

**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**Facultad de Ciencias Naturales y Matemáticas**



**TRABAJO FINAL DE LA MATERIA INTEGRADORA**

**“Aplicación de modelos logísticos y geoestadísticos en  
organismos macroinvertebrados en la cuenca del río  
Guayas”**

Previo a la obtención del Título de:

**INGENIERA EN ESTADÍSTICA INFORMÁTICA**

Presentado por:

**Tatiana Maritza Avilez Arreaga**

**GUAYAQUIL – ECUADOR**

**Año: 2017**

## **AGRADECIMIENTOS**

Al único y sabio Dios, por la bendición y las fuerzas necesarias que me ha dado al permitirme realizar este proyecto. Agradezco infinitamente a mi madre Yancy Arreaga por su constante preocupación, consejos y apoyo incondicional en todas las etapas de mi vida. Gracias a mi padre Milagro Avilez, a mi hermano John Avilez, a mi hermana Soledad y mis sobrinos por su amor, por su motivación para que siga adelante con mis estudios y por su entusiasmo. Gracias familia porque me han dado la mejor herencia y además me han formado en el camino del Señor.

## **DECLARACIÓN EXPRESA**

"La responsabilidad del contenido desarrollado en la presente propuesta de la materia integradora corresponde exclusivamente a Tatiana Maritza Avilez Arreaga y el patrimonio intelectual del mismo a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL"

---

Tatiana Maritza Avilez Arreaga

## RESUMEN

La cuenca del río Guayas está constituida por siete ríos, algunos o parte de ellos han sido obstruidos por los desechos generados por el ser humano, el objetivo de este trabajo es identificar través de ciertas características las zonas contaminadas para que instituciones dedicadas a la protección del medio ambiente puedan tomar cartas en el asunto, el aporte de las especies Baetidae, Chironomidae e Hydropsychidae es fundamental en este estudio debido a que son bioindicadores de calidad del agua.

La metodología fue la siguiente: en el análisis descriptivo se descubrieron las condiciones en las que se desarrollaron estas especies, en el análisis bivariado se manifestaron los parámetros de las variables más importantes en las que estas especies se hacen presentes. En el llamado análisis especializado se utilizaron los Modelos Lineales Generalizados (GLM) para explorar los factores que propician la aparición de las especies, y de qué manera influyeron a las comunidades y el método de Kriging Universal permitió interpolar la abundancia de las especies en los sitios no muestreados.

Las especies Baetidae y Chironomidae se desarrollaron en ambientes con características similares, se han preservado en aguas mansas y además son organismos difíciles de modelar linealmente ya que, el porcentaje de explicación de la variabilidad real de los datos fue baja, mientras que Hydropsychidae se desarrolló en aguas corrientosas y que además se pudo modelar linealmente. Baetidae se desarrolló en la parte centro-oeste de la cuenca del río Guayas, Chironomidae un poco más arriba del área centro-oeste e Hydropsychidae en la parte central y casi toda la orilla de la parte este.

Los GLM son modelos lineales, en este tipo de estudio se incluyeron variables inconstantes, por lo que muchos de los modelos expuestos no resultaron robustos es decir, el porcentaje de explicación de las variables fue bajo.

**Palabras clave:** Regresión Logística, Modelos Lineales Generalizados, Macroinvertebrados, Kriging

## ABSTRACT

*The Guayas river basin is constituted by seven rivers, some or part of them have been obstructed by the wastes generated by the human being, the objective of this work is to identify through certain characteristics the contaminated zones so that institutions dedicated to environmental protection can take action about this, the contribution of species Baetidae, Chironomidae and Hydropsychidae is fundamental in this study because they are bioindicators of water quality.*

*The methodology was as follows: in the descriptive analysis the conditions were found in which these species were developed, in the bivariate analysis the parameters of the most important variables in which these species are present were manifested. In the called specialized analysis, generalized linear models (GLM) were used to explore the factors that favor the emergence of the species, and how they influenced communities and the Universal Kriging method allowed to interpolate abundance of species at sites not sampled.*

*The species Baetidae and Chironomidae were developed in environments with similar characteristics, they have been preserved in calm waters and in addition they are organisms difficult to model linearly as, the percentage of explanation of the real variability of the data was low, whereas Hydropsychidae was developed in corrugated waters and can also be modeled linearly. Baetidae developed in the central-west part of the Guayas River basin, Chironomidae a little above the center-west and Hydropsychidae in the central part and almost all the east shore.*

*GLM are linear models, in this type of study we included inconstant variables, so that many of the models exposed were not robust, that is, the explanation percentage of the variables was low.*

*Keywords: Logistic Regression, Generalized Linear Models, Macroinvertebrates, Kriging*

# ÍNDICE GENERAL

RESUMEN .....	I
ABSTRACT .....	II
ABREVIATURAS .....	VI
SIMBOLOGÍA.....	VII
ÍNDICE DE FIGURAS .....	VIII
ÍNDICE DE TABLAS .....	X
1. INTRODUCCIÓN.....	1
1.1 Descripción del problema.....	1
1.2 Objetivos .....	3
1.2.1 Objetivo General.....	3
1.2.2 Objetivos Específicos .....	3
1.3 Marco Teórico .....	4
1.4 Justificación .....	4
CAPÍTULO 2 .....	6
2. METODOLOGÍA DEL DISEÑO .....	6
2.1 Diagrama de Flujo.....	6
2.2 Área de Estudio .....	7
2.3 Recolección de datos.....	7
2.4 Especies .....	9
2.4.1 Baetidae .....	9
2.4.2 Chironomidae .....	9
2.4.3 Hydropsychidae.....	10
2.5 Descripción de las variables independientes .....	11
2.6 Métodos estadísticos .....	12
2.6.1 Definiciones de estadísticas descriptivas .....	12
Media.....	12
Moda.....	12

Mediana.....	12
Varianza .....	13
Cuantiles.....	13
Histograma .....	14
Diagrama de Cajas.....	14
2.6.2 Modelos Lineales Generalizados (GLM).....	14
2.6.3 Kriging .....	17
Lineal.....	19
Esférico.....	19
Exponencial.....	20
Gaussiano .....	20
Matern .....	20
Circular [15] .....	21
Pentaesférico.....	21
2.7 Construcción, Validación y selección de modelos .....	22
2.7.1 Devianza $D2$ .....	22
2.7.2 Akaike.....	23
2.7.3 Validación cruzada para GLM .....	24
2.7.4 Estrategia de selección del Mejor Modelo de Regresión Logística.....	25
CAPÍTULO 3 .....	26
3. RESULTADOS .....	26
3.1 Análisis descriptivo univariado .....	26
3.2 Análisis descriptivo bivariado .....	36
3.3 Análisis especializado.....	43
3.3.1 Modelo GLM para la Especie Baetidae .....	43
3.3.2 Modelo GLM para la Especie Chironomidae .....	47
3.3.3 Modelo GLM para Especie Hydropsychidae .....	52
3.3.4 Kriging Universal .....	57
CAPÍTULO 4 .....	66
4. DISCUSIONES Y CONCLUSIONES.....	66

4.1	Conclusiones .....	66
4.2	Recomendaciones .....	67
	BIBLIOGRAFÍA .....	68

## ABREVIATURAS

ESPOL	Escuela Superior Politécnica del Litoral
CADS	Centro de Agua y Desarrollo Sustentable
UNICEF	Fondo de las Naciones Unidas para la Infancia
GLM	Modelos Lineales Generalizados
UNESCO	Naciones Unidas para la Educación, la Ciencia y la Cultura
OMS	Organización Mundial de la Salud
PNUMA	Programa de Naciones Unidas para el Medio Ambiente
KFCV	K-Fold Cross-Validation
LOOCV	Leave-one-out Cross-Validation
AIC	Criterio de Información Akaike

## SIMBOLOGÍA

Km	Kilómetro
pH	Potencial de Hidrógeno
m/s	Metros por segundo
m <sup>3</sup> /s	Metros cúbicos por segundo
mg/L	Miligramo por litro
°C	Grados centígrados
mm	Milímetros
cm	Centímetros
FTU/NTU	Unidad nefelométrica de turbidez

## ÍNDICE DE FIGURAS

Figura 2.1. Diagrama de flujo del proceso análisis de variables .....	6
Figura 2.2. Mapa indicando los 120 sitios de muestreos en la cuenca del río Guayas.....	8
Figura 2.3. Especie Baetidae .....	9
Figura 2.4. Especie Chironomidae .....	10
Figura 2.5. Especie Hydropsychidae.....	10
Figura 3.1 Histograma del Tipo de Sedimento .....	27
Figura 3.2. Histograma de Cantidad de Nitrato en el Nitrógeno.....	28
Figura 3.3. Histograma de Variación del ancho del río .....	29
Figura 3.4. Histograma de Turbiedad.....	30
Figura 3.5. Histograma de pH .....	32
Figura 3.6. Histograma de Plantas acuáticas.....	33
Figura 3.7. Histograma de Velocidad.....	34
Figura 3.8. Histograma de Temperatura .....	36
Figura 3.9. Histograma Presencia/Ausencia por Especie .....	37
Figura 3.10. Histograma de la Madera Muerta.....	38
Figura 3.11. Histograma de Tipo de Sedimento por Porcentaje de cobertura .....	39
Figura 3.12. Histograma de Temperatura por Especie .....	40
Figura 3.13. Histograma de pH por Especie .....	41
Figura 3.14. Histograma de Velocidad por Especie .....	42
Figura 3.15. Mapa de Probabilidades de la especie Baetidae .....	45
Figura 3.16. Mapa de Error de la especie Baetidae .....	46
Figura 3.17. Mapa de Probabilidad de la especie Chironomidae .....	49

Figura 3.18. Mapa de Error de la especie Chironomidae .....	50
Figura 3.19. Mapa de Error de la especie Hydropsychidae .....	54
Figura 3.20. Mapa de Error de la especie Hydropsychidae .....	55
Figura 3.21. Semivariograma de la especie Baetidae .....	57
Figura 3.22. Mapa de riqueza de la especie Baetidae generado con Kriging Universal .....	58
Figura 3.23. Mapa de Varianza de la especie Baetidae generado por Kriging Universal .....	59
Figura 3.24. Semivariograma de la especie Chironomidae.....	60
Figura 3.25. Mapa de riqueza de la especie Chironomidae generado con Kriging Universal .....	61
Figura 3.26. Mapa de Varianza de la especie Chironomidae generado con Kriging Universal .....	62
Figura 3.27. Semivariograma de la especie Hydropsychidae .....	63
Figura 3.28. Mapa de riqueza de la especie Hydropsychidae generado con Kriging Universal .....	64
Figura 3.29. Mapa de varianza de la especie Hydropsychidae generado con Kriging Universal .....	65

## ÍNDICE DE TABLAS

Tabla 1. Detalle de las variables independientes .....	11
Tabla 2. Distribución de Frecuencias del Tipo de Sedimento .....	26
Tabla 3. Distribución de Frecuencias de la Cantidad de Nitrato en el Nitrógeno ..	27
Tabla 4. Estadísticas Descriptivas de Cantidad de Nitrógeno en el Nitrato .....	28
Tabla 5. Distribución de Frecuencias de Variación del ancho.....	29
Tabla 6. Distribución de Frecuencias de Turbiedad.....	30
Tabla 7. Estadísticas Descriptivas de Turbiedad .....	30
Tabla 8. Distribución de Frecuencias de pH .....	31
Tabla 9. Estadísticas Descriptivas de pH.....	31
Tabla 10. Histograma de Plantas acuáticas sumergidas .....	32
Tabla 11. Distribución de Frecuencias de Velocidad .....	33
Tabla 12. Estadísticas Descriptivas de Velocidad.....	34
Tabla 13. Distribución de Frecuencias de Temperatura.....	35
Tabla 14. Estadísticas Descriptivas de Temperatura .....	35
Tabla 15. Distribución de Frecuencias de Presencia/Ausencia por Especie.....	37
Tabla 16. Distribución de Frecuencias de la Madera muerta .....	38
Tabla 17. Distribución de Frecuencias de Porcentaje de cobertura por Tipo de Sedimento .....	39
Tabla 18. Distribución de Frecuencias de Temperatura por Especie.....	40
Tabla 19. Distribución de Frecuencias de pH por Especie.....	41
Tabla 20. Distribución de Frecuencias de Velocidad por Especie.....	42
Tabla 21. Coeficientes del Modelo GLM para Baetidae .....	44
Tabla 22. Información del Modelo GLM para Baetidae .....	44
Tabla 23. Tabla de Confusión de Baetidae .....	45

Tabla 24. Coeficientes del Modelo GLM para Chironomidae.....	48
Tabla 25. Información del Modelo GLM de Chironomidae .....	48
Tabla 26. Tabla de Confusión para Chironomidae.....	49
Tabla 27. Coeficientes del Modelo GLM de Hydropsychidae.....	53
Tabla 28. Información del Modelo GLM de Hydropsychidae.....	53
Tabla 29. Tabla de Confusión de Hydropsychidae.....	54

# CAPÍTULO 1

## 1. INTRODUCCIÓN

En este capítulo se presenta estado del proyecto integrador; el mismo que abarca el planteamiento del problema, justificación y objetivos.

### 1.1 Descripción del problema

Con el paso del tiempo la contaminación del líquido vital es más común en el medio en que vivimos, debido a que existen múltiples factores que inciden en la polución de las aguas superficiales, tales como: creciente de urbanización, vertido de residuos industriales y domésticos, escasa depuración de aguas residuales y sedimentación, que conjugados deterioran la calidad del agua de las fuentes de captación de este recurso para la respectiva potabilización.

El agua es indispensable para la vida humana y otras especies por ello la calidad de este recurso debe cumplir con las condiciones de salubridad exigida para el consumo directo, sin embargo, en Guayaquil el agua potabilizada no es confiable de ahí que los ciudadanos recurran a filtros y agua embotellada, en lo que subyacen la insatisfacción y desconfianza en este servicio básico.

Para que el líquido sea considerado aceptable para las diversas actividades humanas, domésticas e industriales debe poseer propiedades químicas, biológicas y físicas, que son determinantes del concepto de “calidad de agua”.

El Fondo de las Naciones Unidas para la Infancia (UNICEF), la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), la Organización Mundial de la Salud (OMS) y el Programa de Naciones Unidas para el Medio Ambiente (PNUMA) han

investigado sobre las condiciones del líquido vital y su vulnerabilidad de cara a los factores antrópicos crecientes, de ahí que sea necesario conocer la realidad de la cuenca del Guayas utilizando métodos estadísticos que permitan modelar la distribución espacial de comunidades de bioindicadores que están relacionados con la polución de las aguas de la cuenca, lo que además proporciona información sobre cuáles son los lugares más contaminados a fin de propiciar futuros programas de remediación ambiental.

Los bioindicadores seleccionados para el estudio de las muestras recolectadas corresponden a la clase de macroinvertebrados *bioindicadores* que suelen estar presentes con mayor frecuencia en ecosistemas fluviales, su presencia está condicionada a los niveles de contaminación del agua y a los cambios ambientales (Instituto Mi Río, 1997; LÓPEZ et al. 1995 en POSADA et al. 1999).

Los macroinvertebrados bentónicos actúan como indicadores debido a que ante la presencia de un tóxico en el agua se estimula una respuesta metabólica que trata de contrarrestar el daño causado por la situación del medio. En el caso de que la toxicidad sea baja, existe la posibilidad de que no se generen cambios en la presencia o abundancia de estos organismos, sin embargo, es posible que se detecte el estrés que el tóxico causa en algunas vías metabólicas o en las propiedades del material hereditario.

Son pocos los bioindicadores utilizados en proyectos de supervisión de cursos de aguas superficiales y algunos han servido para realizar investigaciones en estudios de ecotoxicología, por ejemplo: Hernández en el año 2014 elaboró un trabajo de diploma titulado "*Estudios ecotoxicológicos en diferentes bioindicadores ambientales del bioplaguicida Tricosave-34*", se evaluaron diferentes organismos a dosis única del producto Tricosave-34 desarrollado por el Grupo empresarial

Labiopharm para controlar las plagas producidas por algunos hongos y nematodos.

En nuestro país hay escasa documentación respecto a la aplicación de organismos bentónicos en combinación con un modelado estadístico, por lo que hay que potenciar su aplicación a fin de realizar predicciones en función de la presencia o ausencia de estos organismos que a la vez permiten evaluar y monitorear la condición de lagos y ríos.

## **1.2 Objetivos**

### **1.2.1 Objetivo General**

Identificar cuáles son las zonas idóneas para la aparición de los organismos bioindicadores utilizando Modelos Lineales Generalizados (GLM) y el Método de Kriging para conocer las zonas contaminadas en la cuenca del río Guayas.

