

## **Factores que influyen en el pH del agua mediante la aplicación de modelos de regresión lineal**

### **Factors that influence the pH of water through the application of linear regression models**

Sandra Lorena García  
Alexandra Arguello  
Richard Parra  
*Escuela Superior Politécnica del Litoral, Ecuador*

Marcela Pincay Pilay  
*Universidad Estatal del Sur de Manabí, Ecuador*

Autor para correspondencia: [slgarcia@espol.edu.ec](mailto:slgarcia@espol.edu.ec); [alex\\_yta05@hotmail.com](mailto:alex_yta05@hotmail.com);  
[rparra@espol.edu.ec](mailto:rparra@espol.edu.ec); [marcela.pincay@unesum.edu.ec](mailto:marcela.pincay@unesum.edu.ec)

Fecha de recepción: 31 de Octubre del 2018 - Fecha de aceptación: 1 de abril del 2019

**Resumen:** En este artículo se evalúan algunos factores que pudiesen alterar el pH del agua del río Chimbo ubicado en la provincia de Bolívar en Ecuador. En aguas de ríos no contaminados los valores de pH varían de acuerdo con los cambios en su oxígeno disuelto (od), sulfatos, caudal, cloruro, alcalinidad y aceites. La calidad del agua del Río Chimbo es analizada a través del valor de pH y a fin de estudiar su variabilidad se usan modelos de regresión lineal simple y regresión lineal múltiple, relacionando el pH en función de variables explicativas como: oxígeno disuelto, sulfatos, caudal, cloruro, alcalinidad y aceites. Se hace énfasis en el coeficiente de determinación  $R^2$  ajustado bajo el criterio de normalidad de datos para evaluar la calidad del agua mediante la aplicación de los modelos de regresión simple y múltiple, a fin de determinar los factores más influyentes en el pH. Se exponen y contrastan resultados para los distintos modelos estadísticos obtenidos bajo un software libre “R” y se discute la información de estos para el estudio de la calidad del agua en base a su pH. Finalmente, se establece que las variables que más influyen al pH son la alcalinidad, sulfato y cloruro presentes en el agua y se muestran algunas predicciones del pH del agua en base al mejor modelo obtenido.

**Palabras clave:** pH; calidad del agua; regresión lineal; regresión lineal múltiple; coeficiente de determinación

**Abstract:** In this article some factors that could alter the pH of the Chimbo river water are evaluated, which is in the province of Bolivar in Ecuador. In uncontaminated river waters, pH values vary according to changes in dissolved oxygen (OD), sulphates, flow, chloride, alkalinity and oils. The water quality of the Chimbo River is analyzed through the pH value and to study its variability, the simple linear regression and multiple linear regression models have been used to relate the pH and explanatory variables such as: dissolved oxygen, sulphates, flow, chloride, alkalinity and oils. Emphasis is placed on the coefficient of determination  $R^2$  adjusted under the

criterion of normality of data to evaluate the quality of water through the application of the simple and multiple regression models, in order to determine the most influential factors in the pH. Results for the different statistical models analyzed under free software "R" are exposed and contrasted and their information is discussed for the study of water quality based on its pH. Finally, it establishes that the variables that most influence the pH are the alkalinity, sulfate, and chloride present in the water and we show some predictions of the pH of the water on the basis of the best model obtained.

**Key Words:** pH; water quality; linear regression; multiple linear regression; coefficient of determination

### Introducción

La contaminación se define como la alteración del medio natural producto de factores externos, como son los físicos, químicos o biológicos (Mackenzie & Cornwell, 2012). La contaminación de la ecología acuática ocasiona problemas indestructibles o transitorios en las comunidades biológicas de ríos (Kiely, 2000, p. 309).

