulfatos	,971	-,165	,961	-,332	-,286	-,264	.	1,000	-,783	-,347	-,953	-,549
Arsénico	-,629	,594	-,592	,598	,807	-,239	.	-,783	1,000	,160	,617	,000
Cadmio	-,274	-,669	-,293	-,692	-,256	-,374	.	-,347	,160	1,000	,167	,748
Cobre	-,994	,142	-,991	,373	,122	,542	.	-,953	,617	,167	1,000	,564
Hierro	-,626	-,729	-,661	-,556	-,585	,336	.	-,549	,000	,748	,564	1,000

Ecuación de regresión 2

Tabla 8. Coeficientes ^a

Modelo	Coeficientes no estandarizados	
	B	Error estándar
1 (Constante)	-,631	,767

Sólidos Disueltos Totales	,497	,018	
---------------------------	------	------	--

a. Variable dependiente: Demanda Bioquímica Oxígeno

$$y = B_0 + B_1 \cdot x_1 + B_2 \cdot x_2 + B_n \cdot x_n$$

$$y = -0,631 + 0,497x \quad (2)$$

Validación

Tabla 9. R cuadrado e Independencia

Modelo ^b	R	R cuadrado	R cuadrado ajustado	Durbin-Watson ^c
1	,999 ^a	,997	,996	,89216

a. Predictores: (Constante), Sólidos Disueltos Totales

b. Variable dependiente: Demanda Bioquímica Oxígeno

c. Criterio: Si el valor está entre 1 y 3 aceptamos el supuesto de independencia de errores.

a) **Coefficiente de determinación – R cuadrado:** la variable independiente demanda química de oxígeno, esta explica el 99,7% de la varianza de la variable dependiente (R cuadrado: ,997)

b) **Prueba de Durbin-Watson – Independencia:** El puntaje indica que hay independencia de errores (1,423)

Tabla 10. Colinealidad

Modelo	Estadísticas de colinealidad ^b	
	Tolerancia	VIF
1 Sólidos Disueltos Totales	1,000	1,000

a. Variable dependiente: Demanda Bioquímica Oxígeno

b. Criterio: Ningún valor por encima de 10 y en

conjunto todos los valores cercanos a 1.

c) **Colinealidad:** El factor de varianza inflada (FIV) indica que se cumple el supuesto de no colinealidad (FIV = 1,000).

Tabla 11. Prueba de normalidad

	Shapiro-Wilk		
	Estadístico	gl	Sig.
Standardized Residual	,696	4	,010

d) **Prueba de normalidad:** La prueba de normalidad SW indica que no se cumple el supuesto de normalidad (Estadístico = ,696; gl = 4; Sig. = ,010)

La demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, conforme al parámetro químico de Demanda Química Oxígeno.

Tabla 12. Variables entradas/eliminadas

Modelo	Variables entradas	Método
1	Demanda Química Oxígeno	Por pasos (Criterios: Probabilidad-de-F-para-entrar \leq ,050, Probabilidad-de-F-para-eliminar \geq ,100).
2	Sólidos Disueltos Totales ^b	Entrar

a. Variable dependiente: Demanda Bioquímica Oxígeno

b. Todas las variables solicitadas introducidas.

Por tanto, se aplica el modelo 2 conformada por todas las variables.

Correlaciones

Tabla 13. Correlaciones

		Demanda Bioquímica Oxígeno	Demanda Química Oxígeno	Sólidos Disueltos Totales
Correlación de Pearson	Demanda Bioquímica Oxígeno	1,000	,999	,972
	Demanda Química Oxígeno	,999	1,000	,961
	Sólidos Disueltos Totales	,972	,961	1,000

Ecuación de regresión (unificada)

Tabla 14. Coeficientes ^a

Modelo ^a	Coeficientes no estandarizados	
	B	Error estándar
(Constante)	,635	1,024
2 Demanda Química	,423	,051

Oxígeno		
Sólidos Disueltos Totales	,000	,000

$$y = B_0 + B_1*x_1 + B_2*x_2 + B_n*x_n$$

$$y = 0,635 + 0,423x \quad (3)$$

a. Variable dependiente: Demanda Bioquímica Oxígeno

Validación

Tabla 15. R Cuadrado e Independencia

Modelo ^b	R	R cuadrado	R cuadrado ajustado	Durbin-Watson ^c
2	1,000 ^a	,999	,998	,69326

a Predictores: (Constante), Demanda Química Oxígeno, Sólidos Disueltos Totales

b. Variable dependiente: Demanda Bioquímica Oxígeno

c. Criterio: Si el valor está entre 1 y 3 aceptamos el supuesto de independencia de errores.