### **1.2.2 Objetivos Específicos**

- Determinar los factores ambientales que propician la presencia de los macroinvertebrados.
- Modelar la distribución potencial de las comunidades a partir de un modelo logístico a fin de conocer la probabilidad de presencia/ausencia.
- Predecir la abundancia de las especies bioindicadoras a través de Kriging Universal
- Validar los métodos aplicados.

### **1.3 Marco Teórico**

#### **Evaluación de la calidad de agua a través de macroinvertebrados bentónicos e índices biológicos en ríos tropicales en bosque de Neblina Montana**

Estas especies son indicadores biológicos que se utilizan en la evaluación de ecosistemas lóticos del mundo, debido a los requerimientos especiales y a las adaptaciones evolutivas respecto a ciertas condiciones ambientales, lo que hace que se conviertan en organismos con límites de tolerancia a diferentes alteraciones de su hábitat. El problema es que por lo general estos organismos no se utilizan de manera oficial para la evaluación y monitoreo de la condición de cuerpos de agua. En este estudio se evaluaron tres ríos: Palmeras, Brincador y Guajalito para comparar diferentes índices biológicos desarrollados para zonas templadas y tropicales para evaluar su validez en ríos tropicales, con el fin de que sea una línea base para el manejo y conservación de la reserva Bosque protector río Guajalito.

### **1.4 Justificación**

El Guayas es el río más importante del Ecuador, es el más grande de la costa sudoccidental del Pacífico. Posee una belleza natural y paisajística que le brinda hermosura a la provincia del Guayas además es el hábitat de muchas especies acuáticas, y algunas de ellas son consideradas indicadores de calidad del agua.

En la actualidad, los caudales sufren muchos impactos ambientales y estos especímenes con su presencia o ausencia nos alertan de la condición del agua de cada río.

Esta investigación propone profundizar en el estudio de estos organismos ya que los macroinvertebrados habitan en estos cuerpos de agua y su presencia es el punto focal de fenómenos naturales o antrópicos.

En el desarrollo de este estudio, se podrá evidenciar que estas especies juegan un rol importante en relación a los distintos índices de calidad del agua, no basta con solo prestar atención a los factores físicos y químicos del fluido, el factor biológico también es importante, asequible, diligente y además brinda posibilidades para futuras investigaciones.

Este importante afluente está constituido por siete ríos, que se deben aprovechar al máximo para evaluar la condición del agua, por ello en esta documentación se evidenciará la aplicación de un modelado estadístico que permita en función de estos seres conocer en los lugares muestreados y predecir en los que no han sido muestreados, el estado de calidad del agua en función de la presencia/ausencia o abundancia de estos organismos.

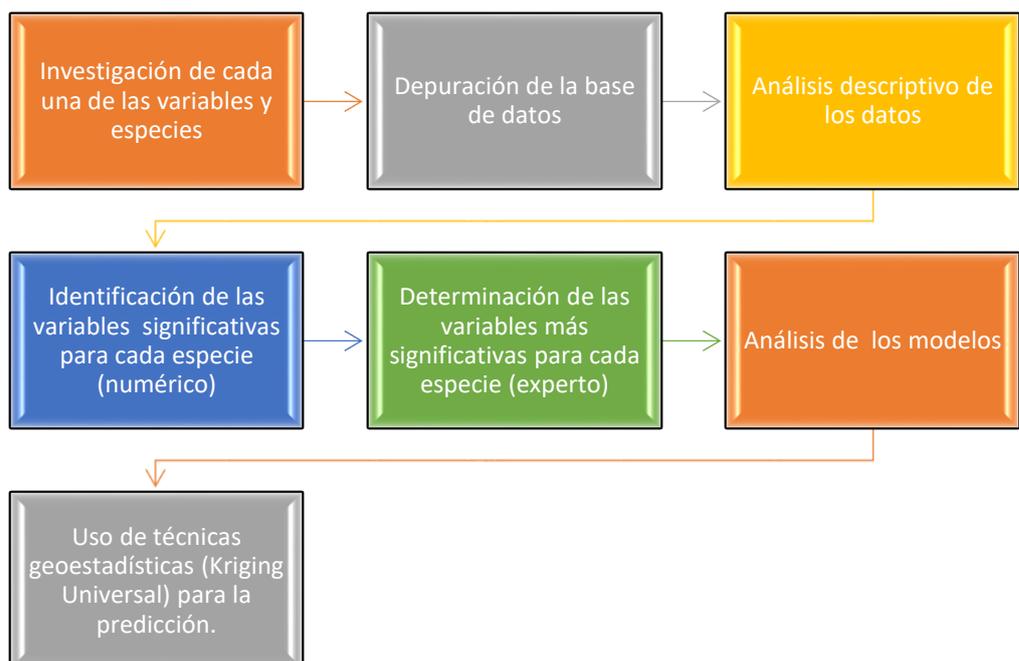
# CAPÍTULO 2

## 2. METODOLOGÍA DEL DISEÑO

En este capítulo se presenta el proceso de análisis de los datos y las metodologías utilizadas para cumplir con el objetivo que se planteó en el Capítulo 1.

### 2.1 Diagrama de Flujo

En esta sección se detallan de manera general los pasos para el análisis de los datos.



**Figura 2.1. Diagrama de flujo del proceso análisis de variables**

*Elaboración propia*

## **2.2 Área de Estudio**

El presente estudio se centra en la distribución geográfica potencial de tres organismos macroinvertebrados que habitan en la cuenca del río Guayas: Baetidae, Chironomidae e Hydropsychidae.

La cuenca del río Guayas está ubicada en la parte centro-occidental del Ecuador, comprende una extensión de 60.000  $Km^2$  y una extensa llanura de 300 Km de largo por 200 Km de ancho. Tiene una precipitación media de 1662 mm, descarga media de 835  $m^3/s$ , capacidad de almacenamiento de 6000 millones de  $m^3$  y 14.350  $m^3/s$  de descarga natural. Integrada por los ríos Pagua, Balao, Naranjal, Chongón, Boliche, Taura, Yaguachi, Pueblo Viejo, Zapotal (Caracol), Balzar, Quevedo, Vinces, Daule y Babahoyo, es la mayor cuenca fluvial de la costa del Pacífico en la América del Sur.

Hace miles de años, las zonas con poco relieve de la cuenca del río Guayas estaban revestidas por el mar, después de que este se retirara dejó zonas con capa de areniscas y arcillas, que se formaron en base a los sedimentos marinos y residuos del volcanismo subacuático. Esta estructura sedimentaria al complementarse con los residuos aluviales procedentes de la cordillera andina, enriqueció su suelo hasta transformarlo en uno de los más aptos para todo tipo de agricultura.

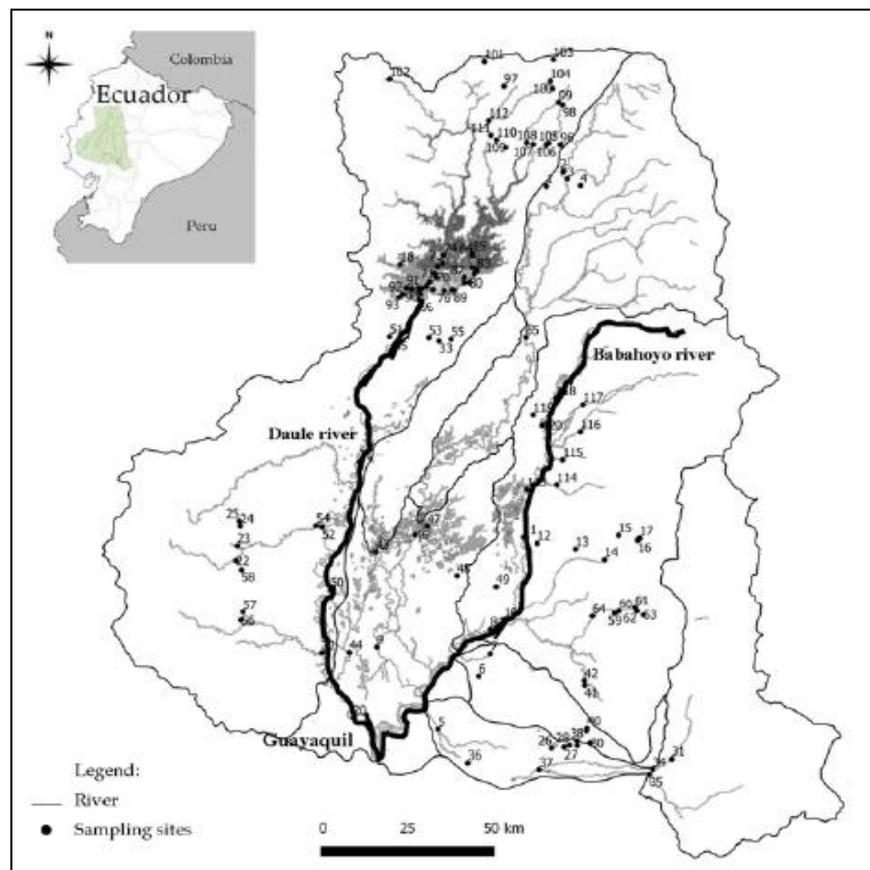
Son 120 sitios de muestreos, de los cuales treinta y dos se localizaron en Daule-Peripa, y los 88 restantes se situaron en ubicaciones ascendentes y descendentes de los ríos de la cuenca del río Guayas.

## **2.3 Recolección de datos**

El proceso de muestreo lo ejecutó el Centro del Agua y Desarrollo Sustentable (CADS) de la Escuela Superior Politécnica del Litoral (ESPOL), al final de la temporada (julio-noviembre) desde el 23 de octubre al 26 de noviembre de 2013.

Debido a que el Ecuador se encuentra en una región tropical, las diferencias estacionales en comparación con las regiones templadas son bajas. El CADS ejecutó una evaluación general para investigar fuentes puntuales de contaminación. Cada variable fue medida con instrumentos especializados y para realizar el proceso de medición, se tomaron muestras de aguas en cada sitio de muestreo y antes de ser analizados, se almacenaron en un lugar fresco y oscuro.

No se pudo medir el error humano, debido a las limitaciones prácticas, a la anchura y la profundidad de los sitios ubicados en el embalse y en los grandes ríos.



**Figura 2.2. Mapa indicando los 120 sitios de muestreos en la cuenca del río Guayas**

*Fuente: Generalized Linear Models to Identify Key Hydromorphological and Chemical Variables Determining the Occurrence of Macroinvertebrates in the Guayas River Basin*

## 2.4 Especies

### 2.4.1 Baetidae

Generalmente viven en aguas lólicas, donde hay mucho oxígeno, debajo de los troncos, rocas, hojas, a veces están adheridas a la vegetación que está sumergida y en fondos arenosos. Son bioindicadores, ayudan a descifrar si el agua es limpia o está ligeramente contaminada. Se alimentan de algas y de tejidos acuáticos y son alimento de los peces.



**Figura 2.3. Especie Baetidae**

*Fuente: Arthropodafotods*

### 2.4.2 Chironomidae

Son muy parecidos a los mosquitos, habitan el ambiente acuático y semiacuático, aguas cloacales, troncos de árboles y vegetales en descomposición. Algunas de las larvas poseen un color rojo brillante, esto se debe a la hemoglobina. Son indicadores de cambios climáticos, la presencia, ausencia de estas larvas en los entornos acuáticos pueden dar señales de grados de contaminación de las aguas.



**Figura 2.4. Especie Chironomidae**

*Fuente: Diptera.info*

### **2.4.3 Hydropsychidae**

Construyen su refugio en rocas livianas hilando una red de captura de seda algunas de ellas, construyen su refugio y redes sobre el nivel del mar y sobre las piedras cuando son zonas de corrientes rápidas. Se alimentan de algas, diatomeas y partículas de detritus.



**Figura 2.5. Especie Hydropsychidae**

*Fuente: Phinizy*

## 2.5 Descripción de las variables independientes

Detalle de las covariables que se van a analizar durante toda la investigación.

**Tabla 1. Detalle de las variables independientes**

VARIABLES	UNIDAD	EXPLICACIÓN
Temperatura	°C	Se tomó una muestra
pH	-	Se tomó una muestra
Turbiedad	FTU/NTU	Se tomó una muestra
Cantidad de Nitrógeno en el Nitrato	mg/L	Evaluado mediante un equipo
Velocidad	m/s	Protocolo
Porcentaje de cobertura de las Macrófitas sumergidas	0	No hay macrófitas
	1	Raro = 1%-5%
	2	Ocasional = 5%-25%
	3	Frecuente = 25%-50%
	4	Común= 50%-75%
	5	Abundante = 75%-100%
Variación del ancho del cuerpo de agua	0	En el embalse
	A	
	B	
	C	
	D	
	E	
Madera muerta		Protocolo
Leña menuda d<3cm	0	Ausente
Rama 3-30 cm	1	Limitada
Troncos d>30 cm	2	Abundante
Tipo de sustratos minerales		Protocolo
Roca	0	Ninguna
Adoquín	1	0%-20%
Grava	2	20%-40%
Arena	3	40%-60%
Barro y arcilla	4	60%-80%
Principal Tipo de Sedimento (Protocolo)	1	Roca
	2	Adoquín
	3	Grava
	4	Arena
	5	Cieno y arcilla

Coordenadas	UTM	Sistema Transversal de Mercator Universal, GPS (Dirección del este y norte).
-------------	-----	--

*Fuente: Centro del Agua y Desarrollo Sostenible (CADS)*

*Elaboración propia*

## 2.6 Métodos estadísticos

### 2.6.1 Definiciones de estadísticas descriptivas

#### **Media**

Es la suma de todos los valores dividido para la cantidad total de la cantidad de valores. Dado  $n$  números, su media aritmética es:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

(2.1)

#### **Moda**

Se denomina así al valor que se repite con mayor frecuencia, se puede hallar tanto en variables cuantitativas como en cualitativas. Pueden existir más de una moda.

#### **Mediana**

Es el número central de los datos, siempre que estos estén ordenados de manera ascendente. En caso de que la cantidad de número sea par, se la calcula con la ayuda del promedio de los dos valores centrales.

## Varianza

Medida de dispersión que no puede ser negativa y su valor será cero, solo cuando, todos los datos tengan el mismo valor, es decir la media es igual a tal valor, por lo que la suma cuadrática es cero.

$$\sigma^2 = \sum_{i=1}^n (X_i - \bar{x})^2$$

(2.2)

Esta medida no está dada en las unidades con la que se midieron los datos por estar elevada al cuadrado, es aquí donde se hace uso de la desviación estándar, la misma que es igual a la raíz cuadrada positiva de la varianza, se denota por  $s$ .

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{x})^2}{n - 1}}$$

(2.3)

## Cuantiles

Estos valores se calculan a partir de una muestra ordenada, juntamente con el gráfico de la Ojiva. La Ojiva es una Figura de la Frecuencia Relativa Acumulada.

Se denota a  $Q_1$  como el primer cuartil de la muestra  $X$ , tal que, valores los menores a  $Q_1$  representan el veinticinco por ciento de la muestra.

Se denota a  $Q_2$  como el segundo cuartil o mediana de la muestra  $X$ , tal que, valores los menores a  $Q_2$  representan el cincuenta por ciento de la muestra.

Se denota a  $Q_3$  como el tercer cuartil de la muestra  $X$ , tal que, valores los menores a  $Q_3$  representan el setenta y cinco por ciento de la muestra.

### **Histograma**

Gráfico bidimensional de “barras”, que se construye a partir de las  $k$  clases obtenidas en la tabla de frecuencia, las mismas que se reflejan en el eje horizontal y las frecuencias relativas se reflejan en el eje vertical.

### **Diagrama de Cajas**

Se grafica después de haber calculado los cuantiles, el valor máximo y mínimo de la muestra. En este gráfico se reflejan varias características importantes de los datos, tiene unas líneas a los extremos de modo cualquier que salga de estos límites, se considera un valor atípico; este valor o valores se denominan así por ser mucho más grandes o mucho más pequeños que el resto de valores.

## **2.6.2 Modelos Lineales Generalizados (GLM)**

Los Modelos Lineales Generalizados (GLM, Generalized Linear Models), son una expansión de los modelos lineales, propuesta hecha por Nelder y Wedderburn (1972), en la actualidad están implementados en el software estadístico profesional de código abierto R junto con su interfaz de manejo fácil RStudio.

A diferencia de los modelos lineales, estos integran a los modelos con variables de respuesta categórica, esto nos permite modelar variables

de respuestas continuas o categóricas ya sean estas binomial, poisson, normal, entre otras. Además de que proponen buenas alternativas al momento de lidiar con variables dependientes que no cumplen con los supuestos de normalidad y homocedasticidad.