Las descargas residuales directas en los ríos representan una causa importante de contaminación (Varea, et al., 1997, p. 36). Por otro lado, la presencia de grandes cantidades de materia orgánica que mediante oxígeno y bacterias dan lugar a su descomposición originan la demanda biológica de oxígeno (DBO) (Kiely, 2000, p. 371). Se produce un incremento de DBO y sólidos en suspensión, también de compuestos tales como nitritos, nitratos, fosfatos y compuestos amoniacales cuando los ríos reciben las descargas de los efluentes (Kiely, 2000, p. 373). Además, Rossen et al. (2009) indican que el aumento del deterioro de la calidad del agua en los embalses y lagos se debe fundamentalmente a las descargas de líquidos de origen cloacal con un inadecuado o nulo tratamiento, constituyendo un importante riesgo de transmisión de enfermedades de origen hídrico.

Según varias publicaciones (Stow et al., 2001; Stanley & Doyle, 2002; Zubala, 2009; Zhang et al., 2013), el embalse de agua en las reservas aumenta las concentraciones de nitrógeno disuelto, amoníaco total más nitrógeno orgánico, nitrógeno total, carbono orgánico disuelto y carbono orgánico total y disminuye las concentraciones de sólidos disueltos y sedimentos suspendidos. Morris & Fan (1998) hicieron énfasis en que la relación entre la descarga y las concentraciones de contaminantes cambian cuando el embalse de agua está en los reservorios.

Los ríos son sistemas abiertos y sus aguas circulan a fin de que los organismos que dependen de él tengan cierto equilibrio. Varios indicadores constituyen una medida de la calidad del agua en un río. Estas medidas incluyen el oxígeno disuelto, la temperatura y el pH, que es un indicador de la concentración de iones hidrógeno. También, Lanza Espino et al. (2000) hacen énfasis en el pH como indicador de calidad de agua.

Otros factores externos que pueden causar fluctuaciones en el pH de un río incluyen los desechos de agricultura, los drenajes ácidos de minería y las emisiones de combustibles fósiles, como el dióxido de carbono, el cual genera un ácido débil cuando se disuelve en el agua del río.

En aguas de ríos no contaminados los valores de pH varían de acuerdo con los cambios en su oxígeno disuelto, sulfatos, caudal, cloruro, alcalinidad y aceites. La ecuación que se utiliza para medir el pH es (Chang R. , 2010):

$$PH = -\log [H_3O^+]$$

La escala del pH va del 1 al 14, mientras más bajo sea el número, más ácido será la solución, por el contrario, mientras más alto sea el número más básica será la solución (Chang R. G., 2013).

Algunos estudios relacionados al análisis del pH en aguas se mencionan a continuación:

Ansari & Salahuddin (2013) recolectaron muestras de agua subterránea de diferentes lugares de la industria de Ghazipur, India y obtuvieron los valores de pH en un laboratorio para realizar un análisis de varianza, llegando a la conclusión que en todos los lugares se presentan valores similares de pH, pero que hay variación de este según el mes de recolección. Esto se debe que el agua en esos meses tiene mayor concentración de iones de hidrógeno afectado por la sedimentación atmosférica.

Algunas técnicas estadísticas multivariantes han sido utilizadas para estudiar la calidad del agua, como en el trabajo de Salim et al. (2014) Donde utilizan regresión lineal, análisis de componentes principales y conglomerados para entender las relaciones de variables que describen las características del agua del lago Wular. Otros trabajos relacionados se pueden encontrar en Viswanath et al. (2015), Zlatanović et al. (2017) , Li et al. (2018), entre otros.

Con estos antecedentes, el pH es un indicador que permite medir la calidad del agua del Río Chimbo, y por tanto será utilizado en este trabajo. A fin de estudiar su variabilidad se usan modelos de regresión lineal simple, y regresión lineal múltiple tomando como variable respuesta el pH en función de variables explicativas como: oxígeno disuelto, sulfatos, caudal, cloruro, alcalinidad y aceites. Para medir el pH del agua se utilizaron herramientas de laboratorios a fin de obtener los datos a ser utilizados en el presente estudio.

El artículo se estructura en tres etapas: la primera presenta las principales características de los modelos de regresión a ser utilizados; la segunda presenta resultados del procesamiento de datos de las mediciones en el pH del agua, a partir del modelo de regresión lineal simple y de regresión múltiple con sus respectivos análisis y la tercera presenta el modelo que mayor ajuste se ha logrado con el pH del agua, por último se entrega conclusiones de la investigación.

