

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE INGENIERÍA CIVIL Y AMBIENTAL

**EVALUACIÓN DE LA RELACIÓN EXISTENTE ENTRE LA
PRESENCIA DE MACROINVERTEBRADOS ACUÁTICOS Y LA
FISICOQUÍMICA DEL AGUA EN RÍOS DE MONTAÑA.**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN MAGÍSTER EN
CIENCIAS DE LA INGENIERÍA PARA LA GESTIÓN DEL RECURSO HÍDRICO**

VELASTEGUI CARRILLO VIVIANA ANDREA

velastegui18@gmail.com

Director: VALENCIA BONILLA NATHALIA TERESA

nathalia.valencia@epn.edu.ec

Codirector: PHD BONILLA BOLAÑOS ANDREA GABRIELA

andrea.bonilla@epn.edu.ec

Quito, febrero 2020

CERTIFICACIÓN

Certificamos que el presente trabajo: “EVALUACIÓN DE LA RELACIÓN EXISTENTE ENTRE LA PRESENCIA DE MACROINVERTEBRADOS ACUÁTICOS Y LA FISICOQUÍMICA DEL AGUA EN RÍOS DE MONTAÑA” fue desarrollado por la Ing. Viviana Velasteguí, estudiante de la Maestría en Ciencias de la Ingeniería para la Gestión del Recurso Hídrico, habiendo supervisado la realización de este trabajo y realizado las correcciones correspondientes, damos por aprobada la redacción final del documento escrito para que prosiga con los trámites correspondientes a la sustentación de la Defensa oral.

Ing. Valencia Bonilla Nathalia MSc.
Director

Ing. Bonilla Bolaños Andrea PhD
Codirector

DECLARACIÓN DE AUTORÍA

Yo, Velasteguí Carrillo Viviana Andrea, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

La Escuela Politécnica Nacional puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

VELASTEGUÍ CARRILLO VIVIANA

DEDICATORIA

Thou art my God, and I will praise thee;

Thou art my God, I will exalt thee.

Psalm 118:28

AGRADECIMIENTO

En primer lugar, deseo agradecer a Dios, ya que su amor y su bondad me acompañaron en este proceso.

A toda todita toda mi familia, el pilar de mi vida y el motivo para seguir creciendo como persona.

A Diego, que me recordó cada día que los planes que Dios tiene para mi son mejores y más grandes que los míos.

A Amy y Elvi, esta experiencia no hubiese sido la misma sin ustedes. A mis viejos amigos y nuevos amigos.

A la Ingeniera Nathalia Valencia, que nuevamente tengo el honor de recibir su guía; por ser una amiga y ejemplo a seguir. A la Dra. Andrea Bonilla, ya que gracias a su experiencia y amplio conocimiento pude culminar con éxito este trabajo. A la Dra Carlita Manciatí, por ser la mejor Coordinadora de Maestría que he conocido y al Dr. Marcos Villacís, porque en su momento, nos permitió cursar esta maestría; gracias a él todo esto comenzó.

ÍNDICE DE CONTENIDO

APROBACIÓN DEL DIRECTOR.....	i
DECLARACIÓN DE AUTORÍA.....	ii
DEDICATORIA.....	iii
AGRADECIMIENTO.....	iv
LISTA DE FIGURAS	vii
LISTA DE TABLAS	x
LISTA DE ANEXOS	xiii
RESUMEN	xiv
ABSTRACT	xv
CAPÍTULO 1	
INTRODUCCIÓN	1
1.1. OBJETIVO GENERAL.....	3
1.2. OBJETIVOS ESPECÍFICOS.....	3
1.3. JUSTIFICACIÓN.....	3
CAPÍTULO 2	
MARCO TEÓRICO Y ESTADO DEL ARTE.....	5
2.1. MARCO TEÓRICO.....	5
2.1.1. Macroinvertebrados Acuáticos.....	5
2.1.2. Parámetros fisicoquímicos e hidráulicos influyentes en la composición de comunidades de macroinvertebrados acuáticos.....	8
2.2. ANÁLISIS DE REGRESIÓN LOGÍSTICA.....	9
2.2.1. Medidas de bondad de ajuste.....	10
2.2.2. Criterios de selección de familias.....	14
2.3. ESTADO DEL ARTE.....	15
CAPÍTULO 3	
METODOLOGÍA.....	17
3.1. ZONA DE ESTUDIO.....	17

3.2. RECOLECCIÓN DE DATOS.....	17
3.3. TOMA DE MUESTRAS Y ANÁLISIS DE LABORATORIO.....	20
3.4. ESPECIES POR MODELAR.....	22
3.5. TRATAMIENTO DE DATOS.....	23
3.6. MODELO DE REGRESIÓN LOGÍSTICA.....	23

CAPÍTULO 4

RESULTADOS Y DISCUSIÓN.....	25
4.1. DISTRIBUCIÓN DE LA COMUNIDAD DE MACROINVERTEBRADOS.....	25
4.2. ANÁLISIS DE CORRELACIÓN DE SPEARMAN.....	29
4.3. MODELO DE REGRESIÓN LOGÍSTICA.....	30
4.3.1. Familia Hydrobiosidae (Sensibilidad a contaminantes: 9).....	30
4.3.2. Familia Simuliidae (Sensibilidad a contaminantes: 8).....	39
4.3.3. Familia Baetidae (Sensibilidad a contaminantes: 7).....	45
4.3.4. Familia Hyalellidae (Sensibilidad a contaminantes: 7).....	48
4.3.5. Familia Glossosomatidae (Sensibilidad a contaminantes: 7).....	53
4.3.6. Familia Planariidae (Sensibilidad a contaminantes: 7).....	56
4.3.7. Familial Hydroptilidae (Sensibilidad a contaminantes: 7).....	61
4.4. RANGOS DE PREFERENCIA DE PARÁMETROS DE CALIDAD.....	66

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES.....	72
CONCLUSIONES.....	72
RECOMENDACIONES.....	72
REFERENCIAS BIBLIOGRÁFICAS.....	76
ANEXOS.....	87

LISTA DE FIGURAS

Figura 1 –	Curvas ROC para dos clasificadores (C1 y C2) en la misma base de datos	14
Figura 2 –	Ubicación geográfica de la zona de estudio Río Jatunhuayco.....	19
Figura 3 –	Ubicación geográfica de la zona de estudio Río Ushimana	20
Figura 4 –	Cantidad total de individuos recolectados (en paréntesis) en las zonas de muestreo	26
Figura 5 –	Número de familias en los puntos de muestreo en el río Jatunhuayco (A, B, C, D) y río Ushimana (E, F, G, H, I, J, K).....	26
Figura 6 –	Número de individuos de la familia Hydrobiosidae por época climática.....	32
Figura 7 –	(a) Probabilidad de encontrar a la familia Hydrobiosidae en relación con nitratos. (b) Familia Hydrobiosidae con sólidos suspendidos. (c) Diagrama de caja, presencia (1) y ausencia (0) de macroinvertebrados de la familia Hydrobiosidae en relación con nitratos. (d) Diagrama de caja Familia Hydrobiosidae con sólidos suspendidos	33
Figura 8 –	(a) Probabilidad de encontrar a la familia Hydrobiosidae en relación con la profundidad. (b) Familia Hydrobiosidae con Ancho del cauce. (c) Diagrama de caja, presencia (1) y ausencia (0) de macroinvertebrados de la familia Hydrobiosidae en relación con la profundidad. (d) Diagrama de caja Familia Hydrobiosidae el ancho del cauce.....	36
Figura 9 –	Curva ROC del modelo estimado con datos del conjunto de prueba.....	39
Figura 10 –	Curva ROC del modelo estimado con datos del conjunto de prueba con parámetros hidráulicos	40
Figura 11 –	Número de individuos de la familia Simuliidae por época climática.....	41
Figura 12 –	(a) Probabilidad de encontrar a la familia Simuliidae en relación con la conductividad. (b) Familia Simuliidae con (NH ₄) ⁺ . (c) Familia	42

	Simuliidae con sólidos totales. (d) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia Simuliidae de acuerdo con la conductividad. (e) Diagrama de caja familia Simuliidae con $(\text{NH}_4)^+$. (f) Diagrama de caja familia Simuliidae con sólidos totales	
Figura 13 –	Curva ROC del modelo estimado con datos del conjunto de prueba.....	46
Figura 14 –	Número de individuos de la familia Baetidae por época climática...	46
Figura 15 –	Curva ROC del modelo estimado con datos del conjunto de prueba.....	49
Figura 16 –	Número de individuos de la familia Hyalellidae por zona de estudio.....	50
Figura 17 –	(a) Probabilidad de encontrar a la familia Hyalellidae en relación con la temperatura. (b) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia Hyalellidae de acuerdo con la Temperatura	51
Figura 18 –	Curva ROC del modelo estimado con datos del conjunto de prueba.....	53
Figura 19 –	Número de individuos de la familia Glossosomatidae por época climática.....	54
Figura 20 –	(a) Probabilidad de encontrar a la familia Glossosomatidae en relación con Coliformes fecales. (b) Familia Glossosomatidae con el pH. (c) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia Glossosomatidae de acuerdo con Coliformes fecales. (d) Diagrama de caja familia Glossosomatidae con el pH	55
Figura 21 –	Curva ROC del modelo estimado con datos del conjunto de prueba.	57
Figura 22 –	Número de individuos de la familia Planariidae por época climática..	58
Figura 23 –	(a) Probabilidad de encontrar a la familia Planariidae en relación con la temperatura. (b) Familia Planariidae con el pH. (c) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de	59

	la familia Planariidae de acuerdo con la temperatura. (d) Diagrama de caja familia Planariidae con el pH	
Figura 24 –	Curva ROC del modelo estimado con datos del conjunto de prueba.	62
Figura 25 –	Número de individuos de la familia Hydroptilidae por época climática.....	62
Figura 26 –	(a) Probabilidad de encontrar a la familia Hydroptilidae en relación con el pH. (b) Familia Hydroptilidae con sólidos totales. (c) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia. (d) Diagrama de caja familia Hydroptilidae con concentración de sólidos totales.....	64
Figura 27 –	Curva ROC del modelo estimado con datos del conjunto de prueba.	66

LISTA DE TABLAS

Tabla 1 –	Características de familias de macroinvertebrados acuáticos	7
Tabla 2 –	Matriz de confusión de un clasificador	12
Tabla 3 -	Escala de interpretación de valore de AUC	14
Tabla 4 -	Resumen de variables fisicoquímicas, hidráulicas, microbiológicas y biológicas recolectadas en los 11 sitios de muestreo del río Jatunhuayco y río Ushimana	21
Tabla 5 -	Parámetros fisicoquímicos de calidad de agua, equipo y método de análisis	22
Tabla 6 -	Valorización de sensibilidad a contaminación en función del índice BMWP-Col	23
Tabla 7 -	Escala de interpretación del coeficiente de Spearman	24
Tabla 8 -	Número de individuos y porcentaje de abundancia (en paréntesis) de las familias en la época seca (I) y época lluviosa (II) en los diferentes puntos de muestreo del río Jatunhuayco	28
Tabla 9 -	Número de individuos y porcentaje de abundancia (en paréntesis) de las familias en la época seca (I) y época lluviosa (II) en los diferentes puntos de muestreo del río Ushimana	29
Tabla 10 -	Coefficiente de correlación de Spearman entre las variables explicativas con su respectivo p-valor	29
Tabla 11 -	Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros fisicoquímicos de calidad del agua correlacionados con la presencia de Hydrobiosidae	30
Tabla 12 -	Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros hidráulicos correlacionados con la presencia de Hydrobiosidae	31
Tabla 13 -	Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba de los parámetros físicoquímicos	32
Tabla 14 -	Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba de parámetros hidráulicos	35

Tabla 15 -	Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de Simuliidae	38
Tabla 16 -	Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba	38
Tabla 17 -	Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de Baetidae	41
Tabla 18 -	Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba	45
Tabla 19 -	Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de Hyalellidae	47
Tabla 20 -	Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba	48
Tabla 21 -	Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de Glossosomatidae	50
Tabla 22 -	Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba	53
Tabla 23 -	Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de Planariidae	54
Tabla 24 -	Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba	57
Tabla 25 -	Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de Hydroptilidae	58
Tabla 26 -	Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba	61
Tabla 27 -	Variación del valor de AIC de acuerdo con el porcentaje de presencia de macroinvertebrados en los puntos de muestreo	63

Tabla 28 –	Rango de temperatura donde aumenta la probabilidad de ocurrencia de macroinvertebrados	66
Tabla 29 –	Rango de valor pH donde aumenta la probabilidad de ocurrencia de macroinvertebrados	67
Tabla 30 –	Rango de valor conductividad donde aumenta la probabilidad de ocurrencia de macroinvertebrados	68
Tabla 31 –	Rango de concentración de sólidos totales donde aumenta la probabilidad de ocurrencia de macroinvertebrados	68
Tabla 32 –	Rango de concentración de sólidos suspendidos donde aumenta la probabilidad de ocurrencia de macroinvertebrados	69
Tabla 33 –	Rango de valores donde aumenta la probabilidad de ocurrencia de macroinvertebrados	69
Tabla 34 –	Rango de valores donde aumenta la probabilidad de ocurrencia de macroinvertebrados	70
Tabla 35 –	Rango de valores donde aumenta la probabilidad de ocurrencia de macroinvertebrados	70
Tabla 36 –	Rango de valores de coliformes fecales donde aumenta la probabilidad de ocurrencia de macroinvertebrados	71
Tabla 37 –	Rango de valores donde aumenta la probabilidad de ocurrencia de macroinvertebrados.....	71
Tabla 38 –	Rango de valores de coliformes fecales donde aumenta la probabilidad de ocurrencia de macroinvertebrados.....	71

LISTA DE ANEXOS

Anexo 1.	Composición y distribución de macroinvertebrados en el río Jatunhuayco	85
Anexo 2.	Composición y distribución de macroinvertebrados en el río Ushimana	87
Anexo 3.	Coeficiente de correlación de Spearman de las variables explicativas..	89
Anexo 4.	Modelo de regresión logística para las familias de macroinvertebrados	91
Anexo 5.	Modelo desarrollado para la familia Hydrobiosidae (Script).....	93

RESUMEN

El uso de poblaciones biológicas, como una herramienta de diagnóstico, es considerado en la actualidad como una opción viable para el monitoreo de la calidad del agua debido a sus bajos costos y facilidad de aplicación. Sin embargo, al ser típicamente especializadas y altamente subjetivas, no se ha podido establecer una metodología que pueda extrapolar a todas las zonas de estudio, como consecuencia de la elevada heterogeneidad que presenta los distintos hábitats. Por ende, es preciso encontrar una metodología que pueda relacionar concepciones importantes como: los parámetros fisicoquímicos del agua y su influencia en la distribución y composición de macroinvertebrados en sistemas lóticos.

En esta investigación, se expondrá una propuesta con miras a aportar nueva información a estos métodos de evaluación de calidad del agua. Empleando un análisis de regresión logística, se evalúa la relación existente entre la presencia de macroinvertebrados acuáticos y los parámetros hidráulicos y fisicoquímicos de calidad del agua, en dos ríos del Ecuador.

Palabras clave: Macroinvertebrados acuáticos, parámetros fisicoquímicos e hidráulicos, análisis de regresión logística.

ABSTRACT

The usage (application) of communities of microorganisms, among other live specimens, is getting a higher consideration in the environmental field, regards to water quality monitoring of systems. However, this is considered a complex study, which analyses a wide variety of species in different habitats, resulting in many subjective considerations of data analysis. A general methodology ought to be proposed, in which the main correlation of physicochemical parameters of lotic systems are related to each kind of microorganisms. This research consists about a correlation of the main water quality parameters and the presence of macroinvertebrates, in order to establish the influence of each of those parameters to the most important living habitant of rivers. In order to sort all the data, a statistical logistic regression analysis is presented to evaluate the relationship between the presence of aquatic macroinvertebrates and the hydraulic and physicochemical parameters of water quality in two rivers of Ecuador

Keywords: Aquatic macroinvertebrates, physicochemical and hydraulic parameters, logistic regression analysis.

CAPÍTULO 1

INTRODUCCIÓN

En América Latina, la investigación sobre el agua ha tomado mayor importancia en las dos últimas décadas debido a la preocupación por la velocidad con la cual se agravan los problemas de escasez de agua, en cantidad y calidad (López-Vera, 2002). De acuerdo con la Organización de las Naciones Unidas, entre los años 2000 y 2017, se produjo un aumento del 61% al 71%, en la proporción de la población mundial que utiliza agua potable gestionada, siendo los países de América Latina y el Caribe los que evidenciaron mayor progreso. (Naciones Unidas, 2019). Esto implica mayor explotación del recurso y aumento en las descargas de agua contaminada en ríos.

La descarga de agua residual de actividades antrópicas degradan la calidad original de los cuerpos de agua, ocasionando alteraciones en los ecosistemas acuáticos, pérdida de hábitats y reducción en la biodiversidad (Díaz-Martínez y Granada-Torres, 2018; Lepom et al., 2009). Dada la complicada situación de disponibilidad de recursos hídricos de buena calidad para abastecimiento, varios investigadores como Mussetta (2013), instan a los actores principales, trabajar en la gestión del recurso hídrico y obtener información sistemática y profunda, para un manejo adecuado del recurso y monitoreo continuo de su calidad.

La recopilación de la información, base para el manejo global del recurso hídrico en el tema de su calidad, se la realiza mediante métodos de detección de contaminantes. De esta manera se puede conocer el nivel de degradación existente en un sitio (Mite et al., 2016). El método de detección de contaminación más empleado es el método analítico, mismo que consiste en aplicar un conjunto de acciones sobre una muestra de agua para determinar la cantidad de un contaminante (analito), lo cual requiere contar con un laboratorio bien equipado y recursos para la adquisición de los reactivos necesarios (Vilanova y Sogorb, 2004).

Si bien, este método tiene gran acogida en el campo de la investigación para medir los niveles de contaminantes químicos, presenta una serie de desventajas: a) el costo del análisis de una amplia gama de contaminantes químicos por cada muestra extraída; b) dificultades físicas cuando grandes volúmenes de agua deben ser recolectados y

extraídos para cuantificar y evaluar y c) la concentración del contaminante refleja únicamente la situación en el momento de recolección de la muestra (Dworak, Gonzalez, Laaser, y Interwies, 2005; Grbovic y Džeroski, 1995; Madrid y Zayas, 2007).

Al evidenciar las limitaciones de un monitoreo fisicoquímico, se exhibe la necesidad de concebir metodologías de detección que puedan brindar un enfoque complementario al monitoreo de la calidad de agua (Dworak et al., 2005). No obstante, las herramientas de monitoreo serán útiles solo si son asequibles, confiables y producen datos que son de calidad comparable entre tiempos, caso contrario se dificultaría su ejecución e incluso podría llegar a ser imposible, lo cual también implica una desventaja en la aplicación de nuevas alternativas (Dworak et al., 2005; Džeroski, Demšar, y Grbović, 2000).

El uso de muestras biológicas (e.j. identificación de macroinvertebrados) se ha contemplado para evidenciar cambios en la calidad de cuerpos de agua ya que pueden reflejar un aumento en la contaminación e indicar las posibles causas o fuentes de contaminación química en un período de tiempo (Grbovic, 1994). Estas muestras no pueden considerarse como métodos de detección de contaminantes, pero la información que proporcionan puede ser de mucha utilidad (Dworak et al., 2005). Por ejemplo, promover a un monitoreo químico selectivo y por ende una reducción en los costos de investigación (Dworak et al., 2005; Grbovic, 1994).

Kolkwitz y Marsson (1902) son los primeros en proponer el uso de la biota como una medida de monitoreo de calidad de agua, y a partir de aquello, se han creado una variedad de índices o categorías de calidad de agua que fundamentan en la identificación de muestras biológicas (Metcalf, 1989). Sin embargo, ninguno es aplicable al 100% en todas las zonas de estudio (Walley, Boyd, y Hawkes, 1992). Es así que estos métodos son típicamente especializados y altamente subjetivos (Walley et al., 1992). Además, el rol ecológico y la importancia en la calidad del agua de los individuos que actúan como indicador es desconocido para muchos y puede asimismo diferir de un país a otro (Grbovic, 1994); lo que lo hace, desde un punto de vista ecológico y de calidad del agua, un tema importante de investigación (Grbovic y Džeroski, 1995).

Esta escasez de conocimiento, referente a la composición y estructura de los macroinvertebrados acuáticos y su relación con los parámetros ambientales de calidad de agua (Arroyo y Encalada, 2009), sumada a la poca accesibilidad de información de contaminantes en campo, no permiten aplicar un monitoreo integral en un cuerpo de agua.

En tal sentido, la presente investigación tiene como objetivo generar nuevos conocimientos; que permitan desarrollar nuevas herramientas para la gestión integral de los recursos hídricos. Para lo cual (1) se analizó la composición y distribución de los macroinvertebrados acuáticos en los ríos Jatunhuayco y Ushimana. Asimismo (2) se construyó, mediante un análisis de regresión logística, modelos de preferencia de macroinvertebrados. Con la finalidad de identificar aquellos parámetros ambientales de calidad de agua considerablemente relacionados. Y por último (3) se estableció los posibles rangos, de parámetros ambientales, en los cuales existe mayor probabilidad de encontrar a los macroinvertebrados estudiados.

1.1. OBJETIVO GENERAL

Estimar la relación existente entre la presencia de macroinvertebrados acuáticos y los parámetros hidráulicos y fisicoquímicos de calidad del agua, mediante un análisis estadístico.

1.2. OBJETIVOS ESPECÍFICOS

- a) Determinar la composición y distribución de macroinvertebrados en el río Jatunhuayco y río Ushimana.
- b) Realizar un análisis de regresión logística con los datos de macroinvertebrados acuáticos encontrados y valores fisicoquímicos e hidráulicos del cuerpo de agua.
- c) Definir, a partir del análisis de regresión logística, rangos de preferencia de calidad del agua para cada familia de macroinvertebrados acuáticos.

1.3. JUSTIFICACIÓN

Actualmente, existe un gran grupo de índices biológicos de calidad de agua. Entre los más utilizados se encuentran: índice biótico extendido (**IBE**), que ha sido acondicionado a varios países europeos (Ghetti, 1986); British Biological Monitoring Working Party score system (**BMWP**) (Armitage, Moss, Wright, y Furse, 1983) el cual ha sido el más estudiado y readaptado a diferentes sistemas fluviales de la Península Ibérica (Alba-Tercedor, 1996; Alba-Tercedor et al., 2002; Alba-Tercedor y Sánchez-Ortega, 1988; Muñoz y Prat, 1996) y en Latinoamérica por Roldán (1999). Los índices de Integridad Biótica (**IBI**), se diseñan para una ecorregión determinada (De Pauw, Gabriels, y Goethals, 2006) y el índice Efemerópteros, Plecópteros y Tricópteros (**EPT**), es un índice más específico que emplea

macroinvertebrados intolerantes a la contaminación (Bournaud, Cellot, Richoux, y Berrahou, 1996; Prat, Ríos, Acosta, y Rieradevall, 2009). Estos índices, brindan una idea global del nivel de contaminación en un cuerpo de agua, sin embargo, a partir de ellos no se puede cuantificar los parámetros físicos y químicos de calidad. Por ende, es preciso encontrar una metodología que pueda relacionar estas concepciones importantes: los parámetros fisicoquímicos del agua y su influencia en los macroinvertebrados (muestras biológicas).

Poco se conoce acerca de la influencia directa de los parámetros fisicoquímicos sobre los macroinvertebrados (Grbovic y Džeroski, 1995), no obstante, en nuestro país, se han desarrollado trabajos enmarcados en esta línea de investigación. En el trabajo de Jerves-Cobo et al (2017), mediante un modelo de regresión logística, se determinó los parámetros fisicoquímicos más influyentes en la presencia de macroinvertebrados pertenecientes a las familias Efemerópteros, Plecópteros o Tricópteros. Este estudio fue desarrollado con información proveniente del Río Machángara, ubicado en la ciudad de Cuenca, Provincia de Azuay Ecuador (Jerves-Cobo et al., 2017a).

Everaert et al. (2014) evaluó, en tres cuencas hidrográficas ubicadas en Ecuador, Etiopía y Vietnam, la preferencia de cinco familias de macroinvertebrados acuáticos a parámetros fisicoquímicos del agua. Mediante un modelo de regresión logístico, se concluyó que las relaciones entre macroinvertebrados y parámetros fisicoquímicos del agua pueden ser específicas para cada cuenca hidrográfica y, por lo tanto, no son transferibles a otras cuencas de los ríos en los trópicos. Esto acarrea a un replanteamiento y verificación de las metodologías ya existentes y proponer nuevas técnicas de detección de contaminantes.

Con estos antecedentes y con el fin de contribuir con este campo del conocimiento, la información aportada por la presente investigación constituirá una herramienta para la gestión de recursos hídricos. Que permita, en primer lugar, abaratar costos al realizar un monitoreo continuo de calidad de agua y a la par, impulse nuevas investigaciones con base en el estudio de macroinvertebrados acuáticos y la relación que existe con los diferentes parámetros ambientales, en ríos de montaña. Los resultados obtenidos aportarán al conocimiento referente a macroinvertebrados acuáticos como indicadores de calidad de agua.

CAPÍTULO 2

MARCO TEÓRICO Y ESTADO DEL ARTE

2.1. MARCO TEÓRICO

2.1.1. MACROINVERTEBRADOS ACUÁTICOS

Macroinvertebrados, en términos generales, se puede definir como aquellos organismos invertebrados que pueden ser observados a simple vista. Estos organismos tienen un tamaño superior a 0,5 mm de largo (Hanson, Springer, y Ramirez, 2010; Roldán y Restrepo, 2008).


Las comunidades de macroinvertebrados presentes en sistemas acuáticos poseen una gran variedad de adaptaciones e importantes diferencias en sus ciclos de vida (Hanson et al., 2010). Existe un gran número de especies de macroinvertebrados acuáticos, así como diversidad en sus formas de vida. Son los principales recicladores de nutrientes en los ecosistemas acuáticos, e intermediarios en el proceso de descomposición de la materia orgánica de origen vegetal (Pérez, 1999).

Al presente, los macroinvertebrados acuáticos son considerados como bioindicadores de la calidad del agua (Ortiz, 2005), debido a su sensibilidad ante el estrés ambiental. Esto implica que ciertas especies de macroinvertebrados requieren buena calidad de agua para sobrevivir, mientras que otras pueden sobrevivir en aguas de baja calidad. (Johnson, Wiederholm, y Rosenberg, 1993)

Características de macroinvertebrados acuáticos (familias)

A continuación, se presenta un resumen de las familias de macroinvertebrados empleadas para este estudio.

Tabla 1 – Características de familias de macroinvertebrados acuáticos.




FAMILIAS	CARATERÍSTICAS	IMAGEN
Hydrobiosidae	<p>Esta familia es más diversa en Australia y en regiones Neotropicales. Mide 22 mm de largo y las primeras patas de los adultos se encuentran muy modificadas. No construyen casas e hilan seda sobre el sustrato para desplazarse. Sus larvas se localizan en ríos de corrientes frías, oxigenadas, de sustrato pedregoso, poco material vegetal y gran cantidad de materia orgánica. Son indicadores aguas oligotróficas (de la Lanza-Espino y Pulido, 2000; Roldán, 1988a).</p>	
Simuliidae	<p>El cuerpo de estos individuos es relativamente pequeño de 0.6 a 2.2 mm de longitud. Presenta manchas dorsales oscuras (Coscarón y C. L. C. Arias, 2007). Habida en ríos con alto nivel de oxígeno, bajo de rocas y troncos (Roldán, 1988a).</p>	
Baetidae	<p>Son insectos de pequeño a mediano tamaño. Las ninfas presentan cuerpos modificados para nadar. Poseen agallas abdominales. Son de color amarillo parduzco. Habitan en aguas rápidas, debajo de troncos, rocas, hojas y se encuentran adheridos a vegetación sumergida. Son Indicadores de aguas de buena calidad. (Domínguez, 2006; Flowers y De la Rosa, 2010; Heckman, 2012).</p>	

Fuente: López J., Soria I. y Carvajal V (2015)

Fuente: López J., Soria I. y Carvajal V (2015)

Fuente: López J., Soria I. y Carvajal V (2015)

Tabla 1 –. Características de familias de macroinvertebrados acuáticos (continuación).


FAMILIAS	CARATERÍSTICAS	IMAGEN
Hyaellidae	Miden entre 5,5 mm y 10,5 mm. Son de coloración blanquecina o amarillenta. Presenta ojos y no tienen palpo mandibular. Viven en aguas corrientes y remansos de quebradas. Su presencia se asocia a la materia orgánica en descomposición, donde se forman densas poblaciones(De Los Ríos, Morrone, y Rivera, 2012; Gonzalez y Watling, 2002; Roldán, 1988a)	
Glossosomatidae	Tiene un tamaño aproximado de 4 mm. Construye capullos con forma de caparazón de tortuga para formar su casa, las cuales se encuentran adheridas a las rocas. Raspan diatomeas y detritus fino de la superficie de las piedras en las corrientes de agua. Habitan en aguas corrientes frías y cálidas, que se encuentren bien oxigenadas(Bueno-Soria y Flint Jr, 1978; Roldán, 1988a).	
Planariidae	Presentan un cuerpo aplanado. Miden entre 2,6 mm y 3,7 mm. Exhiben colores grises, pardos, amarillentos o blancos. Poseen una cabeza destacadamente triangular, con dos manchas oculares. Se las ha encontrado en aguas poco profundas, bien oxigenadas y debajo de piedras, ramas y hojas. Algunas especies pueden resistir cierto grado de contaminación(González y Roldán, 2019; Roldán, 2003a).	

Fuente: López J., Soria I. y Carvajal V (2015)

Fuente: López J., Soria I. y Carvajal V (2015)

Fuente: López J., Soria I. y Carvajal V (2015)

Tabla 1 –. Características de familias de macroinvertebrados acuáticos (continuación).

FAMILIAS	CARATERÍSTICAS	IMAGEN
Hydroptilidae	Son muy pequeños. Miden 3.4 a 4 mm. Presentan agallas al final del abdomen. Viven principalmente en aguas correntosas, sin embargo, pueden sobrevivir en aguas lénticas. Al principio son libres, después tejen unas bolsas aplanadas, cubiertas de partículas finas de arena o material vegetal (Bueno-Soria, 2004; Roldán, 1988a).	

Fuente: López J., Soria I. y Carvajal V (2015)

ELABORADO POR: Velastegui V., 2019

2.1.2. PARÁMETROS FÍSICOQUÍMICOS E HIDRÁULICOS INFLUYENTES EN LA COMPOSICIÓN DE COMUNIDADES DE MACROINVERTEBRADOS ACUÁTICOS.