Su estructura es:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \varepsilon_i$$

(2.4)

Donde:

$$\varepsilon_i \sim N(\mu, \sigma)$$

$Y_i$ : Variable de respuesta

$X_i$ : Covariables, variables independientes

$\beta_i$ : Parámetros del modelo

$\varepsilon_i$ : Error del modelo

Los componentes de GLM son:

#### a. Componente aleatoria

Es un vector aleatorio que involucra a las variables de respuestas ( $Y_i$ ) y su distribución de probabilidad, y que además deben ser independientes.

$Y_i$  toma el valor de 1 con probabilidad  $\pi_i$  y toma el valor de 0 con probabilidad  $1 - \pi_i$ , por lo tanto, la media y la varianza dependen de  $i$  es decir, si algún valor afecta la esperanza también afectará a la varianza.

### **b. Componente sistemático**

Conformada por las variables predictoras  $X_i$  utilizadas en la función predictora lineal.

### **c. Función de enlace**

Es la unión de las dos componentes anteriores, es decir, vincula el valor esperado de las variables de respuesta a las variables explicativas.

$$g(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots$$

(2.5)

De los GLMs, utilizaremos los modelos logísticos.

### **Modelos de Regresión Logística**

Originaria de las décadas de los 60 con el trabajo de Cornfield, Gordon y Smith.

Su objetivo es expresar la probabilidad de que ocurra algún evento en función de otras variables por lo que se tendrá una variable dependiente que en este caso es dicotómica y una o varias covariables, tal como indica.

La variable  $Y$  es de naturaleza dicotómica puede tomar los siguientes valores:

$$Y_i = \begin{cases} 1 ; & \text{si está presente} \\ 0 ; & \text{si no} \end{cases}$$

Las covariables pueden ser de cualquier índole. La probabilidad de que  $Y = 1$  se denota como  $p$ .

La función logística es:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots)}} ; \quad p \in (0,1)$$

(2.6)

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$  son los parámetros del modelo, si el valor de  $\beta$  es mayor a cero y si la variable aumenta entonces  $p$  crece, caso contrario  $p$  disminuye.

### 2.6.3 Kriging

Término que se utiliza para designar al mejor estimador insesgado, lleva el nombre de Kriging en honor a Krige.

G. Matheron en los años 60 desarrolló esta técnica de estimación espacial a partir de los estudios realizados por D. Krige, quién además fue el primero en usar la correlación espacial para propósitos de predicción.

Este método geoestadístico genera a partir de puntos dispersos, una superficie que estima la realidad, difiere de los otros métodos de interpolación porque esta técnica en primera instancia realiza un estudio interactivo de la conducta espacial del fenómeno personificado por los valores de  $Z$  antes de elegir el mejor método de estimación.

Kriging se basa en modelos estadísticos en donde están inmersas las autocorrelaciones es decir, las relaciones entre los puntos medidos, el resultado no solo presenta la predicción del método, sino que también proporciona la precisión del mismo.

Esta herramienta adapta una función matemática a un conjunto de puntos específicos o a todos los puntos dentro de un radio determinado para luego definir el valor de salida para cada sitio.

Fórmula:

$$\hat{Z}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i)$$

(2.7)

Donde:

$Z(s_i)$ : Valor del sitio  $i$

$\lambda_i$ : Ponderación para el valor del sitio  $i$

$s_0$ : Sitio de la predicción

$N$ : Número de valores medidos

Hay dos razones por las que se considera al estimador de Kriging como el óptimo:

- Es insesgado
- La varianza de estimación sea mínima

Entre los métodos de Kriging están: Kriging Simple, Ordinario y Universal. Para efectos de este estudio utilizaremos Kriging Universal.

### **Función variograma**

Modela la continuidad de la variable de estudio, puesto que puede existir correlación entre los diferentes puntos del espacio, por ello es necesario estudiar la velocidad con la que se pierde la correlación al aumentar la distancia que los separa.

Esta función mide la variabilidad espacial, considerando la distancia respectiva ( $h$ ).

$$\gamma(h) = \frac{1}{2} E\{[Z(x) - Z(x+h)]^2\}$$

(2.8)

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2$$

(2.9)

Donde:

$(x)$ : Variable regionalizada

$(z)$ : Función aleatoria asociada

$(h)$ : Número de pares de datos separados por una distancia

$x_i, x_i + h$ : Posición de los pares de datos,  $i = 1, \dots, N(h)$

Existen varios modelos elementales que sirven para ajustar el variograma a la continuidad espacial.

Utiliza las siguientes variables:

$v$ : Varianza

$c_0$ : Nugget

$a$ : Rango

$c_0 + c$ : Varianza asintótica

$h$ : Distancia entre puntos

Algunos de los modelos son:

### **Lineal**

Es un modelo simple, por lo que no presenta silla.

$$\gamma(h) = c_0 + c.h$$

(2.10)

### **Esférico**

Crece rápidamente cerca del origen, pero en distancias mayores decrece, a distancias superiores al rango el crecimiento es nulo.

$$\gamma(h) = \begin{cases} c_0 + c \left[ 1.5 \left( \frac{h}{a} \right) - 0.5 \left( \frac{h}{a} \right)^3 \right], & \text{si } h \leq a \\ c_0 + c, & \text{si } h > a \end{cases}$$

(2.11)

### Exponencial

Se utiliza cuando la dependencia espacial tiene un crecimiento exponencial en base a la distancia entre las observaciones.

$$\gamma(h) = c_0 + c \left( 1 - e^{\left( \frac{-3h}{a} \right)} \right)$$

(2.12)

### Gaussiano

La dependencia espacial desaparece cuando la distancia tiende al infinito.

$$\gamma(h) = c_0 + c \left( 1 - e^{\left( \frac{-h^2}{a^2} \right)} \right)$$

(2.13)

### Matern

$$\gamma(h) = (c_0 + c) \left( 1 - \frac{1}{2^{v-1} \Gamma(v)} \left( \frac{h}{pr} \right)^v K_v \left( \frac{h}{pr} \right) \right)$$

(2.14)

Donde:

pr: Parámetro asociado al rango porque participa la expresión  $\frac{h}{pr}$

$K_v$ : Función de Bessel de segunda clase modificada

### Circular

$$\gamma(h) = \begin{cases} c_0 + c \left[ \frac{2}{\pi} \left( \frac{h}{a} \right) - \left[ 1 - \left( \frac{h}{a} \right)^2 \right] \right]^{1/2} + 2 \text{Sen}^{-1} \left[ \frac{h}{a} \right], & \text{si } h \leq a \\ c_0 + c, & \text{si } h > a \end{cases}$$

(2.15)

### Pentaesférico

Es poco el uso que se le da en el campo de la ecología, y que además no está disponible en todos los programas. Comparado con el esférico este es más gradual y en ocasiones permite un mejor ajuste a los datos.

$$\gamma(h) = \begin{cases} c_0 + c \left[ \left( \frac{15h}{8a} \right) - \frac{5}{4} \left( \frac{h}{a} \right)^3 + \frac{3}{8} \left( \frac{h}{a} \right)^5 \right], & \text{si } 0 < h \leq a \\ c_0 + c, & \text{si } h > a \end{cases}$$

(2.16)

### Kriging Universal

Se utiliza Kriging Universal cuando la función aleatoria no es estacionaria y su valor esperado cambia en el espacio, en algunas ocasiones la variable no satisface estas condiciones, pero puede representarse por su tendencia. Para tratar este tipo de variables se descompone en dos elementos: el primero es la suma de la tendencia y el segundo es un elemento estocástico estacionario con media cero  $R(x)$ .

$$Z(x) = m(x) + R(x),$$

(2.17)

tal que,  $E[R(x)] = 0$ ,  $V[R(x)] = \sigma^2$

$$\rightarrow E[Z(x)] = m(x)$$

(2.18)

Su estimador es:

$$\hat{Z}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i)$$

(2.19)

## 2.7 Construcción, Validación y selección de modelos

Antes de construir el modelo, con el objeto de eliminar las variables que figuren alta colinearidad, se debe analizar la correlación entre ellas para simplificar el modelo, en el caso de que haya correlación alta entre las variables, el GLM perderá su capacidad de predicción.

No existe un modelo único por lo que, se construyeron varios y se eligió el mejor basado en diferentes criterios:

### 2.7.1 Devianza $D^2$

Véase la Ecuación 2.20. la devianza tiene como objetivo determinar si las variables explicativas seleccionadas son importantes para explicar

la variable criterio y la variabilidad de los datos. Cuando la devianza del modelo es pequeña, indica que el ajuste de los datos es bueno.

$$D^2 = \frac{\text{devianza.modelo.nulo} - \text{devianza.residual}}{\text{devianza.modelo.nulo}} * 100$$

(2.20)

### 2.7.2 Akaike

El criterio de información Akaike (AIC) mide la calidad de un modelo estadístico en sentido relativo usualmente se emplea para elegir el mejor modelo entre un grupo de modelos predictores similares con diferentes grados de complejidad. Mientras más complejo sea el modelo, su capacidad de predicción será menos eficiente por lo tanto, un modelo sencillo con variables significativas, tendrá mayor poder de predicción en varias ocasiones aceptando cierto error de precisión.

Cuanto más pequeño sea el AIC, mejor es el ajuste. Para analizar la distribución de los residuos se elabora su histograma y el gráfico de normalidad (QQplot).

$$AIC = 2 * k - 2 * \ln(L)$$

(2.21)

*k*: Número de parámetros.

*L*: Valor máximo en la función de verosimilitud del modelo, y la complejidad a partir el número de parámetros.

Este criterio generalmente se utiliza cuando no se reservan datos para validar el modelo, lo que hace es aproximar la calidad de las

predicciones sobre los mismos datos con los que se ha construido el modelo.

### 2.7.3 Validación cruzada para GLM

Al momento de buscar el mejor modelo que se ajuste a una muestra seleccionada, se puede presentar un problema de sobreajuste, aunque el modelo se ajuste bien a los datos, existe la posibilidad que no se ajuste bien a otros datos.

Es una forma de comprobar que el modelo que se propone predice a un nuevo conjunto de datos correctamente, su trabajo radica en que se divide a la muestra en submuestras, de tal forma que una se utilice para ajustar el modelo y las restantes, para evaluar el modelo.

Existen dos métodos de validación cruzada: K-Fold Cross-Validation (KFCV) y el Leave-one-out Cross-Validation (LOOCV), para efecto de este estudio se empleará K-Fold Cross-Validation (KFCV).

#### **Validación cruzada usando K grupos (K-fold cross validation)**

La muestra se divide en K subconjuntos, K-1 subconjuntos son utilizados para entrenar y el restante se utiliza para probar el modelo. Este proceso se repite K veces, es decir, con cada uno de los posibles subconjuntos de submuestras de pruebas.

Con La media aritmética de los errores de cada iteración, se calcula el error.

$$Cross - Validation_k = \frac{1}{k} \sum_{i=1}^k MSE_i$$

(2.22)

#### 2.7.4 Estrategia de selección del Mejor Modelo de Regresión Logística

Para seleccionar el mejor Modelo Logístico, se evaluó la significancia de cada variable  $X_i$  con la variable  $Y$ , es decir:

$$Y \sim \beta_0 + \beta_1 X_i \text{ para } i = 1, \dots, n$$

(2.23)

Después se ordenaron las variables desde la más significativa a la menos significativa. Luego, se formó el modelo de tal manera que a la variable más significativa se le aumenta la variable significativa siguiente y así sucesivamente, al momento de integrar una variable se observaba si las variables del modelo eran significativas o no, en el caso de no ser significativas no se tomaban en cuenta. Si las variables eran significativas, se tomaba en cuenta el AIC y la Devianza de modelo, dichos valores debían ser más bajos que el *Modelo Nulo*.

Posteriormente, se presentaron los mejores modelos al Ph.D. Luis Domínguez, experto en estas especies y su hábitat, bajo su conocimiento y los resultados numéricos, se seleccionó el mejor modelo para cada especie.

# CAPÍTULO 3

## 3. RESULTADOS

### 3.1 Análisis descriptivo univariado

En esta sección se presentan los siguientes resultados de las estadísticas descriptivas.

Se realizó un análisis exploratorio a través de histogramas, tablas de frecuencias y diagrama de cajas de las covariables más importantes y de las variables dependientes.

#### Principal Tipo de sedimento

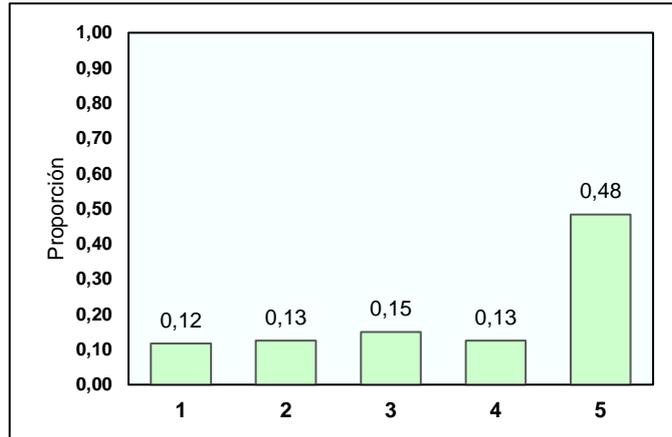
El principal tipo de sedimento que predomina en la cuenca del río Guayas es el Cieno y la Arcilla (48,3%), a la cual le sigue la Grava (15%), algunas de las especies se desarrollan en lugares constituidos por estos dos tipos. Véase *Tabla 2* y *Figura 3.1*.

**Tabla 2. Distribución de Frecuencias del Tipo de Sedimento**

Principal Tipo de sedimento	Proporción
1.Piedra	0,117
2.Adoquín	0,125
3.Grava	0,15
4.Arena	0,125
5.Cieno y Arcilla	0,483
Total	1,000

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia



**Figura 3.1. Histograma del Tipo de Sedimento**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

### Cantidad de Nitrato en el Nitrógeno

La mayoría de los lugares muestreados tienen concentración de Nitrógeno en el Nitrato menor a 0,63 (89,2%), en promedio la cantidad de esta concentración es de 0,37, el 50% de los datos tiene 0,23. Para más detalles *Tabla 3, Tabla 4 y Figura 3.2.*

**Tabla 3. Distribución de Frecuencias de la Cantidad de Nitrato en el Nitrógeno**

Cantidad de Nitrato en el Nitrógeno	Frecuencia Relativa
1.[0,175 - 0,631)	0,892
2.[0,631 - 1,088)	0,075
3.[1,088 - 1,544)	0,017
4.[1,544 - 2,000]	0,017
Total	1,000

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

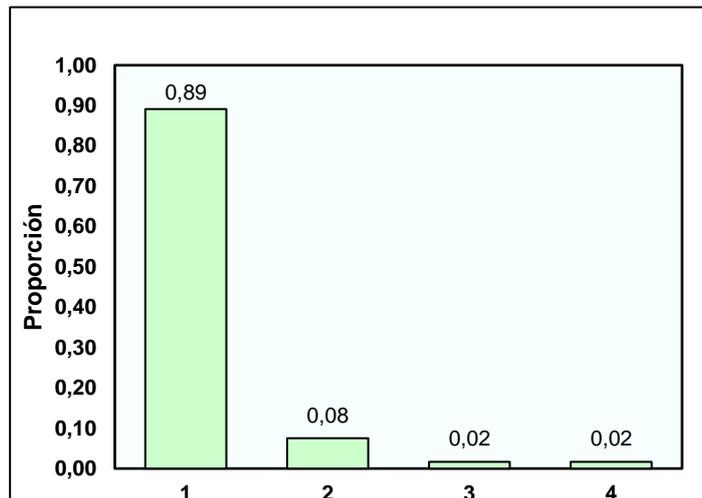
*Elaboración propia*

**Tabla 4. Estadísticas Descriptivas de Cantidad de Nitrato en el Nitrato**

Nitrato en el Nitrógeno	
Media	0,37
Error est. media	0,03
Desv.Estánda	0,29
Mínimo	0,18
Q1	0,23
Mediana	0,23
Q3	0,34
Máximo	2,00
Moda	0,23
N° Moda	62,00

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.2. Histograma de Cantidad de Nitrato en el Nitrógeno**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

### **Variación del ancho del cuerpo de agua**

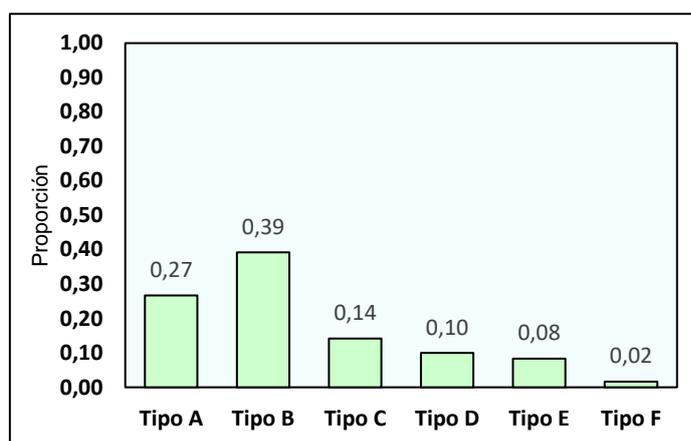
La mayoría de los cuerpos de aguas tienen una variación en el ancho de su estructura Tipo B (39%), para conocer cuál es el tipo B, se puede revisar la sección de la recolección de datos de este trabajo. Véase *Tabla 5* e *Figura 3.3*.