a) Coeficiente de determinación – R cuadrado: la variable independiente demanda química de oxígeno y sólidos disueltos totales, estas explican el 99,9% de la varianza de la variable dependiente (R cuadrado: ,999)

b) Prueba de Durbin-Watson – Independencia: El puntaje indica que hay independencia de errores (2,029)

Tabla 16. Colinealidad

Modelo			Estadísticas de colinealidad ^b	
			Tolerancia	VIF
2	Demanda	Química	,076	13,129
	Oxígeno			
	Sólidos	Disueltos	,076	13,129
	Totales			

a. Variable dependiente: Demanda Bioquímica Oxígeno

b. Criterio: Ningún valor por encima de 10 y en conjunto todos los valores cercanos a 1.

c) **Colinealidad:** El factor de varianza inflada (FIV) indica que no se cumple el supuesto de no colinealidad (FIV = 13,129)

Tabla 17. Prueba de normalidad

	Shapiro-Wilk		
	Estadístico	gl	Sig.
Standardized Residual	,877	4	,328

d) **Prueba de normalidad:** La prueba de normalidad SW indica que se cumple el supuesto de normalidad (Estadístico = ,877; gl = 4; Sig. ,328)

La demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, conforme al parámetro químico demanda química oxígeno y el parámetro físico sólidos disueltos totales.

3. Discusión

Los resultados de las pruebas de hipótesis permitieron elaborar los respectivos modelos predictivos que demuestran lo siguiente:

- El modelo propuesto con el parámetro de demanda bioquímica de oxígeno y los 4 parámetros físicos (ver tabla 1) dio como resultado que existe una relación significativa con el parámetro físico de sólidos disueltos. Esto se puede contrastar con los resultados del estudio de Flores Suárez y Pozo Merejildo (2023), en el que, además, se llega a demostrar que, para la demanda bioquímica y química de oxígeno, la carga contaminante de salida difiere de la carga contaminante de entrada.
- El modelo propuesto que vincula la demanda bioquímica de oxígeno con los 11 parámetros químicos (ver tabla 1) reveló que existe una relación significativa sólo con la Demanda Química Oxígeno. Esta demostración corrobora los hallazgos de García Juárez

y otros (2022), quienes determinan, además, que es viable el tratamiento de aguas en condiciones similares para su reutilización.

- El modelo propuesto que vincula la demanda bioquímica de oxígeno con los parámetros microbiológicos no generó relaciones para un análisis predictivo. Una situación similar concluyen Yamashita y otros (2022) para lo cual sugieren la implementación de un biosensor de la demanda bioquímica de oxígeno para el tratamiento de aguas residuales. Por su parte, Lopez Balladares y Tooth Flores (2022) proponen, para esta situación, la remoción de la demanda química de oxígeno empleando filtros de carbón activado para las aguas residuales.

De acuerdo a García et al (2021), la disparidad en los resultados arrojados por la eliminación de parámetros en los distintos modelos propuestos, estriba en los aspectos de la propia naturaleza o condiciones ambientales y a la necesidad de asignar grados de importancia en los parámetros para que los datos converjan en un modelo con mejores correlaciones y menores desviaciones.

A pesar de ello, Gil Marin (2020) señala que la regresión lineal múltiple es importante para efectuar análisis más aproximados a la realidad de los hechos o procesos naturales, en consecuencia, proporcionó una precisión adecuada para el pronóstico de la calidad de agua. El análisis multivariante, como parte de los métodos cuantitativos de investigación es considerado como uno de los de mayor efectividad al momento de verificar un fenómeno, proceso o recurso (Closas et al., 2013).

Bajo estos argumentos, en la presente investigación se desarrolló además un modelo unificado de todos los parámetros de calidad de agua, aplicando el método de regresión múltiple jerárquico, llegando a resultados similares, a nivel interpretativo, a los logrados por Shenbagalakshmi y otros (2023), quienes aplican un análisis jerárquico profundo de conglomerados (cluster). En el presente estudio se llega a comprobar que la demanda bioquímica de oxígeno determina, en forma significativa, la calidad del agua del estero El Macho de la ciudad de Machala, provincia de El Oro, conforme al parámetro químico demanda química oxígeno y el parámetro físico sólidos disueltos totales. Entonces, mediante la aplicación de la estadística multivariante se puede entender cómo los datos de una variedad de variables se relacionan entre sí y poder efectuar análisis como el de la calidad del agua (Giraud Herrera & Morantes Quintana, 2017).