Las variables ambientales, características ecológicas y perturbaciones físicas en el medio son determinantes en la composición y distribución de la comunidad de macroinvertebrados (Tokeshi, 2009). Las relaciones existentes entre estos factores y la comunidad de macroinvertebrados a menudo han sido materia de investigaciones (Boulaaba, Zrelli, Boumaiza, y Rossaro, 2014; Boumaiza y Thomas, 1994; Depiereux, Feytmans, y Micha, 1983; Ormerod, 1987; Ormerod y Edwards, 1987; Wright, 1995), llevándose a cabo estudios para evaluar la idoneidad de su uso como indicador de buena calidad del agua (De Shon, 1995; Resh, 1995).

Entre las variables más destacadas, se puede considerar a la velocidad de corriente (Hamada y McCreadie, 1999; McCreadie, Adler, y Colbo, 1995) y temperatura (Coscarón y C. L. C. Arias, 2007) como los parámetros del agua más influyentes en la comunidad de macroinvertebrados acuáticos (Lindegaard, 1995). Sin embargo, es pertinente mencionar otros parámetros también importantes como la conductividad (McCreadie y Colbo, 1991),

pH (McCreadie et al., 1995; Ormerod y Edwards, 1987), oxígeno disuelto (McCreadie y Colbo, 1991), DBO, DQO, nutrientes (Johnson et al., 1993) y gradiente altitudinal (Jacobsen y Encalada, 1998a),

El hábitat y la estructura de la vegetación riparia pueden influir en la composición de una comunidad de macroinvertebrados acuáticos. En zonas altas, la vegetación de ribera es menos diversa en comparación con la vegetación de las zonas bajas. Esto ocasiona una producción de material de soporte y alimento idóneo para el asentamiento de comunidades de macroinvertebrados en las zonas bajas (C. P. Villamarín, 2008).

2.2. ANÁLISIS DE REGRESIÓN LOGÍSTICA

El análisis de regresión logística se enmarca en un conjunto de técnicas estadísticas de análisis de datos (Velasco, 1996). Es una herramienta muy útil si se pretende relacionar una variable dependiente cualitativa con una o más variables independientes (Guillén, 1992; Jerves-Cobo et al., 2017a; Velasco, 1996). El análisis de regresión logística tiene como objetivos: a) determinar la existencia o ausencia de relación entre una o más variables independiente (X_i) y una variable dependiente categórica binaria (Y_i); b) examinar el signo de la relación, si es que existe dicha dependencia; y c) predecir la probabilidad de que se produzca el evento definido como $Y_i = 1$ (donde Y_i toma el valor de 1 o de 0 en relación a la presencia o ausencia de una determinada característica), de acuerdo con los valores adoptados por las variables independientes (Velasco, 1996). Para efectos del presente estudio, se denota P_i a la probabilidad de que el macroinvertebrado se encuentre presente en la zona de estudio ($Y_i = 1$), y $1 - P_i$ la probabilidad de que este no se encuentre presente ($Y_i = 0$). Se analiza, en consecuencia, la probabilidad (P_i) en función de las variables explicativas.

Puesto que P_i tiene siempre valores entre 0 y 1, es necesario que el modelo adopte una forma funcional que restrinja los valores predichos a su rango natural ya antes mencionados, así, la regresión logística adopta la forma funcional:

$$P_i = \frac{e^{(\alpha + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \beta_3 \cdot X_{3i} + \dots + \beta_n \cdot X_{ni})}}{1 + e^{(\alpha + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \beta_3 \cdot X_{3i} + \dots + \beta_n \cdot X_{ni})}} \quad (1)$$

Se expresa de igual manera:

$$\frac{P_i}{1 - P_i} = e^{(\alpha + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \beta_3 \cdot X_{3i} \dots \dots + \beta_n \cdot X_{ni})} \quad (2)$$

Las ecuaciones (1) y (2) representan una función logística, en donde X_{ni} representan las variables explicativas, α representa el término independiente o constante; β_i son coeficientes o parámetros de regresión. La función logística es exponencial para la estimación y esta puede convertirse en una función lineal, aplicando el logaritmo natural (Ln) a cada término de esta. (Zuur, Ieno, Walker, Saveliev, y Smith, 2009a)

$$\text{Ln} \left[\frac{P_i}{1 - P_i} \right] = \alpha + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \beta_3 \cdot X_{3i} \dots \dots + \beta_n \cdot X_{ni} \quad (3)$$

Si se conoce que $P_i = P_i(Y = 1)$

$$\text{Ln} \left[\frac{P_i(Y = 1)}{1 - P_i(Y = 1)} \right] = \text{Ln} \left[\frac{P_i(Y = 1)}{P_i(Y = 0)} \right] \quad (4)$$

El cociente $[P_i(Y = 1)/P_i(Y = 0)]$ corresponde a la razón de probabilidad de ocurrencia de un evento (odds). Esta medida, reportada por el modelo, es aplicable para valorar la magnitud de influencia que tienen las variables aleatorias con respecto a un evento posible (Jerves-Cobo et al., 2017a; Zuur et al., 2009a).

2.2.1. MEDIDAS DE BONDAD DE AJUSTE

Criterio de información de Akaike (AIC)

El criterio de información de Akaike es una medida de la calidad relativa de un modelo estadístico, también es conocido como el "Error de predicción Forward". Se deriva del logaritmo de la función de verosimilitud. Como tal, el AIC proporciona una medida para la selección del modelo, mientras menor es el valor de AIC, mejor calidad tiene el modelo estadístico (Canay Pazos, 2008).

El valor de AIC puede calcularse a partir de la siguiente expresión:

$$AIC = -\ln L + \text{parámetros} \quad (5)$$

Donde L es la verosimilitud de un modelo a partir de número de parámetros (García y Pina, 1997)

Matriz de confusión

La matriz de confusión contiene el número de casos clasificados correctamente y el número de casos clasificados de forma errónea. En la diagonal, se encuentran los casos correctamente clasificados. Esto implica que los datos pronosticados y los datos reales son los mismos. Los datos fuera de la diagonal representan los casos clasificados de forma errónea. Mediante la matriz de confusión se puede evaluar la capacidad predictiva de un modelo (Malhotra, 2004).

Tabla 2 –. Matriz de confusión de un clasificador.

	Clasificado positivo	Clasificado negativo
Positivo real	VP	FN
Negativo real	FP	VN

FUENTE: (Liu, 2007a)

ELABORADO POR: Velastegui V., 2019

Donde:

VP: número de clasificaciones correctas de los ejemplos positivos (verdadero positivo)

FN: número de clasificaciones incorrectas de los ejemplos positivos (falso negativo)

FP: número de clasificaciones incorrectas de los ejemplos negativos (falso positivo)

VN: número de clasificaciones correctas de los ejemplos negativos (verdadero negativo)

Con base a la matriz de confusión (Tabla 2), (Liu, 2007) define a la precisión (p) y la sensibilidad (r) de la siguiente manera: La precisión (p) es el número de ejemplos positivos correctamente clasificados dividido por el número total de ejemplos que se clasifican como positivos tal como se precia en la ecuación (6).

$$p = \frac{VP}{VP + FP} \times 100 \quad (6)$$

La sensibilidad (r) es el número de ejemplos positivos clasificados correctamente dividido por el número total de ejemplos positivos reales en el conjunto de datos de prueba (Liu, 2007a).

$$r = \frac{VP}{VP + FN} \times 100 \quad (7)$$

El F-score es el promedio armónico entre la precisión y la sensibilidad. Este puede tomar valores de 0 a 1, donde "1" representa la precisión y sensibilidad perfecta y "0" la peor (Hilbe, 2009; Liu, 2007a).

$$F = \frac{2pr}{p + r} \quad (8)$$

Curvas ROC

Una curva ROC (por sus siglas en inglés: receiver operating characteristic) es un gráfico que representa la tasa de positivos verdaderos versus la tasa de falsos positivos. Consiste en la clasificación de los datos de prueba según las probabilidades que estos casos tengan para pertenecer a la clase positiva, y el caso más probable se identifica en la parte superior de la diagonal presente en el gráfico (Figura 1).

Este gráfico es usado comúnmente para evaluar los resultados de clasificación positiva de un modelo binario (Liu, 2007a). En el ámbito ecológico es considerado como un indicador efectivo del desempeño del modelo, independientemente de la probabilidad de umbral en la que se acepta la presencia del organismo objetivo (Manel, Williams, y Ormerod, 2001).

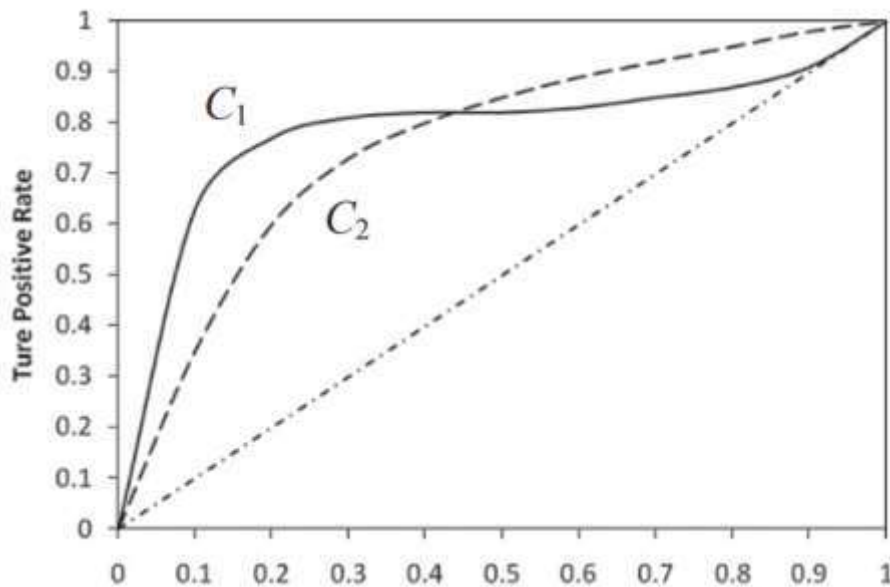


Figura 1 -. Curvas ROC para dos clasificadores (C_1 y C_2) en la misma base de datos.
(Liu, 2007, pág. 84)

Sin embargo, no se puede comparar la idoneidad entre dos clasificadores, de una forma satisfactoria únicamente con el gráfico. Para realizar esta comparación, los investigadores usan el área bajo la curva ROC (AUC). El AUC oscila en valor del 0 al 1. Un modelo cuyas predicciones son 100% incorrectas tiene un AUC de 0.0; otro cuyas predicciones son 100% correctas tiene un AUC de 1.0. En la Tabla 3 se puede apreciar el rango de valores que puede tomar el área bajo la curva ROC y la interpretación relacionada con cada uno de dichos valores (Gironés, 2017).

Tabla 3 -. Escala de interpretación de valore de AUC

RANGO		RELACIÓN
0.5	- 0.6	Test malo
0.65	- 0.75	Test regular
0.76	- 0.9	Test bueno
0.91	- 0.97	Test muy bueno
0.98	- 1	Test excelente

FUENTE: Gironés, (2017)
ELABORADO POR: Velastegui V., 2019

Pseudo R²

En el análisis de regresión logística no se cuenta con un valor que nos indique la variabilidad explicada como en los modelos lineales. No obstante, existe un valor llamado Pseudo R² de McFadden, que nos puede dar una idea de la calidad de ajuste del modelo (McFadden, 1973; Wooldridge, 2006)

$$R_{McFadden}^2 = 1 - \frac{L_{nr}}{L_0} \quad (9)$$

Donde:

L_{nr} : es la función de log verosimilitud para el modelo estimado.

L_0 : es la función de probabilidad log en el modelo con una sola variable.

Si las variables no exhiben poder explicativo, el valor del Pseudo R² será igual a cero, similar a la interpretación que se le da al R² en un modelo de regresión lineal.

2.2.2. CRITERIOS DE SELECCIÓN DE FAMILIAS

Evento de interés por variable (EPV)

La presencia de errores al realizar un análisis de regresión logística multivariada es inevitable (Peduzzi, Concato, Kemper, Holford, y Feinstein, 1996). A menudo, estos errores se producen en el ajuste de datos al desarrollar los modelos. Cierta tipo de error puede surgir al utilizar demasiadas variables de predicción con tamaño de muestra insuficiente. (Harrell, Lee, Matchar, y Reichert, 1985).

Asimismo pueden producirse cuando el ajuste del modelo final se encuentra por debajo del nivel deseado, no incluyéndose así las variables relevantes en el mismo (Ortega y Cayuela, 2002). Para evitar estos tipos de errores, Harrell F. et al (1985) proponen analizar la proporción del número de observaciones con respecto al número de variables predictoras. Según el criterio sugerido por los autores, ésta proporción debe ser mínimo de 10 observaciones por cada variable predictora.

Peduzzi et al (1996), mediante una simulación tipo MonteCarlo, analizaron el efecto que podría tener, la proporción EPV, en los modelos de regresión logística. Concluyeron que,

si el valor obtenido en la proporción es inferior a 10, los coeficientes de regresión se verían claramente sesgados tanto en sentido positivo como negativo. Por esta razón, el EPV es una herramienta que permite limitar el número de variables predictoras, generando un mejor desempeño del modelo de regresión logística. (Peduzzi et al., 1996)

Sensibilidad a contaminantes

Roldán (2003), propone una escala de valores para medir la tolerancia de algunas familias de macroinvertebrados acuáticos a diferentes niveles de contaminación. Estos valores varían del 1 al 10. Aquellas familias más sensibles a la contaminación, se les asigna un valor de 10 y el valor de 1 a aquellas más tolerantes a ecosistemas contaminados. (Roldán, 2003a)

2.3. ESTADO DELARTE

La regresión logística se ha convertido en uno de los procedimientos más empleados en la investigación científica para el análisis de datos de respuesta binaria (Hilbe, 2009; McCullagh y Nelder, 2018). Considerado además, como el método pionero de los modelos lineales generalizados en esta rama de estudio (McCullagh y Nelder, 2018). Vayssières et al (2000), en su trabajo acerca de alternativas de modelos estadísticos para predicción de distribuciones de especies, señalan que métodos paramétricos, como el modelo lineal generalizado (GLM), son más eficientes en contraste con los métodos no paramétricos como los modelos aditivos generalizados (GAM), principalmente, si se cuenta con un conjuntos de datos pequeños (Vayssières, Plant, y Allen-Diaz, 2000). De igual manera, Zuur et al (2019), resaltan el empleo de este tipo de análisis estadístico con variables binarias: presencia o ausencia de especies (Zuur et al., 2009a).

A nivel mundial, se puede subrayar experiencias, en las cuales emplearon un modelo de regresión logística para establecer una relación entre las comunidades de macroinvertebrados acuáticos y parámetros de calidad del agua. Importantes resultados se aprecian en trabajos como el de Eyre et al. (2005), en donde se estudia la relación existente entre macroinvertebrados, la cobertura del suelo y variables ambientales en el río Tyne al norte de Inglaterra (Eyre, Pilkington, McBlane, y Rushton, 2005); y el de Maloney et al. (2009), quien emplea un modelo de regresión logística para predecir la condición biológica de la corriente basada en datos del Índice de integridad biótica de macroinvertebrados bentónicos en un río de Maryland, Estados Unidos (Maloney, Weller,

Russell, y Hothorn, 2009). Manel et al (1999), por su parte, proponen, una comparación entre análisis de regresión logística y redes neuronales artificiales para determinar un patrón de distribución de macroinvertebrados. La conclusión del trabajo es que la regresión logística tiene varias ventajas para estos fines (Manel, Dias, y Ormerod, 1999). Entre otros trabajos, como los realizados por Ferrier (1984) y Bigler et al (2005), modelos de regresión logística han sido utilizado con éxito para clasificar la condición ecológica de un río en base a la presencia de macroinvertebrados acuáticos (Bigler, Kulakowski, y Veblen, 2005; Ferrier, Watson, Pearce, y Drielsma, 2002). Según Roldán, (1999b), la cantidad de información acerca de la ecología y distribución de macroinvertebrados es limitada para América Latina. En cambio, en Norteamérica y Europa existe gran cantidad información detallada al respecto.

A nivel nacional existen, trabajos destacados como el de Jerves-Cobo et al. (2017) y Arreaga (2017), presentan una metodología para construir un modelo de preferencia de macroinvertebrados a parámetros de calidad del agua, empleando un análisis de regresión logística (Jerves-Cobo et al., 2017). Además, Holguin-Gonzalez et al (2013) integran modelos hidráulicos, fisicoquímicos y ecológicos para evaluar la efectividad de las estrategias de gestión de la calidad del agua para el río Cuenca(Holguin-Gonzalez et al., 2013). En fin, Rosero (2011) genera modelos de hábitat como herramienta para la gestión del flujo de agua en ríos de montaña.

Cabe mencionar que, a nivel nacional, el número de publicaciones alusivas a este tema es limitado. La relación macroinvertebrados y parámetros de calidad de agua, es un punto de partida para la construcción de nuevas herramientas con miras a realzar la importancia de la biodiversidad acuática como bioindicador.

CAPÍTULO 3

METODOLOGÍA

A continuación, se detalla el proceso de levantamiento de la información en la zona de estudio. Asimismo, se definen los criterios con los cuales fueron ejecutadas las actividades de campo y de gabinete. Finalmente se describe el desarrollo de modelos de preferencia, para cada uno de los macroinvertebrados acuáticos encontrados como una función de las variables hidráulicas y fisicoquímicos de calidad de agua, con su respectiva validación.

3.1. ZONA DE ESTUDIO

La zona de estudio comprende dos ríos de ecosistema montañoso. El río Jatunhuayco se encuentra en la Provincia de Napo, en la Reserva Ecológica Antisana. Es parte de la Unidad Hidrográfica Jatunhuayco (UHJ). Esta cuenca drena al sistema de abastecimiento o suministro de agua potable para Quito, por lo que se encuentra bajo la intervención de la Empresa Pública Metropolitana de Agua y Saneamiento (EPMAPS). Se ubica a 50 km al sureste de la ciudad de Quito. Esta zona está declarada como Zona de Protección Absoluta (intangible) por la importancia de los procesos hidrológicos y la vulnerabilidad de esta área. (Aguirre y Torres, 2013)

Por otro lado, el río Ushimana se localiza en la provincia de Pichincha, al sur este de la capital. Tiene una longitud aproximada de 12,02 Km (Cabezas, 2015). El río atraviesa un área con asentamientos poblacionales e industria, lo que lo hace el principal receptor de aguas residuales de la zona.

3.2. RECOLECCIÓN DE DATOS

Los datos empleados en esta investigación fueron recolectados y medidos en dos épocas climáticas diferentes. Durante la época seca, con presencia precaria de precipitación, en los meses de octubre de 2015, enero y febrero de 2017. En época lluviosa, con considerable presencia de precipitación, en los meses de enero 2015 y abril 2017. Un total de 11 puntos fueron considerados para la toma de muestras. Cuatro de ellos situados en el río Jatunhuayco (Figura 2) y siete en el río Ushimana (Figura 3).

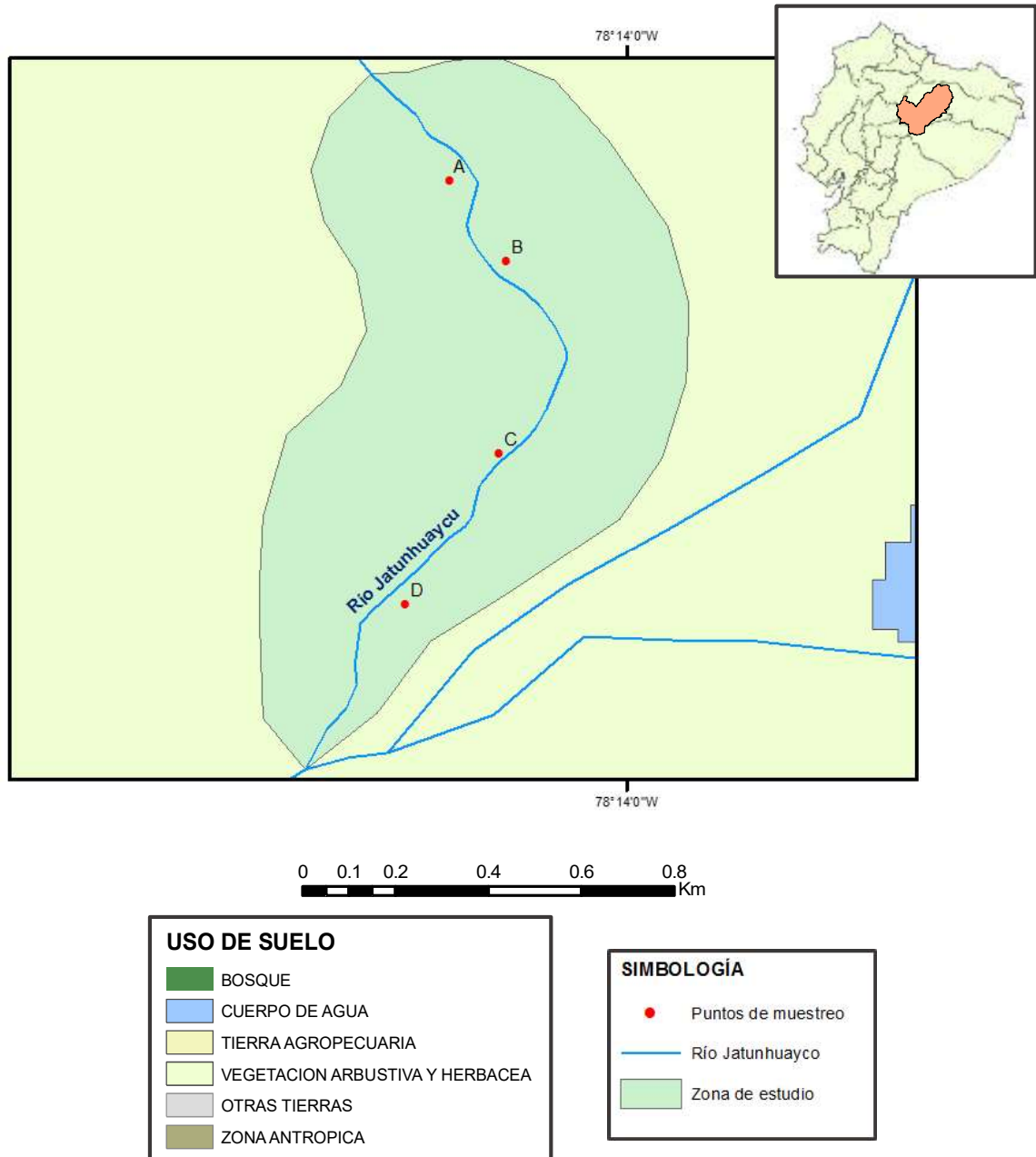


Figura 2 – Ubicación geográfica de la zona de estudio Río Jatunhuayco
ELABORADO POR: Velastegui V., 2019

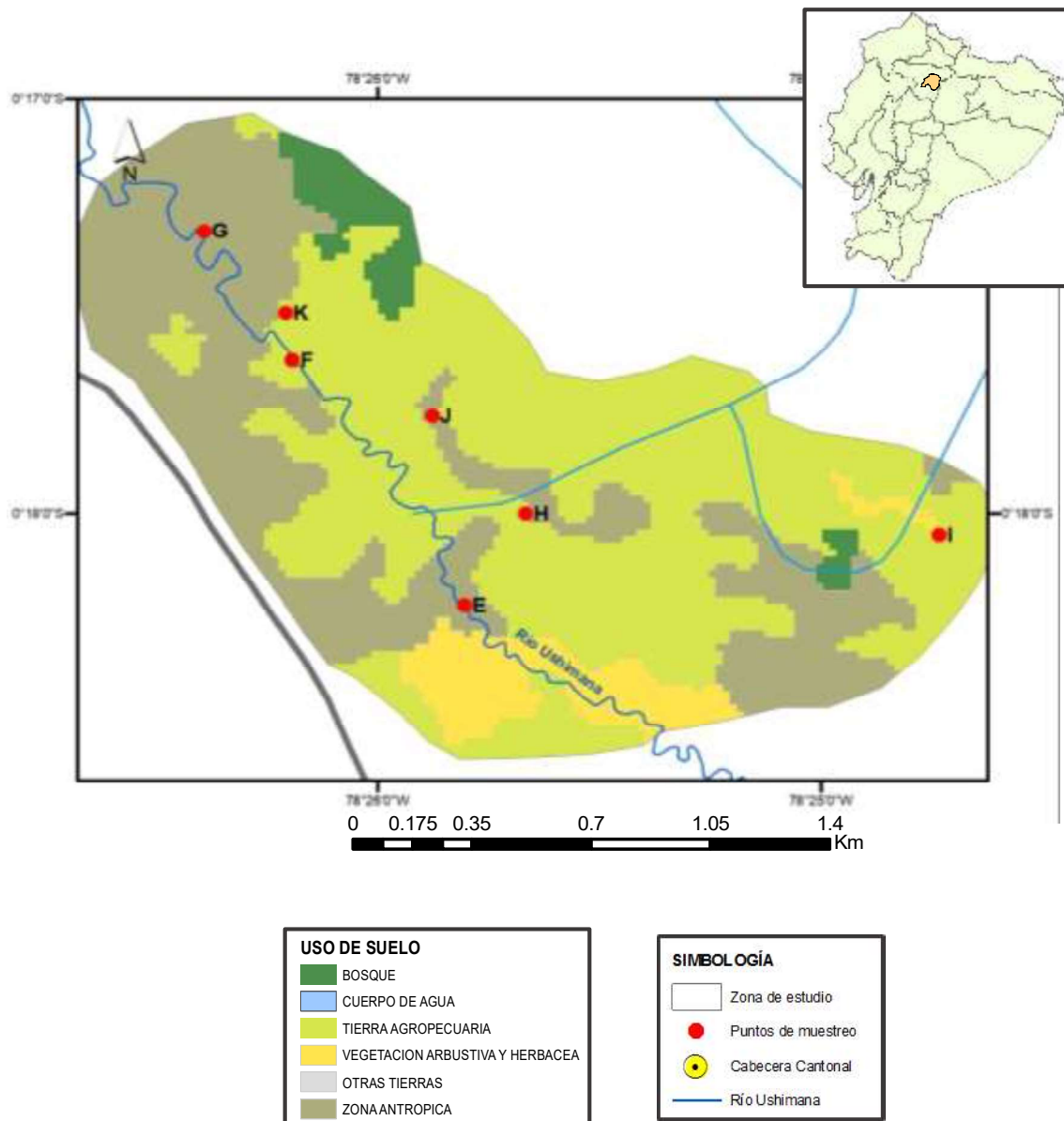


Figura 3 – Ubicación geográfica de la zona de estudio Río Ushimana.
ELABORADO POR: Velastegui V., 2019

El nivel de degradación antrópica fue el criterio de selección de los sitios de muestreo, con la finalidad de contar con un rango más amplio de valores de las variables estudiadas. En cada punto se midió un total de 21 variables, entre fisicoquímicas, hidráulicas, microbiológicas y biológicas (Tabla 4)

Tabla 4 - Resumen de variables fisicoquímicas, hidráulicas, microbiológicas y biológicas recolectadas en los 11 sitios de muestreo del río Jatunhuayco y río Ushimana.

PARÁMETRO	UNID.	MEDIANA	PROMEDIO	DESVIACIÓN ESTÁNDAR	MIN.	MAX.
Temperatura	°C	12.8	13.1	±1.7	10.8	19
Conductividad	us/cm	244.6	258.4	±74.0	177.2	535.7
TDS	mg/L	125.2	138.8	±48.6	101.3	354.9
pH	-	8.3	8.1	±0.6	6.7	9.1
OD	mg/L	9.4	9.1	±1.8	4.1	11.7
OD Sat	% sat	63.4	62.6	-	6.6	96.8
(PO ₄) ³⁻	mg/L	0.8	1.6	±2.5	0.4	14.3
(NO ₃) ⁻	mg/L	0.3	0.4	±0.4	0.1	2.2
(NH ₄) ⁺	mg/L	0.1	0.5	±1.4	0.1	11
Coliformes F	NMP/100mL	72	17988	±57390.7	36	230000
Resistividad	Kohm	4	3.9	±0.9	1.4	7.9
NaCl	ppm	118	129.7	±48.7	11.8	315.8
DBO	mg/L	4.1	6.7	±10.4	0.6	80
DQO	mg/L	50	58.9	±43.8	10	234
S. Totales	mg/L	216	233	±51.5	154.7	461.3
S. Disueltos	mg/L	189	206.9	±52.1	120	421
S. Suspendidos	mg/L	20.7	26.5	±21.8	2	110.7
Velocidad	m/s	0.4	0.8	±2.0	0	11.9
Profundidad	cm	20	20	±11.7	4	43.4
Ancho	m	2.2	2.8	±2.6	0.2	20.7
Altitud	msnm	3919	3685	-	2437	3935

ELABORADO POR: Velastegui V., 2019

3.3. TOMA DE MUESTRAS Y ANÁLISIS DE LABORATORIO.

La recolección de muestras para el análisis de parámetros fisicoquímicos, se realizó conforme a lo establecido en Métodos Normalizados para el Análisis de Aguas Potables y Residuales (Association American Public Health, 1992). Los parámetros se detallan en la Tabla 5.

Tabla 5 -. Parámetros fisicoquímicos de calidad de agua, equipo y método de análisis

PARÁMETRO	EXPRESIÓN	UNIDAD	MÉTODO	EQUIPO
Potencial de Hidrógeno	pH	-	4500 – H B	pH metro ACCUMET
Conductividad	Cond	µs/cm	-	
Resistividad	Res	Kohm	-	
Sólidos Totales Disueltos	STD	mg/L	-	
Temperatura	T	°C	-	Multiparámetro OAKTON PDC 650
Oxígeno Disuelto	OD	mg/L	-	
Porcentaje de saturación de oxígeno disuelto	OD sat	% sat	-	
NaCl		ppm	-	
Fosfatos	(PO ₄) ⁻³	mg/L	45000 P-E	Espectrofotómetro HACH DR 2700
Nitratos	NO ₃	mg/L	4500 NO ₃ -E	
Ion Amonio	(NH ₄) ⁺	mg/L	4599 NH ₃ -B	Digestor reactor digital, HACH, modelo DRB200
DBO		mg/L	(Método Winkler-CICAM)	
DQO		mg/L	5220	Incubador IF55 MEMMERT
Coliformes Fecales	NMP	NMP/100mL	9222 C	
S. Totales	ST	mg/L	2540 B	Horno Ney Vulcan 3-130
S. Suspendidos	SS	mg/L	2540 D	Cámara de calor Serie FD Classic.Line

FUENTE: (Association American Public Health, 1992)
ELABORADO POR: Velasteguí V., 2019

En la medición de velocidad, se utilizó un micromolinetete C2 OTT con contador digital OTT Z400. La Hélice N°5, código 249845 fue empleada para esta actividad. En cada punto de muestreo se identificaron microhábitats. En cada microhábitat encontrado, se realizó el muestreo de macroinvertebrados acuáticos. El proceso de recolección se basa en la remoción del sustrato del río para capturar los macroinvertebrados acuáticos en una red Surber de 0.33 m x 0.33 m de área colectora y 250 µm de apertura de malla. Bandejas de color blancas fueron empleadas para distinguir los organismos capturados. Todos los

individuos fueron preservados con alcohol etílico al 70% y transportados en fundas herméticas para su posterior identificación en el laboratorio.