**Tabla 5. Distribución de Frecuencias de Variación del ancho**

Variación del Ancho	Proporción
Tipo A	0,267
Tipo B	0,392
Tipo C	0,142
Tipo D	0,100
Tipo E	0,083
Tipo F	0,017
Total	1,000

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.3. Histograma de Variación del ancho del río**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

## **Turbiedad**

Las aguas superficiales muchas veces presentan turbiedad, el 98% de los puntos de muestreos presentan turbiedad menor a 88,91 unidades de turbidez nefelométricas. En promedio los cuerpos de aguas presentaron 9,8 y el 50% tiene menos de 3,38 unidades de turbidez nefelométricas, lo que indica que el 50% de las masas de aguas tienen una cantidad moderada de plantas y animales. Véase *Tabla 6, Tabla 7 y Figura 3.4.*

**Tabla 6. Distribución de Frecuencias de Turbiedad**

Turbiedad	Proporción
1.[0,00 - 88,91)	0,983
2.[88,91 - 177,82)	0,008
3.[177,82 - 266,73)	0,000
4.[266,73 - 355,63]	0,008
Total	1,000

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

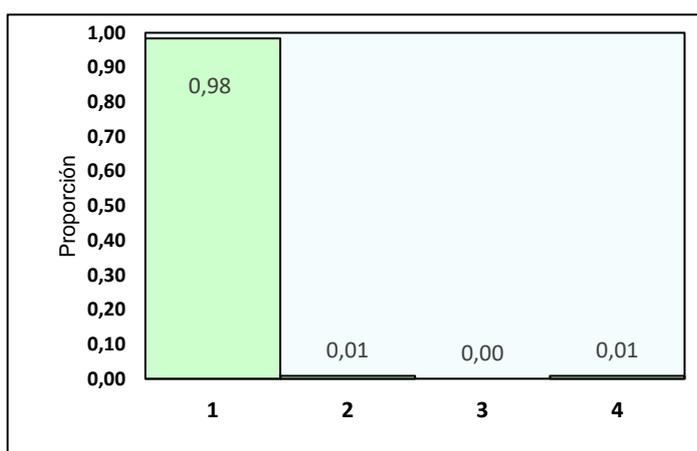
Elaboración propia

**Tabla 7. Estadísticas Descriptivas de Turbiedad**

Turbiedad	
Media	9,82
Error est. media	3,20
Desv.Estándar	35,09
Mínimo	0,00
Q1	1,37
Mediana	3,38
Q3	6,05
Máximo	355,63
Modo	0,825
N° Moda	3,00

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia



**Figura 3.4. Histograma de Turbiedad**

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia

## pH

El 10% de los ríos tienen un pH mayor a 8,41 esto indica que el agua es muy ácida debido a esto, muchos animales acuáticos les será imposible vivir en estos lugares. El 50% de las masas de agua tienen más de 7,02 niveles de pH y menos de 8,41, esto muestra que la mayoría de las aguas son básicas. Para más detalles *Tabla 8 Tablas 9 y Figura 3.5.*

**Tabla 8. Distribución de Frecuencias de pH**

pH	Proporción
1.[6.56 - 7.02)	0,08
2.[7.02 - 7.48)	0,33
3.[7.48 - 7.95)	0,27
4.[7.95 - 8.41)	0,22
5.[8.41 - 8.87]	0,1

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

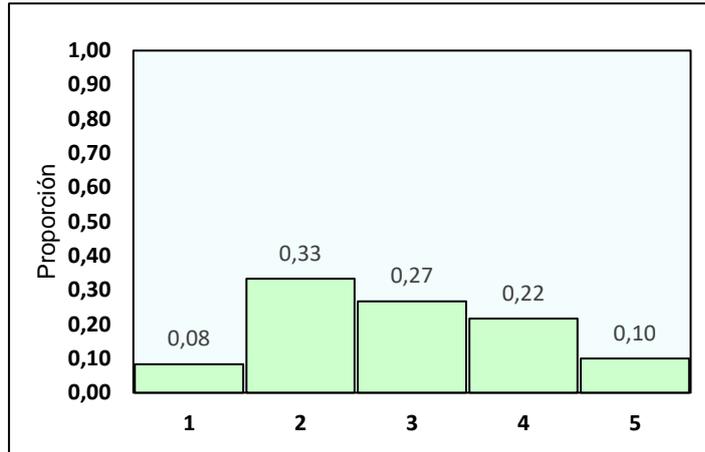
*Elaboración propia*

**Tabla 9. Estadísticas Descriptivas de pH**

pH	
Media	7,66
Error est. media	0,05
Desv.Estándar	0,53
Mínimo	6,56
Q1	7,21
Mediana	7,58
Q3	8,11
Máximo	8,87
Moda	8,42
N° Moda	3,00

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.5. Histograma de pH**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

### **Plantas acuáticas sumergidas**

El 95,8% de los cuerpos de agua no están cubiertos por plantas acuáticas o macrófitas, en los restantes, estas plantas ocupan hasta el 25% del lugar.

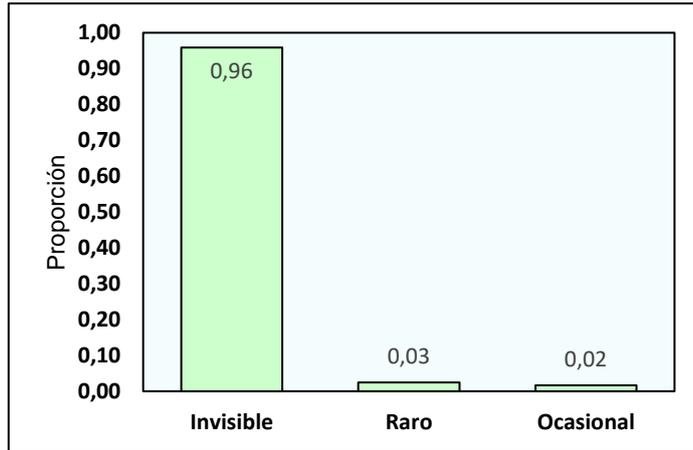
Véase *Tabla 10* y *Figura 3.6*.

**Tabla 10. Histograma de Plantas acuáticas sumergidas**

Plantas acuáticas sumergidas	Proporción
Invisible	0,958
Raro (1%-5%)	0,025
Ocasional (5%-25%)	0,017
Total	1,000

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.6. Histograma de Plantas acuáticas**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)  
Elaboración propia*

### **Velocidad de los cuerpos de agua**

La velocidad del río es muy importante ya que si estos son correntosos pueden degradar con mayor facilidad los desechos, lo que además ayuda a la limpieza de las aguas. De acuerdo a los resultados, el 72,5% de los puntos observados tienen una velocidad menor a 0,375 m/s. Véase *Tabla 11, Tabla 12 e Figura 3.7.*

**Tabla 11. Distribución de Frecuencias de Velocidad**

Velocidad	Proporción
1.[0,000 - 0,375)	0,725
2.[0,375 - 0,750)	0,233
3.[0,750 - 1,125)	0,025
4.[1,125 - 1,500)	0,017

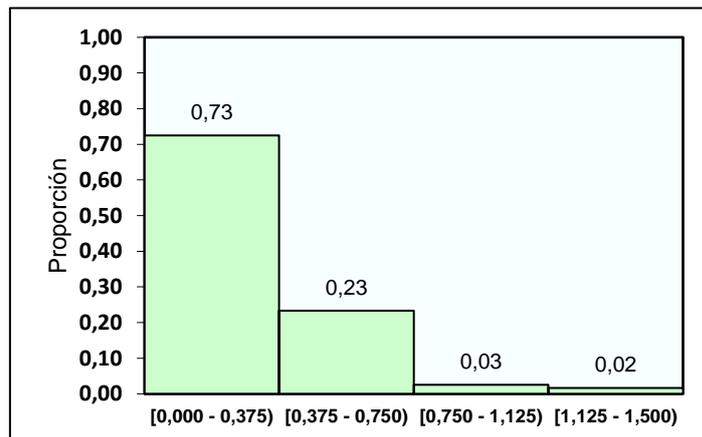
*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)  
Elaboración propia*

**Tabla 12. Estadísticas Descriptivas de Velocidad**

Velocidad	
Media	0,234
Error est. media	0,026
Desv.Estándar	0,281
Mínimo	0,000
Q1	0,000
Mediana	0,150
Q3	0,400
Máximo	1,500
Moda	0,000
N° Moda	44,000

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.7. Histograma de Velocidad**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

## Temperatura

La temperatura que predomina en estos lugares está entre 25,45 °C y 27,59 °C, representada por el 38,3%. La temperatura media es de 25,98 °C. El 75% de los cuerpos de aguas tiene una temperatura menor a 27,37 °C. Para más detalles *Tabla 13, Tabla 14 y Figura 3.8.*

**Tabla 13. Distribución de Frecuencias de Temperatura**

Temperatura	Proporción
1.[19.04 - 21.17)	0,042
2.[21.17 - 23.31)	0,058
3.[23.31 - 25.45)	0,292
4.[25.45 - 27.59)	0,383
5.[27.59 - 29.72)	0,175
6.[29.72 - 31.86)	0,025
7.[31.86 - 34.00]	0,025

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

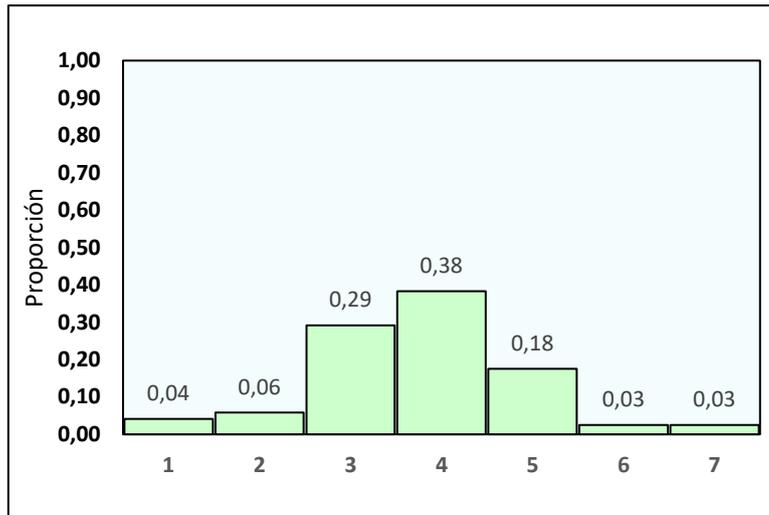
*Elaboración propia*

**Tabla 14. Estadísticas Descriptivas de Temperatura**

Temperatura	
Media	25,981
Error est. media	0,227
Desv.Estándar	2,484
Mínimo	19,036
Q1	24,754
Mediana	26,037
Q3	27,37
Máximo	34
Moda	25,64
N° Moda	3

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.8. Histograma de Temperatura**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

### 3.2 Análisis descriptivo bivariado

Se realizó el análisis bivariado de algunas de los factores importantes para la presencia de cada especie, a continuación, se muestran los resultados por cada factor.

#### **Ausencia y Presencia por especie**

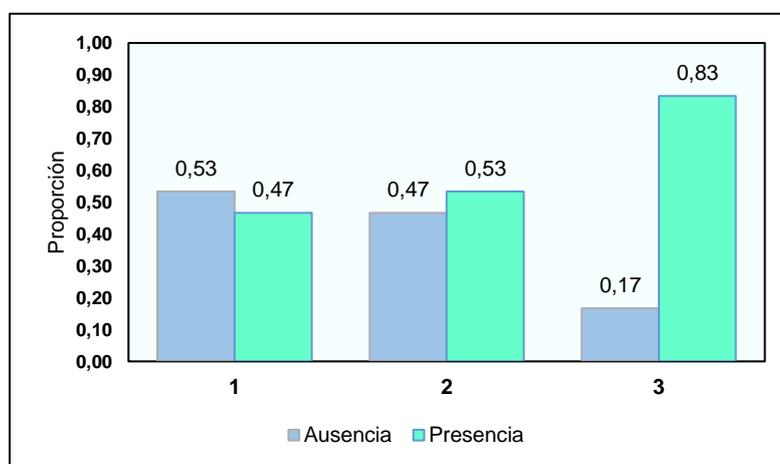
En este caso en particular, los resultados nos enseñan que la presencia de Baetidae (53%) y Chironomidae (83%) es notoria en comparación con su ausencia, mientras que en Hydropsychidae sucede lo contrario ya que su ausencia es del 74%. Más detalles en la *Tabla 15 y Figura 3.9.*

**Tabla 15. Distribución de Frecuencias de Presencia/Ausencia por Especie**

MacroInvertebrados	Ausencia	Presencia	Total
1.Baetidae	0,467	0,533	1,000
2.Chironomidae	0,167	0,833	1,000
3.Hydropsychidae	0,742	0,258	1,000

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.9. Histograma Presencia/Ausencia por Especie**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

### **Madera Muerta**

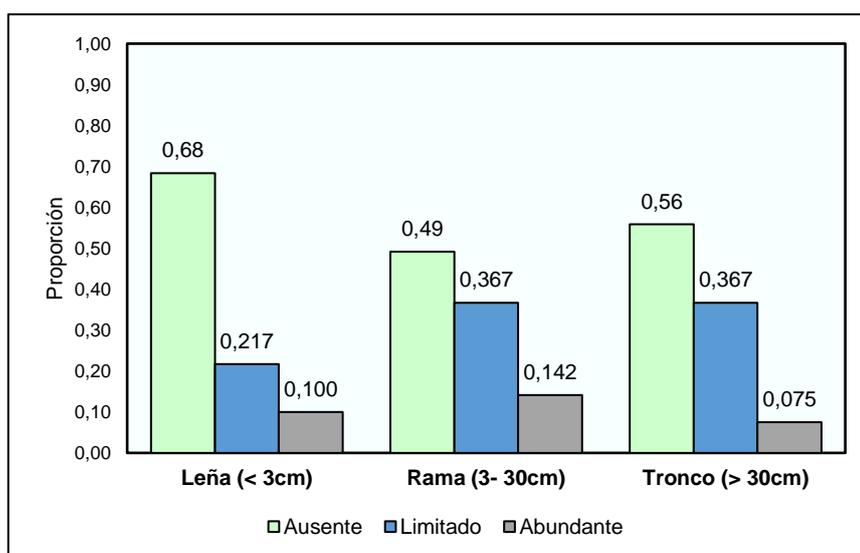
Según la clasificación de la madera muerta en este estudio, se detallan los siguientes resultados: ausencia de la leña (68,3%), ausencia de las ramas (49,2%) y ausencia de los troncos (55,8%). Más detalles en la *Tabla 16* y *Figura 3.10*.