## **Metodología**

Antes de realizar el modelo de regresión lineal simple se analizará la correlación de las variables objeto de estudio para descubrir la asociatividad lineal entre variables de tipo continuo. Aparicio et al. (2004) Señalan que un análisis de correlación “consta de dos partes: gráficos de dispersión y análisis de correlación propiamente dicho, que consiste en el cálculo de las correlaciones y la realización de los contrastes pertinentes”.

### **Correlación Simple**

Un análisis de correlación servirá para determinar si existe relación de tipo lineal entre las variables. Si tal relación lineal existe se tendría la justificación para ajustar el modelo de regresión lineal simple.

Se define el coeficiente de correlación (p) de Pearson como (Aparicio et al., 2004):

$$p = \frac{cov(x, y)}{\sqrt{var(x)var(y)}}$$

Donde Var identifica la varianza y COV la covarianza

Dada una muestra  $(x_1, y_1), \dots, (x_n, y_n)$  de dos variables X e Y, se calcula el coeficiente de correlación lineal simple (r) de Pearson de la siguiente manera:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Cuando existe correlación lineal entre dos variables, este coeficiente es útil para cuantificar el grado de asociación entre ellas, es decir el coeficiente de correlación de Pearson es una medida de asociación lineal libre de escala, con valores comprendidos entre -1 y 1, invariante a transformaciones lineales de las variables que se detallan en la Tabla 1.

**Tabla 1**  
**Valoración Numérica de correlación simple**

Valorización numérica	Significancia
$r = 0$	Nada de asociación lineal
$r = 1$ o $-1$	Asociación lineal perfecta
$r < 0$	Correlación negativa
$r > 0$	Correlación positiva

Fuente: adaptado de Aparicio et al. (2004)

### Modelo de regresión lineal simple

Los modelos de regresión lineal son ampliamente usados en la ingeniería ya que sirven para analizar el procedimiento de las variables de entrada y salida estableciendo predicciones y estimaciones (Montgomery et al., 2006). La ecuación (1) muestra la representación de un modelo de regresión lineal simple, donde Y es la variable de salida, X la variable de entrada,  $\beta_0$  y  $\beta_1$  son los parámetros del modelo o coeficientes de regresión y  $\epsilon$  es el error del modelo

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (1)$$

La ecuación (2) muestra el modelo ajustado de regresión lineal simple, donde  $\hat{y}$  es el valor estimado o esperado de la variable de salida para la variable de entrada X. Los estimadores del modelo son calculados por el método de mínimos cuadrados usando las ecuaciones (3) y (4)

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X \quad (2)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3)$$

$$\widehat{B}_0 = \bar{y} + \widehat{\beta}_1 \bar{x} \quad (4)$$

### Modelo de regresión lineal Múltiple

En la ecuación (5), Montgomery et al. (2006) describen el modelo de regresión lineal múltiple y en la ecuación (6) se muestra el modelo ajustado de regresión lineal múltiple. En este modelo la variable de salida Y se puede relacionar con K variables de entrada. Los estimadores de los coeficientes de este modelo se obtienen a partir de la ecuación (7).

$$Y = \beta_0 + B_1X + B_2X_2 + \dots + B_kX_k + \varepsilon \quad (5)$$

$$\widehat{Y} = \widehat{\beta}_0 + \widehat{\beta}_1X + \widehat{\beta}_2X_2 + \dots + B_kX_k \quad (6)$$

$$\widehat{\beta} = (X'X)^{-1}X'Y \quad (7)$$

Donde  $\varepsilon_i$  es una variable aleatoria que se supone normal con media cero, varianza constante y no correlacionada. Para probar normalidad se usará la prueba de Shapiro Wilk, mientras que para probar homocedasticidad (varianza constante) y no correlación se usarán las pruebas de Breusch-Pagan y Durbin-Watson, respectivamente (Aparicio et al., 2004)