El reconocimiento taxonómico, a nivel de familia, se efectuó con claves taxonómicas disponibles y se registró la abundancia numérica de cada grupo. (Domínguez y Fernández, 2009; Prat, Rieradevall, Acosta, y Villamarín, 2011; Roldán, 1988b) Además, un experto en biología fluvial aportó, a partir de su experiencia, juicio en el proceso de identificación de macroinvertebrados.

3.4. ESPECIES POR MODELAR

El criterio de selección de familias se realizó en base a la cantidad de eventos de interés por variable (EPV) y sensibilidad a la contaminación. En consecuencia, de los 44 taxones identificados en la zona de estudio, se decidió trabajar con 7 (Tabla 6). Para cada uno de ellos, se determinó el EPV y el puntaje de tolerancia a la contaminación de acuerdo con el trabajo de Roldán (2003). Fueron seleccionadas aquellas familias que presentaban un número adecuado de observaciones en función de las variables predictoras, y alta sensibilidad a la contaminación, ya que los modelos de predicción de ocurrencia de estos macroinvertebrados son más confiables (Forio et al., 2016; Roldán, 1988b). Esto ayudó a limitar el número de familias de macroinvertebrados acuáticos.

Tabla 6 -. Valorización de sensibilidad a contaminación en función del índice BMWP-Col

FAMILIA	VALORACIÓN Roldán 2003 BMWP-Col
Hydrobiosidae	9
Simuliidae	8
Baetidae	7
Hyaellidae	7
Glossosomatidae	7
Planariidae	7
Hydroptilidae	7

FUENTE: (Roldán, 1988b)
ELABORADO POR: Velasteguí V., 2019

3.5. TRATAMIENTO DE DATOS

En primera instancia, se procedió a realizar una exploración de los datos recolectados, para identificar tendencias y patrones (Zuur et al., 2009a). La distribución de familias encontradas a lo largo del gradiente altitudinal fue determinada a partir de la abundancia relativa de cada una de ellas (Campbell, 2001, p 1165).

Y, finalmente, mediante una prueba de correlación de Spearman se determinó aquellas variables explicativas que presentan una relación lineal con el resto de las variables, y posteriormente excluirlas del análisis. El tratamiento de datos se realizó en el lenguaje de programación R, con la interfaz gráfica de R-Studio (Ihaka y Gentleman, 1996).

Tabla 7 -. Escala de interpretación del coeficiente de Spearman

RANGO			RELACIÓN
0	–	0.25	Escasa o nula
0.26	–	0.50	Débil
0.51	–	0.75	Entre moderada y fuerte
0.76	–	1.00	Entre fuerte y perfecta

FUENTE: (Cabrera, 2009)

3.6. MODELO DE REGRESIÓN LOGÍSTICA

El enfoque para el desarrollo de los modelos, en esta investigación, fue el análisis de regresión logística. Este análisis se enmarca en el conjunto de modelos lineales generalizados (GLM).

En total, se construyó 7 modelos de preferencia, correspondiente a cada una de las familias seleccionados. La variable respuesta fue considerada como una variable categórica, es decir se evaluó la presencia y ausencia de los macroinvertebrados asociados a parámetros hidráulicos, fisicoquímicos y microbiológicos de calidad de agua. De acuerdo con el lenguaje de programación R el nivel 1 ($Y = 0$), correspondiente “ausencia” de macroinvertebrados. Para el nivel 2 ($Y = 1$) corresponde a la “presencia” de macroinvertebrados, que será el nivel de referencia en el modelo.

Todos los datos recopilados, durante la estación seca como durante la estación húmeda, se utilizaron para construir los modelos. El conjunto de datos se fraccionó en dos subconjuntos, de los cuales uno se empleó como subconjunto de entrenamiento y el otro subconjunto se utilizó para validar el modelo.

Para establecer los conjuntos de datos, entrenamiento y prueba del modelo, se empleó criterios de Validación Cruzada de K iteraciones o K-Fold Cross Validation, con un valor de $K = 10$ (Steyerberg et al., 2001; Zuur, Ieno, Walker, Saveliev, y Smith, 2009b). Para evaluar el rendimiento del modelo y seleccionar el más idóneo, se consideraron los criterios: criterio de información Akaike (AIC) (Agresti, 2003), curvas ROC para determinar la precisión en la predicción (Torres, 2010) y la matriz de Confusión para evaluar la capacidad predictiva o de clasificación del modelo (Liu, 2007a).

CAPÍTULO 4

RESULTADOS Y DISCUSIÓN

En este capítulo se presentan los resultados de la estimación de los modelos generados para cada una de las familias de macroinvertebrados acuáticos. Posterior a ello, se analiza las medidas de bondad de ajuste y finalmente, se interpretan los resultados obtenidos en el modelo logístico estimado.

4.1. DISTRIBUCIÓN DE LA COMUNIDAD DE MACROINVERTEBRADOS.

Fueron identificadas 44 familias en las zonas de estudio. El mayor número de familias se encontró en el río Jatunhuayco (Figura 5). En cuanto a la época climática, el mayor número de individuos fue recolectado en época seca (Figura 4), en comparación con la época lluviosa. Esto, en correspondencia con lo presentado en el trabajo de Jacobsen (2003) y Encalada (1997).

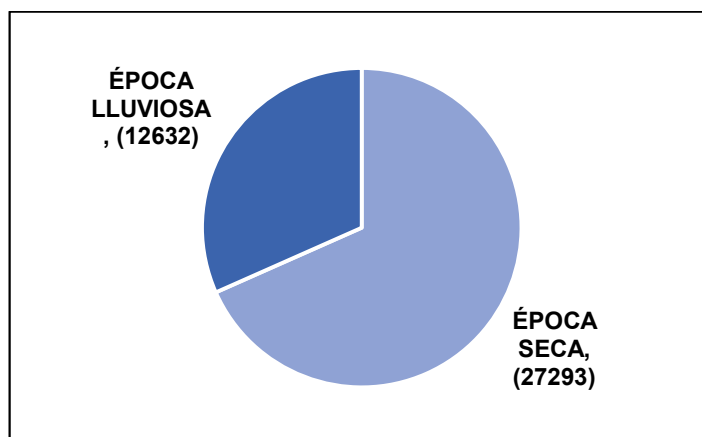


Figura 4 –. Cantidad total de individuos recolectados (en paréntesis) en las zonas de muestreo.

ELABORADO POR: Velastegui V., 2019

La diversidad de familias decrece con el gradiente altitudinal como se puede apreciar en la Figura 5. De acuerdo con trabajos realizados en ríos del Ecuador, el comportamiento de la diversidad debería ser contrario, es decir, incrementar la diversidad de familia a medida que disminuye la altitud (Encalada, 1997a; Jacobsen, Rostgaard, y Vásconez, 2003a). Sin

embargo, en este caso las zonas de estudio que se encuentran en cotas más bajas (2443 – 2470 msnm) presentan un alto grado de contaminación (Cabezas, 2015) lo que repercute en el número de familias encontradas (ver Tabla 10).

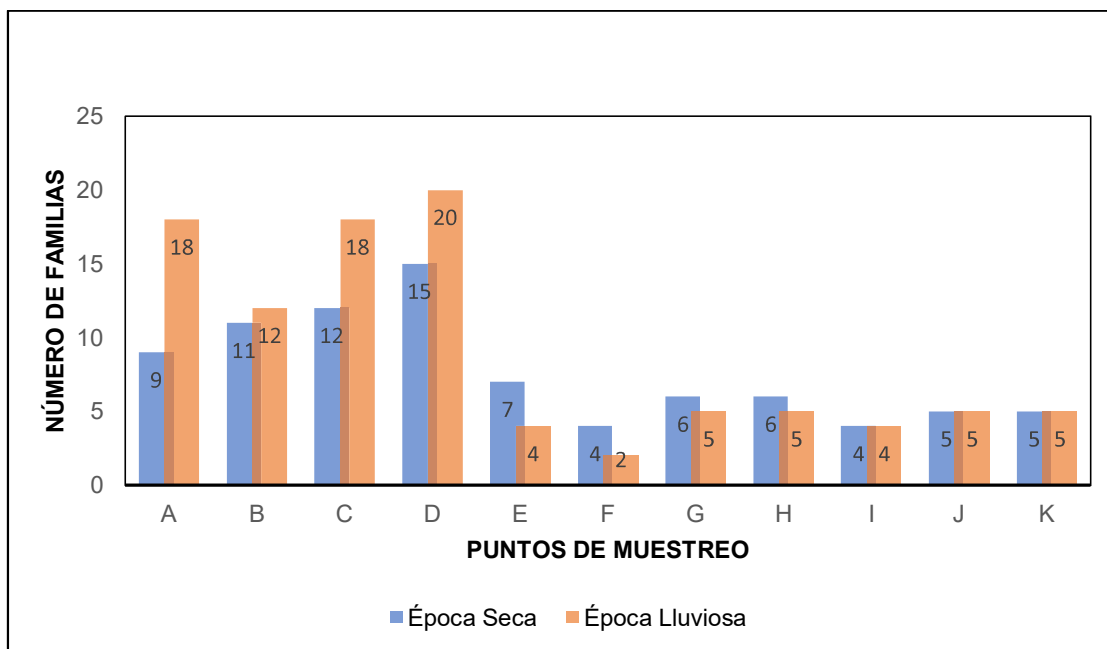


Figura 5 –. Número de familias en los puntos de muestreo en el río Jatunhuayco (A, B, C, D) y río Ushimana (E, F, G, H, I, J, K)
ELABORADO POR: Velastegui V., 2019

La familia con mayor número de individuos es la *Chironomidae* (27904). Esta familia se encontró en las dos zonas de estudio, las cuales presentan una diferencia en el grado de contaminación. La segunda familia con el mayor número es la *Hyaletellidae* (3052). Esta familia se encontró en mayor medida en el río Jatunhuayaco al igual que la familia *Elmidae*. Un patrón similar fue observado en los trabajos de Villamarín (2008), Jacobsen (2003) y Jascobsen et al. (2007). La familia *Haplotaxidae* (2225) por el contrario se encontró únicamente en la zona con menor altitud, en el río Ushimana. Aquellas familias que se encontraron en un número mínimo fueron: *Hydrachnidae*, *Phoridae*, *Limnichyidae*, *Ephydriidae*, *Hidracarina*, *Mesoveliidae*, *Gastropoda*, *Aractidae*, *Hidrophilidae* y *Gerridae*. En la Tabla 8 y Tabla 10 se puede apreciar la cantidad de individuos identificados por familias.

Tabla 8 -. Número de individuos y porcentaje de abundancia (en paréntesis) de las familias en la época seca (I) y época lluviosa (II) en los diferentes puntos de muestreo del río Jatunhuayco.

	A		B		C		D	
	I	II	I	II	I	II	I	II
Glossiphoniidae	-	10 (0.3)	-	-	-	-	-	-
Simuliidae	1 (0.1)	12 (0.4)	1 (0.1)	45 (2.2)	6 (0.4)	27 (0.9)	11 (1)	3 (0.2)
Chironomidae	476 (61)	1599 (53.1)	790 (65)	869 (41.9)	630 (37.9)	1142 (38.6)	360 (32.9)	875 (46.7)
Tipulidae	-	-	-	-	3 (0.2)	1 (0.03)	4 (0.4)	6 (0.3)
Muscidae	-	2 (0.1)	-	-	-	1 (0.03)	-	1 (0.1)
Baetidae	3 (0.4)	10 (0.3)	5 (0.4)	30 (1.4)	8 (0.5)	2 (0.1)	9 (0.8)	9 (0.5)
Empididae	4 (0.5)	13 (0.4)	1 (0.1)	6 (0.3)	10 (0.6)	54 (1.8)	10 (0.9)	47 (2.5)
Hyaellidae	247 (31.7)	927 (30.8)	312 (25.7)	903 (43.6)	187 (11.2)	119 (4)	226 (20.7)	129 (6.9)
Scirtidae	2 (0.3)	3 (0.1)	4 (0.3)	2 (0.1)	1 (0.1)	5 (0.2)	5 (0.5)	13 (0.7)
Sphaeriidae	-	6 (0.2)	-	-	-	-	1 (0.1)	3 (0.2)
Glossosomatidae	1 (0.1)	24 (0.8)	3 (0.2)	26 (1.3)	337 (20.3)	228 (7.7)	128 (11.7)	50 (2.7)
Leptoceridae	4 (0.5)	-	8 (0.7)	1 (0.05)	2 (0.1)	6 (0.2)	3 (0.3)	-
Elmidae	1 (0.1)	2 (0.1)	4 (0.3)	1 (0.05)	106 (6.4)	29 (1)	61 (5.6)	12 (0.6)
Planariidae	5 (0.6)	35 (1.2)	19 (1.6)	45 (2.2)	127 (7.6)	128 (4.3)	91 (8.3)	97 (5.2)
Limnephilidae	9 (1.2)	103 (3.4)	31 (2.5)	59 (2.8)	194 (11.7)	393 (13.3)	126 (11.5)	305 (16.3)
Turbificidae	24 (3.1)	232 (7.7)	23 (1.9)	69 (3.3)	20 (1.2)	5 (0.2)	24 (2.2)	67 (3.6)
Hirudinea	3 (0.4)	15 (0.5)	10 (0.8)	10 (0.5)	3 (0.2)	2 (0.1)	-	2 (0.1)
Hydroptilidae	-	3 (0.1)	-	-	8 (0.5)	314 (10.6)	17 (1.6)	41 (2.2)
Hydrobioscidae	-	-	-	1 (0.05)	17 (1)	59 (2)	12 (1.1)	33 (1.8)
Hydropsychidae	-	-	-	-	2 (0.1)	-	-	-
Limnichyidae	-	-	1 (0.1)	-	-	-	-	-
Ephydriidae	-	-	1 (0.1)	-	-	-	-	-
Sphaeriidae	-	-	3 (0.2)	-	-	-	-	1 (0.1)
Nematoda	-	1 (0.03)	-	-	-	2 (0.1)	5 (0.5)	2 (0.1)
Hidracarina	-	-	-	-	1 (0.1)	-	-	-
Mesoveliidae	-	-	-	-	1 (0.1)	-	-	-

Tabla 9 -. Número de individuos y porcentaje de abundancia (en paréntesis) de las familias en la época seca (I) y época lluviosa (II) en los diferentes puntos de muestreo del río Jatunhuayco (continuación).

	A		B		C		D	
	I	II	I	II	I	II	I	II
Lombriculidae	-	3 (0.1)	-	-	-	46 (1.6)	-	3 (0.2)
Ptylodactylidae	-	9 (0.3)	-	4 (0.2)	-	396 (13.4)	-	174 (9.3)
Gastropoda	-	-	-	1 (0.05)	-	-	-	-
Aractidae	-	-	-	-	-	1 (0.03)	-	-
Collembola	-	-	-	-	-	1 (0.03)	-	2 (0.1)

ELABORADO POR: Velasteguí V., 2019

Tabla 10 -. Número de individuos y porcentaje de abundancia (en paréntesis) de las familias en la época seca (I) y época lluviosa (II) en los diferentes puntos de muestreo del río Ushimana.

	E		F		G		H		I		J		K	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II
Haplotaixidae	27 (2.6)	1 (0.2)	44 (0.4)	45 (4.8)	482 (30.7)	69 (27.2)	181 (8.1)	130 (37.29)	724 (21.2)	131 (31)	99 (5.6)	131 (45.2)	151 (33.1)	10 (32.3)
Glossiphoniidae	-	-	-	-	5 (0.3)	174 (68.5)	108 (4.8)	46 (13.2)	111 (3.3)	37 (8.8)	195 (11)	141 (48.6)	38 (8.3)	4 (12.9)
Simuliidae	440 (42.6)	113 (26.8)	-	-	-	1 (0.4)	71 (3.29)	2 (0.6)	31 (0.9)	-	5 (0.3)	-	-	-
Chironomidae	541 (52.4)	104 (24.6)	11990 (99.5)	902 (95.2)	1054 (67.1)	-	1869 (83.3)	162 (46.4)	2542 (74.5)	246 (58.3)	1468 (82.5)	9 (3.1)	265 (58.1)	11 (35.5)
Tipulidae	2 (0.2)	-	-	-	-	-	-	-	1 (0.03)	-	-	-	-	-
Psychodidae	-	-	-	-	12 (0.8)	1 (0.4)	-	-	1 (0.03)	-	3 (0.2)	1 (0.3)	1 (0.2)	2 (6.5)
Syrphidae	-	-	-	-	1 (0.1)	-	-	1 (0.3)	-	-	-	2 (0.7)	-	-
Muscidae	-	-	-	-	-	2 (0.8)	2 (0.1)	-	-	-	-	-	-	-
Baetidae	19 (1.8)	14 (3.3)	-	-	-	2 (0.8)	-	7 (2)	-	6 (1.4)	-	-	-	3 (9.7)
Empididae	-	-	-	-	-	-	1 (0.05)	-	-	-	-	-	-	-
Pyralidae	-	-	-	-	1 (0.1)	-	-	-	-	-	1 (0.1)	-	-	-

Tabla 11 -. Número de individuos y porcentaje de abundancia (en paréntesis) de las familias en la época seca (I) y época lluviosa (II) en los diferentes puntos de muestreo del río Ushimana (continuación).

	E		F		G		H		I		J		K	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II
Sminthuridae	-	-	-	-	-	1 (0.4)	-	-	1 (0.03)	-	-	-	-	-
Hydrachnidae	-	-	-	-	1 (0.1)	-	-	-	-	-	-	-	-	-
Hyaellidae	-	-	-	-	2 (0.1)	-	-	-	-	-	-	-	-	-
Physidae	2 (0.2)	3 (0.7)	11 (0.1)	-	10 (0.6)	-	-	1 (0.3)	1 (0.03)	-	1 (0.1)	-	-	-
Lymnaeidae	1 (0.1)	-	-	-	1 (0.1)	-	-	-	-	-	-	-	-	-
Sphaeriidae	1 (0.1)	-	-	-	-	3 (1.2)	-	-	-	-	-	-	-	1 (3.2)
Turbificidae	-	1 (0.2)	-	-	-	-	11 (0.5)	-	-	2 (0.5)	8 (0.4)	5 (1.7)	-	-
Hirudinea	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Hydroptilidae	-	-	-	-	-	-	-	-	-	-	-	-	1 (0.2)	-
Phoridae	-	-	-	-	1(0.1)	-	-	-	-	-	-	-	-	-
Staphylinidae	-	-	2 (0.02)	-	-	-	-	-	-	-	-	-	-	-
Hidrophilidae	-	1 (0.2)	-	-	-	-	-	-	-	-	-	-	-	-
Gerridae	-	1 (0.2)	-	-	-	-	-	-	-	-	-	-	-	-
Ceratopogonidae	-	-	-	-	-	1 (0.4)	-	-	-	-	-	1 (0.3)	-	-

ELABORADO POR: Velasteguí V., 2019

4.2. ANÁLISIS DE CORRELACIÓN DE SPEARMAN

De acuerdo con el análisis de correlación, tres variables fueron omitidas. Estas presentaron un p-valor inferior a 0.05, suficiente para rechazar la hipótesis nula ya que presentaron alta correlación entre ellas. Estas variables son: concentración de NaCl, TDS (sólidos disueltos totales) y Resistividad. En la Tabla 12 se resume los valores obtenidos en el análisis de

correlación de Spearman para estos parámetros. En el Anexo 1 se encuentra la información detallada de todas las variables explicativas.

Tabla 12 -. Coeficiente de correlación de Spearman entre las variables explicativas con su respectivo p-valor.

Variables explicativas		Coefficientes	p-valor
TDS	Conductividad	0.96	0.0071
NaCl	Conductividad	0.89	0.0426
NaCl	TDS	0.87	0.0578
OD	Altitud	0.79	0.1252
Resistividad	Conductividad	-0.91	0.0297
Resistividad	TDS	-0.89	0.0417

ELABORADO POR: Velasteguí V., 2019

4.3. MODELO DE REGRESIÓN LOGÍSTICA

Para la construcción de los modelos fueron considerados además criterios empíricos provenientes de referencias bibliográficas como (Collier, Croker, Hickey, Quinn, y Smith, 1995; Encalada, 1997b; Jacobsen y Encalada, 1998b; Jacobsen, Rostgaard, y Vásconez, 2003b; Jacobsen, Schultz, y Encalada, 1997a; Merritt y Cummins, 1996; Ríos-Touma, Encalada, y Prat Fornells, 2011) . A continuación, se exponen los resultados de los modelos logísticos estimados para cada una de las familias seleccionadas.

4.3.1. FAMILIA *HYDROBIOSIDAE* (SENSIBILIDAD A CONTAMINANTES: 9)

Un total de 122 individuos pertenecientes a esta familia fueron identificados. La mayoría fue encontrada en el río Jatunhuayco en la época lluviosa, como se puede apreciar en la Figura 6.

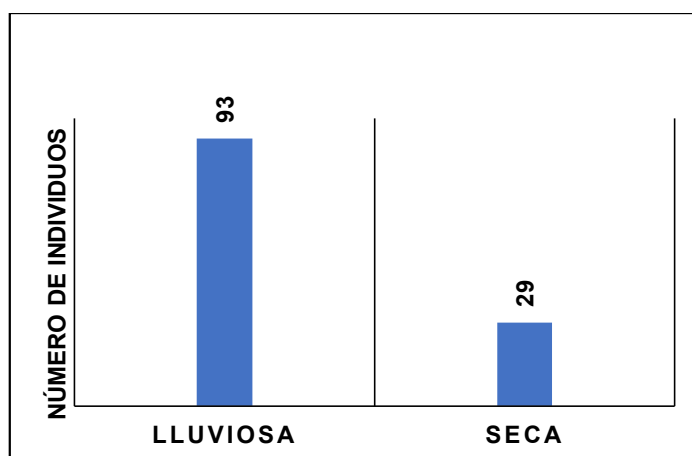


Figura 6 -. Número de individuos de la familia *Hydrobiosidae* por época climática
ELABORADO POR: Velastegui V., 2019

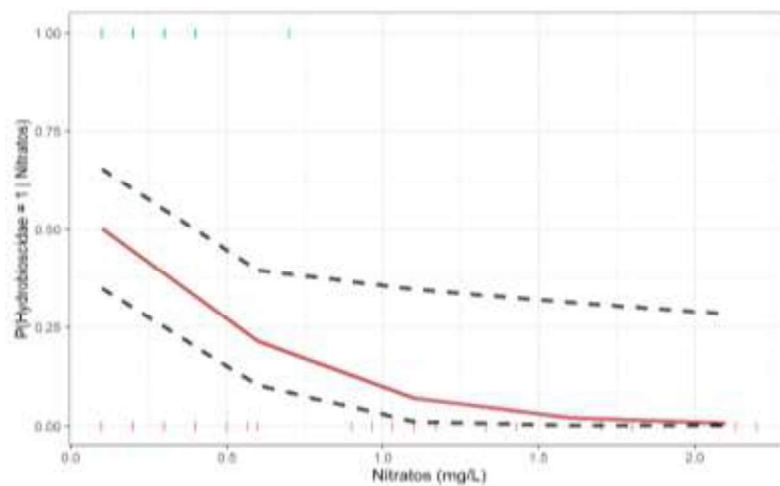
En cuanto al análisis de regresión logística los resultados obtenidos para la familia *Hydrobiosidae* se puede apreciar en la Tabla 13. El efecto de los parámetros tales como: concentración de Nitratos (p -valor < 0.05) y sólidos suspendidos (p -valor < 0.05) están asociados con la presencia de esta taxa.

Tabla 13 -. Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros fisicoquímicos de calidad del agua correlacionados con la presencia de *Hydrobiosidae*.

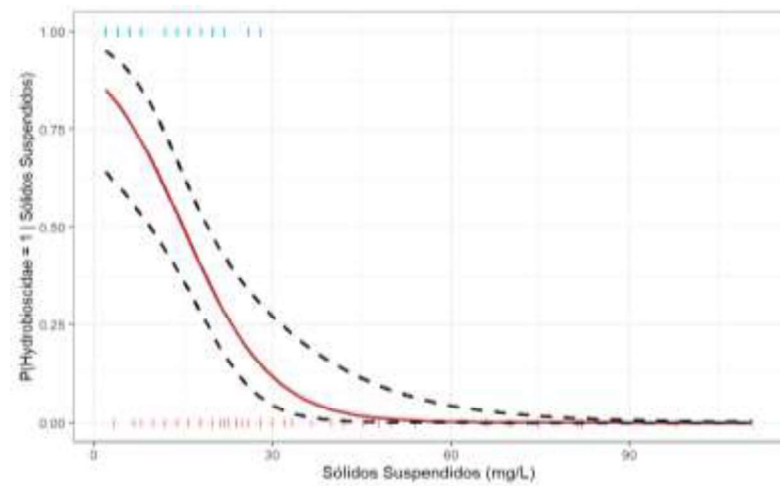
Variabes predictoras	Parámetros de regresión	Coefficientes	p-valor
	α	8.2919	0.000193
(NO ₃) ⁻	β_1	-6.8325	0.019425
ln_Sólidos S	β_3	-2.4985	6.29e-05
AIC		67.33	
Pseudo R²		0.61	

ELABORADO POR: Velastegui V., 2019

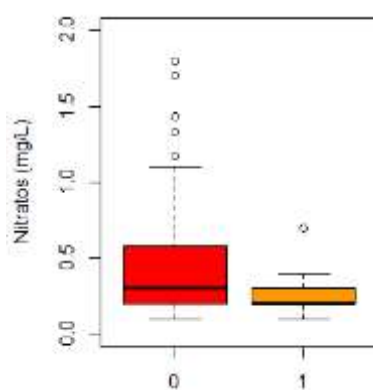
Se puede apreciar la relación negativa entre los parámetros analizados y la presencia del macroinvertebrado. En los dos casos la probabilidad de encontrar a la familia *Hydrobiosidae* disminuye con el aumento de la concentración de los parámetros relacionados. (Ver Figura 7)



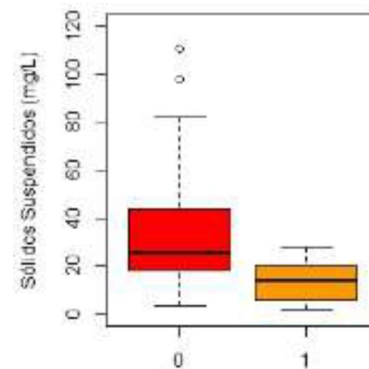
(a)



(b)



(c)



(d)

Figura 7 – (a) Probabilidad de encontrar a la familia *Hydrobiosidae* en relación con nitratos. **(b)** Familia *Hydrobiosidae* con sólidos suspendidos. **(c)** Diagrama de caja, presencia (1) y ausencia (0) de macroinvertebrados de la familia *Hydrobiosidae* en relación con nitratos. **(d)** Diagrama de caja Familia *Hydrobiosidae* con sólidos suspendidos.

ELABORADO POR: Velastegui V., 2019

En la Figura 7c – color rojo, correspondiente a la ausencia de la familia *Hydrobiosidae*, el valor de la mediana de la concentración de nitratos es igual a 0.2 mg/L. Los valores atípicos causan sesgo en la distribución de los datos. El rango donde se presentan la mayoría de los datos es de 0.2 mg/L a 0.6 mg/L.

En cuanto al diagrama de caja en presencia de la familia *Hydrobiosidae* (Figura 7c – color naranja), los datos se encuentran menos dispersos. La mediana coincide con el primer cuartil, lo que denota asimetría y sesgo. El 50% de los datos de concentración de nitratos se encuentran en el rango de 0.15 mg/L a 0.4mg/L.

Aunque no existen estudios relacionados con los efectos directos que tienen los nitratos sobre la comunidad de macroinvertebrados, es conocido que puede afectar la capacidad de los animales acuáticos para sobrevivir, crecer y reproducirse, lo que resulta en la toxicidad directa (aguda o crónica) de este compuesto nitrogenado inorgánico (Camargo y Alonso, 2006). Es importante recalcar que la respuesta de la comunidad de macroinvertebrados a ciertos contaminantes difiere de una población a otra (Roldán, 2016), por lo que una cierta concentración de nitratos en el agua puede causar la disminución de la abundancia relativa de individuos de una familia o, en el peor de los casos, la desaparición de un grupo taxa en particular.

En el diagrama de caja, correspondiente a la concentración de sólidos suspendidos, en ausencia de la familia *Hydrobiosidae* (Figura 7d – color rojo) la mediana es de aproximadamente 25 mg/L. La mayoría de las muestras tomadas presentan un rango de concentración de sólidos suspendidos de 20 mg/L a 45 mg/L. Se evalúa, además, una distribución asimétrica de los datos, influenciada por los valores atípicos presentes. Cabe mencionar, para comprender el contexto de los resultados, que se realizó un muestreo de parámetros fisicoquímicos en dos cuerpos de agua con diferentes grados de contaminación, por lo que posiblemente sea la causa de los valores atípicos.

En contraste con el diagrama de caja en presencia de la familia *Hydrobiosidae* (Figura 7d – color naranja), la mediana es igual a 18 mg/L menor a la que se observó en ausencia de esta familia. La mayoría de los valores de sólidos suspendidos se encuentran entre 2 mg/L y 19 mg/L.