**Tabla 16. Distribución de Frecuencias de la Madera Muerta**

Madera Muerta	Ausente	Limitado	Abundante	Total
Leña (< 3cm)	0,683	0,217	0,100	1,000
Rama (3- 30cm)	0,492	0,367	0,142	1,000
Tronco (> 30cm)	0,558	0,367	0,075	1,000

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.10. Histograma de la Madera Muerta**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

### Tipo de Sedimento

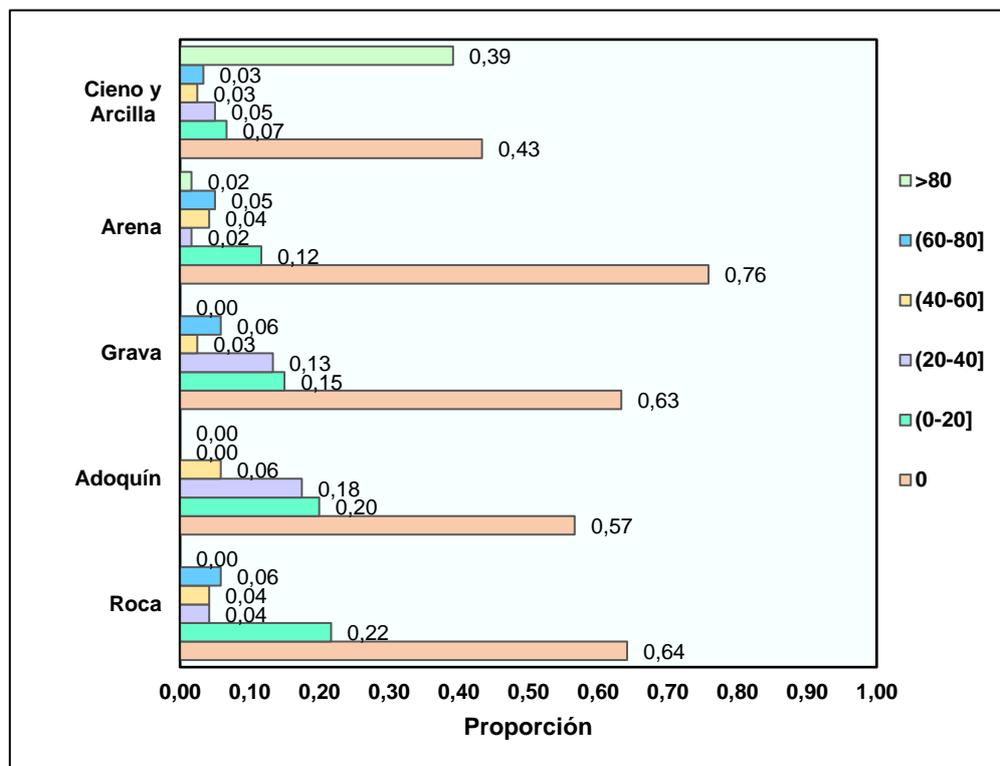
Los lugares muestreados en su mayoría no tienen ningún tipo de sedimento (0%), son pocos los que están cubiertos por más del 60% de su estructura por algún tipo de sedimento. Véase *Tabla 17* y *Figura 3.11*.

**Tabla 17. Distribución de Frecuencias de Porcentaje de cobertura por Tipo de Sedimento**

Tipo de sedimento	Porcentaje						Total
	0	(0-20]	(20-40]	(40-60]	(60-80]	>80	
Roca	0,642	0,217	0,042	0,042	0,058	0,000	1,000
Adoquín	0,567	0,200	0,175	0,058	0,000	0,000	1,000
Grava	0,633	0,150	0,133	0,025	0,058	0,000	1,000
Arena	0,758	0,117	0,017	0,042	0,050	0,017	1,000
Cieno y Arcilla	0,433	0,067	0,050	0,025	0,033	0,392	1,000

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia



**Figura 3.11. Histograma de Tipo de Sedimento por Porcentaje de cobertura**

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia

## Temperatura

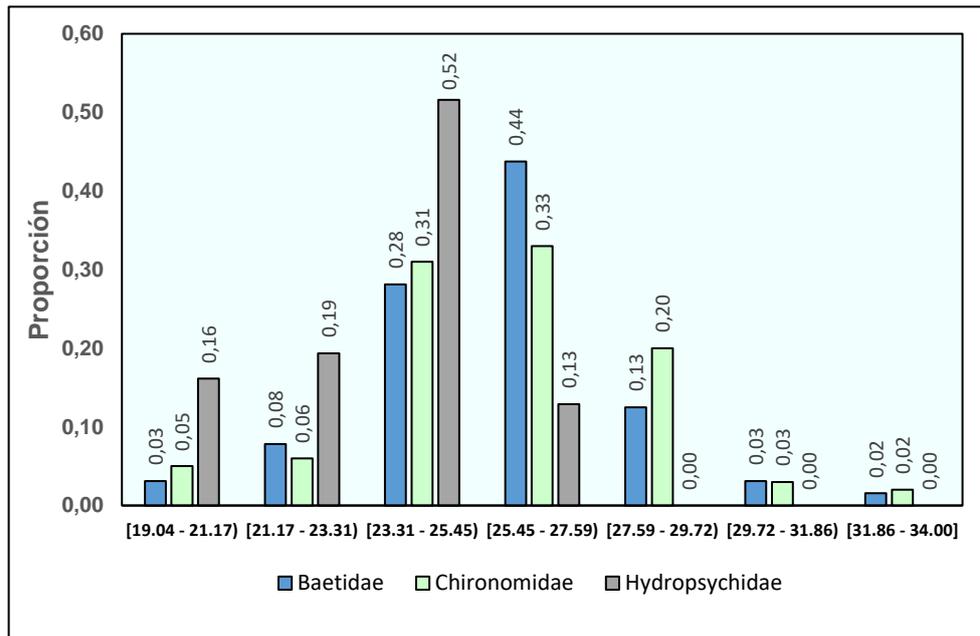
La mayoría de las especies de Baetidae (44%) y Chironomidae (33%) prefieren los lugares que tengan entre 25,45°C y 27,59°C, mientras más del 50% de Hydropsychidae prefiere que la temperatura sea mayor a 23,31°C y no más de 25,45°C. Véase *Tabla 18 e Figura 3.12*.

**Tabla 18. Distribución de Frecuencias de Temperatura por Especie**

Especie	Temperatura							
	1,00	2,00	3,00	4,00	5,00	6,00	7,00	Total
Baetidae	0,03	0,08	0,28	0,44	0,13	0,03	0,02	1,00
Chironomidae	0,05	0,06	0,31	0,33	0,20	0,03	0,02	1,00
Hydropsychidae	0,16	0,19	0,52	0,13	0,00	0,00	0,00	1,00

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia



**Figura 3.12. Histograma de Temperatura por Especie**

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia

## pH

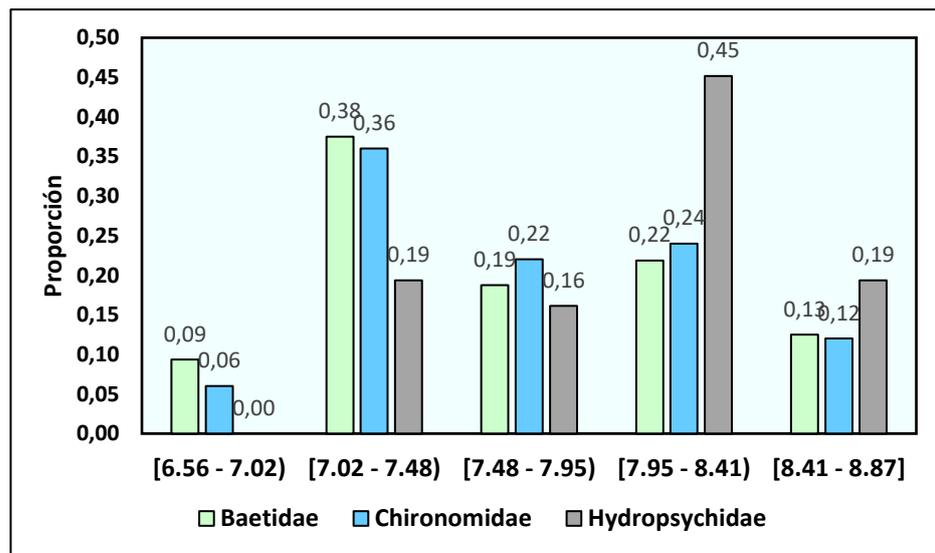
Las especies Baetidae y Chironomidae se presentan en los lugares con valores de pH en un intervalo de 7,02 y 7,48, sin embargo Hydropsychidae se desempeña mejor cuando este factor está entre 7,95 y 8,41. Véase *Tabla 19 y Figura 3.13*.

**Tabla 19. Distribución de Frecuencias de pH por Especie**

Especie	pH					Total
	1	2	3	4	5	
Baetidae	0,090	0,380	0,190	0,220	0,130	1,000
Chironomidae	0,060	0,360	0,220	0,240	0,120	1,000
Hydropsychidae	0,000	0,190	0,160	0,450	0,190	1,000

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



**Figura 3.13. Histograma de pH por Especie**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

## Velocidad

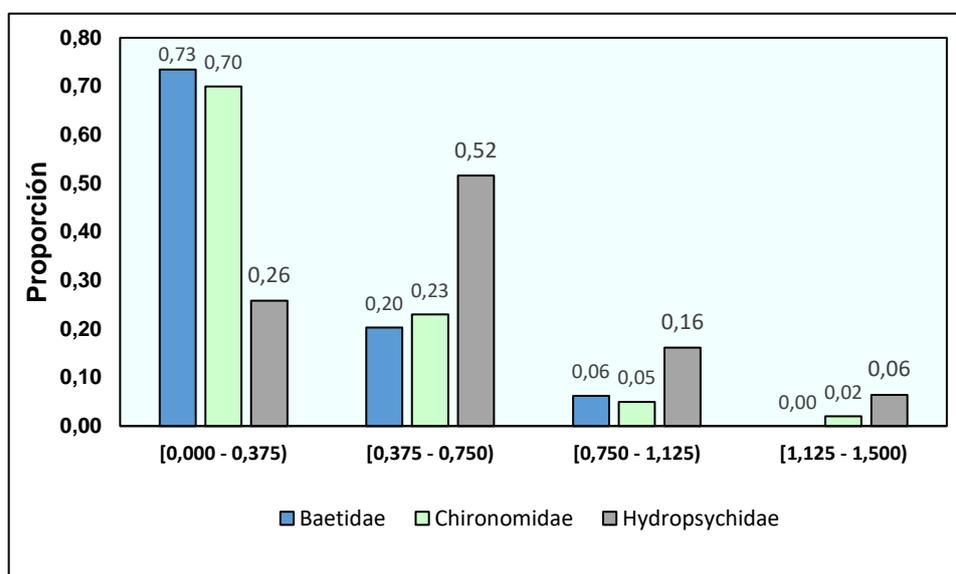
Las especies Baetidae (73%) y Chironomidae (70%) optan por cuerpos de agua que tienen una velocidad comprendida en el intervalo de 0 a 0,38 m/s, para el caso de Hydropsychidae esta prefiere estar en donde la velocidad es mayor a 0,37 y menor 0,75 m/s. Para más detalle *Tabla 20* y *Figura 3.14*.

**Tabla 20. Distribución de Frecuencias de Velocidad por Especie**

Especie	Velocidad				Total
	[0,000 - 0,375)	[0,375 - 0,750)	[0,750 - 1,125)	[1,125 - 1,500)	
Baetidae	0,730	0,200	0,060	0,000	1,000
Chironomidae	0,700	0,230	0,050	0,020	1,000
Hydropsychidae	0,260	0,520	0,160	0,060	1,000

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia



**Figura 3.14. Histograma de Velocidad por Especie**

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia

### 3.3 Análisis especializado

#### 3.3.1 Modelo GLM para la Especie Baetidae

El “mejor” modelo para la especie Baetidae es el siguiente:

$$Y_1 \sim \beta_0 + \beta_1 \text{Adoq} + \beta_2 \text{Nitrato} + \beta_3 \text{Varanch}$$

Donde

Adoq = Adoquín

Nitrato = Nitrato en el Nitrógeno

Varanch= Variación en el ancho

Al observar la *Tabla 21* correspondiente a los coeficientes del modelo propuesto, se observa en base al valor p que las variables seleccionadas son de gran importancia para este modelo.

Se puede observar que cuando el adoquín ocupa entre el 40% y el 60% del cuerpo de agua, aumenta la posibilidad de que esta especie esté presente, sin embargo cuando el adoquín se sitúa en un nivel 1 y 2 en el que ocupan hasta un 20%, disminuye la posibilidad de que estén presentes.

A mayor concentración de Nitrato, la especie desaparece. La variación del ancho del lugar de muestreo que favorece a la presencia de baetidae es el nivel 4, los niveles 2 y 3 también la favorecen, pero no tanto como el nivel 4 por otra parte, los niveles restantes provocan que esta especie se ausente.

**Tabla 21. Coeficientes del Modelo GLM para Baetidae**

Coeficientes	Estimado
Intercepto	1,010
Adoq1	-0,163
Adoq2	-1,349
Adoq3	0,836
Nitrato	-2,465
Varanch1	-0,249
Varanch2	1,071
Varanch3	0,455
Varanch4	2,046
Varanch5	-15,328

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Tabla 22*, se muestra la información del modelo que tiene un error del 40% de predicción. A pesar de que el criterio de información Akaike es bajo comparado con el modelo completo, el  $R^2$  anuncia que el modelo elegido explica el 57,7% de la variabilidad de los datos, es decir, no es un buen ajuste, pero es el que mejor explica la realidad de esta especie además de que al momento de elegir el modelo se tomaron en cuenta varios criterios.

**Tabla 22. Información del Modelo GLM de Baetidae**

Información del Modelo	
Porcentaje de error	40%
AIC	161,85
Devianza Nula	165,82
Devianza Residual	70,18
$R^2$	57,7%

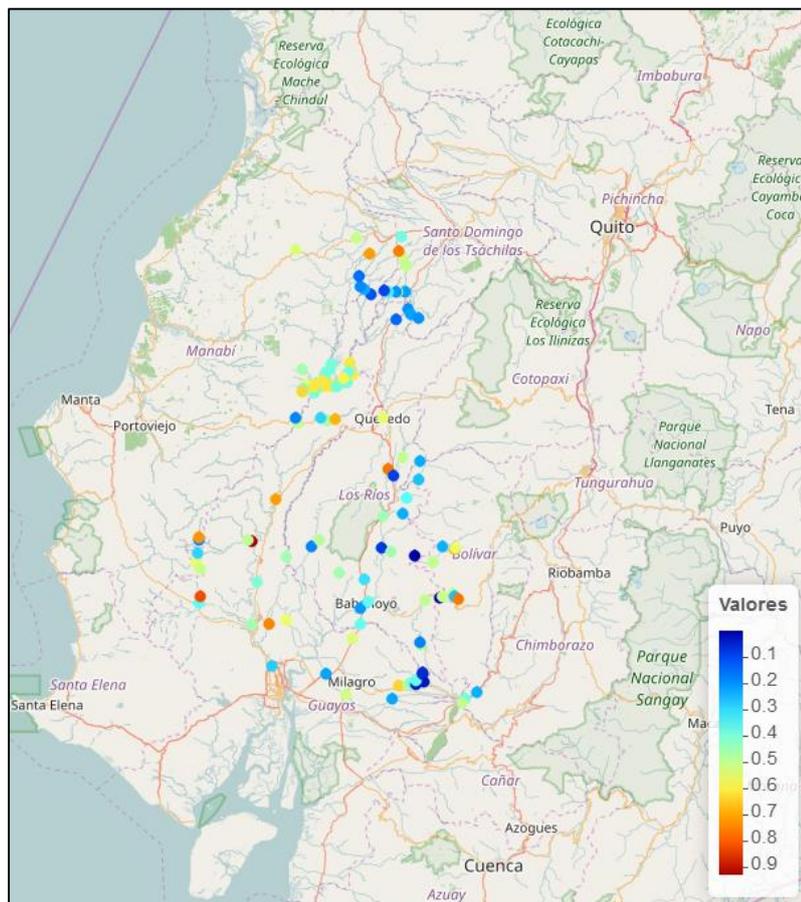
*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*



Utilizando la probabilidad ( $Y_1$ ) de que Baetidae esté presente en cada uno de los lugares de muestreo en base a las variables expuestas anteriormente, se muestra la *Figura 3.15*, donde 0,1 es cuando la especie tiene menos probabilidad de estar presente y 0,9 es cuando la especie tiene mayor probabilidad de presencia. Se puede decir que están cerca de las provincias de Santo Domingo de Tsáchilas y de Los Ríos.

### Mapa de Error de Baetidae



**Figura 3.16. Mapa de Error de la especie Baetidae**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Figura 3.16* se comparan los valores reales de Baetidae comparados con los valores estimados por el modelo, la escala de valores indica el porcentaje de error, donde 0,1 es cuando el error es del 10% y 0,9 cuando el error es del 90%. Este modelo en gran parte de los sitios muestreados el error es bajo y pocos son los que tienen un error mayor al 80%.

### 3.3.2 Modelo GLM para la Especie Chironomidae

$$Y_2 \sim \beta_0 + \beta_1 Turb + \beta_2 pH + \beta_3 AcSum$$

Donde:

Turb: Turbiedad

pH: pH

AcSum: Plantas acuáticas sumergidas

Al observar la Tabla 24, correspondiente a los coeficientes del modelo propuesto, podemos observar en base al valor p que las variables seleccionadas son de gran importancia para este modelo.

Al aumentar las unidades de Turbidez, se incrementa la probabilidad de que esta especie desaparezca, sin embargo, a mayor nivel de pH aumenta la posibilidad de que esta especie este presente.

La probabilidad de presencia de este organismo aumenta significativamente cuando los cuerpos de agua son cubiertos en un intervalo del 1% al 5% de plantas acuáticas sumergidas, sin embargo, cuando los cuerpos de agua están cubiertos en un intervalo de 5% al 25% por las macrófitas sumergidas disminuye la presencia.