En los modelos de regresión lineal múltiple se debe evaluar la presencia de la multicolinealidad que aparece cuando existe una dependencia casi lineal entre las variables de entrada X del modelo. El problema de la multicolinealidad afecta a la estimación de las varianzas de los estimadores de los coeficientes del modelo, por tanto afectaría a sus pruebas de significancia (Kutner et al., 2004)

### Coefficiente de Determinación $R^2$

Montgomery et al. (2006) especifican que este coeficiente mide la proporción de la variación de la respuesta Y que es explicada por el modelo de regresión. En la ecuación (8) se detalla cómo se obtiene este coeficiente donde  $SS_R$  es la medida de variabilidad del modelo de regresión y  $SS_T$  corresponde a la medida de variabilidad de Y sin considerar el efecto de la variable de entrada X.

$$R^2 = \frac{SS_R}{SS_T} = \frac{\sum_{i=1}^n (\widehat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad 0 \leq R^2 \leq 1 \quad (8)$$

De acuerdo a la ecuación citada si el valor de  $R^2$  se acerca a 1, entonces se tiene un buen ajuste y el modelo de regresión es capaz de explicar adecuadamente la variación de la variable de salida Y.

### Multicolinealidad

Kutner et al. (2004) presentan la ecuación (9) en donde se puede determinar la multicolinealidad de las variables de entrada X a partir de factores de inflación de la varianza.

$$FIV_j = \frac{1}{1 - R_j^2} \quad (9)$$

Donde  $R_j^2$  es el coeficiente de determinación múltiple entre  $X_j$  y las restantes p-1 variables, si el valor de  $FIV_j > 10$  da indicio de problema de multicolinealidad.

Es importante tener en cuenta que, las técnicas existentes para detectar y tratar multicolinealidad no siempre logran el éxito teniéndose que recurrir a otros parámetros de análisis (Del Valle Moreno & Guerra Bustillo, 2012).

### Descripción de los datos obtenidos

Los datos obtenidos de la presente investigación son de tipo experimental, con el propósito de medir diversas variables que caracterizan al agua del río Chimbo en 15 puntos distintos tomados de manera aleatoria, para lo cual se tomaron muestras que fueron llevadas a laboratorio a fin de medir los valores de pH, cloruro, sulfatos, aceites, alcalinidad y caudal.

### Resultados

#### Valores de pH del agua

En la Tabla 2 se puede visualizar los 15 puntos de referencia del río Chimbo donde se midió el valor de pH. Se puede apreciar que los valores de pH de cada punto están dentro del rango de 6 a 8.5 que puede considerarse aceptable para ser potabilizado de acuerdo a lo que se indica en APHA et al. (1992)

**Tabla 2**  
**Valores de pH del agua en el río Chimbo**

Nº	pH
1	8.22
2	7.86
3	7.58
4	8.18
5	8.14
6	8.75
7	8.31
8	7.34
9	8.12
10	8.17
11	7.99
12	7.37
13	6.95
14	7.98
15	7.82

Fuente: elaboración propia

### Correlación simple de las variables objeto de estudio

En la Figura 1 se observa que  $r_{y,od} = 0.45$   $r_{y,Cloruro} = -0.25$ ;  $r_{ySulfato} = 0.55$   
 $r_{y,aceites} = -0.23$   $r_{y.alcalinidad} = -0.57$   $r_{y,Caudal} = 0.29$ , por lo que se podría decir que la variable que se encuentra más relacionada con el pH del agua para la posible construcción del nuevo modelo es la alcalinidad pues presenta el mayor valor de correlación en valor absoluto. La alcalinidad es la cantidad de carbonatos de calcio (CaCO<sub>3</sub>) que hay en el agua, también se puede definir como la cantidad de ácido necesaria para disminuir el pH de una muestra.