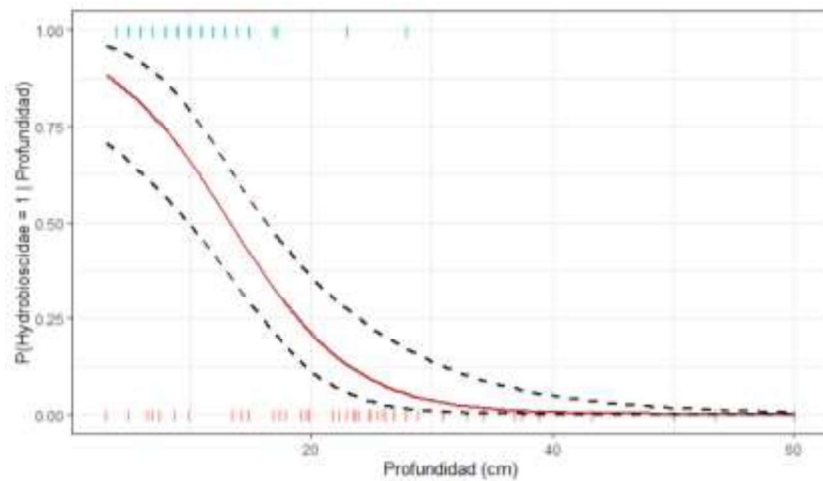
Dentro de los sistemas de agua corriente, la existencia de una elevada concentración de material en suspensión repercute de mayor forma sobre ciertos grupos de macroinvertebrados, específicamente en aquellos que basan su alimentación en el consumo de otras larvas de insectos acuáticos (depredadores). Esto se debe principalmente a que un aumento de los sólidos acarreados por un cuerpo de agua puede disminuir la visibilidad de los cazadores y, como consecuencia, minimizar el índice de éxito de atrapar presas (Cretaz y Barten, 2007). De acuerdo a Tomanova, Goitia, & Helešic (2006), la familia *Hydrobiosidae* pertenece al grupo funcional de los depredadores, lo cual explicaría su relación negativa con la concentración de sólidos suspendidos del agua con un p-valor menor al 1%.

Simultáneamente se realizó el análisis de regresión logística para los parámetros hidráulicos, y los resultados se presentan a continuación:

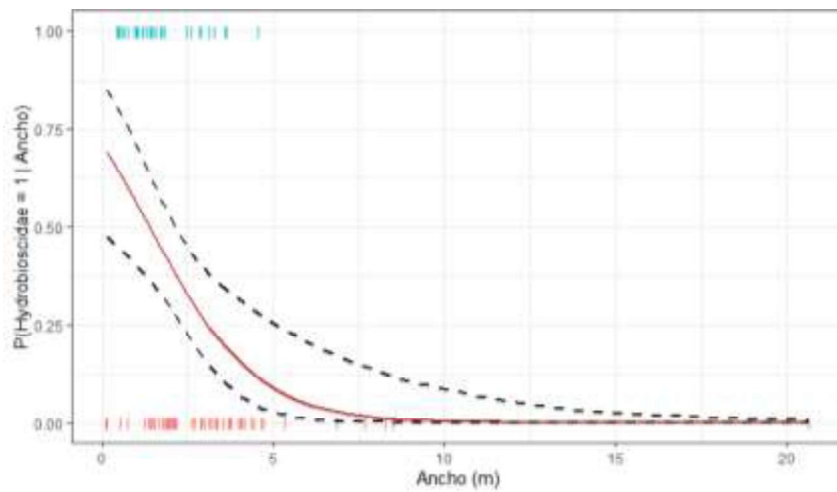
Tabla 14 -. Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros hidráulicos correlacionados con la presencia de *Hydrobiosidae*.

Variables predictoras	Parámetros de regresión	Coefficientes	p-valor
	α	4.57394	5.66e-05
Profundidad	β_1	-0.20493	7.82e-05
Ancho	β_3	-0.83281	0.00493
AIC		60.25	
Pseudo R²		0.46	

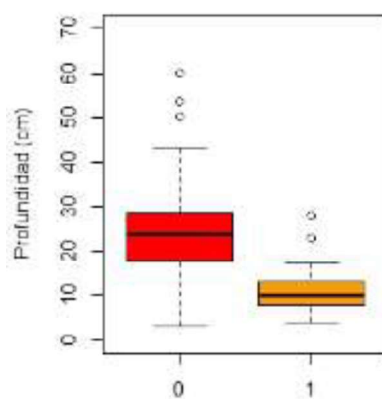
ELABORADO POR: Velastegui V., 2019



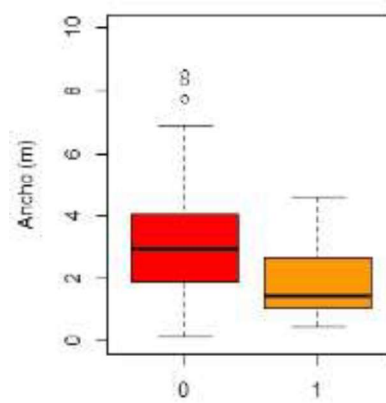
(a)



(b)



(c)



(d)

Figura 8 – (a) Probabilidad de encontrar a la familia *Hydrobiosidae* en relación con la profundidad. (b) Familia *Hydrobiosidae* con Ancho del cauce. (c) Diagrama de caja, presencia (1) y ausencia (0) de macroinvertebrados de la familia *Hydrobiosidae* en relación con la profundidad. (d) Diagrama de caja Familia *Hydrobiosidae* el ancho del cauce.

ELABORADO POR: Velastegui V., 2019

En la Figura 8 literal a y b, se observa la relación negativa existente entre la profundidad y el ancho del cauce, información corroborada con los diagramas de caja presentados en el literal c y d. En el diagrama de caja de la profundidad, en ausencia de la familia *Hydrobiosidae*, el valor de la mediana es igual a 24 cm que en este caso es cercano al valor de la media. La mayoría de los valores se encuentran en el intervalo de 17,6 cm a 28.8 cm. Al comparar con el diagrama en presencia de esta familia, se puede apreciar una disminución en la mediana (10 cm) de los valores de profundidad. El rango intercuartil se encuentra entre 8 cm y 13.3 cm. Con estos resultados se puede concluir a que mayor profundidad en un cuerpo de agua, disminuye la probabilidad de existencia de la familia *Hydrobiosidae*. El mismo comportamiento se manifiesta con el ancho del cauce, la probabilidad disminuye a medida que el cauce aumente sus dimensiones.

Es conocido que sistemas fluviales de alta montaña, se ven influenciados por la pendiente, modificando las características de los cauces. Esto afecta a la velocidad de flujo del agua, la misma que a su vez da forma a los ríos. Los ríos asociados al ritrón presentan características similares, independiente de la región en la cual se sitúan, ya que el ancho y la profundidad de los cauces es menor en comparación con aquellos sistemas fluviales de llanura. Del mismo modo, en sistemas de alta montaña la pendiente propicia velocidades de flujo más altas en comparación con las zonas más bajas (Roldán & Ramírez, 2008; Munro, 2010).

La microdistribución de la familia *Hydrobiosidae* dentro de los cuerpos de agua corriente está ligada de cierto modo a sus hábitos alimenticios, para lo cual está equipada con dos pseudopatas anales que le permiten aferrarse fuertemente al sustrato y además tiene la capacidad de fabricar un hilo de seda que usa para adherirse a las rocas y cazar (J. Thorp y Rogers, 2014). Frecuentemente se encuentran en ríos con elevadas velocidades, baja profundidad, ancho de cauce pequeño y un sustrato rocoso (Mosquera, 2018)

(Statzner, 1981) de igual forma señala que hay una enorme cantidad de estudios que correlacionan la microdistribución de los invertebrados bentónicos con la velocidad, la profundidad, el tipo de sustrato o una combinación de estos, ya que estos tres factores también determinan la naturaleza y el alcance de la capa límite bentónica (De Deckker y Williams, 2012). Esto contrasta los resultados obtenidos en el presente estudio para esta familia, donde se encontró una relación inversa entre la preferencia de este grupo con ambos parámetros hidráulicos con un p-valor mucho menor al 1%.

Matriz de confusión

En la Tabla 15 se presenta la matriz de confusión, en donde se puede observar la clasificación realizada por el modelo con los datos de entrenamiento y prueba correspondientes a la familia *Hydrobiosidae*.

Tabla 15 -. Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba de los parámetros físicoquímicos.

	DATOS OBSERVADOS			
	ENTRENAMIENTO		PRUEBA	
	Ausencia	Presencia	Ausencia	Presencia
Ausencia	7	0	5	1
Presencia	6	18	0	2

ELABORADO POR: Velasteguí V., 2019

La precisión obtenida de acuerdo con la ecuación (6), para los datos de prueba, es igual a 100% y la sensibilidad igual a 83%. El F-score = 0.91, es satisfactoria con más de 90% de su valor. (Hilbe, 2009; Liu, 2007a)

Se realizó el diagnóstico del rendimiento del modelo obtenido a partir de los parámetros hidráulicos. En la Tabla 16 se puede apreciar el poder de clasificación del modelo con los parámetros profundidad y ancho del cauce.

Tabla 16 -. Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba de parámetros hidráulicos.

	DATOS OBSERVADOS			
	ENTRENAMIENTO		PRUEBA	
	Ausencia	Presencia	Ausencia	Presencia
Ausencia	8	0	4	1
Presencia	7	21	1	2

ELABORADO POR: Velasteguí V., 2019

La precisión obtenida de acuerdo con la ecuación (6), para los datos de prueba, es igual a 80% y la sensibilidad igual a 80%. El F-score = 0.80, es muy buena su clasificación. (Hilbe, 2009; Liu, 2007a)

Curva ROC

El valor del área bajo la curva ROC, para este modelo con los datos de prueba, es de 0.966 (ver Figura 9). De acuerdo con Liu (2007) y Gironés (2017), el modelo es un clasificador muy bueno; es decir clasifica correctamente la presencia y ausencia de esta taxa de acuerdo con los parámetros fisicoquímicos del agua.

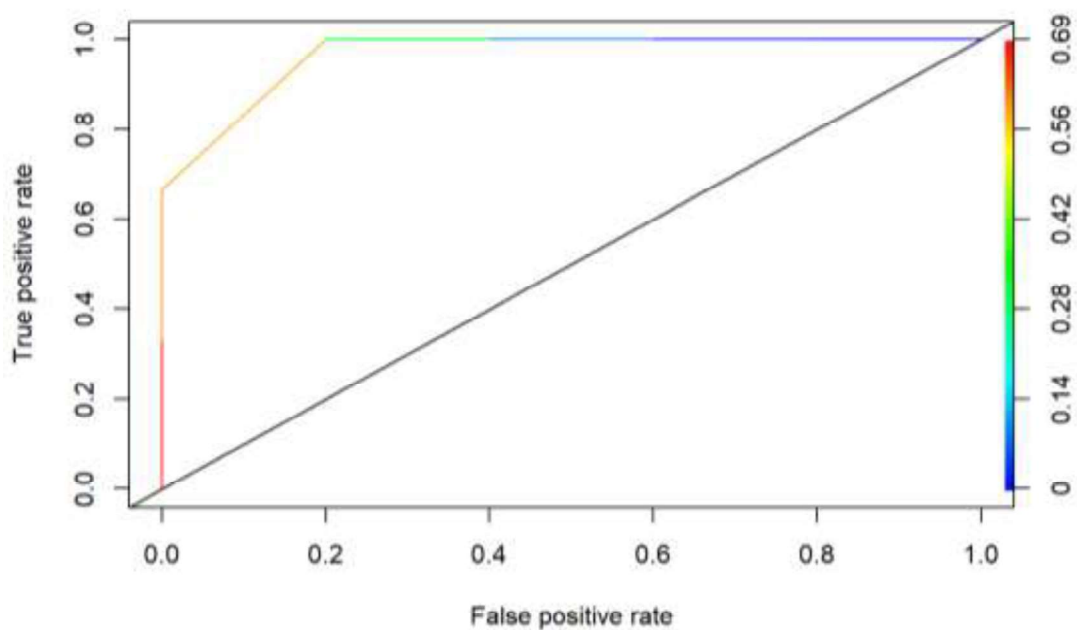


Figura 9 – Curva ROC del modelo estimado con datos del conjunto de prueba
ELABORADO POR: Velastegui V., 2019

La curva correspondiente al modelo desarrollado con los parámetros hidráulicos se presenta en la Figura 10. En ella podemos ver que la clasificación realizada es buena con un valor de AUC igual a 0.866.

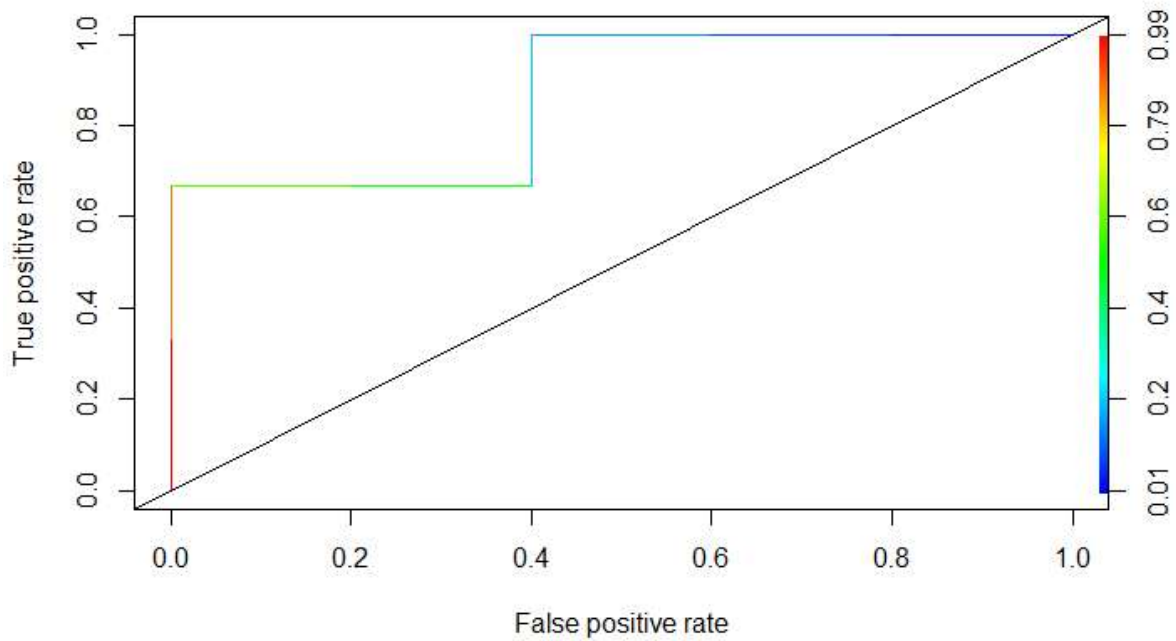


Figura 10 – Curva ROC del modelo estimado con datos del conjunto de prueba con parámetros hidráulicos.

ELABORADO POR: Velastegui V., 2019

4.3.2. FAMILIA *SIMULIIDAE* (SENSIBILIDAD A CONTAMINANTES: 8)

En 32 sitios muestreados, de un total de 86, se identificaron individuos pertenecientes a la familia *Simuliidae*. La mayor cantidad se encontró en época seca, el 86% de los cuales fueron identificados en el río Ushimana y el 14% en el río Jatunhuayco.

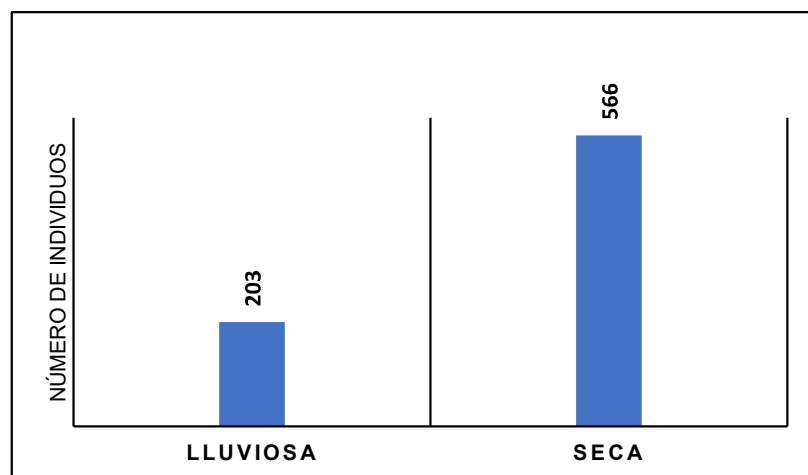


Figura 11 –. Número de individuos de la familia *Simuliidae* por época climática
ELABORADO POR: Velastegui V., 2019

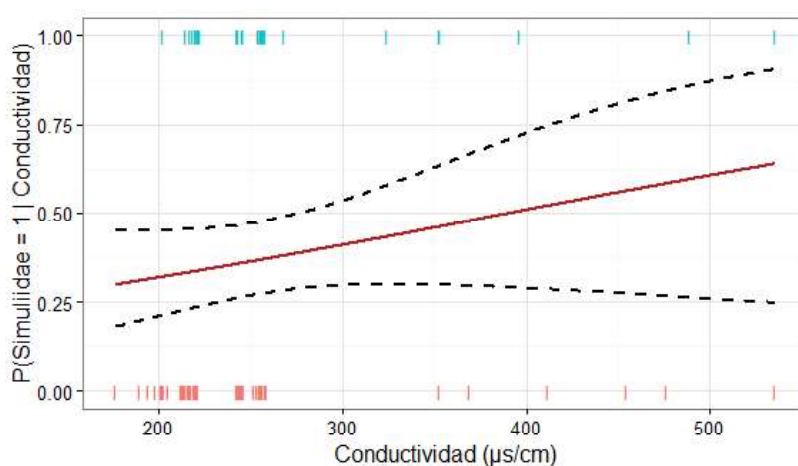
Las variables significativas obtenidas a partir del análisis de regresión logística se presentan en la Tabla 17.

Tabla 17 -. Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de *Simuliidae*.

Variables predictoras	Parámetros de regresión	Coefficientes	p-valor
	α	2.092732	0.19124
Conductividad	β_1	0.019075	0.00721
(NH ₄) ⁺	β_2	0.588477	0.03156
Sólidos Totales	β_3	-0.033715	0.00763
AIC		99.57	
Pseudo R²		0.11	

ELABORADO POR: Velastegui V., 2019

Respecto con la predicción del modelo, los parámetros tales como la conductividad, ion amonio y la concentración de sólidos totales son significativos con la existencia de la familia *Simuliidae*. Se puede observar, para cada parámetro la relación existente (Figura 12).



(a)

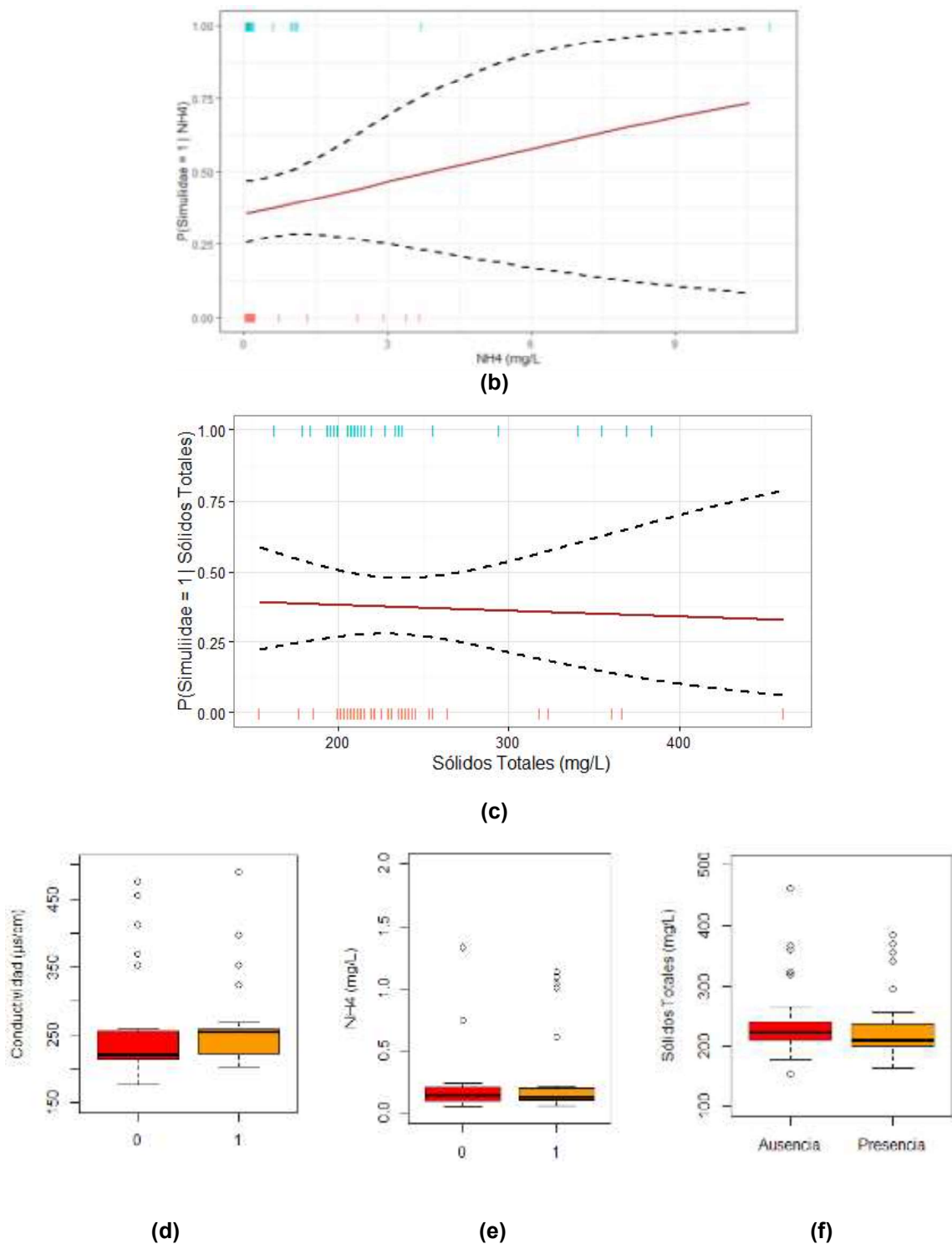


Figura 12 –. (a) Probabilidad de encontrar a la familia *Simuliidae* en relación con la conductividad. (b) Familia *Simuliidae* con $(\text{NH}_4)^+$. (c) Familia *Simuliidae* con sólidos totales. (d) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia *Simuliidae* de acuerdo con la conductividad. (e) Diagrama de caja familia *Simuliidae* con $(\text{NH}_4)^+$. (f) Diagrama de caja familia *Simuliidae* con sólidos totales.

ELABORADO POR: Velasteguí V., 2019

En la Figura 12d, se puede comparar los diagramas de cajas de la relación entre la conductividad y la familia *Simuliidae*. Para el primer diagrama de caja (rojo), correspondiente a la ausencia de la familia *Simuliidae*, el valor de la mediana es de 221 $\mu\text{s}/\text{cm}$. La mayoría de los datos se encuentran en el rango de 214 $\mu\text{s}/\text{cm}$ a 254 $\mu\text{s}/\text{cm}$. En el segundo caso (naranja) el valor de la mediana es 255 $\mu\text{s}/\text{cm}$, mayor en relación con el diagrama de caja correspondiente a la ausencia de *Simuliidae*. La mayoría de los datos presentan valores entre 222 $\mu\text{s}/\text{cm}$ y 258 $\mu\text{s}/\text{cm}$. En los dos casos existen valores atípicos de conductividad.

La conductividad afecta la distribución y composición de macroinvertebrados ya que propicia la desaparición de especímenes que carecen de un sistema osmorregulador que les permite adaptarse a medios con una alta conductividad (Roldán, 2003b). Para el presente estudio, de acuerdo a los datos estipulados en la Tabla 17, esta familia tiene una relación directa con este parámetro; además, presenta un p-valor menor al 1% por lo que se puede afirmar que este grupo tiene preferencia por ambientes salinos. De acuerdo a Coscarón & Arias (2007a), la familia *Simuliidae* prefiere normalmente los cuerpos de agua oligosalinos, es decir, que presentan una conductividad baja; sin embargo, existen algunas subcategorías que pueden soportar ambientes con alta conductividad.

En la Figura 12e, se expone la presencia y ausencia de *Simuliidae* en función de la concentración del ion amonio (NH_4^+). Es mínima la diferencia que existe entre las medianas de los datos de concentración del ion amonio cuando la familia está ausente (0.14 mg/L) y presente (0.13 mg/L). En los dos casos el rango intercuartil se encuentra entre 0.09 mg/L y 0.21 mg/L. Así mismo, el valor mínimo es de 0.05 mg/L y el máximo 0.24 mg/L; valores superiores al valor máximo son considerados como valores atípicos, los cuales se observan en mayor medida cuando individuos de la familia *Simuliidae* se encuentran presentes.

Del mismo modo, se puede observar que existe una relación directa entre la presencia de esta familia con la concentración de amonio en el cuerpo de agua analizado con un p-valor menor al 5%. Aunque no existen estudios relacionados con los efectos directos que tienen el amonio sobre la comunidad de macroinvertebrados, es conocido que puede lesionar las agallas respiratorias de los organismos acuáticos (Lewis, 1998). Sin embargo, esta familia

parece presentar una cierta afinidad con la existencia de este compuesto en el medio llevándonos a formular 2 hipótesis:

- H_1 La familia Simuliidae se desarrolla mejor en medios con concentraciones de amonio elevadas.
- H_2 : La familia Simuliidae no se ve afectada por el aumento en la concentración de amonio en el medio.

De la experiencia se ha observado que este tipo de organismos requieren de sustratos estables para desarrollarse, esto debido a que permanecen pegados a la superficie de las rocas, troncos u hojas mediante un colchón de seda que es producido por un órgano anal similar al de las arañas (Ampyx, 2001). Esta condición hace que deban desarrollarse en la capa límite laminar (Moss, 2013), donde generalmente decrece la tasa de difusión de gases disueltos, por lo que su distribución y patrones comportamentales talvez estén ligados a sus hábitos alimenticios y no a una dependencia de los gases disueltos en el medio. Debido a ello, la hipótesis más razonable a la cual debería atribuirse su preferencia por ambientes con presencia de amonio, es la segunda.

El diagrama de caja correspondiente a la concentración de sólidos totales (Figura 12f) en ausencia de *Simuliidae*, presenta una mediana igual a 222 mg/L. La mayoría de los datos toman valores entre 210 mg/L y 240 mg/L. Referente a la presencia, existe dispersión y asimetría en los datos en comparación con la ausencia de esta familia. La mediana es igual a 210 mg/L y el rango intercuartil se encuentra entre los valores de 200mg/L y 236 mg/L.

Esta familia además presenta una relación inversa con los sólidos totales del medio acuático en el que se desarrollan con un p-valor menor al 1%. Al ser organismos catalogados como colectores-filtradores (Tomanova, Goitia, y Helešic, 2006b), su alimentación se basa en la captación de partículas orgánicas que transporta la corriente; sin embargo, el material particulado acarreado de manera excesiva puede ser depositado en los espacios intersticiales del sustrato rocoso disminuyendo así, el área óptima para el establecimiento de esta familia (Gómez y Salazar, 2017).

Matriz de confusión

En la Tabla 18 se presenta la matriz de confusión, en donde se puede observar la clasificación realizada por el modelo con los datos de entrenamiento y datos de prueba. La precisión es igual a 80% y la sensibilidad igual a 80% para los datos de prueba. El F-score = 0.8. (Hilbe, 2009; Liu, 2007a)

Tabla 18 -. Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba.

	DATOS OBSERVADOS			
	ENTRENAMIENTO		PRUEBA	
	Ausencia	Presencia	Ausencia	Presencia
Ausencia	7	1	4	1
Presencia	6	11	1	2

ELABORADO POR: Velasteguí V., 2019

Curva ROC

El valor del área bajo la curva ROC, para este modelo con los datos de prueba, es de 0.8 (ver Figura 13). De acuerdo con Liu (2007) y Gironés (2017), el modelo es un clasificador muy bueno.

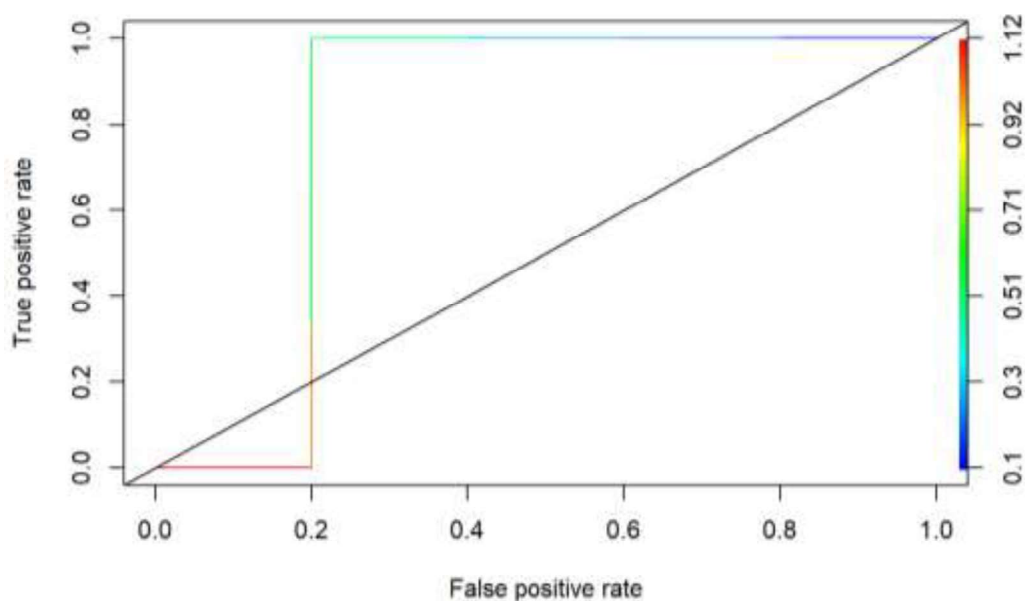


Figura 13 – Curva ROC del modelo estimado con datos del conjunto de prueba
ELABORADO POR: Velastegui V., 2019

4.3.3. FAMILIA *BAETIDAE* (SENSIBILIDAD A CONTAMINANTES: 7)

Individuos de esta familiar fueron encontrados en mayor cantidad en época lluviosa. De los cuales, se identificaron el 60% en el río Jatunhuayco y el 40% en el río Ushimana.

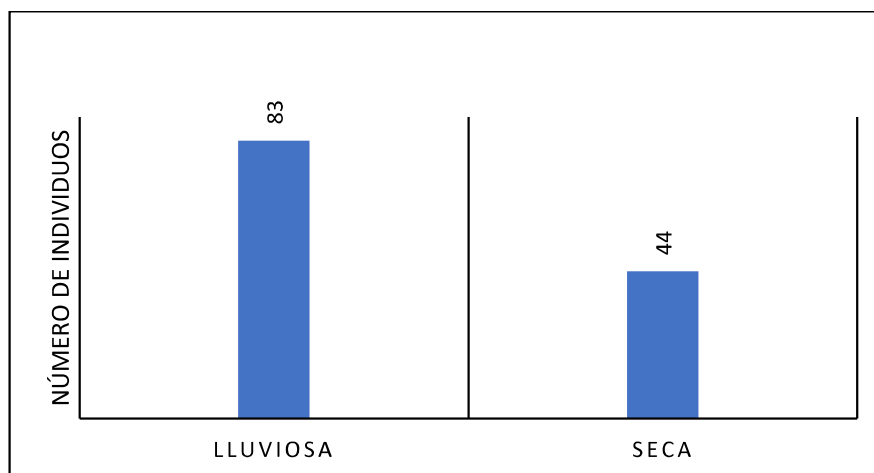


Figura 14 – Número de individuos de la familia *Baetidae* por época climática
ELABORADO POR: Velastegui V., 2019

En la Tabla 19 se puede evaluar el resultado del análisis de regresión logística.