**Tabla 24. Coeficientes del Modelo GLM para Chironomidae**

Coeficientes	Estimado
Intercepto	-7,862
Turb	-0,057
pH	1,332
AcSum1	13,969
AcSum2	-2,697

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Tabla 25* se muestra la información del modelo que tiene un error del 15% de predicción. El criterio de información Akaike es bajo comparado con el modelo completo, el  $R^2$  anuncia que el modelo elegido explica el 63,6% de la variabilidad de los datos es decir, no es un buen ajuste, pero es el que mejor explica la realidad de esta especie además de que fue elegido tomando en cuenta varios criterios.

**Tabla 25. Información del Modelo GLM de Chironomidae**

Información del Modelo	
Porcentaje de error	15%
AIC	98,22
Devianza Nula	110,5
Devianza Residual	40,22
$R^2$	63,60%

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

La *Tabla 26* muestra que la predicción de ausencia de la especie no es confiable debido a que el 85% de los datos ausentes están mal ubicados dentro de la tabla de confusión.

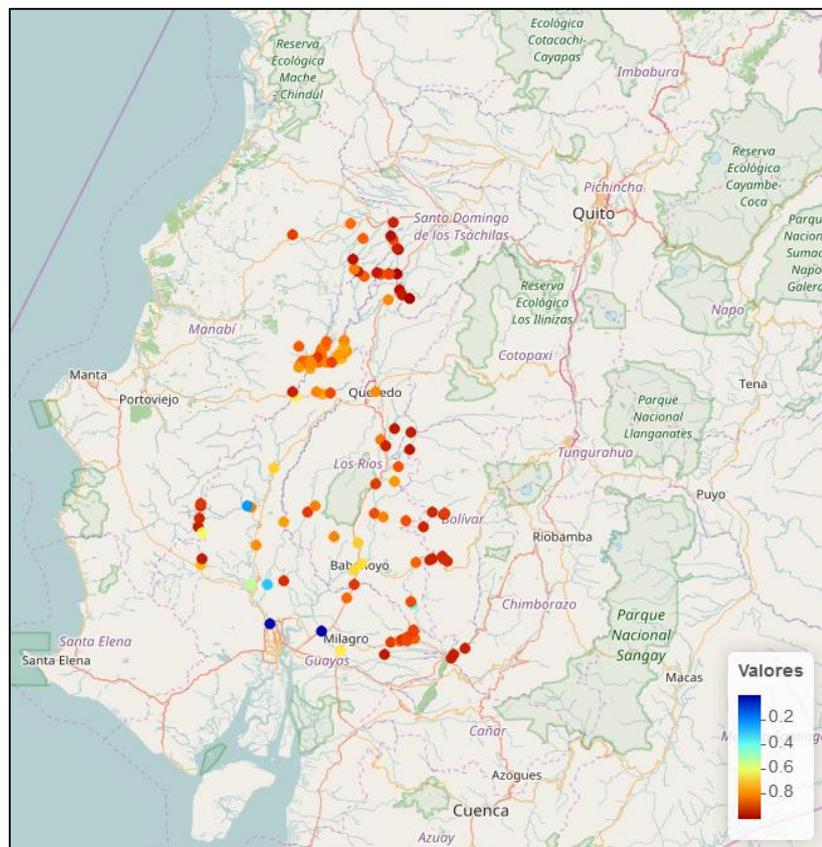
**Tabla 26. Tabla de Confusión para Chironomidae**

Chironomidae	Ausencia	Presencia
Ausencia	0,15	0,85
Presencia	0,01	0,99

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia

**Mapa de Probabilidad de Chironomidae**



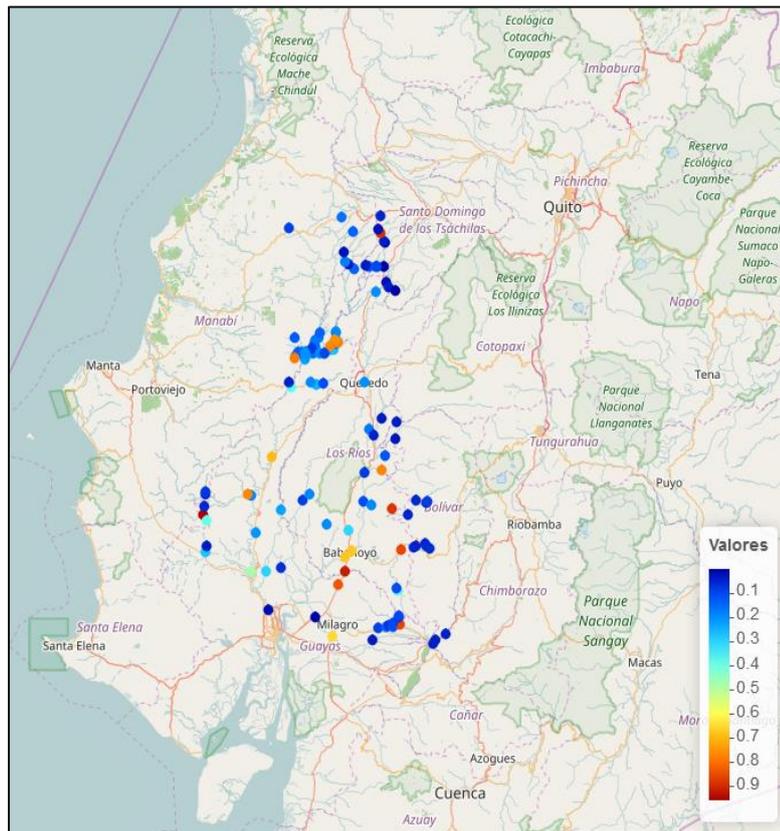
**Figura 3.17. Mapa de Probabilidad de la especie Chironomidae**

Fuente: Centro del Agua y Desarrollo Sustentable (CADS)

Elaboración propia

Utilizando la probabilidad ( $Y_2$ ) de que Chironomidae esté presente en cada uno de los lugares de muestreo en base a las variables expuestas anteriormente, se muestra la *Figura 3.17*, donde 0,2 es cuando la especie tiene menos probabilidad de estar presente y 0,8 es cuando la especie tiene mayor probabilidad de presencia. Este organismo es el más difícil de situar en algún lugar en específico debido a que en más del 50% de los sitios de muestreos tiene más del 80% de probabilidad de estar presente.

### Mapa de Error de Chironomidae



**Figura 3.18. Mapa de Error de la especie Chironomidae**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Figura 3.18* se comparan los valores reales de Chironomidae comparados con los valores estimados por el modelo, la escala de valores indica el porcentaje de error, donde 0,1 es cuando el error es del 10% y 0,9 cuando el error es del 90%. Gran parte de los sitios muestreados tienen un error menor al 30% y pocos son los que tienen un error mayor al 80%.

### 3.3.3 Modelo GLM para Especie Hydropsychidae

$$Y_3 \sim \beta_0 + \beta_1 Vel + \beta_2 Temp + \beta_3 Tron + \beta_4 Grav$$

Donde:

Temp: Temperatura

Tron: Tronco

Grav: Grava

Al observar la Tabla 27, correspondiente a los coeficientes del modelo propuesto, se puede observar en base al valor p que las variables seleccionadas son de gran importancia para este modelo.

La Velocidad de la corriente de las masas de agua influyen de manera positiva sobre la probabilidad de presencia de la especie ya que al aumentar la velocidad aumenta la posibilidad de presencia de la especie, esto es cierto, debido a que esto provoca que el agua esté libre de desechos.

La Temperatura influye negativamente sobre la probabilidad de presencia de Hydropsychidae es decir, que a medida que la temperatura aumenta la presencia disminuye.

La aparición de estos organismos aumenta en el caso de que haya troncos ya sea en cantidades limitadas o abundantes y cuando los cuerpos de agua contienen en un intervalo del 40% al 60% de grava disminuye lentamente la probabilidad de aparición.

**Tabla 27. Coeficientes del Modelo GLM de Hydropsychidae**

Coeficientes	Estimado
Intercepto	22,548
Vel	5,551
Temp	-1,098
Tron1	0,825
Tron2	2,942
Grav1	3,004
Grav2	1,542
Grav3	-15,094
Grav4	1,834

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Tabla 28* se muestra la información del modelo que tiene un error del 14% de predicción. El criterio de información Akaike es bajo comparado con el modelo completo, el  $R^2$  anuncia que el modelo elegido explica el 70,5% de la variabilidad de los datos es decir, tiene buen ajuste y es razonablemente bueno.

**Tabla 28. Información del Modelo GLM de Hydropsychidae**

Información del Modelo	
Porcentaje de error	14%
AIC	65,42
Devianza Nula	137,11
Devianza Residual	40,42
$R^2$	70,50%

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

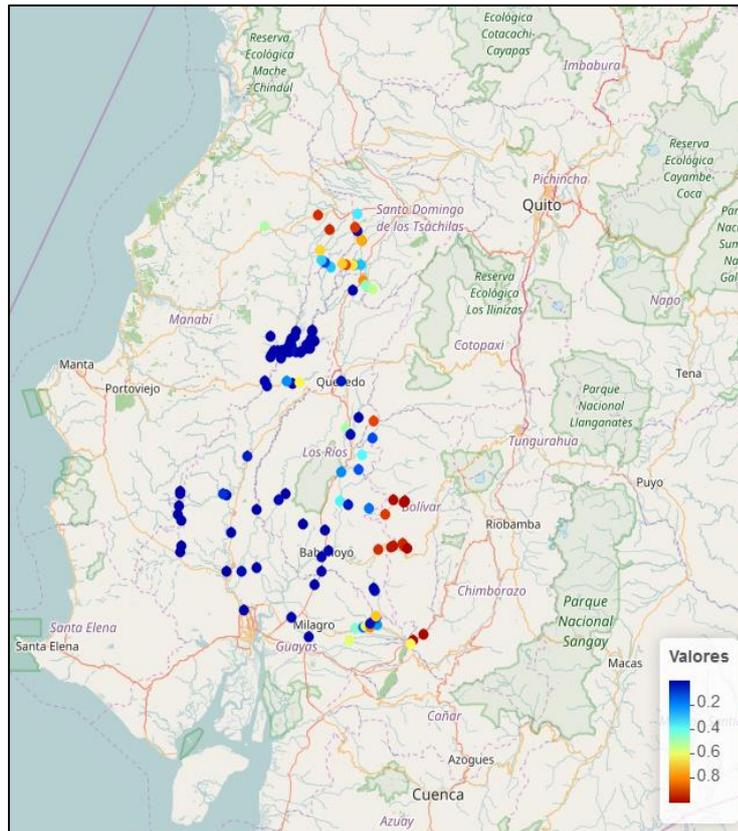
La *Tabla 29* muestra que la predicción de este modelo es buena ya que el 87,5% de los datos se clasifican correctamente.

**Tabla 29. Tabla de Confusión de Hydropsychidae**

Hydropsychidae	Ausencia	Presencia
Ausencia	0,91	0,09
Presencia	0,16	0,84

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)  
Elaboración propia*

### Mapa de Probabilidad de Hydropsychidae

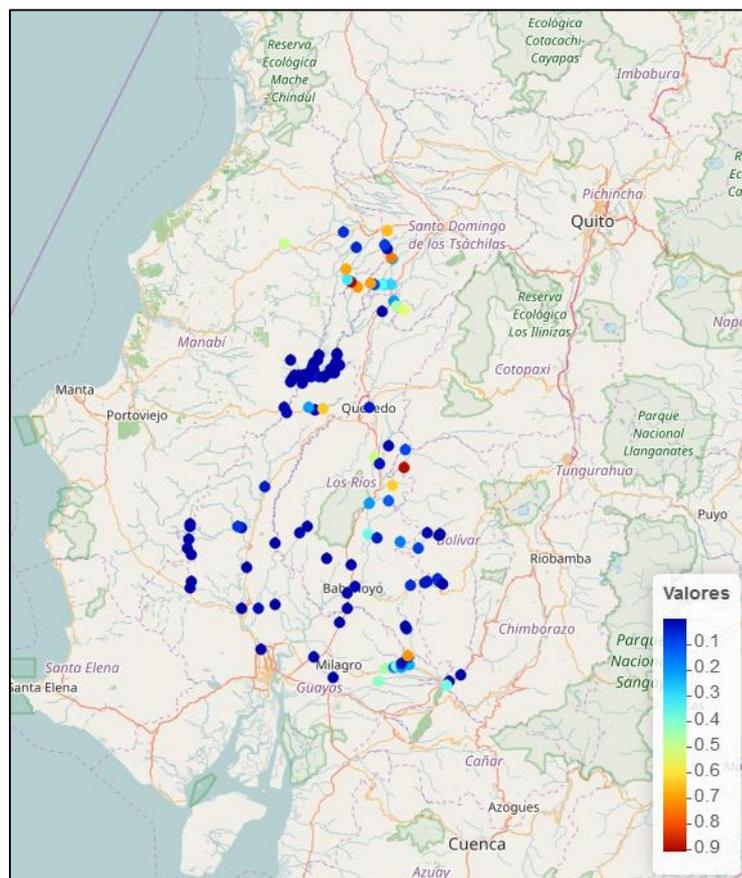


**Figura 3.19. Mapa de Error de la especie Hydropsychidae**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)  
Elaboración propia*

Utilizando la probabilidad ( $Y_3$ ) de que Hydropsychidae esté presente en cada uno de los lugares de muestreo en base a las variables expuestas anteriormente, se muestra la *Figura 3.19*, donde 0,2 es cuando la especie tiene menos probabilidad de estar presente y 0,8 es cuando la especie tiene mayor probabilidad de presencia. De acuerdo a los resultados este organismo se lo puede localizar cerca de la provincia de Santo Domingo de los Tsáchilas y de la provincia del Bolívar, debido a que tiene una probabilidad de presencia mayor al 80%.

### Mapa de Error de Hydropsychidae



**Figura 3.20. Mapa de Error de la especie Hydropsychidae**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Figura 3.20* se comparan los valores reales de Hydropsychidae comparados con los valores estimados por el modelo, la escala de valores indica el porcentaje de error, donde 0,1 es cuando el error es del 10% y 0,9 cuando el error es del 90%. La mayoría de los sitios muestreados tienen un error menor al 10% y pocos son los que tienen un error mayor al 65%.

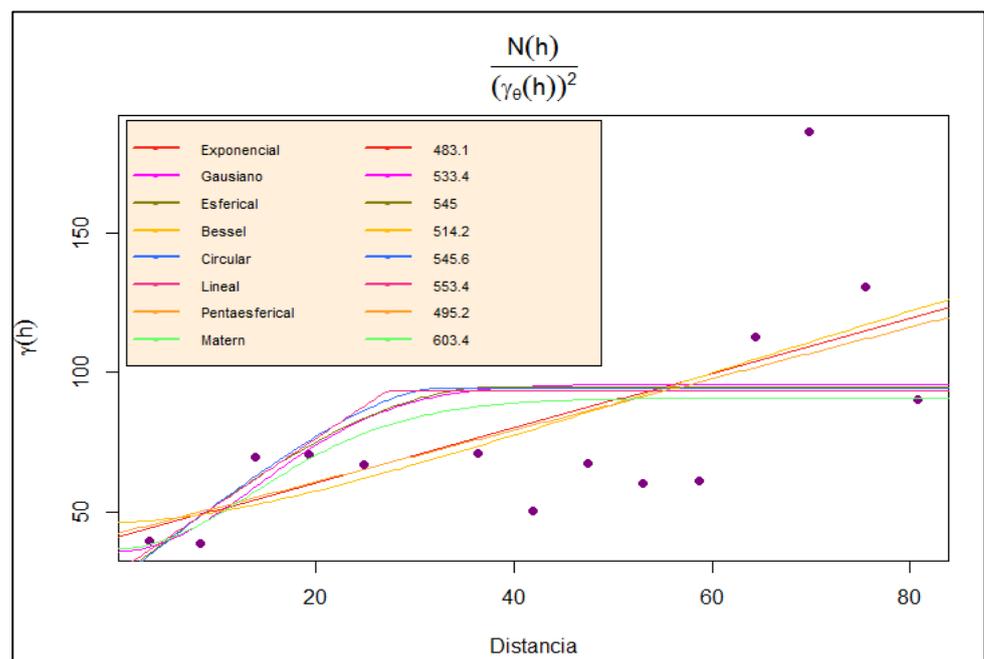
### 3.3.4 Kriging Universal

El método de Kriging Universal para la abundancia de cada especie en concordancia con la variable *Elevación*, se exploraron cada uno de los variogramas para luego escoger el que mejor se ajusta al modelo que se ha propuesto. El krigeado en todas las especies se realizó con un radio de 1,5 Km alrededor de cada uno de los puntos de estudio.

#### Especie Baetidae

Modelo:  $Y_B \sim \text{Elevación}$ , donde  $Y_B$  = Cantidad de especies Baetidae

La *Figura 3.21* muestra que el modelo de variograma que mejor se ajusta es el Exponencial, debido a que la varianza del error es menor en comparación a los otros modelos aplicados.