	PH	od	cloruro	sulfato	aceites	alcalinidad	caudal
PH	1.0000000	0.4550172	-0.25372200	0.5538856	-0.23339691	-0.57159316	0.29679762
od	0.4550172	1.0000000	-0.47810473	0.5104306	-0.68192795	-0.55173974	0.27995360
cloruro	-0.2537220	-0.4781047	1.00000000	-0.5000115	0.07081722	0.81020768	-0.05587916
sulfato	0.5538856	0.5104306	-0.50001146	1.0000000	-0.27887675	-0.44309496	0.70363887
aceites	-0.2333969	-0.6819279	0.07081722	-0.2788768	1.00000000	0.11783796	-0.43867412
alcalinidad	-0.5715932	-0.5517397	0.81020768	-0.4430950	0.11783796	1.00000000	0.01099304
caudal	0.2967976	0.2799536	-0.05587916	0.7036389	-0.43867412	0.01099304	1.00000000

Figura 1 Determinación de correlaciones individuales

Se puede además visualizar que existe una alta relación lineal positiva entre alcalinidad y cloruro por lo que las aguas que mostraron menor alcalinidad también presentaron menos cloruro y las que presentaron mayor alcalinidad también tenía mayores cantidades de cloruro.

Para aplicar los modelos de regresión y determinar los factores que influyen en la medición del pH de la calidad del agua, se utilizó el software de distribución libre R (R Code Team, 2018)

### Modelo de regresión Lineal Simple

En la Figura 2 se puede visualizar el comportamiento gráfico del modelo de regresión lineal simple construido entre las variables pH y alcalinidad, de lo que se puede decir que en las muestras de agua del río Chimbo se encontró mayor alcalinidad para valores menores de pH.

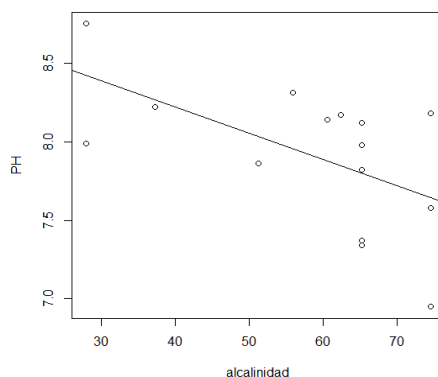


Figura 2 Modelo de regresión simple entre alcalinidad y pH

Se construyeron modelos de regresión lineal simple entre pH como variable de salida y las demás variables, pero con la única variable de entrada que mejoró el modelo es con la alcalinidad dado que la correlación entre estas dos variables es la más representativa.

**Tabla 3**  
**Resultados de ajuste del Modelo de Regresión Lineal**

Modelo	Ecuación	R <sup>2</sup>	Valor p
Regresión	pH=8.88 – 0.016*alcalinidad mg/l	0.326	0.026

Fuente: elaboración propia

Con respecto al análisis de este modelo se puede decir que la media de  $\beta_0 = 8.88$  del pH del agua cuando la alcalinidad tiene valor de 0. Por otro lado, el valor de  $\beta_1 = - .,016$  quiere decir que por cada mg de CaCO<sub>3</sub> por litro de agua el pH va a disminuir en 0.016. En cuanto al valor p se comprueba que es menor a 0.05, por tanto, el aumento de alcalinidad influye inversamente en el pH del agua.

En la figura 3 se puede visualizar la normalidad de los residuos del modelo, para lo cual se utilizó la prueba de *Shapiro-wilk*.

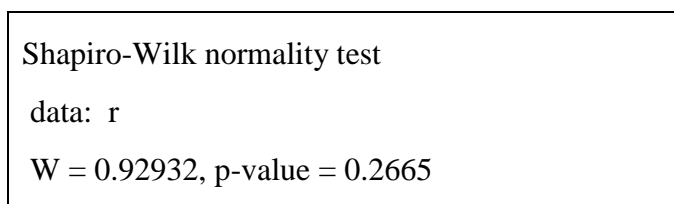


Figura 3 Prueba de Shapiro Wilk para verificar normalidad de residuos

El valor p valor de 0.26 mayor a 0.05, da indicio de que residuos provienen de una distribución normal. Además, los supuestos de no correlación de residuos y de homocedasticidad son contrastados con las pruebas de Durbin-Watson y Breusch-Pagan, obteniéndose valores p que indican que se cumplen con estos supuestos.