Tabla 19 -. Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de *Baetidae*.

Variabes predictoras	Parámetros de regresión	Coefficientes	p-valor
	α	51.16272	0.01018
ln_pH	β_1	-19.96125	0.01948
(NH ₄) ⁺	β_2	-35.48809	0.00477
Sólidos disueltos	β_3	-0.02647	0.02872
AIC		81.41	
Pseudo R²		0.17	

ELABORADO POR: Velastegui V., 2019

De acuerdo a los datos contenidos en la Tabla 19, existe una relación inversa entre la presencia de este organismo y el pH con un p-valor menor al 5%. Este parámetro es uno de los más importantes reguladores de la vida acuática ya que indica las proporciones de las distintas formas de carbono inorgánico que los organismos fotosintéticos utilizan dentro

de sus procesos vitales; aunque es difícil distinguir entre los efectos específicos de este parámetro con los que resultan de la acción de otros (Margalef, 1986). La familia Baetidae pertenece al grupo funcional alimenticio de los raspadores y también de los colectores-recolectores (Tomanova et al., 2006b); en ambos casos, su dieta depende del perifiton pegado al sustrato y de las partículas orgánicas depositadas en medio de las rocas. Aunque no se conoce los efectos directos de este parámetro sobre la comunidad acuática, los cambios de pH pueden afectar directamente el suministro alimenticio para los invertebrados acuáticos al limitar la reproducción de organismos fotosintéticos y, consecuentemente, afectar indirectamente sobre su población (Rosemond, Reice, Elwood, y Mulholland, 1992).

También se puede evidenciar una relación inversa entre la presencia del radical amonio y esta familia. Aunque no existen estudios relacionados con los efectos directos que tiene el amonio sobre esta familia, se sabe que puede lesionar las agallas respiratorias de los organismos acuáticos (Lewis, 1998). Estructuralmente, este grupo de invertebrados presentan agallas laminadas insertadas lateralmente las cuales pueden ser visibles a simple vista en algunos casos (Bauernfeind y Soldan, 2012); al encontrarse expuestas estas estructuras respiratorias al medio circundante probablemente, debido a su mayor área de contacto, sean más susceptibles de ser afectadas por este ion. Además, Jacobsen (2008) sostiene que las branquias son importantes reguladoras iónicas y no solamente actúan como órganos respiratorios, por lo que la presencia de este radical podría estar impidiendo el adecuado funcionamiento de estas estructuras.

En la Tabla 19 se puede evidenciar que los Sólidos disueltos se relacionan inversamente con la presencia de esta familia. En un estudio realizado por Mosquera (2018) se puede evidenciar un patrón similar para este organismo, donde su abundancia se ve afectada negativamente a medida que aumenta la conductividad del agua. Cabe mencionar que, si bien es cierto que la conductividad y los TDS son parámetros diferentes de calidad del agua, existe una estrecha relación entre ellos ya que los TDS son una medida de las sales inorgánicas disueltas en el agua pero, por motivos de facilidad, se mide la conductividad eléctrica (Roldán y Ramírez, 2008).

Matriz de confusión

En la Tabla 20 se presenta la matriz de confusión, en donde se puede observar la clasificación realizada por el modelo con los datos de entrenamiento y prueba.

Tabla 20 -. Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba.

	DATOS OBSERVADOS			
	ENTRENAMIENTO		PRUEBA	
	Ausencia	Presencia	Ausencia	Presencia
Ausencia	6	0	2	2
Presencia	3	15	2	0

ELABORADO POR: Velasteguí V., 2019

De acuerdo con la matriz, tiene una precisión de 50% y la sensibilidad de 50% para los datos de prueba. El F-score = 0.5. (Hilbe, 2009; Liu, 2007a)

Curva ROC

El valor del área bajo la curva ROC, para este modelo con los datos de prueba, es de 0.5 (ver Figura 15). De acuerdo con Liu (2007) y Gironés (2017), el modelo es un mal clasificador.

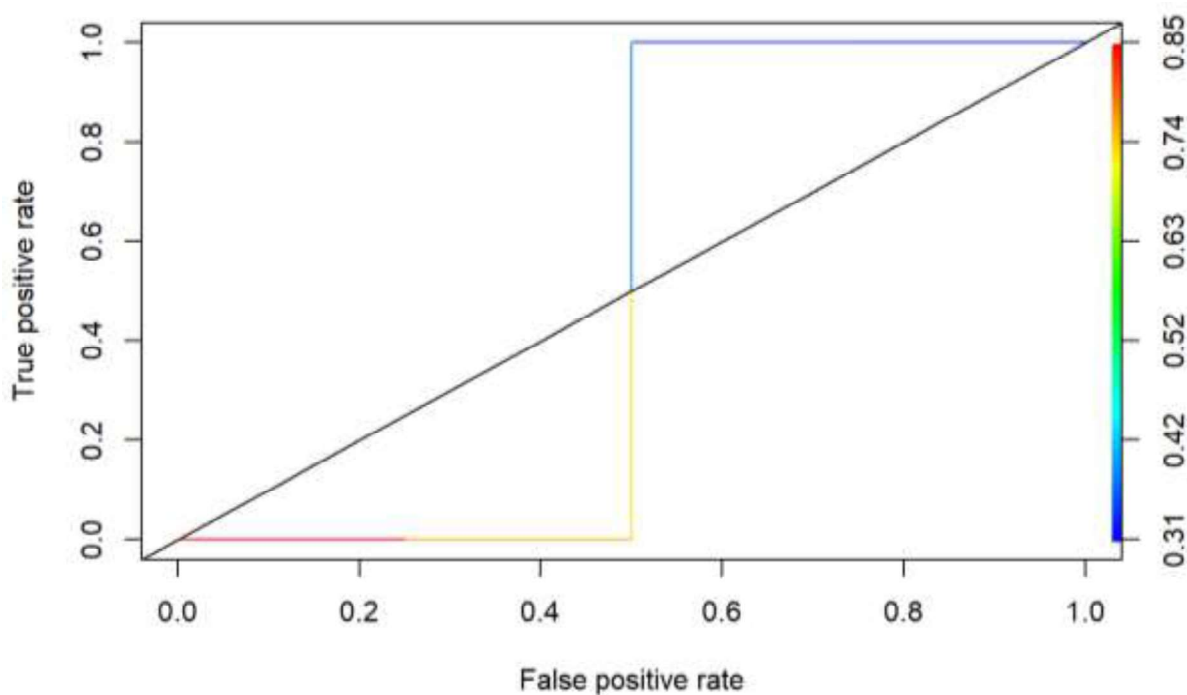


Figura 15 –. Curva ROC del modelo estimado con datos del conjunto de prueba
ELABORADO POR: Velastegui V., 2019

A pesar de obtener parámetros de calidad de agua significativos (p -valor < 0.05) en este modelo de regresión, el análisis de las medidas de ajuste denota una capacidad pobre de clasificación.

4.3.4. FAMILIA *HYALELLIDAE* (SENSIBILIDAD A CONTAMINANTES: 7)

El 99% de individuos fue identificado en el río Jatunhuayco. De acuerdo con la estación climática, se contabilizó mayor número de individuos en la época lluviosa en comparación con la época seca.

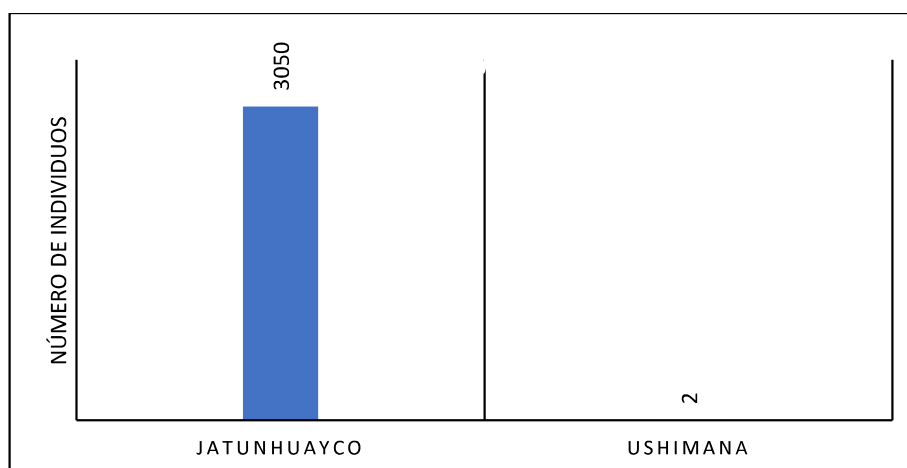


Figura 16 – Número de individuos de la familia *Hyalellidae* por zona de estudio.
ELABORADO POR: Velastegui V., 2019

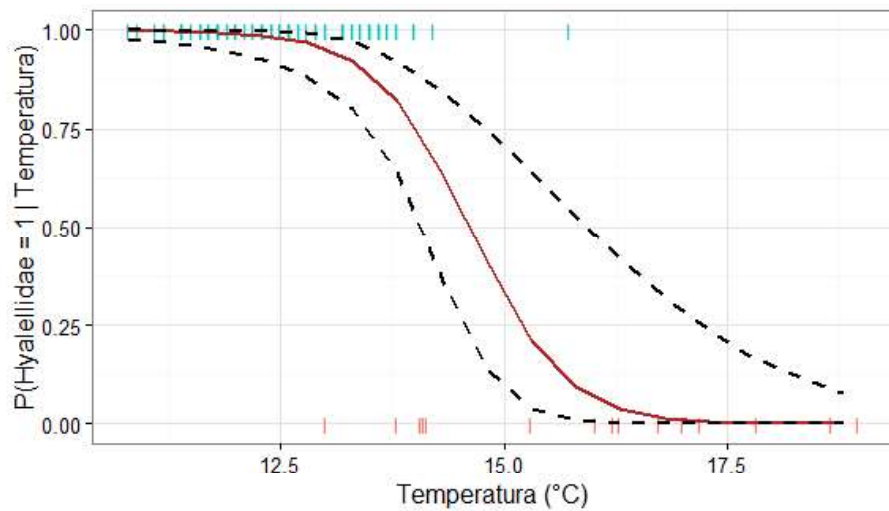
La variable significativa para esta familia *Hyalellidae* es la Temperatura como se aprecia en la Tabla 21.

Tabla 21 -. Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de *Hyalellidae*.

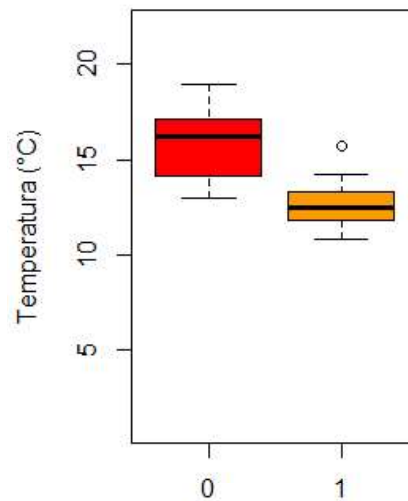
Variables predictoras	Parámetros de regresión	Coefficientes	p-valor
	α	72.254	0.000877
In_Temperatura	β_1	-27.005	0.001111
AIC		32.67	
Pseudo R²		0.61	

ELABORADO POR: Velastegui V., 2019

Con respecto a la temperatura, este parámetro tiene una influencia inversamente proporcional a la probabilidad de encontrar individuos de *Hyalellidae* (ver Figura 17a).



(a)



(b)

Figura 17 – (a) Probabilidad de encontrar a la familia *Hyalellidae* en relación con la temperatura.

(b) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia *Hyalellidae* de acuerdo con la Temperatura.

ELABORADO POR: Velastegui V., 2019

La temperatura afecta directamente a la probabilidad de encontrar a la familia *Hyalellidae* como se observa en la Figura 17b. El diagrama de color rojo, en ausencia de la familia *Hyalellidae*, presenta mayor dispersión en los datos. La mediana tiene un valor de 16°C. El 50% de los datos se encuentran en el rango de 14°C a 17°C. El valor mínimo de temperatura, cuando *Hyalellidae* está ausente, es de 13°C y el máximo valor es de 19°C. En este diagrama no se observan datos atípicos.

Referente a diagrama en presencia de la familia *Hyaellidae*, la mediana es igual 12.5 °C, el cual coincide con el valor de la media. El rango intercuartil se encuentra entre 11.8°C y 13.3°C, con un valor mínimos de temperatura de 10.8°C y máximo de 14.2°C. En este diagrama se observa un valor atípico igual a 15.7°C.

En la Tabla 21 se puede evidenciar la existencia de una relación inversa entre la temperatura del cuerpo acuático y la presencia de la familia *Hyaellidae* con un p-valor menor al 1%. En estudios realizados por Villamarín (2008) en los Andes ecuatorianos, se puede observar un patrón similar de distribución de estos organismos en el páramo y puna; es decir, la familia *Hyaellidae* está limitada solamente a ecosistemas de altura. Entre las posibles causas que pueden esclarecer este comportamiento están la disminución de la temperatura por efecto del incremento de la altitud, y la disminución de la presión parcial de oxígeno mediada por esta misma causa (Jacobsen, 2000).

Durante un estudio realizado por Madsen et al. (2015), se evidenció que una reducción en la temperatura incrementa insignificadamente la actividad de la familia *Hyaellidae*, mientras que en otras familias de macroinvertebrados reduce su actividad de manera considerable. En ríos de gran altitud, las bajas temperaturas pueden producir estrés térmico sobre algunos organismos, lo que podría derivarse en una hipoxia sistémica y causar su disminución y/o desaparición de un hábitat en particular (Jacobsen, 2003). Mientras que en otros, como es el presente caso, las bajas temperaturas brindarían un entorno idóneo para que se desarrollen adecuadamente, justificando así su preferencia por este tipo de ambientes.

Matriz de confusión

En la Tabla 22 se presenta la matriz de confusión, en donde se puede observar la clasificación realizada por el modelo con los datos de entrenamiento y de prueba. La precisión, en este caso es igual a 100% y la sensibilidad igual a 100% para los datos de prueba. El F-score = 1, es satisfactoria con un valor 100%. (Hilbe, 2009; Liu, 2007a)

Tabla 22 -. Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba.

	DATOS OBSERVADOS			
	ENTRENAMIENTO		PRUEBA	
	Ausencia	Presencia	Ausencia	Presencia
Ausencia	0	1	1	0
Presencia	1	59	0	7

ELABORADO POR: Velasteguí V., 2019

Curva ROC

El valor del área bajo la curva ROC, para este modelo con los datos de prueba, es de 1 (ver Figura 18). Clasifica correctamente la presencia y ausencia de esta taxa de acuerdo con los parámetros fisicoquímicos del agua.

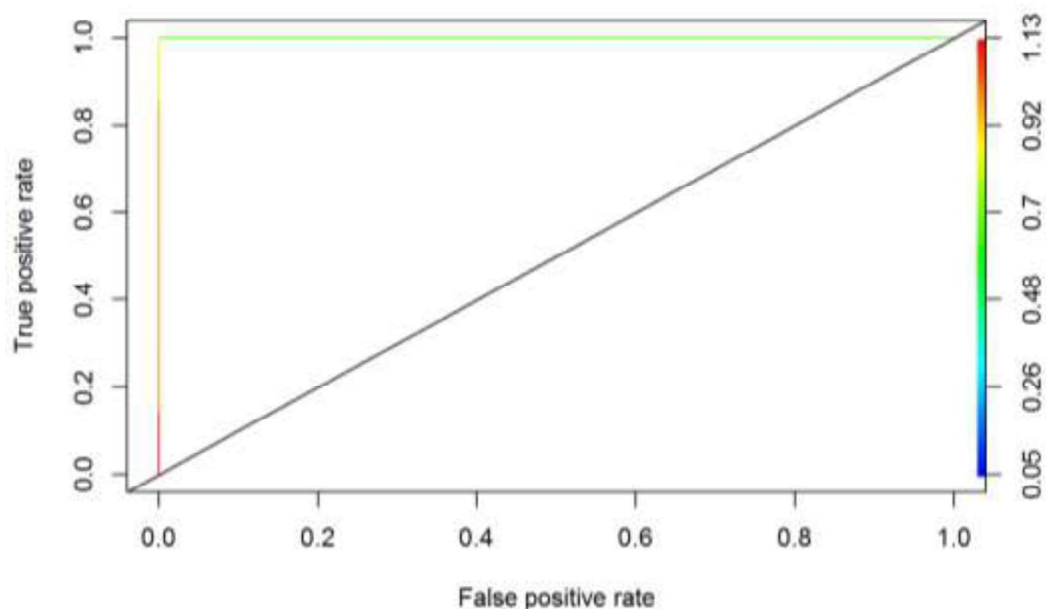


Figura 18 -. Curva ROC del modelo estimado con datos del conjunto de prueba
ELABORADO POR: Velastegui V., 2019

4.3.5. FAMILIA *GLOSSOSOMATIDAE* (SENSIBILIDAD A CONTAMINANTES: 7)

La mayor parte de los individuos fue encontrada en el río Jatunhuayco. En cuanto a la distribución de acuerdo con la época climática, en la Figura 19 se puede apreciar un incremento en la cantidad del número de macroinvertebrados en la época seca.

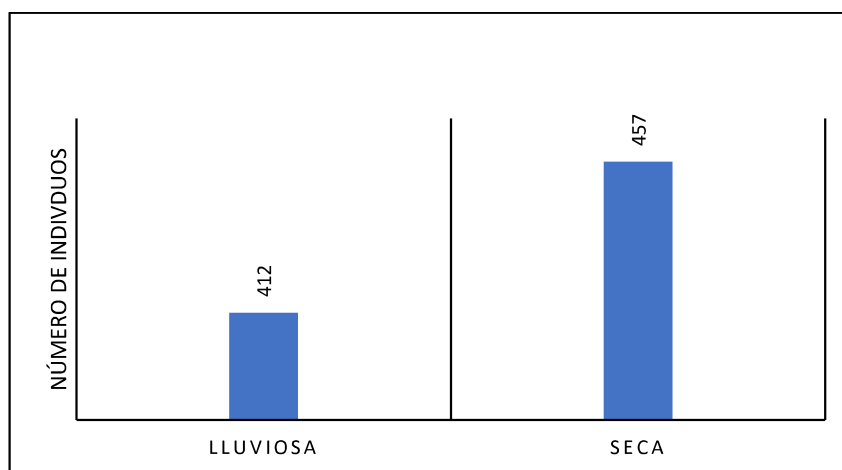


Figura 19 –. Número de individuos de la familia *Glossosomatidae* por época climática.
ELABORADO POR: Velastegui V., 2019

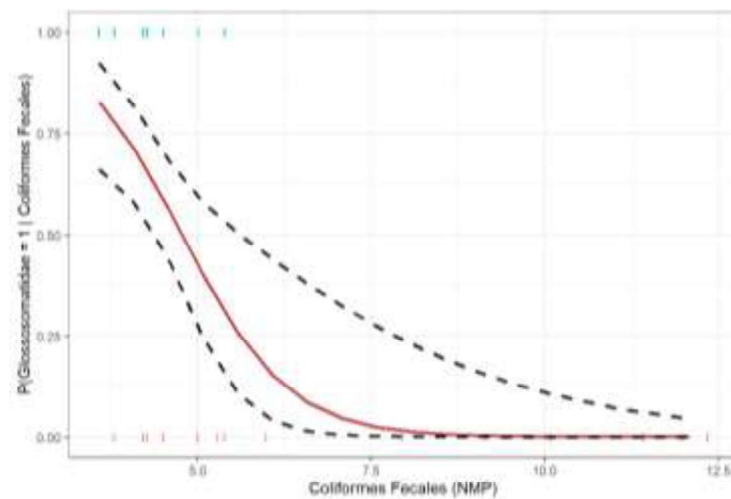
La presencia de Coliformes Fecales y el pH presentaron mayor significancia al relacionarse con la familia *Glossosomatidae* como se presenta en la Tabla 23.

Tabla 23 -. Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de *Glossosomatidae*.

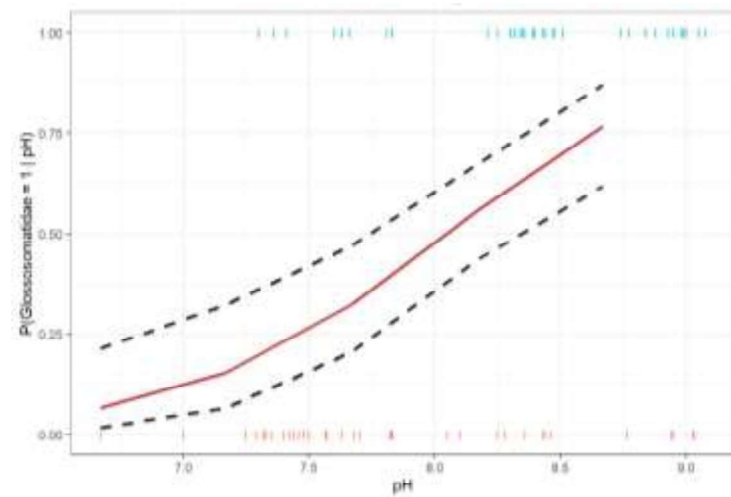
Variables predictoras	Parámetros de regresión	Coefficientes	p-valor
	α	-7.5280	0.074543
In_Coliformes F	β_1	-1.3697	0.008149
pH	β_2	1.7240	0.000798
AIC		77.7	
Pseudo R²		0.34	

ELABORADO POR: Velastegui V., 2019

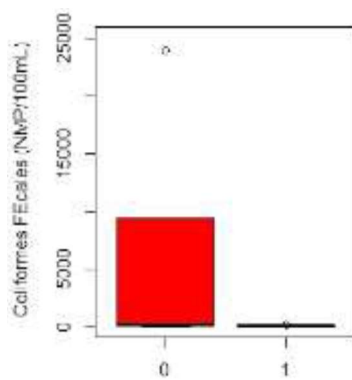
En la Figura 20 se aprecia la relación existente entre las variables significativas y la presencia de la familia *Glossosomatidae*.



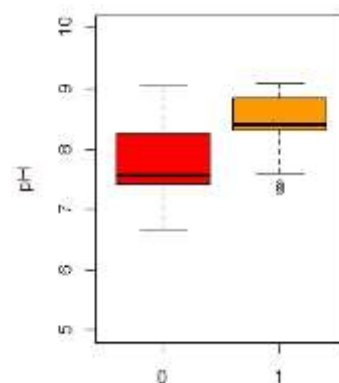
(a)



(b)



(c)



(d)

Figura 20 – (a) Probabilidad de encontrar a la familia *Glossosomatidae* en relación con Coliformes fecales. **(b)** Familia *Glossosomatidae* con el pH. **(c)** Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia *Glossosomatidae* de acuerdo con Coliformes fecales. **(d)** Diagrama de caja familia *Glossosomatidae* con el pH.

ELABORADO POR: Velastegui V., 2019

Uno de los parámetros significativos relacionados con la presencia de la familia *Glossosomatidae* es la cantidad de coliformes fecales. En la Figura 20c, se observa el diagrama de caja correspondiente a la presencia y ausencia de la familia *Glossosomatidae* en relación con este parámetro de calidad. Una distribución asimétrica de los datos es visible en el gráfico, esto debido a valores altos de Coliformes fecales reportados en el río Ushimina. En ausencia de esta familia, el valor mínimo de coliformes fecales es 45 NMP/100mL y el máximo es 230000 NMP/100mL. La mediana es igual a 9400 NMP/100mL. En cuanto al diagrama de caja en presencia de *Glossosomatidae*, el rango de datos mínimo y máximo disminuye en comparación a la ausencia de esta familia. El valor mínimo es igual a la mediana es igual 36 NMP/100mL y el valor máximo es 91 NMP/100mL. La mayoría de los datos se encuentran en el rango de 45 NMP/100mL a 82 NMP/100mL.

Al analizar el gráfico de pH (Figura 20d), es comparable la diferencia entre las medianas en ambos casos. En ausencia de *Glossosomatidae* la mediana tiene p-valorH igual 7.5 y el rango intercuartil se encuentra entre 7.4 y 8.2. En cuanto a la presencia, se observa mayor valor en la mediana, igual a 8.4, en comparación con la ausencia de esta familia. El rango intercuartil se encuentra entre 8.28 y 8.86, con la presencia de valores atípicos inferiores a pH de 7.4.

De acuerdo a los datos contenidos en la Tabla 23, existe una relación inversa entre la presencia de este organismo y el pH con un p-valor menor al 1%. Esta familia, al igual que los Baetidos, pertenecen al grupo funcional alimenticio de los raspadores y colectores-recolectores (Tomanova et al., 2006b); su dieta depende del perifiton pegado al sustrato y de las partículas orgánicas depositadas en medio de las rocas. A diferencia de la familia *Baetidae*, estos organismos permanecen pegados al sustrato la mayor parte del tiempo, ya que construyen estuches a partir de los granos finos depositados entre las rocas (Gillott, 2005). Además, debido a que su capacidad de moverse es menor a causa de su forma de vida y adaptaciones estructurales, su presencia podría verse afectada por una disminución en el suministro de perifiton que crezca en el sustrato (Rosemond et al., 1992) ya que el pH puede alterar las propiedades electroquímicas del perifiton (Gari y Corigliano, 2006)

Matriz de confusión

En la Tabla 24 se presenta la matriz de confusión, en donde se puede observar la clasificación realizada por el modelo con los datos de entrenamiento y de prueba.

Tabla 24 -. Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba.

	DATOS OBSERVADOS			
	ENTRENAMIENTO		PRUEBA	
	Ausencia	Presencia	Ausencia	Presencia
Ausencia	7	1	4	0
Presencia	11	35	0	4

ELABORADO POR: Velasteguí V., 2019

La precisión, en este caso es igual a 100% y la sensibilidad igual a 100% para los datos de prueba. El F-score = 1, es satisfactoria con un valor superior al 90%. (Hilbe, 2009; Liu, 2007a)

Curva ROC

El valor del área bajo la curva ROC, para este modelo con los datos de prueba, es de 1 (ver Figura 21).

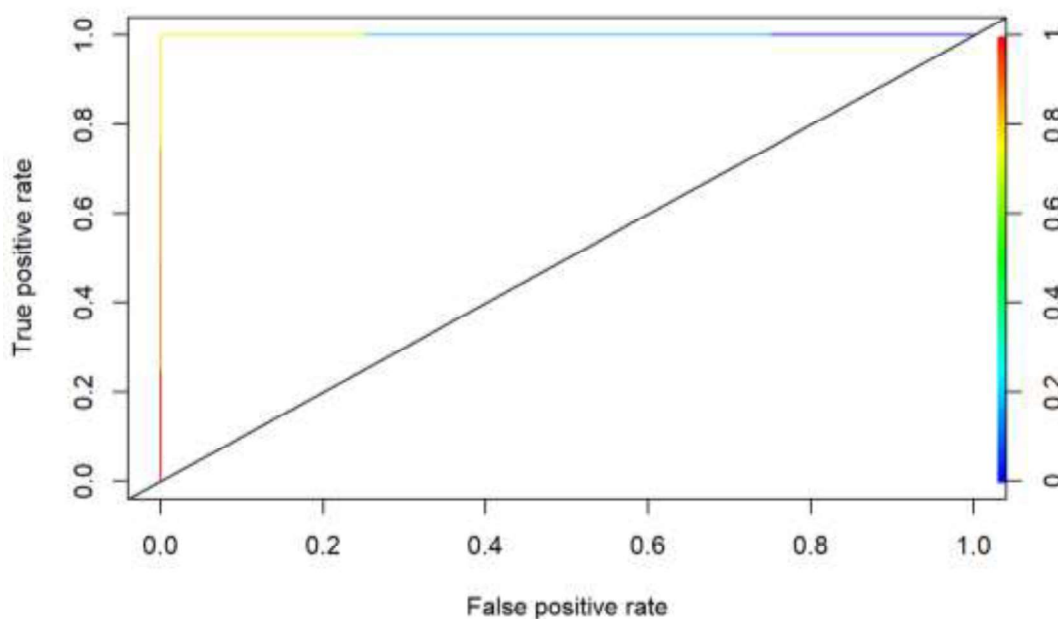


Figura 21 –. Curva ROC del modelo estimado con datos del conjunto de prueba
ELABORADO POR: Velastegui V., 2019

4.3.6. FAMILIA *PLANARIIDAE* (SENSIBILIDAD A CONTAMINANTES: 7)

Individuos de la familia *Planariidae* fueron encontrados en mayor medida en el río Jatunhuayco. La época lluviosa es la que presenta un número superior de taxas como se visualiza en la Figura 22.

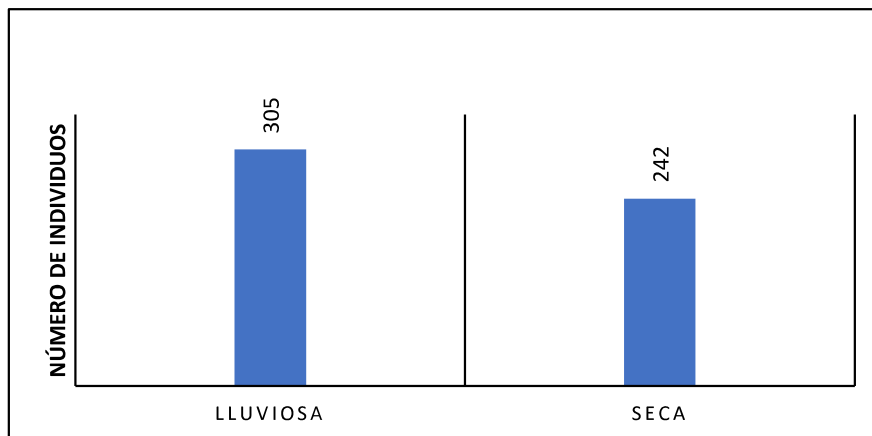


Figura 22 –. Número de individuos de la familia *Planariidae* por época climática.
ELABORADO POR: Velastegui V., 2019

Se detalla a continuación, las variables significativas encontradas de acuerdo con el análisis de regresión logística.

Tabla 25 -. Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de *Planariidae*.