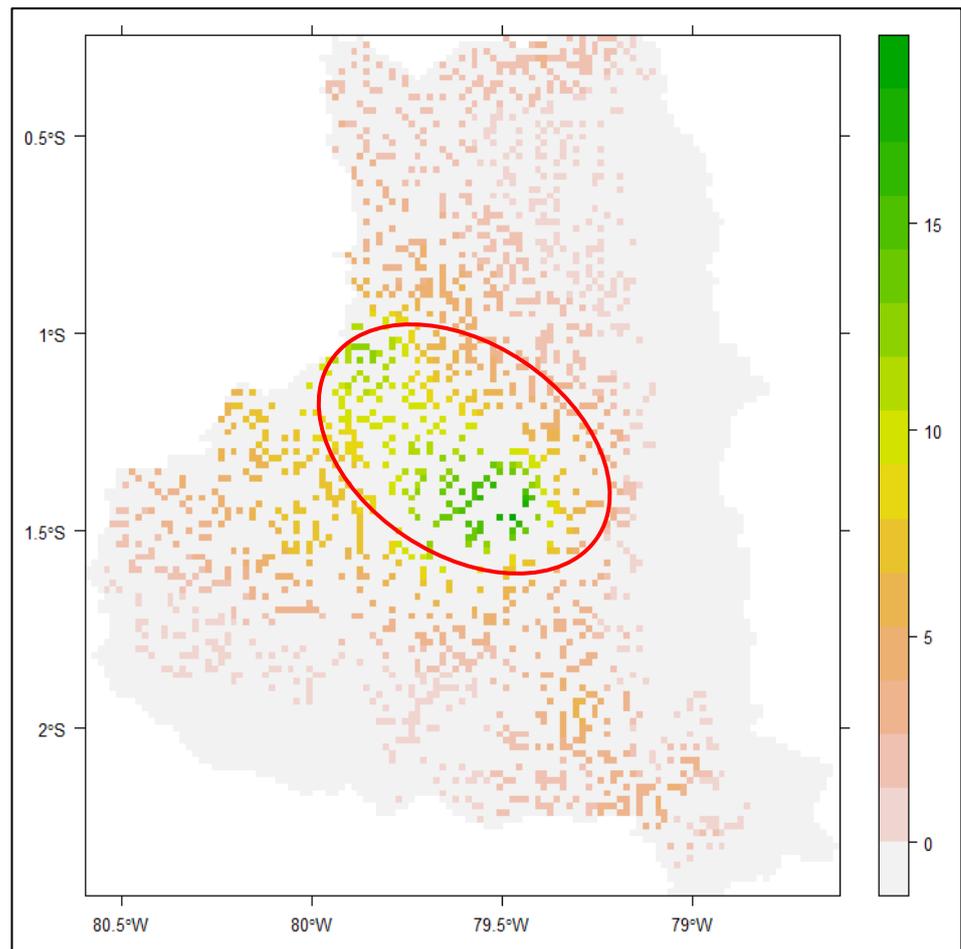


**Figura 3.21. Semivariograma de la especie Baetidae**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

Luego de hallar el modelo de variograma, se procede con su aplicación y con ello cumplir con uno de los fines principales de esta memoria, la construcción de una malla que provea los valores de abundancia en varios puntos desconocidos, es decir que no han sido muestreados. La región marcada con el óvalo rojo es donde según Kriging puede existir más de 10 unidades de esta especie. Véase la *Figura 3.22*.

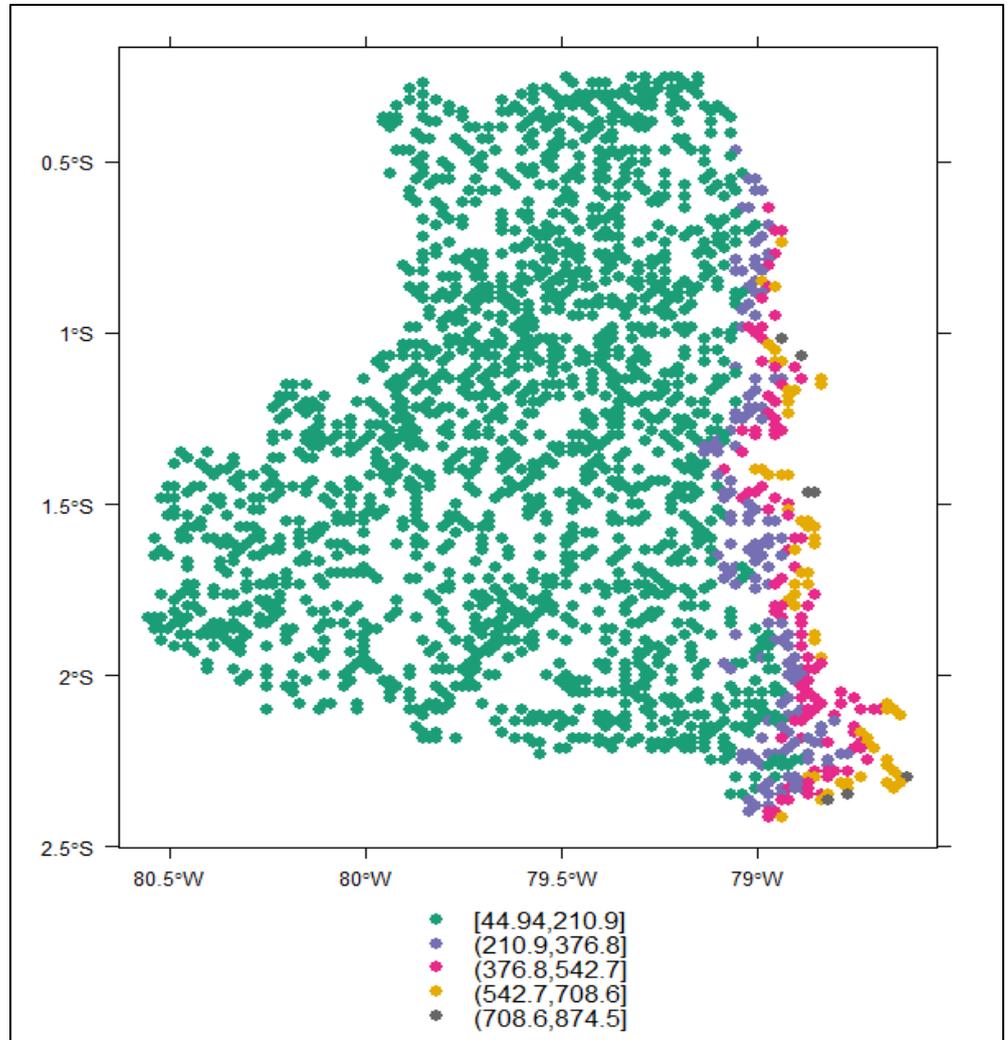


**Figura 3.22. Mapa de riqueza de la especie Baetidae generado con Kriging Universal**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Figura 3.23* se observa la distribución de la varianza de este modelo en el espacio geográfico de la cuenca del río Guayas.



**Figura 3.23. Mapa de Varianza de la especie Baetidae generado por Kriging Universal**

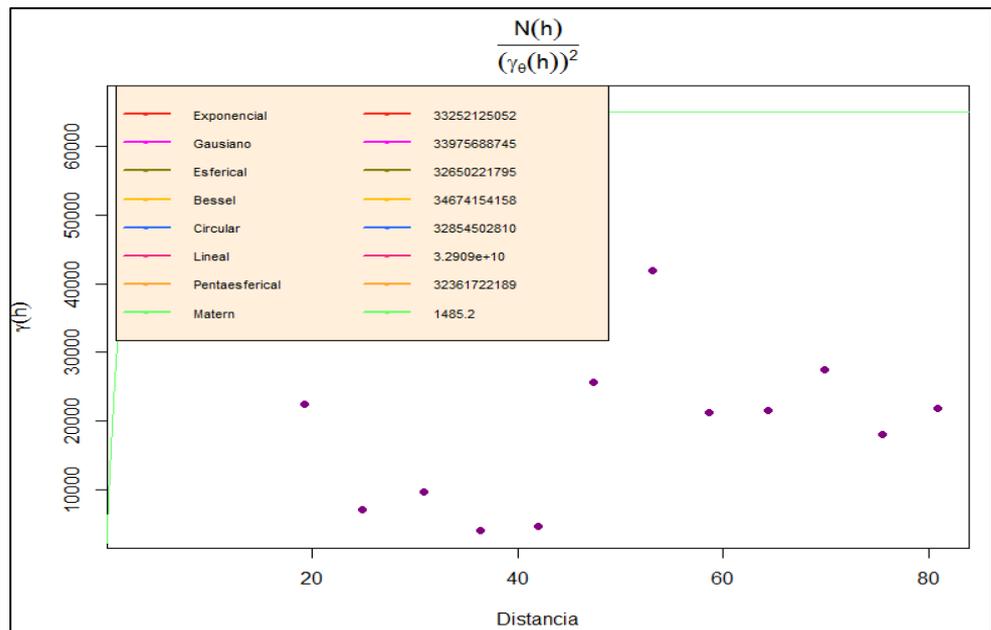
*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

## Especie Chironomidae

Modelo:  $Y_C \sim \text{Elevación}$ , donde  $Y_C$  = Cantidad de especies Chironomidae

La *Figura 3.24* muestra que el modelo de variograma que mejor se ajusta es el Matern, debido a que la varianza del error es menor en comparación a los otros modelos aplicados.

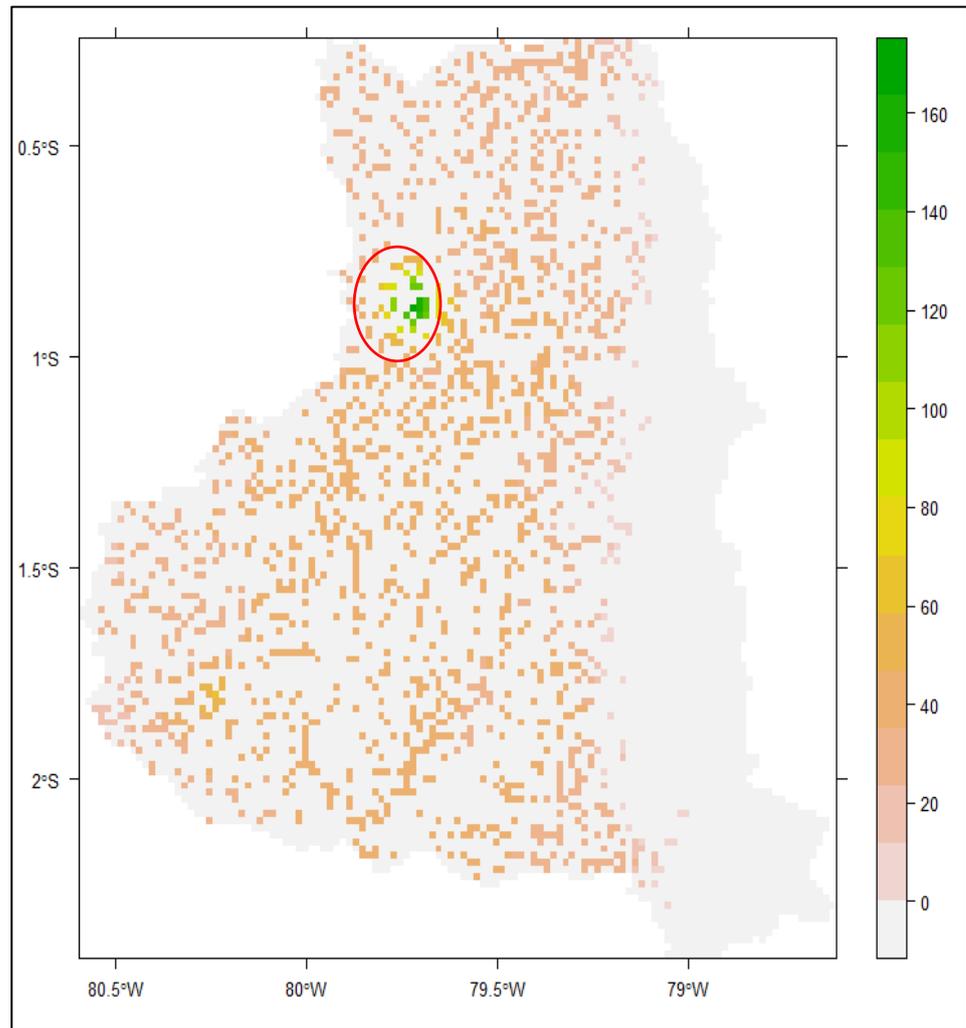


**Figura 3.24. Semivariograma de la especie Chironomidae**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

Luego de hallar el modelo de variograma, se procede con su aplicación y la construcción de una malla que provea los valores de abundancia en varios puntos desconocidos, es decir que no han sido muestreados. La región marcada con el óvalo rojo es donde según Kriging puede existir más de 100 unidades de esta especie. Véase la *Figura 3.25*.

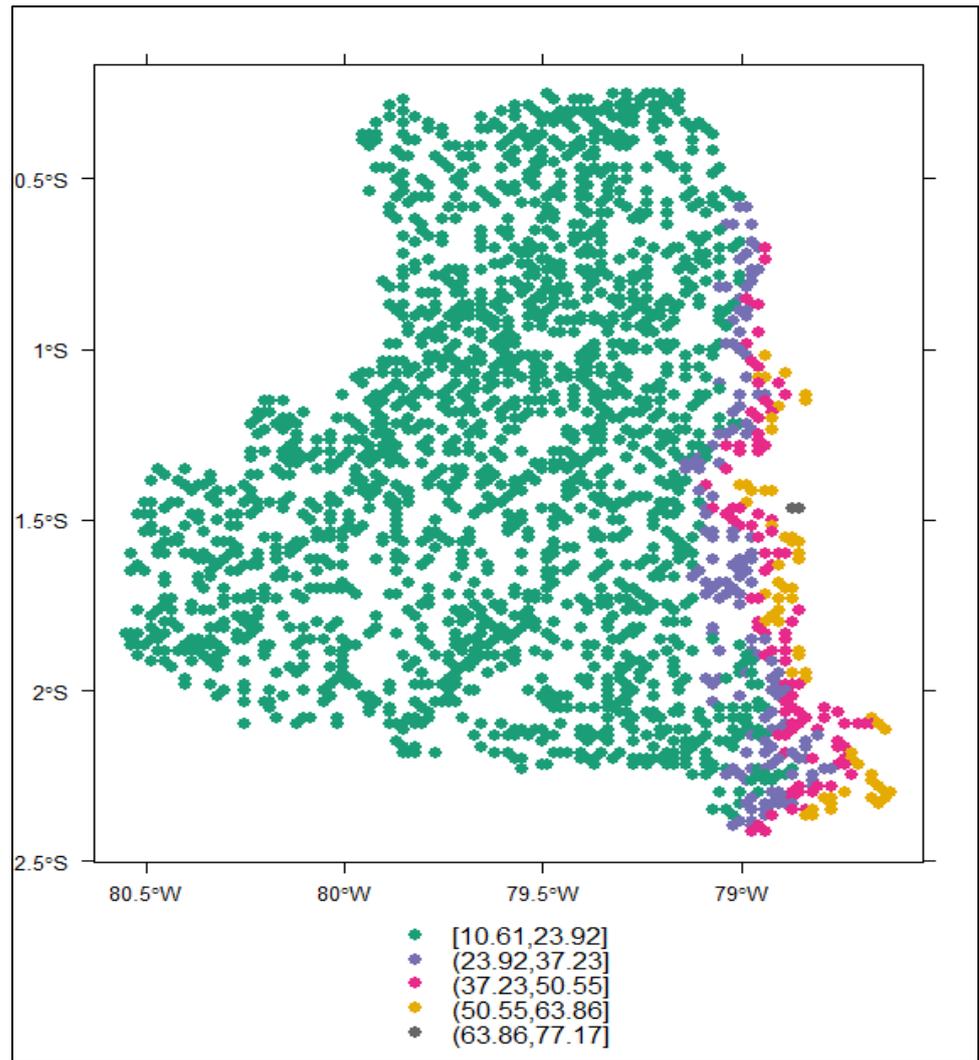


**Figura 3.25. Mapa de riqueza de la especie Chironomidae generado con Kriging Universal**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Figura 3.26* se observa la distribución de la varianza en el espacio geográfico de la cuenca del río Guayas.