### Modelo de regresión Múltiple

En la Figura 4 se puede apreciar que los factores que influyen en la determinación del pH son, la alcalinidad, el sulfato, seguidos por oxígeno disuelto y el caudal, de los cuales la variable del sulfato denota una clara tendencia de una distribución normal en sus datos, mientras que las otras variables denotan sesgos.



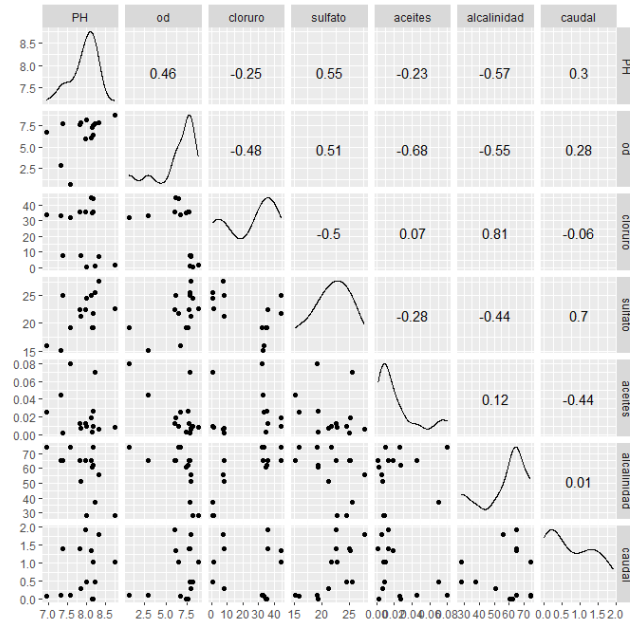


Figura 4 Gráficos de densidad de todas las variables objeto de estudio y valores de correlación entre las variables

En la tabla 4, se presentan los mejores modelos de regresión múltiple obtenidos con las variables analizadas. Se puede apreciar que todos los modelos cumplen con los supuestos de los residuos, sin embargo, el modelo con el  $R^2$  más alto es el modelo 2.

Tabla 4: Resultados de los Modelo de Regresión Múltiple

Modelo	Modelo 1	Modelo 2	Modelo 3
Variables analizadas	Oxígeno disuelto, sulfato, caudal, cloruro, alcalinidad, aceites	sulfato, cloruro, alcalinidad	cloruro, alcalinidad
Ho: Varianza constante vs H1: Varianza No constante	Valor p: 0.42 Homocedasticidad	0.24 Homocedasticidad	0.80 Homocedasticidad
$R^2$	0.3993	0.5605	0.3634
Durbin-Watson test	0.22 No existe correlación	0.31 No existe correlación	0.83 No existe correlación
Shapiro-Wilk normality test residuos	0.65 Normalidad en Residuos	0.41 Normalidad en Residuos	0.31 Normalidad en Residuos

Fuente: elaboración propia

### Multicolinealidad de las variables de entrada del modelo de regresión Múltiple

Se obtuvo un  $R^2$  de las variables del modelo 2 de 0.5605, aplicando la fórmula de factores de inflación de la varianza como se detalla:

$$FIV_j = \frac{1}{1 - 0.5605} = 2.27$$

Lo que indica que no existe multicolinealidad en las variables que influyen en la determinación del pH del agua, porque el valor es menor a 10 (Aparicio et al., 2004); además el índice de inflación entre variables explicativas del modelo es menor a 10. Por lo que se considera como mejor modelo el 2:

**Tabla 5: Resultados de ajuste del Modelo de Regresión Múltiple**

Modelo	Ecuación	R <sup>2</sup>	P valor
Regresión lineal múltiple	pH=7.65 +0.066*sulfato (mg/l)+0.022*cloruro (mg/l) – 0.029*alcalinidad (mg/l)	0.56	0.0068