Variabes predictoras	Parámetros de regresión	Coefficientes	p-valor
	α	-8.75209	0.21731
Temperatura	β_1	-2.1179	0.00203
pH	β_2	4.5678	0.00150
AIC		44.78	
Pseudo R²		0.60	

ELABORADO POR: Velastegui V., 2019

La temperatura y el pH son parámetros que afectan la presencia de macroinvertebrados de la familia *Planariidae*. En la Figura 23a, se presenta un gráfico en el cual se observa la tendencia que decrece de la posibilidad de la presencia de un macroinvertebrado de la familia *Planariidae*. En la Figura 23b, se presenta la relación con el pH en el cuerpo de agua.

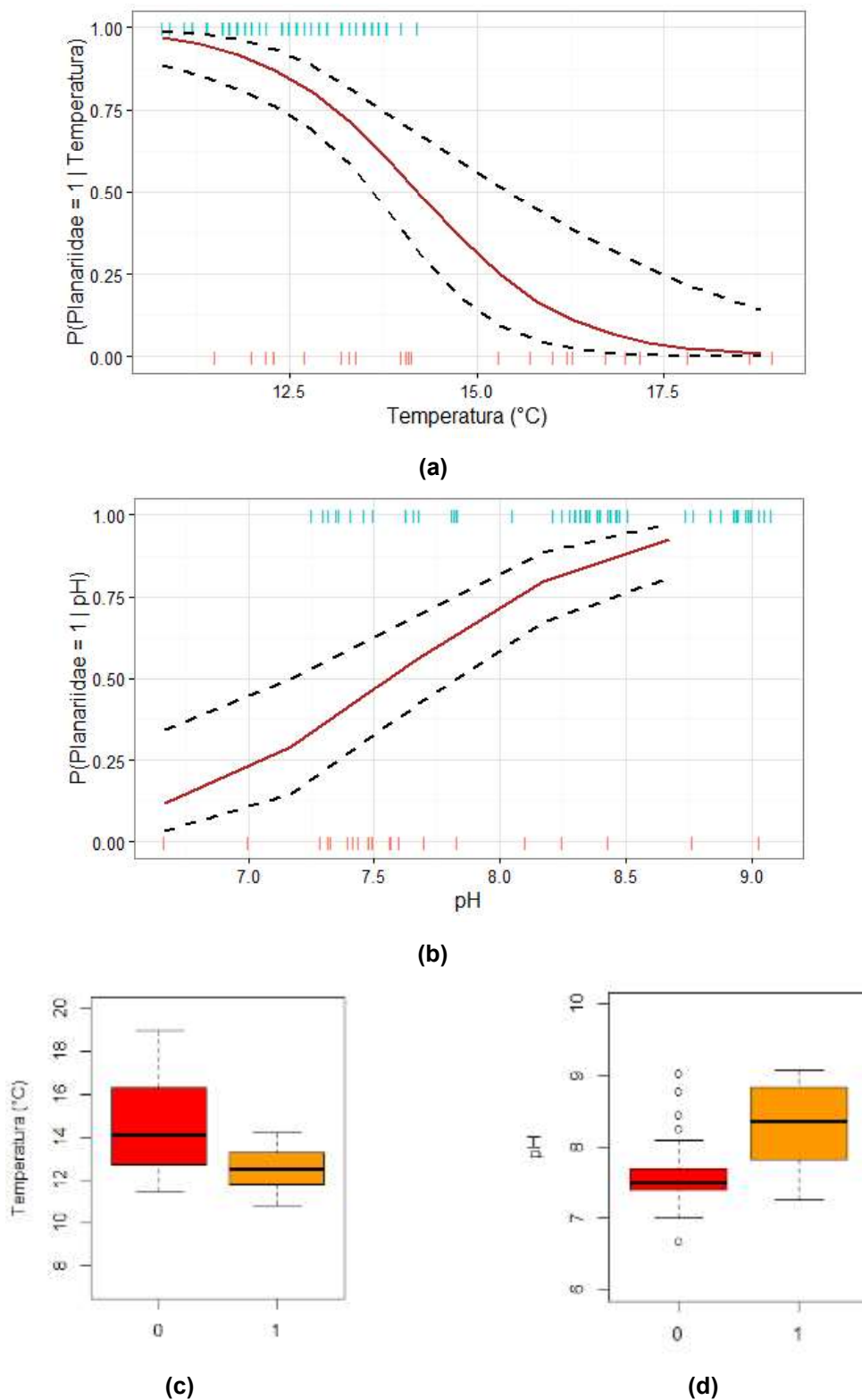


Figura 23 – (a) Probabilidad de encontrar a la familia *Planariidae* en relación con la temperatura. (b) Familia *Planariidae* con el pH. (c) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia *Planariidae* de acuerdo con la temperatura. (d) Diagrama de caja familia *Planariidae* con el pH.

ELABORADO POR: Velastegui V., 2019

En la Figura 23c, se exhibe el diagrama de caja de la familia *Planariidae* y la temperatura. Se aprecia que, a menores valores de temperatura, existe mayor presencia de individuos de esta familia. El valor de la mediana en caso de ausencia de esta familia es igual a 14.1°C. La mayoría de los datos se encuentran entre los valores de 12.5°C y 16.5°C. Los datos se encuentran más dispersos en comparación con los datos cuando la familia *Planariidae* se encuentra presente. En presencia de *Planariidae* la mediana de los valores de temperatura es 12.5°C y el rango intercuartil se encuentra entre 11.7°C y 13.3°C.

Los resultados mostrados en la Tabla 25 muestran la existencia de una relación inversa entre la temperatura del cuerpo acuático y la presencia de la familia *Planariidae* con un p-valor menor al 1%. De acuerdo a Jacobsen & Dangles (2017), estos organismos se presentan entre los 2100 y 2700 m.s.n.m donde la temperatura de los cuerpos de agua oscilan entre los 15 y 17°C (Jacobsen, 2000). De la experiencia, se puede corroborar esta información ya que las planarias prefieren aguas bien oxigenadas y con temperaturas de entre los 16 y 20°C; frecuentemente se presentan en ríos pequeños aferradas a los sotaventos de las rocas y donde existen velocidades de flujo entre bajas y moderadas. Jacobsen, Rostgaard, & Vásquez (2003b) sostienen que los macroinvertebrados son poiquiloterms, debido a ello la temperatura tiene incidencia sobre su capacidad de captar oxígeno, al limitar y/o reducir su metabolismo para adaptarse al medio. En el caso de la presente familia, Wieser (2012), describe la existencia de rangos óptimos de reproducción para este grupo los cuales se encuentran entre los 15 y los 18°C, razón por la cual la temperatura influenciaría sobre su tasa reproductiva y capacidad de supervivencia. Este efecto se extiende hasta el nivel de población de la familia *Planariidae* y se expresa como un cambio en los patrones conductuales.

Con relación al pH (Figura 23d) vemos mayor dispersión en el diagrama color naranja, el cual corresponde a la presencia de *Planariidae*. En este diagrama el valor de la mediana es igual a 8.36 y la mayoría de los datos de pH se encuentran entre 7.8 y 8.8. Para el caso de ausencia de esta familia de acuerdo con el pH, se observa que la mediana disminuye a un p-valorH igual a 7.50, además se estrecha el rango intercuartil, el cual va de 7.3 a 7.7.

Las planarias presentan una amplia distribución en sistemas acuáticos tropicales de altura ya que pueden vivir tanto en aguas corrientes como en estancadas, pegadas fuertemente al sustrato rocoso (Roldán y Ramírez, 2008). Los cambios de pH en el medio acuático pueden afectar directamente a nivel fisiológico a ciertas familias de macroinvertebrados, al

impulsar procesos de regulación de su entorno interno para armonizarlos con su entorno externo (Powers, 1930). Mientras que en otros casos puede afectar indirectamente su distribución y abundancia al influenciar sobre el crecimiento de organismos fotosintéticos que son la base alimenticia para macroinvertebrados pastoreadores. A diferencia de otros grupos taxa, la familia Planariidae es sensible a los cambios de pH en el medio acuático, debido a que la concentración de iones hidrógeno en el entorno puede afectar la tasa de regeneración de sus tejidos (Owen, Weiss, y Prince, 1939). La acidificación del agua desequilibra el balance de iones hidrógeno entre el interior y el exterior de los organismos, causando una reducción de la tasa de oxidación respiratoria e impidiendo que se oxiden adecuadamente los materiales celulares (Alper, 1941).

Matriz de confusión

En la Tabla 26 se presenta la matriz de confusión, en donde se puede observar la clasificación realizada por el modelo con los datos de entrenamiento y prueba.

Tabla 26 -. Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba.

	DATOS OBSERVADOS			
	ENTRENAMIENTO		PRUEBA	
	Ausencia	Presencia	Ausencia	Presencia
Ausencia	4	1	1	0
Presencia	6	52	1	6

ELABORADO POR: Velasteguí V., 2019

La precisión, en cuanto a los datos de prueba, es igual al 50 % y la sensibilidad igual a 100% para los datos de prueba. El F-score = 0.6, lo cual es bajo. Considérese además que la cantidad total de casos es igual a dos, al estar uno de esos casos bien clasificado y el otro no, se obtiene el valor tan bajo de F-score (Hilbe, 2009; Liu, 2007a)

Curva ROC

A pesar del antecedente, en cuanto a clasificación, el valor del área bajo la curva ROC, para este modelo con los datos de prueba, es de 1 (ver Figura 24). De acuerdo con Gironés

(2017), el modelo es un clasificador muy bueno; clasifica satisfactoriamente la presencia y ausencia de esta taxa.

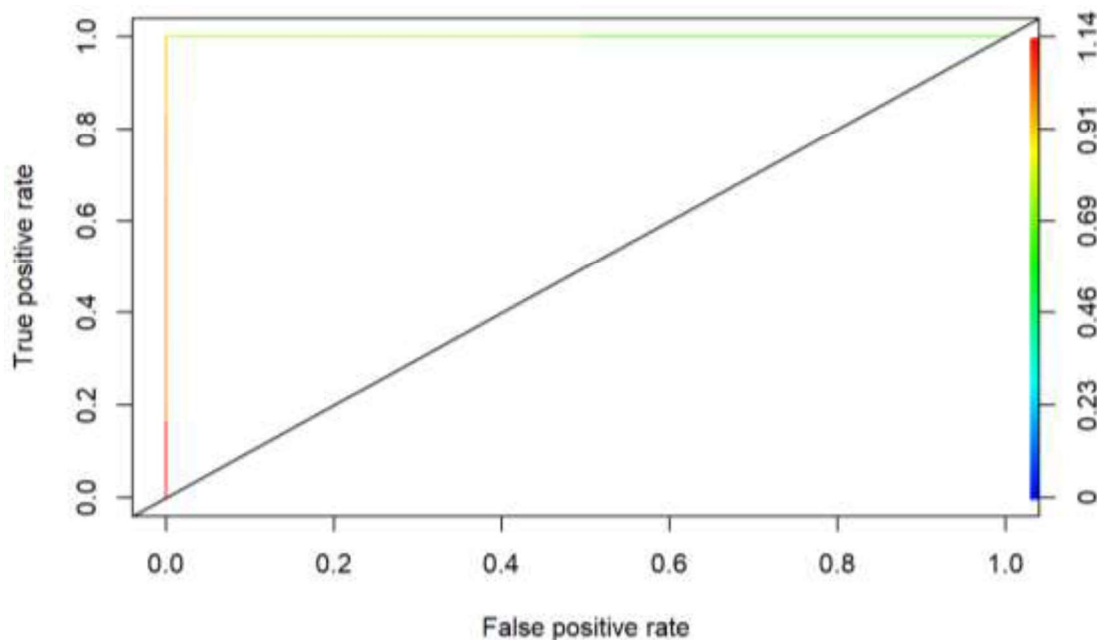


Figura 24 –. Curva ROC del modelo estimado con datos del conjunto de prueba
ELABORADO POR: Velastegui V., 2019

4.3.7. FAMILIAL *HYDROPTILIDAE* (SENSIBILIDAD A CONTAMINANTES: 7)

En 22 de un total de 86 puntos de muestreo fueron identificados individuos de la familia *Hydroptilidae*. La mayoría de estos puntos de muestreo se encuentran en el río Jatunhuayco. De acuerdo con la época climática, la época lluviosa presentó mayor número de individuos.

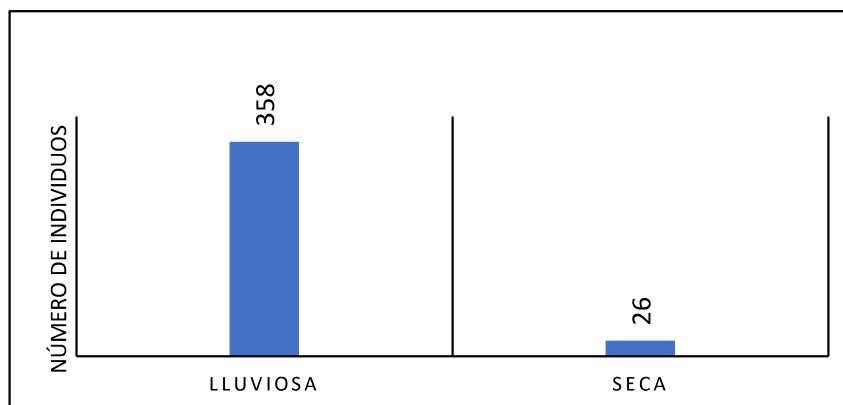


Figura 25 –. Número de individuos de la familia *Hydroptilidae* por época climática.
ELABORADO POR: Velastegui V., 2019

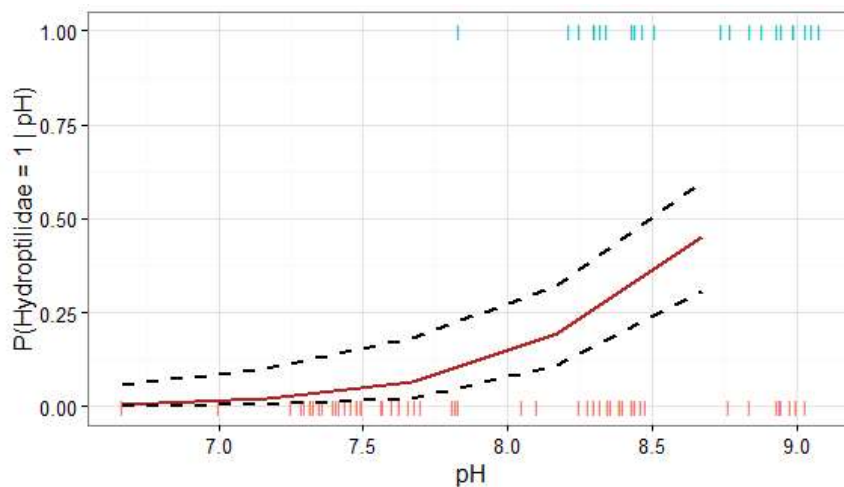
Los parámetros relacionados con la presencia de la familia *Hydroptilidae* son: pH y sólidos totales. El detalle de estas variables se presenta en la Tabla 27.

Tabla 27 -. Parámetros de regresión del modelo con mejor ajuste para la predicción de parámetros más correlacionados con la presencia de *Hydroptilidae*.

Variables predictoras	Parámetros de regresión	Coefficientes	p-valor
	α	-55.36072	0.001506
ln_pH	β_1	30.76341	0.000978
Sólidos Totales	β_2	-0.05056	0.009156
AIC		64.41	
Pseudo R²		0.34	

ELABORADO POR: Velastegui V., 2019

En la Figura 26, se presenta la tendencia de las variables con respecto a la existencia de la familia *Hydroptilidae*.



(a)

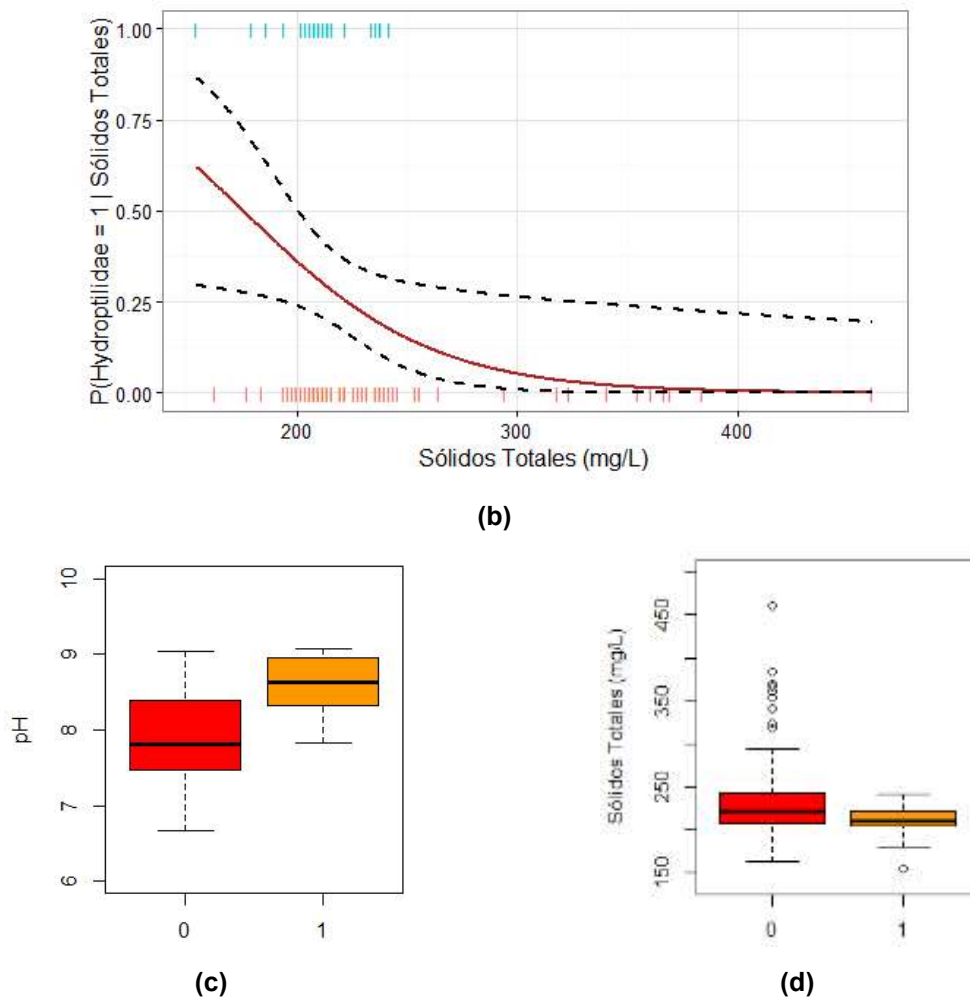


Figura 26 –. (a) Probabilidad de encontrar a la familia *Hydroptilidae* en relación con el pH. (b) Familia *Hydroptilidae* con sólidos totales. (c) Diagrama de caja de la presencia (1) y ausencia (0) de macroinvertebrados de la familia. (d) Diagrama de caja familia *Hydroptilidae* con concentración de sólidos totales.

ELABORADO POR: Velastegui V., 2019

La relación entre la ausencia de la familia y el pH se aprecia en la Figura 26c, el cual presenta un valor de la mediana igual a 7.82 y un rango intercuartil entre 7.5 y 8.4. Existe dispersión en los datos, en donde el valor mínimo es de 6.7 y el máximo de 9.03. No se aprecia valores atípicos en el diagrama de caja. Referente a la presencia de *Hydroptilidae*, existe mayores registros de individuos al presentar pH mayores. La mediana es igual a 8.6 y el rango intercuartil toma valores entre 8.32 y 8.9, es decir la mayoría de los datos se encuentran en este intervalo.

De acuerdo a estudios realizados por (Everaert et al., 2014) en el Ecuador, esta familia presentó una relación positiva con medios acuáticos con tendencia a la basicidad, lo cual es contrastado con los resultados hallados en el presente estudio. Esta familia, al igual que los Baetidos y Glossosomatidos pertenecen al grupo funcional alimenticio de los raspadores y colectores-recolectores (Tomanova et al., 2006b); su dieta depende del perifiton pegado al sustrato y de las partículas orgánicas depositadas en medio de las rocas. Estos organismos, al igual que la familia *Glossosomatidae*, son del orden Trichoptera y permanecen pegados al sustrato la mayor parte del tiempo ya que construyen estuches a partir de los granos finos depositados entre las rocas (Gillott, 2005). Debido a que su capacidad de movilizarse es menor a causa de su forma de vida y adaptaciones estructurales, su relación con el pH puede deberse a un efecto indirecto de este parámetro sobre el establecimiento de organismos fotosintéticos, ya que su dieta alimenticia está basada en el consumo de diatomeas y perifiton (J. H. Thorp y Rogers, 2014).

En la

Figura 26d, se presenta el diagrama correspondiente a la concentración de sólidos totales y la familia *Hydroptilidae*. En el primer diagrama (color rojo), la mediana es igual a 222 mg/L de sólidos totales. La mayor cantidad de datos de concentración de sólidos totales se encuentran 208mg/L y 243 mg/L. En el segundo caso (color naranja), la mediana es 211 mg/L y el rango intercuartil toma valores entre 203 mg/L y 225mg/L.

De acuerdo a lo observado en la Tabla 27, esta familia presenta una relación inversa con los sólidos totales del medio acuático en el que se desarrollan con un p-valor menor al 1%. Al ser organismos catalogados como raspadores y colectores-filtradores (Tomanova et al., 2006b), su dieta se basa en el consumo de partículas orgánicas finas depositadas en los espacios intersticiales de las rocas (Gómez y Salazar, 2017). Pese a que esta familia depende de la deposición de gránulos finos para construir sus estuches, una tasa excesiva de deposición de sólidos podría provocar su desaparecimiento al provocar una pérdida del área efectiva de establecimiento para sus casas portátiles.

Matriz de confusión

En la Tabla 28 se presenta la matriz de confusión, en donde se puede observar la clasificación realizada por el modelo con los datos de entrenamiento y prueba. La precisión, en este caso es igual a 100% y la sensibilidad igual a 100% para los datos de prueba. El F-score = 1 es satisfactoria con un valor superior al 90%. (Hilbe, 2009; Liu, 2007a)

Tabla 28 -. Matriz de confusión del conjunto de datos de entrenamiento y datos de prueba.

	DATOS OBSERVADOS			
	ENTRENAMIENTO		PRUEBA	
	Ausencia	Presencia	Ausencia	Presencia
Ausencia	8	0	6	0
Presencia	8	9	0	2

ELABORADO POR: Velasteguí V., 2019

Curva ROC

El valor del área bajo la curva ROC, para este modelo con los datos de prueba, es de 0.9285714 (ver Figura 27). De acuerdo Gironés (2017), el modelo es un clasificador muy bueno.

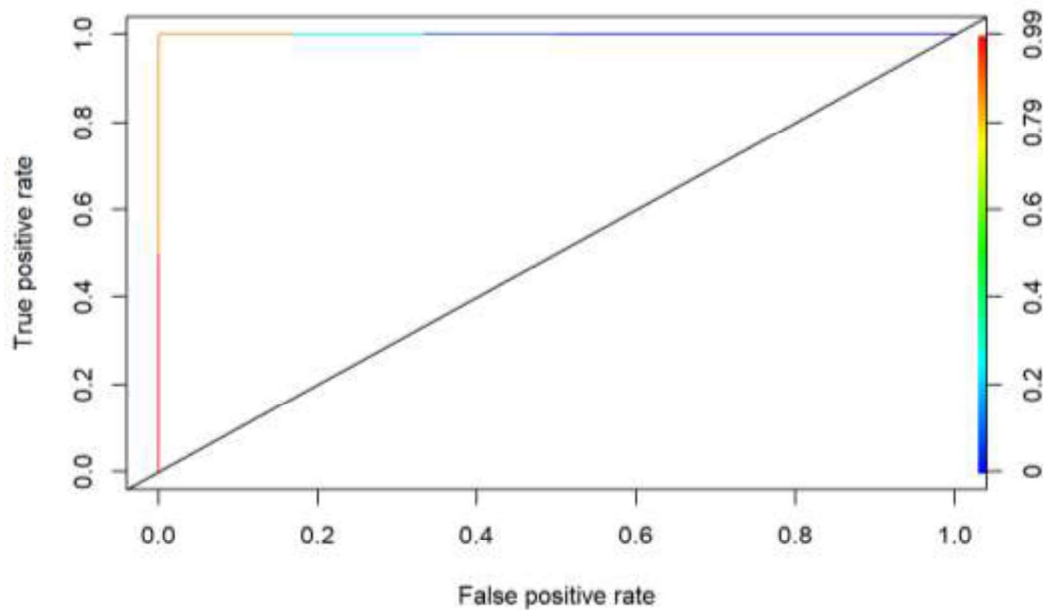


Figura 27 —. Curva ROC del modelo estimado con datos del conjunto de prueba
ELABORADO POR: Velastegui V., 2019

4.4. RANGOS DE PREFERENCIA DE PARÁMETROS DE CALIDAD

De los modelos obtenidos para cada una de las familias seleccionadas, aquel que presentó el menor valor de AIC corresponde a la familia *Hyaellidae* (AIC: 35.77). La matriz de confusión obtenida denota una clasificación adecuada, tanto para los datos de entrenamiento como para los datos de prueba (ver Tabla 21). El área bajo la curva ROC (AUC) y el F-score son igual a 1, que de acuerdo con Gironés (2017), el modelo es un predictor satisfactorio. Cabe mencionar, que para esta familia de macroinvertebrados la cantidad de individuos presentes en los 86 punto de muestreo, es superior al resto de familias, es decir que, de los 86 puntos de muestreo, el 83% de ellos se encontraron presentes individuos de la familia *Hyaellidae* (ver Tabla 29).

Tabla 29 -. Variación del valor de AIC de acuerdo con el porcentaje de presencia de macroinvertebrados en los puntos de muestreo.

Familia	AIC	Porcentaje de presencia
Hydrobiosidae	67.33	35%
Simuliidae	99.6	37%
Baetidae	-	-
Hyaellidae	32.67	83%
Glossosomatidae	77.72	52%
Planariidae	44.78	71%
Hydroptilidae	64.41	26%

ELABORADO POR: Velastegui V., 2019

De acuerdo a los resultados obtenidos en el análisis de regresión logística, las variables significativas presentes en los modelos son: temperatura, pH, conductividad, sólidos totales, sólidos disueltos, nitratos (NO_3^-), ion amonio (NH_4^+) y coliformes fecales.

La **temperatura** es la variable en común entre la familia *Hyaellidae* y *Planariidae*. La relación presente es negativa. El rango de temperatura en la cual existe la mayor presencia de estas familias se encuentra en la Tabla 30. A partir de los 14°C la probabilidad de encontrar individuos de esta familia decrece.

Tabla 30 –. Rango de temperatura donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	Temperatura °C		
Hyaellidae	12.1	-	14.6
Planariidae	10.8	-	13.3

ELABORADO POR: Velastegui V., 2019

Potencial de hidrógeno (pH) denota la acidez o alcalinidad de un cuerpo de agua o sustancia, el cual indica la cantidad de iones de hidronio presentes (Cueva, 2003).

Las familias *Glossosomatidae*, *Planariidae* y *Hydroptilidae* están relacionadas con el pH del agua, de acuerdo con estos resultados. En todos los casos, pH superior a 8, existe una mayor ocurrencia de estos individuos. El intervalo de valores de pH se presenta en la Tabla 31.

Tabla 31 –. Rango de valor pH donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	pH		
Glossosomatidae	7.8	-	8.9
Planariidae	7.8	-	9.08
Hydroptilidae	8.4	-	9.08

ELABORADO POR: Velastegui V., 2019

La **conductividad** es la medida de la capacidad de una sustancia para conducir la corriente eléctrica, lo que implica el movimiento de los iones de las impurezas presentes en el agua (Lapeña y Rigola, 1989). La **conductividad** es parte del modelo de predicción correspondiente a la familia *Simuliidae*. A partir del análisis, existe mayor ocurrencia a valores altos de conductividad. Los rangos de conductividad se detallan en la Tabla 32.

En trabajos realizados por Królak et al (2007) y McCreddie et al (1992), se estudia el comportamiento de especies de macroinvertebrados pertenecientes a la familia *Simuliidae*. En estas investigaciones se analiza la distribución, de estos macroinvertebrados, asociada a factores medio ambientales y calidad de agua. En sus resultados determinan, cada uno

con sus respectivas metodologías (mediante análisis de varianza (ANOVA) y análisis de componentes principales PCA, respectivamente), una relación directa y proporcional a la presencia de *Simuliidae* con la conductividad (Królak y Korycińska, 2008; McCreadie et al., 1995).

Tabla 32 –. Rango de valor conductividad donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	Conductividad ($\mu\text{s/cm}$)		
Simuliidae	202	-	535

ELABORADO POR: Velastegui V., 2019

Se encontró que tres familias se relacionan con **sólidos totales**. Estas familias son *Simuliidae* e *Hydroptilidae*. Los sólidos totales son indicadores de contaminación de actividades domésticas, agrícolas e industriales (Delgadillo, 2010).

En la familia *Hydroptilidae* la probabilidad de existencia disminuye a medida que aumenta la cantidad de sólidos totales en el agua, limitando su presencia a valores superiores a 250 mg/L. Del mismo modo, los modelos de la familia *Simuliidae* presentan el mismo comportamiento en cuanto a esta variable, disminuyendo su ocurrencia a mayores concentraciones de sólidos totales (ver Tabla 33).

Tabla 33 –. Rango de concentración de sólidos totales donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	Sólidos Totales (mg/L)		
Simuliidae	163	-	384
Hydroptilidae	154	-	242

ELABORADO POR: Velastegui V., 2019

Dentro de los sólidos totales existe una clasificación, la cual diferencia el material disuelto en el agua (sólidos disueltos) y el material suspendido (sólidos suspendidos). Los **sólidos suspendidos** presentes en el agua corresponden a la fracción de partículas orgánicas e inorgánicas no disueltas (Gómez, 2000; Ramos, Sepúlveda, y Villalobos, 2003a). La

cantidad excesiva de sólidos suspendidos puede ser peligrosa para las formas de vida acuática (Ramos, Sepúlveda, y Villalobos, 2003b).

En cuanto a la familia *Hydrobiosidae*, presentó una relación significativa con este parámetro. La posibilidad de encontrar a esta familia disminuye con el aumento de la cantidad de sólidos suspendidos.

En la Tabla 34 se detalla el rango de valores para el cual aumenta la posibilidad de encontrar a individuos de esta familia.

Tabla 34 – Rango de concentración de sólidos suspendidos donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	Sólidos suspendidos (mg/L)		
Hydrobiosidae	2	-	18

ELABORADO POR: Velastegui V., 2019

Los parámetros hidráulicos relacionados con la familia *Hydrobiosidae* son profundidad y ancho de cauce. Esto afecta al régimen hidráulico y al sustrato del cauce. El rango más probable se presenta en la Tabla 35.