Conclusión

Una vez finalizada la investigación se pudo comprobar :

- La verificación de la hipótesis 1, permitió establecer una correlación fuerte de ,972 entre la variable dependiente Demanda Bioquímica Oxígeno y una sola variable independiente Sólidos Disueltos Totales. El coeficiente de determinación señala que el modelo explica el 94,4% de la varianza de la variable dependiente.
- La verificación de la hipótesis 2, determinó una correlación fuerte de ,999 entre la variable dependiente Demanda Bioquímica Oxígeno y una sola variable independiente Demanda Química Oxígeno. El coeficiente de determinación indica que el modelo explica el 99,7% de la varianza de la variable dependiente.
- En relación con la hipótesis 3, se determinó que no hay correlaciones significativas entre las variables Demanda Bioquímica Oxígeno con las variables Coliformes Fecales y Coliformes Totales.
- Se implementó un modelo unificado con las variables que presentaron mayor correlación respecto a la variable dependiente: demanda bioquímica de oxígeno. Por lo tanto, se concluye, en que la demanda bioquímica de oxígeno determina en forma significativa la calidad del agua del estero, conforme establece el parámetro químico demanda química oxígeno y el parámetro físico sólidos disueltos totales. El coeficiente de determinación indica que estas variables explican el 99,9% de las alteraciones/variaciones de la variable dependiente.
- La validación mediante pruebas de independencia y normalidad respalda la confiabilidad del modelo estimado para evaluar la calidad del agua, asegurando resultados confiables y estimaciones precisas en futuras aplicaciones.

Referencias

Aguas Machala EP. (2022). *Transparencia*. Ley Orgánica de Transparencia y Acceso a la Información Pública. <https://www.aguasmachala.gob.ec/transparencia>

Álvarez, R., Vergara, E., Corral, M., & Navarro, Á. (2012). Análisis multivariante de la línea de aguas consuntivas en la ETAP "Río Iregua" de Logroño. *XVI Congreso Internacional de Ingeniería de Proyectos*, 846-857.

Bluhm Gutiérrez, J. (2008). *Modelación unidimensional de la calidad del agua en embalses. Análisis comparativo de modelos y multivariantes* [Http://purl.org/dc/dcmitype/Text, Universitat Politècnica de València]. <https://dialnet.unirioja.es/servlet/tesis?codigo=18239>

Bravo Moreno, J. I. (2022). *Modelación Matemática del tratamiento Anaerobio de Aguas Residuales Urbanas para la Hacienda Teresita, sector Agrícola de Milagro*. En: <https://repositorio.unemi.edu.ec/handle/123456789/6050>

Closas, A., Arriola, E. A., Zening, C. I. K., Amarilla, M. R., & Jovanovich, E. C. (2013). Análisis multivariante, conceptos y aplicaciones en Psicología Educativa y Psicometría. *Enfoques: Revista de la Universidad Adventista del Plata*, 25(1), 65-92.

Flores Suárez, Á. G., & Pozo Merejildo, E. J. (2023). *Evaluación estadística de los parámetros de análisis físico, químico y biológico de los afluentes y efluentes del sistema de tratamiento de aguas residuales regulado por la Empresa Pública Mancomunada Aguapen-EP, ubicado en la cabecera cantonal del cantón Santa Elena, provincia de Santa Elena*.

García Juárez, H. D., Mendoza Zuta, J. C., Armas Juárez, R. A., & Cruz Salinas, L. E. (2022). Tratamiento de aguas residuales provenientes del proceso de curtido de pieles. *Alfa Revista de Investigación en Ciencias Agronómicas y Veterinaria*, 6(18), 423-435.

Gil-Marín, J. A., & González, A. (2020). Modelo de calidad del agua de un río mediante el uso combinado de análisis de componentes principales (ACP) y regresiones lineales múltiples (RLM). Caso de estudio: Cuenca del río Guarapiche, Maturín, Monagas, Venezuela. *Anales Científico* 81(1), 152-172.

Giraud Herrera, L. M., & Morantes Quintana, G. R. (2017). Aplicación del análisis multivariante para la sostenibilidad ambiental urbana. *Bitácora Urbano Territorial*, 27(1), 89-100.