Fuente: elaboración propia

Para el diseño del modelo de regresión múltiple se tomaron como variables influyentes en la determinación de la calidad del agua al sulfato, cloruro y alcalinidad que dieron como origen un R<sup>2</sup> de 0.56. A partir del diseño de este modelo se puede interpretar que el valor  $\beta_0 = 7.65$  q es el promedio de pH ante la ausencia de sulfatos, cloruro y alcalinidad. Por otro lado el valor de  $\beta_1 = 0.066$  quiere decir que por cada mg/l de sulfato el pH va aumentar en 0.066, así mismo el  $\beta_2 = 0.022$  da indicio que por cada mg/l de cloruro el pH va aumentar 0.022 y el  $\beta_3 = -0.029$  implica que cada mg/l de alcalinidad el pH va a disminuir 0.029. Con respecto a su p valor del estadístico de F se comprueba que es menor a 0.05 se rechaza Ho, por lo tanto, las variables de entrada influyen en la determinación del pH del agua. A continuación, se presentan los valores de pH predichos por el modelo presentado en la Tabla 5.

**Tabla 6: Valores predichos de pH**

pH observado	pH predicho
8.22	8.27
7.86	7.73
7.58	7.45
8.18	7.88
8.14	7.92
8.75	8.37
8.31	8.01
7.34	7.48
8.12	8.39
8.17	7.88
7.99	8.47
7.37	7.57
6.95	7.26
7.98	8.02
7.82	8.02

Fuente: elaboración propia

## Conclusiones

Se determinó la relación de las variables de entrada y salida a partir de los modelos de regresión lineal simple y modelo de regresión múltiple con los indicadores de medición del pH del agua a partir de factores como alcalinidad, sulfatos, aceites, cloruros, oxígeno disuelto.

Se comprobó en los modelos realizados con respecto al pH del agua, que mientras más se vaya mejorando su  $R^2$ , esas variables de entrada serán los factores más influyentes en la determinación del pH del agua.

Las variables de entrada que más influyen en la determinación del pH del agua en los dos modelos de regresión (lineal – múltiple) es la alcalinidad. Sin embargo, el comportamiento conjunto de alcalinidad, sulfato y cloruro afecta significativamente al valor de pH del agua del Río Chimbo. Adicionalmente, cabe mencionar que los modelos construidos cumplen con los supuestos de Regresión Lineal, por lo que todas las inferencias de este estudio obtenidas mediante esta metodología son confiables.

Se presentan además estimaciones del pH del agua del Río Chimbo a través del mejor modelo descrito por la relación  $\text{pH} = 7.65 + 0.066 * \text{sulfato (mg/l)} + 0.022 * \text{cloruro (mg/l)} - 0.029 * \text{alcalinidad (mg/l)}$ . Estas estimaciones son muy cercanas a los valores reales de pH del río Chimbo de la muestra estudiada. A través de los coeficientes del modelo se puede apreciar que hay una relación positiva entre las variables sulfato, cloruro y pH, mientras que la relación entre alcalinidad y pH sería negativa.

Se recomienda establecer mecanismos de control en las instituciones de regulación y cuidado del agua, para evitar cambios en las variables antes mencionadas (alcalinidad, sulfato y cloruro) para evitar variaciones significativas en el pH del agua, puesto que esto podría ocasionar el aumento en la acidez o alcalinidad del mismo dejando a la región sin esa fuente de consumo del líquido vital.

Se recomienda tener mucho cuidado al beber el agua del Río Chimbo sin un proceso de purificación, puesto que, según los resultados obtenidos el nivel de pH bordea el nivel aceptable del líquido vital siendo este un valor cercano a 7, pero el índice del valor del pH se ve muy influenciado por los valores que tomen sus variables de salida, alcalinidad, sulfato y cloruro.

## Bibliografía

- Ansari, F., & Salahuddin. (2013). Statistical Analysis for the Presence of pH Content of Ground Water at Different Locations of Industrial area at Ghazipur in India. *Global Journal of Science Frontier Research Mathematics and Decision Sciences*, 13(9).
- Aparicio, J., Martínez, M. M., & Morales, J. (2004). *Modelos aplicados en R*. Universidad Miguel Hernández.
- APHA, AWWA & WPCF (1992). *Métodos normalizados para el Análisis de Aguas Potables y Residuales*. España: Díaz de Santos.