Tabla 35 – Rango de valores donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	Profundidad (cm)	Ancho (m)
Hydrobiosidae	11.2 - 19.2	0.15 - 3.15

ELABORADO POR: Velastegui V., 2019

La concentración de **nitratos (NO_3^-)** e **ion amonio (NH_4^+)**, se relaciona con la familia *Hydrobiosidae* y *Simuliidae*, respectivamente. A mayores valores de nitratos disminuye la probabilidad de encontrar individuos de la familia *Hydrobiosidae*. En contraste, la relación existente entre la familia *Simuliidae* y el ión amonio es positiva.

Tabla 36 – Rango de valores donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	Ion amonio (NH ₄) ⁺ mg/L
Simuliidae	0.05 - 1.55

ELABORADO POR: Velastegui V., 2019

Los nitratos e ion amonio componen una fuente importante para los organismos, sin embargo, altas concentraciones pueden provocar implicaciones ecológicas desfavorables (Roldán y Restrepo, 2008), ocasionando deterioro en las comunidades acuáticas.

Tabla 37 – Rango de valores donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	Nitratos (NO ₃) ⁻ mg/L
Hydrobiosidae	0.1 - 0.4

ELABORADO POR: Velastegui V., 2019

Los **coliformes fecales** instituyen un indicador de contaminación por materia de origen fecal. La especie predominante es la *Escherichia coli*, que es una bacteria común en el tracto digestivo y puede causar enfermedades (World Health Organization Pan American Health, 1985).

En el análisis de regresión logística la familia Glossosomatidae presenta una relación negativa con valores altos de coliformes fecales.

Tabla 38 – Rango de valores de coliformes fecales donde aumenta la probabilidad de ocurrencia de macroinvertebrados.

Familia	Coliformes Fecales NMP/100mL
Glossosomatidae	36 - 90

ELABORADO POR: Velastegui V., 2019

Pese a las relaciones identificadas en el análisis de regresión logística, para la familia *Baetidae*, el bajo rendimiento de este imposibilitó considerar el resultado como adecuado. El número limitado de muestras y la alta capacidad de adaptarse a varias condiciones ecológicas complican la identificación de relaciones existentes de esta taxa con parámetros físicoquímicos del agua. Villamarín (2008), hace alusión al comportamiento de individuos de la familia *Baetidae*. Recalca, en su trabajo, la no tan clara tendencia que estos tienen en cuanto a los cambios de abundancia a lo largo del gradiente altitudinal medido. Además, señala que dependiendo de los géneros de la familia *Baetidae*, estos pueden ser más abundantes en zona bajas (*Baetodes*, *Comelobaetidius*, *Mayobaetis* y *Vaipes*), y en zonas altas (*Andesiops* y *Vernerius*) (C. P. Villamarín, 2008).

Así mismo, trabajos de Forero-Céspedes A. M. et al. (2016) y Flowers et al. (2010), recalcan la amplia tolerancia a la temperatura y a medios con alta carga de materia orgánica (nitratos) (Flowers y De la Rosa, 2010).

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

- En el río Jatunhuayco y en el río Ushimana, la familia más abundante corresponde a la familia Chironomidae, en las dos épocas climáticas estudiadas.
- En total, 8 parámetros de calidad de agua y 2 parámetros hidráulicos fueron identificados, en el análisis de regresión logística, como variables significativas de la ocurrencia de las familias analizadas. Estas variables son temperatura, pH, conductividad, sólidos totales, sólidos suspendidos, coliformes fecales, nitratos, ion amonio, profundidad y ancho de cauce.
- Fue posible encontrar una relación entre los macroinvertebrados acuáticos y los parámetros hidráulicos y fisicoquímicos de calidad del agua. De esto, se obtuvo 8 modelos de preferencia empleando el análisis de regresión logística. El modelo con el mejor rendimiento es el de la familia *Hyalellidae*, ya que presentó el menor valor de AIC = 32.67, AUC = igual a 1, el F-score = 1 y pseudo R² = 0.61.
- El modelo obtenido de la familia *Baetidae* es inexacto. Se puede apreciar en la matriz de confusión, la capacidad limitada que tiene el modelo para clasificar correctamente los datos de calentamiento como los datos de prueba. La presencia de esta familia es bastante dispersa en los dos ríos estudiados y no es clara la relación con los parámetros de calidad del agua. De acuerdo con Villamarín (2008), una de las causas para tal comportamiento es el nivel taxonómico analizado. La preferencia de ciertos géneros de la familia *Baetidae* a las variables fisicoquímicas e hidráulicas puede variar ampliamente (C. P. Villamarín, 2008).
- En el análisis de regresión logística, la familia *Hydrobiosidae*, fue la única que presentó parámetros hidráulicos como variables significativas. Los parámetros de calidad del agua conexos son: concentración de nitratos (p-valor = 0.019425) y sólidos suspendidos (p-valor = 6.29e-05). La familia *Hydrobiosidae* presentó una relación negativa con los dos parámetros de calidad del agua. Es mayor la probabilidad de hallar a la familia *Hydrobiosidae* cuando los nitratos (NO₃⁻) se encuentran en el rango de 0.1 mg/L a 0.4 mg/L y la concentración de sólidos suspendidos (mg/L) entre 2mg/L y 18 mg/L. Referente a los parámetros hidráulicos

relacionados, estos comprenden la profundidad y el ancho del cauce. Al aumentar, tanto de la profundidad como del ancho del cauce, la probabilidad de encontrar individuos de esta familia disminuye. El rango de profundidad más probable es entre 3.3 cm a 6.3 cm. Por lo que concierne al ancho de cauce, el rango varía entre 0.15 m y 3.14 m.

Las medidas de ajuste, matriz de confusión y curva ROC, denotan un rendimiento en la clasificación muy bueno. Sin embargo, la cantidad de datos empleados para evaluar el modelo es limitado, lo que hace que una sola clasificación errónea penalice en gran medida el diagnóstico de ajuste.

- La familia *Simuliidae*, presentó relaciones significativas con la conductividad (p-valor = 0.00721), ión amonio (p-valor = 0.03156) y la concentración de sólidos totales (p-valor = 0.00763). Presenta una tendencia directamente proporcional con la conductividad, ion amonio y la probabilidad de presencia e inversa para la concentración de sólidos totales. Con una clasificación muy buena de los datos de prueba, se pudo establecer rangos de probabilidad de presencia. Existe mayor probabilidad en el rango de conductividad de 202 $\mu\text{s}/\text{cm}$ a 535 $\mu\text{s}/\text{cm}$, ion amonio de entre 0.05 mg/L y 1.55 mg/L y concentración de sólidos totales 163 mg/L a 384 mg/L.
- La familia *Glossosomatidae* presentó relación con un indicador de contaminación de origen fecal (Coliformes Fecales) y pH. La relación con los Coliformes Fecales es negativa y el rango de mayor probabilidad de presencia es entre 36 NMP/100mL y 90 NMP/100mL. Este resultado es relevante para posteriores investigaciones, ya que la familia *Glossosomatidae* podría emplearse como indicador indirecto en una evaluación preliminar de la contaminación de patógenos en ríos. Al respecto, no existe información bibliográfica en la cual se profundice la relación existente con este parámetro. La relación con el pH es positiva. Mayor probabilidad de presencia en un intervalo de pH de 7.8 a 8.9.
- *Planariidae* es relacionada con parámetros como la temperatura y el pH. La probabilidad de su presencia aumenta con valores de temperatura entre 10.8°C y 13.3°C. Por lo que concierne al pH, la familia *Planariidae* aumenta con pH básico. El rango de pH en el cual se reportó la mayor probabilidad de encontrar individuos de esta familia va desde 7.8 a 9.1.
- La familia *Hydroptilidae* se relaciona con el pH y concentración de sólidos totales. La relación con el pH es directamente proporcional. El rango con mayor

probabilidad se encuentra entre 8.1 y 9.4. En relación con la concentración de sólidos totales, Hydroptilidae disminuye su presencia con valores superiores a 242 mg/L e inferiores a 154 mg/L.

- Los valores obtenidos de pseudo R^2 más bajos corresponden a los modelos de las familias: *Simuliidae* (pseudo $R^2 = 0.11$), *Baetidae* (pseudo $R^2 = 0.17$), *Glossosomatidae* (pseudo $R^2 = 0.34$) y *Hydroptiliida* (pseudo $R^2 = 0.34$) los cuales denotan la capacidad predictiva pobre de las variables en los modelos estimados. Los modelos de las familias: *Planariidae* (pseudo $R^2 = 0.60$), *Hydrobioscidae* (pseudo $R^2 = 0.61$), *Hyalellidae* (pseudo $R^2 = 0.61$) presentan valores de pseudo R^2 comunes encontrados en modelos ecológicos (Jerves-Cobo et al., 2017b)
- El nivel altitudinal es considerado uno de los parámetros más importantes (Jacobsen y Encalada, 1998a; Jacobsen et al., 2003a; Jacobsen, Schultz, y Encalada, 1997b), ya que revela la variabilidad de parámetros como temperatura, presión atmosférica, oxígeno disuelto y materia orgánica. Sin embargo, no se encontró una relación directa con las familias estudiadas, así como con otros parámetros encontrados en literatura como el oxígeno disuelto y la velocidad de flujo (Holguin-Gonzalez et al., 2013).

RECOMENDACIONES

- De los 8 modelos desarrollados, se recomienda trabajar con 3 de ellos, los cuales corresponden las familias: *Planariidae* (pseudo $R^2 = 0.60$), *Hydrobiosidae* (pseudo $R^2 = 0.61$), *Hyalellidae* (pseudo $R^2 = 0.61$). Para el resto de los modelos que no alcanzaron el rendimiento deseado, es conveniente trabajar con más información tanto de parámetros fisicoquímicos como hidráulicos.
- Procurar emplear, para un análisis de regresión logística con la función logit, aquellas taxas que cuenten con un mayor porcentaje de datos asociados a la presencia. Aquellos modelos con mayor porcentaje de presencia obtuvieron el mejor rendimiento en esta investigación, como el modelo obtenido de *Hyalellidae* (83% de presencia) y *Planariidae* (71% de presencia)
- Realizar una investigación más detallada con la familia *Glossosomatidae*, ya que presentó una relación considerable con coliformes fecales, indicador de contaminación fecal, ya que en la bibliografía es limitada al respecto.

- Buscar indicadores (taxones) a nivel de género, ya que por la adaptabilidad de ciertos individuos puede causar complicaciones al estimar variables relacionadas, como es el caso de la familia *Baetidae*.
- Ampliar el número de zonas de estudio, para obtener mayor variedad en el rango de valores de parámetros fisicoquímicos de calidad de agua para que su análisis sea aplicable a otros ríos del Ecuador. Considerar, además, la abundancia de individuos como una variable respuesta en próximos análisis de regresión, ya que para el presente trabajo bastó con que un individuo sea encontrado en el punto de muestreo para ser considerado como presente. Podría emplearse un modelo de regresión logística multinomial, de esa manera identificar el comportamiento de los macroinvertebrados ante los parámetros fisicoquímicos del agua.
- Indudablemente, la escasez de datos sobre los regímenes de flujo bentónico y la hidrodinámica de macroinvertebrados bentónicos se debe en parte a las dificultades involucradas en la medición de las velocidades del agua adyacentes al sustrato de la corriente (Davis, 1986). Por lo que, se recomienda, para futuros estudios, un muestreo especializado para determinar la hidrodinámica de los macroinvertebrados en un cuerpo de agua.

REFERENCIAS BIBLIOGRÁFICAS

- Agresti, A. (2003). *Categorical data analysis* (2.^a ed.). Florida, Estados Unidos: John Wiley & Sons.
- Aguirre, N., y Torres, J. (2013). Informe de la línea base del estado de degradación de los páramos de Jatunhuayco: escenario de degradación y caracterización de sistemas de referencia. *FONAG (Fondo para la Protección del Agua)*. Quito, Ecuador, 5,6.
- Alba-Tercedor, J. (1996). Macroinvertebrados acuáticos y calidad de las aguas de los ríos. *IV Simposio del agua en Andalucía (SIAGA)*. Almería (Vol. 2, pp 203–213).
- Alba-Tercedor, J., Jáimez-Cuéllar, P., Álvarez, M., Avilés, J., Bonada i Caparrós, N., Casas, J., Mellado, A., et al. (2002). Caracterización del estado ecológico de ríos mediterráneos ibéricos mediante el índice IBMWP (antes BMWP'). *Limnetica*, 2002, vol. 21, num. 3-4, p. 175-185.
- Alba-Tercedor, J., y Sánchez-Ortega, A. (1988). Un método rápido y simple para evaluar la calidad biológica de las aguas corrientes basado en el de Hellawell (1978). *Limnetica*, 4(5), 1–56.
- Alper, C. (1941). The Effect of the Hydrogen Ion Concentration in the Regeneration of *Planaria Dorocephala*. *Bios*, 12(4), 264-270. Recuperado a partir de <https://www.jstor.org/stable/4604612>
- Ampyx. (2001). *Studia dipterologica*. Ampyx.
- Armitage, P. D., Moss, D., Wright, J. F., y Furse, M. T. (1983). The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. *Water research*, 17(3), 333–347.
- Arroyo, C., y Encalada, A. C. (2009). Evaluación de la calidad de agua a través de macroinvertebrados bentónicos e índices biológicos en ríos tropicales en bosque de neblina montano. *ACI Avances en Ciencias e Ingenierías*, 1(1).
- Association American Public Health. (1992). *Métodos normalizados para el análisis de aguas potables y residuales*. Madrid, España.
- Avilez, T. M. (2017). *“Aplicación de modelos logísticos y geoestadísticos en organismos macroinvertebrados en la cuenca del río Guayas* (Bachelor’s Thesis, Ingeniero en Estadística, Guayaquil, 2017). ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.
- Bauernfeind, E., y Soldan, T. (2012). *The Mayflies of Europe (Ephemeroptera)*. BRILL.

- Bigler, C., Kulakowski, D., y Veblen, T. T. (2005). Multiple disturbance interactions and drought influence fire severity in Rocky Mountain subalpine forests. *Ecology*, 86(11), 3018–3029.
- Boulaaba, S., Zrelli, S., Boumaiza, M., y Rossaro, B. (2014). Relationships between physical and chemical factors and aquatic macroinvertebrates in perennial streams in the arid northern mountain basin El Batinah, Oman.
- Boumaïza, M., y Thomas, A. G. B. (1994). Premier inventaire faunistique et biogéographique des Baetidae (Insecta, Ephemeroptera) de Tunisie. *Bull. Soc. Sc. Nat., Tunisie*, 23, 26–30.
- Bournaud, M., Cellot, B., Richoux, P., y Berrahou, A. (1996). Macroinvertebrate community structure and environmental characteristics along a large river: congruity of patterns for identification to species or family. *Journal of the North American Benthological Society*, 15(2), 232–253.
- Bueno-Soria, J. (2004). Trichoptera (Insecta). *Artrópodos de Chamela/Chamela Arthropods*, 171.
- Bueno-Soria, J., y Flint Jr, O. S. (1978). Catálogo sistemático de los tricópteros de México (Insecta: Trichoptera), con algunos registros de Norte, Centro y Sudamérica. *Anales del Instituto de Biología, Universidad Nacional Autónoma de México, Serie Zoología* (Vol. 49, pp 189–218).
- Cabezas, V. S. (2015). *Evaluación de la calidad del agua del río Ushimana en el área de influencia de la parroquia Alangasí y propuesta del plan de gestión de las descargas contaminantes de la zona de estudio*. (The Bachelor's Thesis). Quito, Ecuador: Escuela Politécnica Nacional.
- Cabrera, E. (2009). El coeficiente de correlación de los rangos de Spearman Caracterización. *Rev haban cienc méd La Habana*, 8(2).
- Camargo, J., y Alonso, Á. (2006). Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: A global assessment. *Environment International*, 32(6), 831-849. doi:10.1016/j.envint.2006.05.002
- Campbell, N. A. (2001). *Biología: Conceptos y relaciones*. Ciudad de México, México: Pearson Educación.
- Canay Pazos, J. R. (2008). El uso de entornos virtuales de aprendizaje en las universidades presenciales: un análisis empírico sobre la experiencia del Campus Virtual de la USC.
- Collier, K. J., Croker, G. F., Hickey, C. W., Quinn, J. M., y Smith, B. S. (1995). Effects of hydraulic conditions and larval size on the microdistribution of Hydrobiosidae

- (Trichoptera) in two New Zealand rivers. *New Zealand Journal of Marine and Freshwater Research*, 29(4), 439–451.
- Coscarón, S., y Arias, C. (2007). *Neotropical Simuliidae (Diptera: Insecta)*. Pensoft Publishers.
- Coscarón, S., y Arias, C. L. C. (2007). *Aquatic Biodiversity in Latin America Volume 3: Neotropical Simuliidae (Diptera: Insecta)* (Vol. 3). Pensoft publishers.
- Cretaz, A. L. de la, y Barten, P. K. (2007). *Land Use Effects on Streamflow and Water Quality in the Northeastern United States*. CRC Press.
- Cueva, A. (2003). *Química 1*. Ediciones Umbral.
- Davis, J. A. (1986). Boundary layers, flow microenvironments and stream benthos. *Limnology in Australia* (pp 293–312). Springer.
- De Deckker, P., y Williams, W. D. (2012). *Limnology in Australia* (Vol. 61). Springer Science & Business Media.
- De Los Ríos, P., Morrone, J., y Rivera, R. (2012). Distributional patterns of the South American species of *Hyaella* (Amphipoda: Hyaellidae).
- De Pauw, N., Gabriels, W., y Goethals, P. L. (2006). *River monitoring and assessment methods based on macroinvertebrates* (1.^a ed.). Chichester, Inglaterra: John Wiley and Son, Ltd.: Chichester, UK.
- Delgadillo, O. (2010). *Depuración de aguas residuales por medio de humedales artificiales*. Nelson Antequera.
- Depiereux, E., Feytmans, E., y Micha, J.-C. (1983). Utilisation critique de l'analyse en composantes principales et du cluster analysis pour la description d'échantillons d'invertébrés benthiques en eau douce. *Oikos*, 81–94.
- Díaz-Martínez, J. A., y Granada-Torres, C. A. (2018). Efecto de las actividades antrópicas sobre las características fisicoquímicas y microbiológicas del río Bogotá a lo largo del municipio de Villapinzón, Colombia. *Revista de la Facultad de Medicina*, 66(1), 45–52.
- Domínguez, E. (2006). *Ephemeroptera de América del sur* (Vol. 2). Pensoft Publishers.
- Domínguez, E., y Fernández, H. R. (2009). *Macroinvertebrados bentónicos sudamericanos: sistemática y biología* (1.^a ed.). Tucumán, Argentina: Fundación Miguel Lillo Tucumán.
- Dworak, T., Gonzalez, C., Laaser, C., y Interwies, E. (2005). The need for new monitoring tools to implement the WFD. *Environmental Science & Policy*, 8(3), 301–306.
- Džeroski, S., Demšar, D., y Grbović, J. (2000). Predicting chemical parameters of river water quality from bioindicator data. *Applied Intelligence*, 13(1), 7–17.

- Encalada, A. (1997). *Densidad y abundancia de macroinvertebrados en relación a factores físico químicos y fuentes de alimento en dos tipos de ríos de páramo del Ecuador* (Masters thesis. Pontificia Universidad Católica del Ecuador, Quito.). Quito.
- Everaert, G., Neve, J. D., Boets, P., Dominguez-Granda, L., Mereta, S. T., Ambelu, A., Hoang, T. H., et al. (2014). Comparison of the Abiotic Preferences of Macroinvertebrates in Tropical River Basins. *PLOS ONE*, 9(10), e108898. doi:10.1371/journal.pone.0108898
- Eyre, M. D., Pilkington, J. G., McBlane, R. P., y Rushton, S. P. (2005). Macroinvertebrate species and assemblages in the headwater streams of the River Tyne, northern England in relation to land cover and other environmental variables. *Hydrobiologia*, 544(1), 229–240.
- Ferrier, S., Watson, G., Pearce, J., y Drielsma, M. (2002). Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodiversity & Conservation*, 11(12), 2275–2307.
- Flowers, R. W., y De la Rosa, C. (2010). Capítulo 4: Ephemeroptera. *Revista de Biología Tropical*, 58, 63–93.
- Forio, M. A. E., Van Echelpoel, W., Dominguez-Granda, L., Mereta, S. T., Ambelu, A., Hoang, T. H., Boets, P., et al. (2016). Analysing the effects of water quality on the occurrence of freshwater macroinvertebrate taxa among tropical river basins from different continents. *AI Communications*, 29(6), 665–685.
- García, M. A., y Pina, J. A. L. (1997). *IV Simposio de Metodología de las Ciencias del Comportamiento*. EDITUM.
- Gari, N., y Corigliano, M. del C. (2006). El efecto de arroyos tributarios de menor orden sobre el perifiton de un río de llanura. *Biología Acuática*.
- Ghetti, P. F. (1986). I Macroinvertebrati nell'Analisi di qualità dei corsi d'aqua. Stazione Sperimentale Agraria Foréstale. *Parma, Università di Parma*, 57(1), 105.
- Gillott, C. (2005). *Entomology*. Springer Science & Business Media.
- Gironés, J. (2017). *Minería de datos: modelos y algoritmos*, Editorial UOC, 2017. *ProQuest Ebook Central*.
- Gómez. (2000). *Saneamiento ambiental*. EUNED.
- Gómez, S. J., y Salazar, C. (2017). *Abundancia y biomasa de macroinvertebrados en relación de cuatro tipos de sustratos y de variables fisicoquímicas en La Laguna La Virginia, Páramo de Sumapaz* (Master's Thesis). Universidad de Bogotá Jorge Tadeo Lozano.

- Gonzalez, E. R., y Watling, L. (2002). Redescription of *Hyaella azteca* from its type locality, Vera Cruz, Mexico (Amphipoda: Hyaellidae). *Journal of Crustacean Biology*, 22(1), 173–183.
- González, E., y Roldán, G. (2019). Monitoreo biológico de la calidad del agua en las Américas. *Calidad del Agua en las Américas*, 39.
- Grbovic, J. (1994). *Applicability of various procedures for the assessment of quality of Torrential streams* (PhD Thesis). Eslovenia: PhD Thesis, Biotechnical Faculty, University of Ljubljana.
- Grbovic, J., y Džeroski, S. (1995). Knowledge discovery in a water quality database. *Proc 1st Intl Conf Knowl. Discov. Data Min. KDD95 AAAI Press Menlo Park CA*, 16(1), 81-86.
- Guillén, M. F. (1992). *Análisis de regresión múltiple* (2.^a ed., Vol. 4). Madrid, España: Centro de Investigaciones Sociológicas Madrid.
- Hamada, N., y McCreddie, J. W. (1999). Environmental factors associated with the distribution of *Simulium perflavum* (Diptera: Simuliidae) among streams in Brazilian Amazonia. *Hydrobiologia*, 397, 71–78.
- Hanson, P., Springer, M., y Ramirez, A. (2010). Capítulo 1: Introducción a los grupos de macroinvertebrados acuáticos. *Revista de Biología Tropical*, 58, 3–37.
- Harrell, J. F., Lee, K. L., Matchar, D. B., y Reichert, T. A. (1985). Regression models for prognostic prediction: advantages, problems, and suggested solutions. *Cancer treatment reports*, 69(10), 1071–1077.
- Heckman, C. W. (2012). *Encyclopedia of South American aquatic insects: Ephemeroptera: illustrated keys to known families, genera, and species in South America*. Springer Science & Business Media.
- Hilbe, J. M. (2009). *Logistic regression models*. Chapman and Hall/CRC.
- Holguin-Gonzalez, J. E., Boets, P., Alvarado, A., Cisneros, F., Carrasco, M. C., Wyseure, G., Nopens, I., et al. (2013). Integrating hydraulic, physicochemical and ecological models to assess the effectiveness of water quality management strategies for the River Cuenca in Ecuador. *Ecological modelling*, 254, 1–14.
- Ihaka, R., y Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, 5(3), 299–314.
- Jacobsen, D. (2000). Gill size of trichopteran larvae and oxygen supply in streams along a 4000-m gradient of altitude. *Journal of the North American Benthological Society*, 19(2), 329–343.

- Jacobsen, D. (2003). Altitudinal changes in diversity of macroinvertebrates from small streams in the Ecuadorian Andes. *Archiv für Hydrobiologie*, 158(2), 145–167.
- Jacobsen, D. (2008). Low oxygen pressure as a driving factor for the altitudinal decline in taxon richness of stream macroinvertebrates. *Oecologia*, 154(4), 795–807.
- Jacobsen, D., y Dangles, O. (2017). *Ecology of High Altitude Waters*. Oxford University Press.
- Jacobsen, D., y Encalada, A. (1998). The macroinvertebrate fauna of Ecuadorian highland streams in the wet and dry season. *Archiv für Hydrobiologie*, 142, 53–70.
- Jacobsen, D., Rostgaard, S., y Vásquez, J. J. (2003). Are macroinvertebrates in high altitude streams affected by oxygen deficiency? *Freshwater Biology*, 48(11), 2025–2032. doi:10.1046/j.1365-2427.2003.01140.x
- Jacobsen, D., Schultz, R., y Encalada, A. (1997). Structure and diversity of stream invertebrate assemblages: the influence of temperature with altitude and latitude. *Freshwater Biology*, 38(2), 247–261.
- Jerves-Cobo, R., Everaert, G., Iñiguez-Vela, X., Córdova-Vela, G., Díaz-Granda, C., Cisneros, F., Nopens, I., et al. (2017). A methodology to model environmental preferences of EPT taxa in the Machangara River basin (Ecuador). *Water*, 9(3), 195.
- Johnson, R. K., Wiederholm, T., y Rosenberg, D. M. (1993). Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates. *Freshwater biomonitoring and benthic macroinvertebrates*, 40–158.
- Królak, E., y Korycińska, M. (2008). Taxonomic Composition of Macroinvertebrates in the Liwiec River and its Tributaries (Central and Eastern Poland) on the Basis of Chosen Physical and Chemical Parameters of Water and Season. *Polish Journal of Environmental Studies*, 17(1).
- de la Lanza-Espino, G., y Pulido, S. H. (2000). *Organismos indicadores de la calidad del agua y de la contaminación (bioindicadores)*. Plaza y Valdés.
- Lapeña, M. R., y Rigola, M. (1989). *Tratamiento de aguas industriales: aguas de proceso y residuales* (Vol. 27). Marcombo.
- Lepom, P., Brown, B., Hanke, G., Loos, R., Quevauviller, P., y Wollgast, J. (2009). Needs for reliable analytical methods for monitoring chemical pollutants in surface water under the European Water Framework Directive. *Journal of Chromatography A*, 1216(3), 302–315.
- Lewis, R. A. (1998). *Lewis' Dictionary of Toxicology*. CRC Press.

- Lindegaard, C. (1995). Chironomidae (Diptera) of European cold springs and factors influencing their distribution. *Journal of the Kansas Entomological Society*, 108–131.
- Liu, B. (2007). *Web data mining: exploring hyperlinks, contents, and usage data* (2.^a ed.). Chicago, Estados Unidos: Springer Science & Business Media.
- López-Vera, F. (2002). Tres cuestiones sobre la sostenibilidad del aprovechamiento de los recursos hídricos. *Encuentros multidisciplinares*, 10, 2.
- Madrid, Y., y Zayas, Z. P. (2007). Water sampling: Traditional methods and new approaches in water sampling strategy. *TrAC Trends in Analytical Chemistry*, 26(4), 293–299.
- Madsen, P. B., Morabowen, A., Andino, P., Espinosa, R., CAUVY-FRAUNIÉ, S., Dangles, O., y Jacobsen, D. (2015). Altitudinal distribution limits of aquatic macroinvertebrates: an experimental test in a tropical alpine stream. *Ecological entomology*, 40(5), 629–638.
- Malhotra, N. K. (2004). *Investigación de mercados: un enfoque aplicado*. Pearson educación.
- Maloney, K. O., Weller, D. E., Russell, M. J., y Hothorn, T. (2009). Classifying the biological condition of small streams: an example using benthic macroinvertebrates. *Journal of the North American Benthological Society*, 28(4), 869–884.
- Manel, S., Dias, J.-M., y Ormerod, S. J. (1999). Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological modelling*, 120(2-3), 337–347.
- Manel, S., Williams, H. C., y Ormerod, S. J. (2001). Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of applied Ecology*, 38(5), 921–931.
- Margalef, R. (1986). *Ecología*. Omega.
- McCreadie, J. W., Adler, P. H., y Colbo, M. H. (1995). Community structure of larval black flies (Diptera: Simuliidae) from the Avalon Peninsula, Newfoundland. *Annals of the Entomological Society of America*, 88(1), 51–57.
- McCreadie, J. W., y Colbo, M. H. (1991). Spatial distribution patterns of larval cytotypes of the *Simulium venustum/verecundum* complex (Diptera: Simuliidae) on the Avalon Peninsula, Newfoundland: factors associated with occurrence. *Canadian Journal of Zoology*, 69(10), 2651–2659.
- McCullagh, P., y Nelder, J. A. (2018). *Generalized linear models*. Routledge.
- McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

- Merritt, R. W., y Cummins, K. W. (1996). *An introduction to the aquatic insects of North America*. Kendall Hunt.
- Metcalfe, J. L. (1989). Biological water quality assessment of running waters based on macroinvertebrate communities: history and present status in Europe. *Environmental pollution*, 60(1-2), 101–139.
- Mite, R. B., Ochoa, L. S., Osorio, B. G., Suatunce, P., Ocampo, E. D., y Arevalo, L. C. (2016). Calidad del agua destinada al consumo humano en un cantón de Ecuador/Quality of water intended for human consumption in a canton of Ecuador. *Ciencia Unemi*, 9(20), 109–117.
- Mosquera, H. (2018). Evaluación de la distribución de macroinvertebrados bentónicos en un río altoandino. Caso de estudio: Microcuenca del río Blanco, 192.
- Moss, B. (2013). *Ecology of Fresh Waters: A View for the Twenty-First Century*. John Wiley & Sons.
- Munro, E. (2010). *El río y la forma. Introducción a la geomorfología fluvial*. RIL Editores.
- Muñoz, I., y Prat, N. (1996). Effects of water abstraction and pollution on macroinvertebrate community in a Mediterranean river. *Limnetica*, 12, 9–16.
- Naciones Unidas. (2019). *Informe de los Objetivos de Desarrollo Sostenible 2019*. Nueva York. Recuperado a partir de https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019_Spanish.pdf
- Ormerod, S. J. (1987). The influences of habitat and seasonal sampling regimes on the ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye, Wales. *Hydrobiologia*, 150(2), 143–151.
- Ormerod, S. J., y Edwards, R. W. (1987). The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. *Freshwater Biology*, 17(3), 533–546.
- Ortega, M., y Cayuela, A. (2002). Regresión logística no condicionada y tamaño de muestra: una revisión bibliográfica. *Revista Española de Salud Pública*, 76(1), 85–93.
- Ortiz, L. L. (2005). La bioindicación de la calidad del agua: Importancia de los macroinvertebrados en la cuenca alta del Río Juan Amarillo, Cerros Orientales de Bogotá. *Umbral científico*, (7), 5–11.
- Owen, S. E., Weiss, H. A., y Prince, L. H. (1939). Carcinogens and planarian tissue regeneration. *The American Journal of Cancer*, 35(3), 424–426.

- Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., y Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*, 49(12), 1373–1379.
- Pérez, G. R. (1999). Los macroinvertebrados y su valor como indicadores de la calidad del agua. *Academia Colombiana de Ciencia*, 23(88), 375–387.
- Powers, E. B. (1930). The Relation between pH and Aquatic Animals. *The American Naturalist*, 64(693), 342-366. doi:10.1086/280321
- Prat, N., Rieradevall, M., Acosta, R., y Villamarín, C. (2011). Guía para el reconocimiento de las larvas de Chironomidae (díptera) de los ríos altoandinos de Ecuador y Perú. *Grupo de Investigación FEM Departamento de Ecología, Universidad de Barcelona, España*, 78(5), 36-42.
- Prat, N., Ríos, B., Acosta, R., y Rieradevall, M. (2009). *Los macroinvertebrados como indicadores de calidad de las aguas* (1.ª ed.). Tucumán, Argentina: Fundación Miguel Lillo, 2009.
- Ramos, R., Sepúlveda, R., y Villalobos, F. (2003). El agua en el medio ambiente: muestreo y análisis. *Mexicali, Mexico*.
- Ríos-Touma, B., Encalada, A. C., y Prat Fornells, N. (2011). Macroinvertebrate assemblages of an Andean high-altitude tropical stream: the importance of season and flow. *International review of hydrobiology*, 96(6), 667–685.
- Roldán, G. (1988). Guía para el estudio de los macroinvertebrados acuáticos del Departamento de Antioquia. *Universidad de Antioquia, Fondo FEN, Medellín*.
- Roldán, G. (1999). Los macroinvertebrados y su valor como indicadores de la calidad del agua. *Academia Colombiana de Ciencia*, 23(88), 375–387.
- Roldán, G. (2003). *Bioindicación de la calidad del agua en Colombia: Propuesta para el uso del método BMWP Col* (1.ª ed.). Medellín, Colombia: Universidad de Antioquia.
- Roldán, G. (2016). Los macroinvertebrados como bioindicadores de la calidad del agua: cuatro décadas de desarrollo en Colombia y Latinoamérica. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, 40(155), 254. doi:10.18257/raccefyn.335
- Roldán, y Ramírez, J. (2008). Fundamentos de limnología neotropical. 2da. *Universidad de Antioquia*.
- Rosemond, A. D., Reice, S. R., Elwood, J. W., y Mulholland, P. J. (1992). The effects of stream acidity on benthic invertebrate communities in the south-eastern United States. *Freshwater Biology*, 27(2), 193–209.

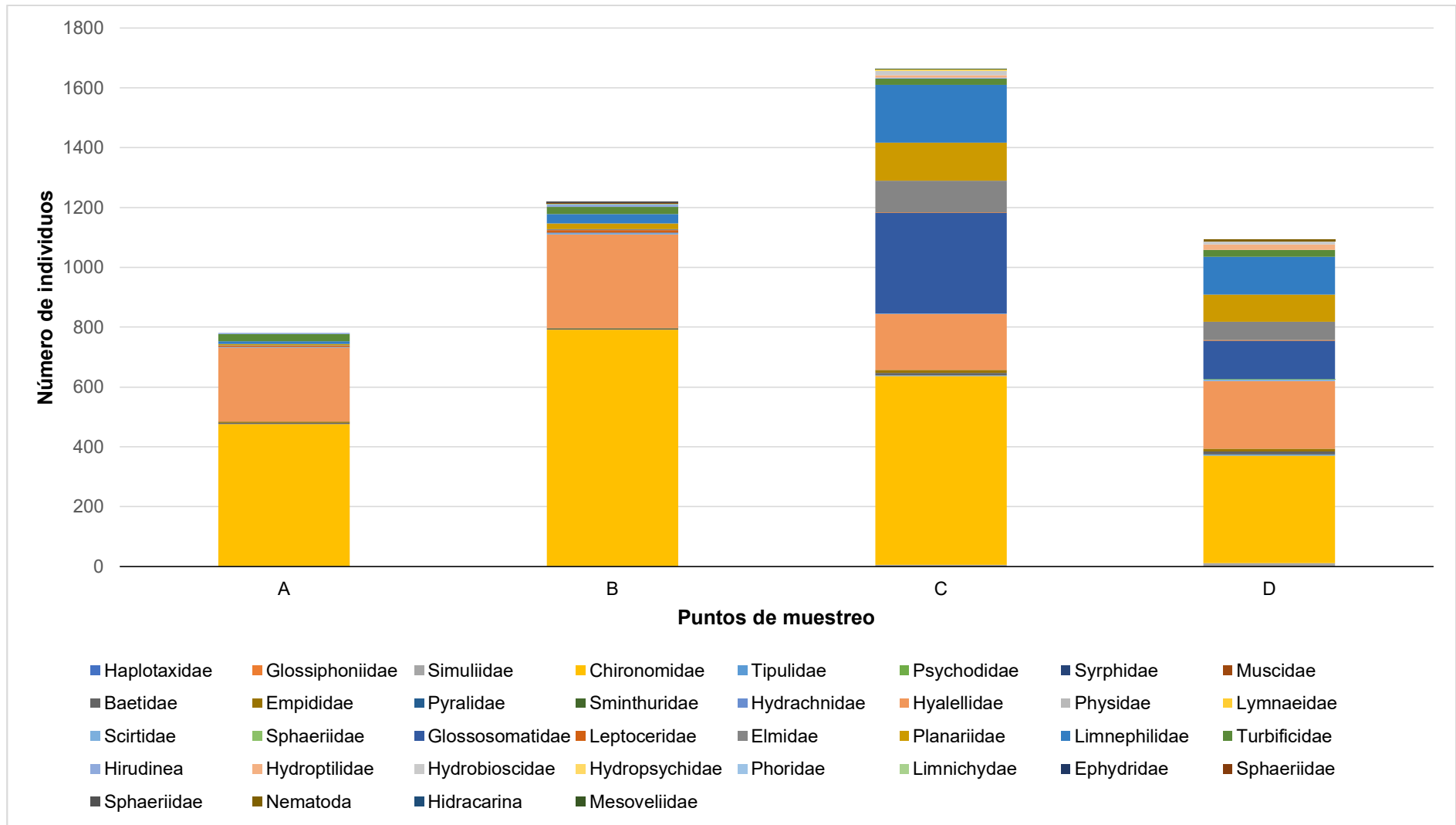
- Rosero López, D. (2011). *Impacto de las capacitaciones de agua en la integridad ecológica y posibles soluciones para el manejo del flujo de agua en ríos de páramo* (Master's thesis, Maestra en Ecología con mención en Manejo de Recursos Naturales, Quito, 2011). Quito: USFQ, 2011.
- Statzner, B. (1981). The relation between "hydraulic stress" and microdistribution of benthic macroinvertebrates in a lowland running water system, the Schierenseebrooks (North Germany). *Archiv fur Hydrobiologie*, 91, 192–218.
- Steyerberg, E. W., Harrell Jr, F. E., Borsboom, G. J., Eijkemans, M. J. C., Vergouwe, Y., y Habbema, J. D. F. (2001). Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *Journal of clinical epidemiology*, 54(8), 774–781.
- Thorp, J. H., y Rogers, D. C. (2014). *Thorp and Covich's Freshwater Invertebrates: Ecology and General Biology*. Elsevier.
- Thorp, J., y Rogers, C. (2014). *Thorp and Covich's Freshwater Invertebrates: Ecology and General Biology*. Elsevier.
- Tokeshi, M. (2009). *Species coexistence: ecological and evolutionary perspectives*. John Wiley & Sons.
- Tomanova, S., Goitia, E., y Helešic, J. (2006). Trophic Levels and Functional Feeding Groups of Macroinvertebrates in Neotropical Streams. *Hydrobiologia*, 556(1), 251-264. doi:10.1007/s10750-005-1255-5
- Torres, A. (2010). Curvas ROC para Datos de Supervivencia. *Aplicación a Datos Biomédicos*.
- Vayssières, M. P., Plant, R. E., y Allen-Diaz, B. H. (2000). Classification trees: An alternative non-parametric approach for predicting species distributions. *Journal of vegetation science*, 11(5), 679–694.
- Velasco, M. S. (1996). La regresión logística. Una aplicación, a la demanda de estudios universitarios. *Estadística Española*, 141, 193–217.
- Vilanova, E., y Sogorb, M. A. (2004). *Técnicas analíticas de contaminantes químicos: Aplicaciones toxicológicas, medioambientales y alimentarias* (1.^a ed.). Madrid, España: Ediciones Díaz de Santos.
- Villamarín, C. (2008). Estructura y composición de las comunidades de macroinvertebrados acuáticos en ríos altoandinos del Ecuador y Perú. Diseño de un sistema de medida de la calidad del agua con índices multimétricos, 207.
- Walley, W. J., Boyd, M., y Hawkes, H. (1992). An expert system for the biological monitoring of river pollution. *Lakes and Reservoirs: Research and Management*, 17(2), 73-80.

- Wieser, W. (2012). *Effects of Temperature on Ectothermic Organisms: Ecological Implications and Mechanisms of Compensation*. Springer Science & Business Media.
- Wooldridge, J. M. (2006). *Introducción a la econometría: un enfoque moderno*. Editorial Paraninfo.
- World Health Organization Pan American Health. (1985). *Guías para la calidad del agua potable* (Vol. 3). Pan American Health Org.
- Wright, J. (1995). Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. *Australian Journal of Ecology*, 20(1), 181–197.
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., y Smith, G. M. (2009a). *Mixed effects models and extensions in ecology with R*. Gail M, Krickeberg K, Samet JM, Tsiatis A, Wong W, editors (1.ª ed.). California, Estados Unidos: Springer.
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., y Smith, G. M. (2009b). Statistics for biology and health. Gail M, Krickeberg K, Samet JM, Tsiatis a, Wong W (eds) R. Springer, New York, NY.

ANEXOS

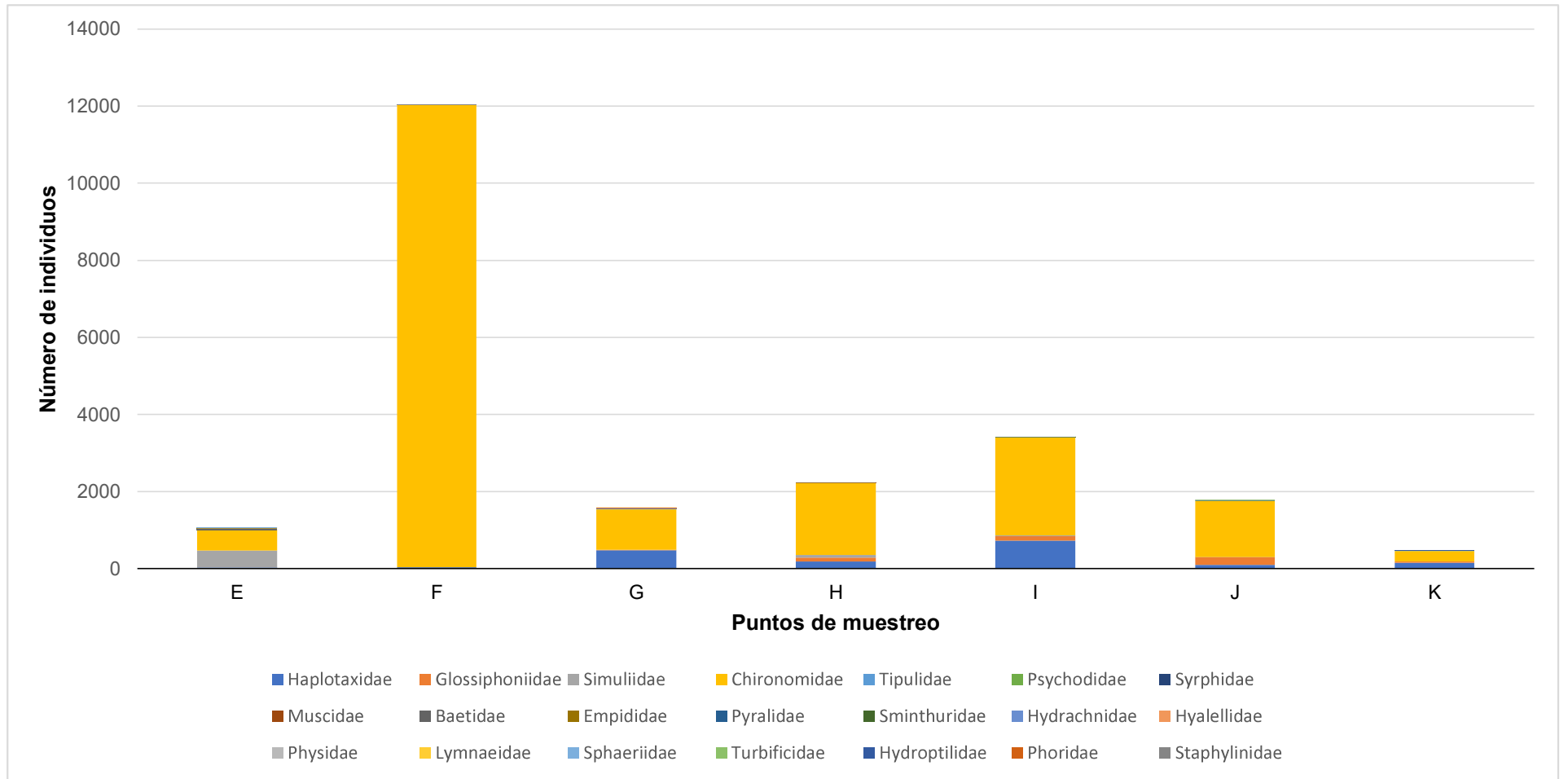
ANEXO 1:

Composición y distribución de macroinvertebrados en el río Jatunhuayco



ANEXO 2:

Composición y distribución de macroinvertebrados en el río Ushimana



ANEXO 3:

Coeficiente de correlación de Spearman de las variables explicativas.

RIABLES	Temp.	Conduc.	TDS	pH	OD	(PO ₄) ⁻³	(NO ₃) ⁻	(NH ₄) ⁺	Colif. F.	Resist.	NaCl	DBO	DQO	S. Total	S. Disuel.	S. Suspen.	Vel.	Prof.	Ancho	Altitud
Temperatura	1.0																			
Conductividad	0.5	1.0																		
TDS	0.4	1.0	1.0																	
pH	0.0	0.3	0.3	1.0																
OD	-0.3	-0.6	-0.6	-0.1	1.0															
(PO ₄) ⁻³	0.6	0.5	0.5	-0.3	-0.2	1.0														
(NO ₃) ⁻	0.4	0.3	0.3	-0.2	-0.4	0.4	1.0													
(NH ₄) ⁺	0.3	0.0	0.0	-0.7	-0.1	0.5	0.5	1.0												
Coliformes F	0.3	0.3	0.3	-0.3	-0.2	0.5	0.3	0.5	1.0											
Resistividad	-0.6	-0.9	-0.9	-0.3	0.6	-0.5	-0.3	0.0	-0.4	1.0										
NaCl	0.5	0.9	0.9	0.3	-0.6	0.5	0.3	0.1	0.3	-0.9	1.0									
DBO	0.5	0.6	0.6	0.2	-0.2	0.6	0.1	0.0	0.3	-0.7	0.6	1.0								
DQO	0.2	0.3	0.3	0.3	0.2	0.3	-0.2	-0.4	0.2	-0.4	0.3	0.7	1.0							
S. Totales	0.4	0.3	0.3	-0.1	-0.1	0.5	0.2	0.3	0.3	-0.4	0.4	0.2	0.2	1.0						
S. Disueltos	0.3	0.3	0.2	-0.3	-0.1	0.4	0.2	0.3	0.5	-0.3	0.3	0.4	0.2	0.5	1.0					
S. Suspendidos	0.4	0.2	0.2	-0.3	0.0	0.5	0.0	0.2	0.2	-0.3	0.2	0.3	0.4	0.1	0.1	1.0				
Velocidad	-0.1	-0.2	-0.2	-0.2	0.5	0.1	-0.3	0.0	-0.1	0.2	-0.2	-0.1	0.2	-0.1	-0.1	0.2	1.0			
Profundidad	0.2	-0.2	-0.2	-0.5	0.3	0.2	-0.1	0.4	0.1	0.2	-0.1	0.1	0.1	0.0	0.0	0.4	0.3	1.0		
Ancho	0.2	0.0	0.0	-0.3	-0.1	0.2	0.3	0.4	0.1	0.1	0.0	0.0	-0.2	-0.1	0.0	0.3	-0.5	0.2	1.0	
Altitud	-0.2	-0.6	-0.6	-0.3	0.8	-0.2	-0.4	0.0	-0.3	0.6	-0.6	-0.1	0.2	-0.3	-0.3	0.3	0.5	0.4	0.1	1.0

ANEXO 4:

Modelo de regresión logística para las familias de macroinvertebrados

Variable explicativa	Hydrobiosidae		Simuliidae		Hyaellidae		Glossosomatidae		Planariidae		Hydroptilidae	
	Coeficiente	p-valor	Coeficiente	p-valor	Coeficiente	p-valor	Coeficiente	p-valor	Coeficiente	p-valor	Coeficiente	p-valor
α	8.2919	0.000193	2.09	0.19124	72.254	0.000877	-7.528	0.074543	-8.75209	0.21731	-55.36072	0.001506
Temperatura					-27.005	0.001111			-2.1179	0.00203		
pH							1.724	0.000798	4.5678	0.0015	30.76341	0.000978
Conductividad			0.019075	0.00721								
Sólidos Totales			-0.033715	0.00763							-0.05056	0.009156
Sólidos Susp.	-2.4985	6.29E-05										
Coliformes fecales							-1.3697	0.008149				
(NO ₃) ⁻	-6.8325	0.019425										
(NH ₄) ⁺			0.588477	0.03156								
AIC	67.33		99.6		32.67		77.6		44.78		64.41	

ANEXO 5:

Modelo desarrollado para la familia *Hydrobiosidae* (Script)

HYDROBIOSCIDAE
 Viviana Velasteguí
 01/7/2019

LIBRERIAS

```
library(tidyverse)
library(caret)
library(readxl)
library(MASS)
library(broom)
library(ggplot2)
library(modelr)
library(ggfortify)
library(ROCR)
library(psc1)
theme_set(theme_classic())
```

ANÁLISIS DE DATOS

```
## Classes 'tbl_df', 'tbl' and 'data.frame': 86 obs. of 17 variables
:
## $ Temperature : num 14 13.3 12.6 12.3 12.2 12.3 13.4 12.8 12
.7 11.6 ...
## $ Conductivity : num 202 203 199 203 190 ...
## $ pH : num 7.57 7.6 7.68 7.4 7.42 7.48 7.44 7.5 7.4
8 7.32 ...
## $ OxigenoD : num 11.3 11.5 11.7 10.3 10.2 ...
## $ PO4 : num 1.006 0.526 0.8 0.85 0.62 ...
## $ NO3 : num 0.6 0.2 0.3 0.1 0.3 0.1 0.2 0.4 0.4 0.6
...
## $ NH4 : num 0.17 0.18 0.21 0.22 0.2 0.21 0.24 0.21 0
.19 0.22 ...
## $ Coliformes_F : num 150 150 150 150 150 150 150 150 150 91 .
..
## $ DBO : num 4.1 4.1 4.1 4.1 4.1 4.1 4.1 4.1 4.1 1.21
...
## $ DQO : num 50 50 50 50 50 50 50 50 50 38 ...
## $ Solidos_Totales : num 214 196 204 222 208 ...
## $ Solidos_Disueltos : num 192 210 216 198 190 ...
## $ Solidos_Suspendidos: num 18 20 20 26 28 ...
## $ Velocidad : num 0.305 0.38 0.539 0.609 0.897 ...
## $ Profundidad : num 24 25 39 20 31 27 20 18 26 17 ...
## $ Ancho : num 5.36 4.13 1.87 3.22 1.41 3.38 4.09 2.75
2.94 8.33 ...
## $ Hydrobioscidae : num 0 0 0 0 0 0 0 0 0 0 ...
```

DATOS DE ENTRENAMIENTO Y TEST

```
table(Hydrobioscidaedatos1$Hydrobioscidae)
contrasts(Hydrobioscidaedatos1$Hydrobioscidae)
```

```

sample_n(Hydrobioscidaedatos1,3)

#split data into training and test set
set.seed(123)
training.sample.Hydrob<-Hydrobioscidaedatos1$Hydrobioscidae %>% createDataPartition(p=0.9,list=FALSE)
train.data.Hydrob<-Hydrobioscidaedatos1[training.sample.Hydrob,]
test.data.Hydrob<-Hydrobioscidaedatos1[-training.sample.Hydrob,]
table(train.data.Hydrob$Hydrobioscidae)

```

CONSTRUCCIÓN DE MODELO1

```

modelo1<-glm(Hydrobioscidae ~., data= train.data.Hydrob,family=binomial)
%>% stepAIC(trace=FALSE)
summary(modelo1)

hist(modelo1$residuals)

anova(modelo1, test="Chisq")

probabilities1<-modelo1%>% predict(test.data.Hydrob,type="response")
predicted.classes1<-ifelse(probabilities1>0.5,"pos","neg")
mean(predicted.classes1==test.data.Hydrob$Hydrobioscidae)

```

CONSTRUCCIÓN DE MODELO2

```

modelo3<-glm(Hydrobioscidae~ N03+log(Solidos_Suspendidos), data= train.data.Hydrob,family=binomial)

summary(modelo3)

hist(modelo3$residuals)

```

BONDAD DE AJUSTE DEL MODELO DEVIANZA Y CHI2

```

dev <- modelo3$deviance
nullDev <- modelo3>null.deviance
modelChi <- nullDev - dev
modelChi

chigl <- modelo3$df.null - modelo3$df.residual
chisq.prob <- 1 - pchisq(modelChi, chigl)
chisq.prob

R2.h1 <- modelChi/modelo3>null.deviance
R2.h1

pR2(modelo3)

```

DIAGNÓSTICO DEL MODELO

SUPUESTOS DEL MODELO LINEALIDAD

```
Hydrobioscidaedatos1$logNO3Int <- log(Hydrobioscidaedatos1$N03) * Hydrobioscidaedatos1$N03
Hydrobioscidaedatos1$logSolidos_SuspendidosInt <- log(Hydrobioscidaedatos1$Solidos_Suspendidos) * Hydrobioscidaedatos1$Solidos_Suspendidos

linealidad <- glm(Hydrobioscidae ~ N03 + Solidos_Suspendidos + logNO3Int + logSolidos_SuspendidosInt, data = Hydrobioscidaedatos1, family = binomial())
summary(linealidad)
```

MULTICOLINEALIDAD

```
library(car)
vif(modelo3)
```

MATRIZ DE CONFUSIÓN2

```
#Cálculo de La probabilidad predicha por el modelo con los datos de test
prob.modelo3 <- predict(object = modelo3, newdata = test.data.Hydrob,
                        type = "response")
```

```
# Vector de caracteres "Down"
pred.modelo3 <- rep("0", length(prob.modelo3))
# Sustitución de "Ausencia" por "Presencia" si La probabilidad a posteriori > 0,5
pred.modelo3[prob.modelo3 > 0.5] <- "1"
```

```
# Matriz de confusión
table(pred.modelo3, test.data.Hydrob$Hydrobioscidae)
```

```
# Test error rate
mean(pred.modelo3 != test.data.Hydrob$Hydrobioscidae)
```

CURVAS ROC

```
pred2 <- prediction(prob.modelo3, test.data.Hydrob$Hydrobioscidae)
perf2 <- performance(pred2, "tpr", "fpr")
plot(perf2, colorize=TRUE)
abline(a=0, b= 1)
```

```
auc2 <- performance(pred2, measure = "auc")
auc3 <- auc2@y.values[[1]]
auc3
```

GRÁFICO DE LA REGRESIÓN

NITRATOS

```

# Se crea un vector con nuevos valores interpolados en el rango de obser-
vaciones.
nuevos_puntos.NO3 <- seq(from = min(Hydrobioscidaedatos2$NO3), to = max(
Hydrobioscidaedatos2$NO3), by = 0.5)

# Predicciones de los nuevos puntos según el modelo.
# Si se indica se.fit = TRUE se devuelve el error estándar de cada predi-
cción
# junto con el valor de la predicción (fit).
mod.NO3 <- glm(Hydrobioscidae ~ NO3, data = Hydrobioscidaedatos2, family
= "binomial")
predicciones.NO3 <- predict(mod.NO3, data.frame(NO3 = nuevos_puntos.NO3)
,
                        se.fit = TRUE)

# Mediante la función logit se transforman los log_ODDs a probabilidades
predicciones_logit.NO3 <- exp(predicciones.NO3$fit) / (1 + exp(prediccio-
nes.NO3$fit))

# Se calcula el límite inferior y superior del IC del 95% sustrayendo e
# incrementando el logODDs de cada predicción 1.95*SE. Una vez calculado
s los
# logODDs del intervalo se transforman en probabilidades con la función
logit.
limite_inferior.NO3      <- predicciones.NO3$fit - 1.95 *predicciones.N
O3$se.fit
limite_inferior_logit.NO3 <- exp(limite_inferior.NO3) / (1 + exp(limite_
inferior.NO3))
limite_superior.NO3      <- predicciones.NO3$fit + 1.95 *predicciones.N
O3$se.fit
limite_superior_logit.NO3 <- exp(limite_superior.NO3) / (1 + exp(limite_
superior.NO3))

# Se crea un dataframe con los nuevos puntos y sus predicciones
datos_curva.NO3 <- data.frame(NO3 = nuevos_puntos.NO3, probabilidad_pres
encia.NO3 = predicciones_logit.NO3, limite_inferior_logit.NO3 =limite_in
ferior_logit.NO3, limite_superior_logit.NO3 =limite_superior_logit.NO3)

ggplot(Hydrobioscidaedatos2, aes(x = NO3, y = Hydrobioscidae)) +
  geom_point(aes(color = as.factor(Hydrobioscidae)), shape = "I", size =
3) +
  geom_line(data = datos_curva.NO3, aes(y = probabilidad_presencia.NO3),
           color = "firebrick", lwd =1) +
  geom_line(data = datos_curva.NO3, aes(y = limite_inferior_logit.NO3),
           linetype = "dashed", lwd =1) +
  geom_line(data = datos_curva.NO3, aes(y = limite_superior_logit.NO3),
           linetype = "dashed", lwd =1) +
  theme_bw() +
  labs(title = "PRESENCIA ~ NITRATOS", type="l", lwd =1,
       y = "P(Hydrobioscidae = 1 | Nitratos)", x = "Nitratos (mg/L)") +

```

```
theme(legend.position = "null") +
theme(plot.title = element_text(hjust = 0.5))
```

SÓLIDOS SUSPENDIDOS

```
# Se crea un vector con nuevos valores interpolados en el rango de obser-
vaciones.
```

```
nuevos_puntos.ss <- seq(from = min(Hydrobioscidaedatos2$Solidos_Suspendi-
dos), to = max(Hydrobioscidaedatos2$Solidos_Suspendidos), by = 0.5)
```

```
# Predicciones de los nuevos puntos según el modelo.
```

```
# Si se indica se.fit = TRUE se devuelve el error estándar de cada predi-
cción
```

```
# junto con el valor de la predicción (fit).
```

```
mod.ss <- glm(Hydrobioscidae ~ Solidos_Suspendidos, data = Hydrobioscidae
datos2, family = "binomial")
```

```
predicciones.ss <- predict(mod.ss, data.frame(Solidos_Suspendidos = nuev-
os_puntos.ss), se.fit = TRUE)
```

```
# Mediante la función Logit se transforman los Log_ODDs a probabilidades
```

```
predicciones_logit.ss <- exp(predicciones.ss$fit) / (1 + exp(prediccione-
s.ss$fit))
```

```
# Se calcula el límite inferior y superior del IC del 95% sustrayendo e
# incrementando el LogODDs de cada predicción 1.95*SE. Una vez calculado
s los
```

```
# LogODDs del intervalo se transforman en probabilidades con la función
Logit.
```

```
limite_inferior.ss <- predicciones.ss$fit - 1.95 *predicciones.ss$
se.fit
```

```
limite_inferior_logit.ss <- exp(limite_inferior.ss) / (1 + exp(limite_in-
ferior.ss))
```

```
limite_superior.ss <- predicciones.ss$fit + 1.95 *predicciones.ss$
se.fit
```

```
limite_superior_logit.ss <- exp(limite_superior.ss) / (1 + exp(limite_su-
perior.ss))
```

```
# Se crea un dataframe con los nuevos puntos y sus predicciones
```

```
datos_curva.ss <- data.frame(Solidos_Suspendidos = nuevos_puntos.ss, pro-
babilidad_presencia.ss = predicciones_logit.ss, limite_inferior_logit.ss
=limite_inferior_logit.ss, limite_superior_logit.ss =limite_superior_log-
it.ss)
```

```
ggplot(Hydrobioscidaedatos2, aes(x = Solidos_Suspendidos, y = Hydrobiosc-
idae)) +
```

```
geom_point(aes(color = as.factor(Hydrobioscidae)), shape = "I", size =
3) +
```

```
geom_line(data = datos_curva.ss, aes(y = probabilidad_presencia.ss),
color = "firebrick", lwd = 1) +
```

```
geom_line(data = datos_curva.ss, aes(y = limite_inferior_logit.ss),
```

```
linetype = "dashed",lwd =1) +  
geom_line(data = datos_curva.ss, aes(y = limite_superior_logit.ss),  
linetype = "dashed",lwd =1) +  
theme_bw() +  
labs(title = "PRESENCIA ~ SÓLIDOS SUSPENDIDOS",type="l",lwd =1,  
y = "P(Hydrobioscidae = 1 | Sólidos Suspendidos (mg/L))", x = "Sólidos Suspendidos (mg/L)") +  
theme(legend.position = "null") +  
theme(plot.title = element_text(hjust = 0.5))
```