Gómez Miranda, I. N., & Peñuela Mesa, G. A. (2017). Revisión de los métodos estadísticos multivariados usados en el análisis de calidad de aguas. *Revista Mutis; Vol. 6, Núm. 1 (2016); 54-63*. <https://doi.org/10/647>

Lopez Balladares, J. C. J., & Tooth Flores, F. M. (2022). *Remoción de demanda química de oxígeno de aguas residuales empleando carbón activado en la laguna de oxidación, Nuevo Chimbote-2022*. En: <https://repositorio.ucv.edu.pe/handle/20.500.12692/97332>

Mendoza, K. S., Acosta, D. S., Torres, L. F., & Parra, J. M. (2023, mayo 2). Modelo de regresión lineal múltiple sobre el PIB. *RPubs*. https://rpubs.com/Karen_Mendoza/1038030

Montero Granados, R. (2016). Modelos de regresión lineal múltiple. *Documentos de Trabajo en Economía Aplicada*. En: https://www.ugr.es/~montero/matematicas/regresion_lineal.pdf

Morantes-Quintana, G. R., Rincón-Polo, G., & Pérez-Santodomingo, N. A. (2019). Modelo de regresión lineal múltiple para estimar concentración de PM1. *Revista Internacional de Contaminación Ambiental*, 35(1), 179-194.

Nel, T., Clarke, C. E., & Hardie, A. G. (2022). Evaluation of simple and multivariate linear regression models for exchangeable base cation conversion between seven measurement

techniques on South African soils. *Geoderma Regional*, 30, e00571. <https://doi.org/10.1016/j.geodrs.2022.e00571>

Osorio-Ortega, M., García-González, J., Saquicela-Rojas, R., & Cadme, M. (2021). Determinación del índice de calidad del agua en ríos de Santo Domingo de los Tsáchilas, Ecuador. *Ingeniería del Agua*, 25(2), 115-126.

Quiñones Huatangari, L. (2021). *Estimación de la calidad del agua, mediante el desarrollo de un modelo matemático dinámico, Río Utcubamba-Perú*. En: <https://renati.sunedu.gob.pe/handle/sunedu/3204966>

Samboni, N., Trujillo, A. R., & Carvajal, Y. (2011). Aplicación de los indicadores de calidad y contaminación del agua en la determinación de la oferta hídrica neta. *Ingeniería y Competitividad*, 13(2), 49-60.

Sarie Palas, A. D. (2020). Análisis de la vulnerabilidad de las familias en el sector denominado Manglar del Afro de las Riberas del Canal El Macho de la Ciudad de Machala. *Repositorio Digital FLACSO Ecuador*, 50. <http://hdl.handle.net/10469/16543>

Shen, Y., Zhang, H., & Tang, J. (2022). Hydrodynamics and water quality impacts of large-scale reclamation projects in the Pearl River Estuary. *Ocean Engineering*, 257, 111432. <https://doi.org/10.1016/j.oceaneng.2022.111432>

Shenbagalakshmi, G., Shenbagarajan, A., Thavasi, S., Gomathy Nayagam, M., & Venkatesh, R. (2023). Determination of water quality indicator using deep hierarchical cluster analysis. *Urban Climate*, 49, 101468. <https://doi.org/10.1016/j.uclim.2023.101468>

Sokolov, S., & Black, K. P. (1996). Modelling the time evolution of water-quality parameters in a river: Yarra River, Australia. *Journal of Hydrology*, 178(1), 311-335. [https://doi.org/10.1016/0022-1694\(95\)02797-1](https://doi.org/10.1016/0022-1694(95)02797-1)

Sotomayor, G. (2016). *Evaluación de la calidad de las aguas superficiales mediante técnicas de estadística multivariante: Un estudio de caso en la cuenca del Río Paute, al sur de Ecuador*. En: <https://sedici.unlp.edu.ar/handle/10915/53959>

Syeed, M. M. M., Hossain, M. S., Karim, M. R., Uddin, M. F., Hasan, M., & Khan, R. H. (2023). Surface water quality profiling using the water quality index, pollution index and statistical methods: A critical review. *Environmental and Sustainability Indicators*, 18, 100247. <https://doi.org/10.1016/j.indic.2023.100247>

Torres, P., Cruz, C. H., & Patiño, P. J. (2009). Índices de calidad de agua en fuentes superficiales utilizadas en la producción de agua para consumo humano: Una revisión crítica. *Revista Ingenierías Universidad de Medellín*, 8(15), 79-94.

Yamashita, T., Hasegawa, T., Hayashida, Y., Ninomiya, K., Shibata, S., Ito, K., Mizuguchi, H., & Yokoyama, H. (2022). Energy savings with a biochemical oxygen demand (BOD)- and pH-based intermittent aeration control system using a BOD biosensor for swine wastewater treatment. *Biochemical Engineering Journal*, 177, 108266. <https://doi.org/10.1016/j.bej.2021.108266>