- Chang, R. (2010). *Química*. México DF: McGraw-Hill.
- Chang, R. G. (2013). *Química* (11a ed. edición. México; Madrid: MacGraw-Hill. ISBN 978-607-15-0928-4.
- Del Valle Moreno, J., & Guerra Bustillo, W. (2012). La Multicolinealidad en modelos de Regresión Lineal Múltiple. *Revista Ciencias Técnicas Agropecuarias*, 80-83.
- Kiely, G. (2000). *Ingeniería Ambiental: Fundamentos, entornos, tecnologías y sistemas de gestión*. España, Madrid.
- Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied Linear Regression Models*. McGraw-Hill Irwin.
- Lanza, E. G., Hernández, P. S., & Carbajal, P. J. (2000). Comisión Nacional del Agua de México, & Universidad Nacional Autónoma de México. Organismos indicadores de la calidad del agua y de la contaminación: Bioindicadores.
- Li, J., Liu, H., & Chen, J. (2018). Microplastics in freshwater systems: A review on occurrence, environmental effects, and methods for microplastics detection. *Water Research*, 362-374.
- Mackenzie, D., & Cornwell, D. (2012). *Introduction to Environmental Engineering (The McGraw-hill Series in Civil and Environmental Engineering)* 5th Edition.
- Montgomery, D., Peck, E., & Vining, G. (2006). *Introducción al análisis de regresión lineal*. México: Ed. Limussa.
- Morris, G. L., & Fan, J. (1998). *En Reservoir sedimentation handbook: Design and management of dams, reservoirs, and watershed for sustainable use*. New York: McGraw-Hil.
- R Code Team. (8 de septiembre de 2018). R: Un Lenguaje y Entorno para la Estadística Informática. En F. R. Estadística. Viena, Austria. Obtenido de <http://www.rproject.org>
- Rossen, A., Calvo, D., Rodríguez, M., Bustamante, A., Korol, S., & Angelaccio, C. (2009). Evaluación de la Calidad del agua mediante modelación lineal de los indicadores de contaminación fecal en el embalse San Roque. XXII Congreso Nacional del Agua.
- Salim, A., Gowhar, M., Sayar, Y., & Ashok, K. (2014). Statistical Assessment of Water Quality Parameters for Pollution Source Identification in Sukhnag Stream: An Inflow Stream of Lake Wular (Ramsar Site). *Journal of Ecosystems*. doi:<https://doi.org/10.1155/2014/898054>.
- Stanley, E. H., & Doyle, M. W. (2002). A geomorphic perspective. *BioScience*, 52(8), 693-701.
- Stow, C. A., Borsuk, M. E., & Stanley, D. W. (2001). Long-term changes in watershed nutrient inputs and riverine exports in the Nesuse River. *Water Research*, 35(6), 1489-1499.
- Varea, A., Suárez, L., Chávez, G., Cordero, M., Álvarez, N., & Espinosa, F. (1997). Biodiversidad, bioprospección y bioseguridad. Quito, Ecuador.
- Viswanath, N. C., Dileep Kumar, P. G., Ammad, K. K., & Usha Kumari, E. R. (2015). Statistical Analysis of Quality of Water in Various Water Shed for Kozhikode City, Kerala, India. *Aquatic Procedia*, 1078-1085.

- Zhang, Y., Xia, J., Shao, Q., & Zhai, X. (2013). Water quantity and quality simulation by improved SWAT in highly regulated Huai River Basin of China. *Stochastic Environmental Research & Risk Assessment*, 27(1).
- Zlatanović, L., van der Hoek, J. P., & Vreeburg, J. H. (2017). An experimental study on the influence of water stagnation and temperature change on water quality in a full-scale domestic drinking water system. *Water Research*, 761-772.
- Zubala, T. (2009). Influence of dam reservoir on the water. *Ecohydrology & Doyle*, 9(2-4), 165-173.