**Figura 3.26. Mapa de Varianza de la especie Chironomidae generado con Kriging Universal**

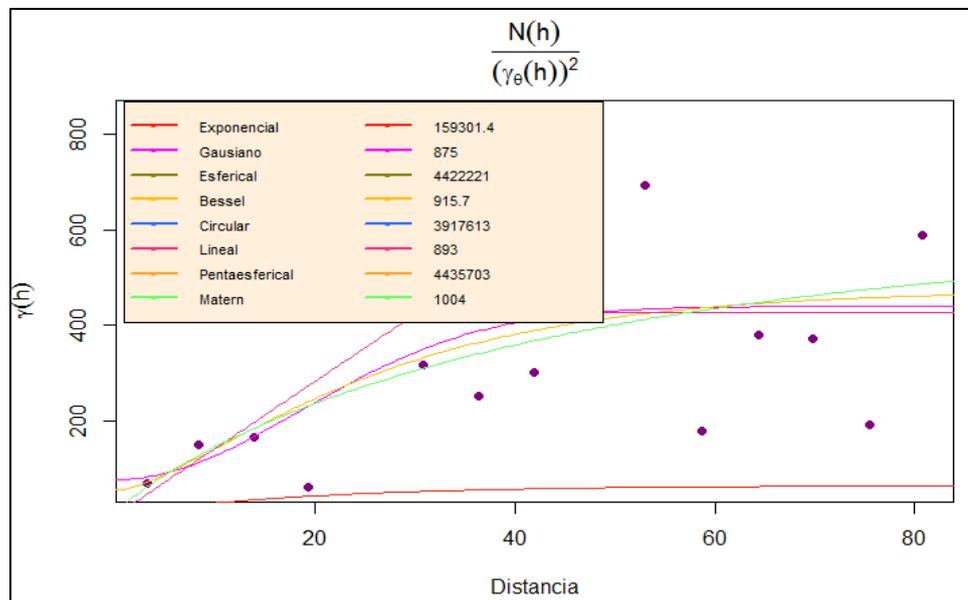
*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

## Especie Hydropsychidae

Modelo:  $Y_H \sim \text{Elevación}$ , donde  $Y_H$  = Cantidad de especies Hydropsychidae

La *Figura 3.27* muestra que el modelo de variograma que mejor se ajusta es el Gausiano, debido a que la varianza del error es menor en comparación a los otros modelos aplicados.



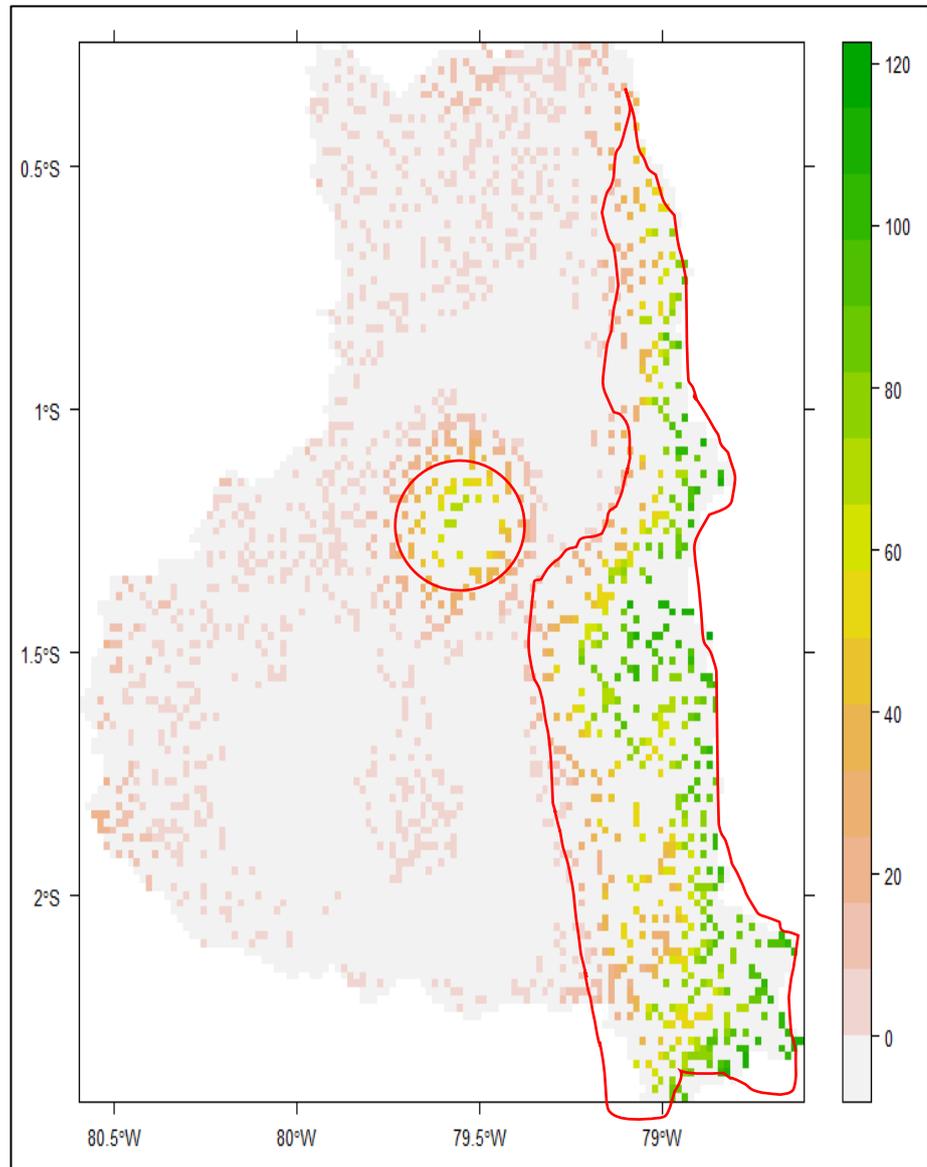
**Figura 3.27. Semivariograma de la especie Hydropsychidae**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

Luego de hallar el modelo de variograma, se procede con su aplicación y la construcción de una malla que provea los valores de abundancia en varios puntos desconocidos, es decir que no han sido muestreados. Las regiones marcadas con las formas rojas son las

áreas en donde según Kriging Universal pueden existir más de 60 unidades de esta especie. Véase *Figura 3.28*.

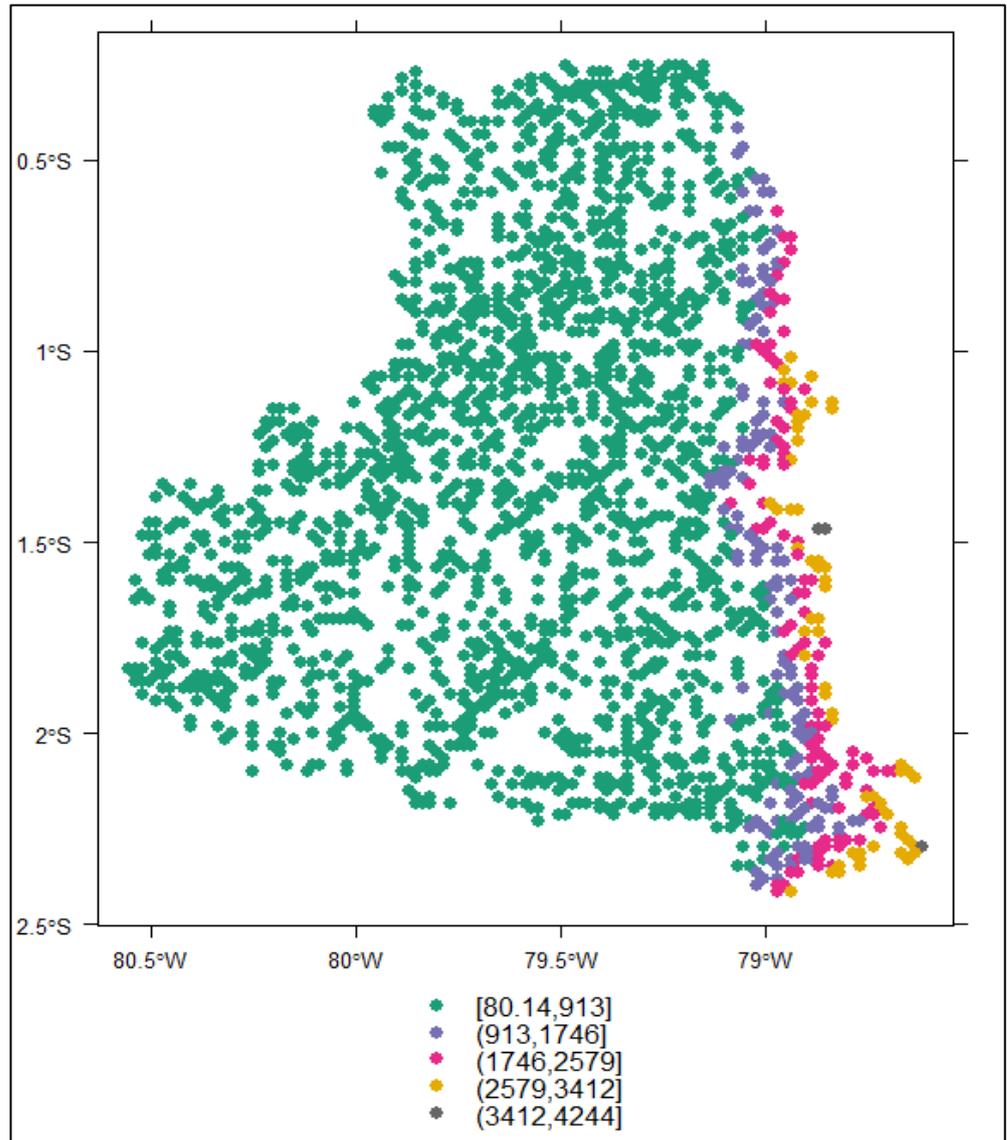


**Figura 3.28. Mapa de riqueza de la especie *Hydropsychidae* generado con Kriging Universal**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

En la *Figura 3.29* se observa la distribución de la varianza en el espacio geográfico de la cuenca del río Guayas.



**Figura 3.29. Mapa de varianza de la especie Hydropsychidae generado con Kriging Universal**

*Fuente: Centro del Agua y Desarrollo Sustentable (CADS)*

*Elaboración propia*

# CAPÍTULO 4

## 4. DISCUSIONES Y CONCLUSIONES

En esta sección se resalta la información más importante para la toma de decisiones. Cabe recalcar que los datos fueron evaluados mediante equipos especializados y profesionales en este campo.

### 4.1 Conclusiones

En el estudio descriptivo se encontró que las especies Baetidae y Chironomidae comparten algunos factores, es decir se desarrollan en ambientes con características similares.

En el análisis bivariado se evidenció que Baetidae y Chironomidae se preservan en aguas mansas. La especie Hydropsychidae se presenta cuando los niveles de pH son mayores a 7,95 siendo el valor máximo 8,87.

Los Modelos GLM indican que las especies Baetidae y Chironomidae son organismos difíciles de modelar linealmente ya que, el porcentaje de explicación de la variabilidad real de los datos es baja, mientras que en Hydropsychidae se hallaron características que definen el ambiente propicio para el desarrollo de esta especie.

El Kriging Universal expresa que Baetidae se desarrolla en la parte centro-oeste de la cuenca del río Guayas, Chironomidae un poco más arriba del área centro-oeste e Hydropsychidae en la parte central y casi toda la orilla de la parte este.

## 4.2 Recomendaciones

El proceso de muestreo para cualquier tipo de estudio se debe definir de tal manera que la cantidad de ausencia y presencia de las especies sean iguales, para evitar el sesgo.

Se recomienda utilizar otros Modelos que puedan explicar con mayor precisión y fiabilidad el entorno en el que se desarrollan estas especies, tales como: el Modelo Aditivo Generalizado (GAM).

Mientras más exploraciones se realicen en las aguas superficiales y en estos organismos, ayudaremos en mayor medida el manejo de este recurso y así garantizar la calidad del agua.

## BIBLIOGRAFÍA

- Aguirre Andrade, J. (2011). *Validación de los indicadores biológicos (Macroinvertebrados) para el monitoreo de la cuenca del río Yanuncay*. Recuperado el 2016, de <http://dspace.ups.edu.ec/bitstream/123456789/1197/14/UPS-CT002208.pdf>
- Alonso et al. (2001). Regresión Logística. Un ejemplo de su uso en Endocrinología. *Revista Cubana de Endocrinología*, 12(1). Obtenido de [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1561-29532001000100007](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1561-29532001000100007)
- ArcGIS Pro. (s.f.). *Esri: GIS Mapping Software, Spatial Data Analytics & Location Platform*. Recuperado el 2017, de <https://pro.arcgis.com/es/pro-app/tool-reference/3d-analyst/how-kriging-works.htm>
- Coello. (s.f.). *Corporación Autónoma Regional de Tolima*. Recuperado el 2016, de [https://www.cortolima.gov.co/sites/default/files/images/stories/centro\\_documentos/pom\\_coello/diagnostico/apendices/invertebrados/fichas\\_efemeropteros.pdf](https://www.cortolima.gov.co/sites/default/files/images/stories/centro_documentos/pom_coello/diagnostico/apendices/invertebrados/fichas_efemeropteros.pdf)
- Coello. (s.f.). *Corporación Autónoma Regional del Tolima*. Recuperado el 2016, de [https://www.cortolima.gov.co/sites/default/files/images/stories/centro\\_documentos/pom\\_coello/diagnostico/apendices/invertebrados/fichas\\_trichoptera.pdf](https://www.cortolima.gov.co/sites/default/files/images/stories/centro_documentos/pom_coello/diagnostico/apendices/invertebrados/fichas_trichoptera.pdf)
- Damanik, M., Everaert, G., Eurie, M., Tien, T., Lock, K., Sasha, P., . . . Goethals, P. (2016). Generalized Linear Models to Identify Key Hydromorphological and Chemical Variables Determining the Occurrence of Macroinvertebrates in the Guayas River Basin (Ecuador). *Water*.
- Díaz, M., & Casar, R. (2009). *Geoestadística Aplicada. Estimación Espacial*. México: Universidad Nacional Autónoma de México. Obtenido de

[http://mmc2.geofisica.unam.mx/cursos/geoest/Presentaciones/CG6\\_2009.pdf](http://mmc2.geofisica.unam.mx/cursos/geoest/Presentaciones/CG6_2009.pdf)

Gallardo, A. (2008). Métodos Geoestadísticos para el análisis de datos ecológicos espacialmente explícitos. En *Ecología Espacial*. España. Recuperado el 2017, de [http://www.escet.urjc.es/biodiversos/espa/investigacion/intercambio/doc/publicaciones/Ecologia\\_espacial\\_book\\_Cap6.pdf](http://www.escet.urjc.es/biodiversos/espa/investigacion/intercambio/doc/publicaciones/Ecologia_espacial_book_Cap6.pdf)

Giraldo Henao, R. (s.f.). *Introducción a la Geoestadística*. Bogotá: Universidad Nacional de Colombia. Recuperado el 2017, de [ftp://ftp.ciat.cgiar.org/DAPA/projects/Cursos\\_Talleres/Curso\\_R/DOCUMENTOS/LIBRO%20DE%20GEOESTADISTICA.pdf](ftp://ftp.ciat.cgiar.org/DAPA/projects/Cursos_Talleres/Curso_R/DOCUMENTOS/LIBRO%20DE%20GEOESTADISTICA.pdf)

González, K. (s.f.). *Modelos Lineales Generalizados*. Recuperado el 2016, de Universidad Nacional Mayor de San Marcos: [http://sisbib.unmsm.edu.pe/bibvirtualdata/tesis/Basic/Gonzales\\_K\\_K/cap1.pdf](http://sisbib.unmsm.edu.pe/bibvirtualdata/tesis/Basic/Gonzales_K_K/cap1.pdf)

Hernández-Sori, L. (2014). *Estudios ecotoxicológicos en diferentes bioindicadores ambientales del bioplaguicida Tricosave-34*. Santa Clara. Recuperado el 2016, de [https://www.researchgate.net/publication/283015416\\_Estudios\\_ecotoxicologicos\\_en\\_diferentes\\_bioindicadores\\_ambientales\\_del\\_bioplaguicida\\_Tricosave-34](https://www.researchgate.net/publication/283015416_Estudios_ecotoxicologicos_en_diferentes_bioindicadores_ambientales_del_bioplaguicida_Tricosave-34)

Muñoz Tolosa, L. (2015). *Uso de Kriging Universal en la Simulación Condicional de Leyes*. Obtenido de Universidad de Chile: <http://repositorio.uchile.cl/bitstream/handle/2250/134975/Uso-de-Kriging-universal-en-la-simulacion-condicional-de-leyes.pdf?sequence=1>

Pintos, S. (2012). *Variograma Matern*. Recuperado el 2016, de Petrociencia: <http://petrociencia.com.ve/blogs/geoestadistica/wp-content/uploads/2014/04/maternvario.pdf>

- Reyes, J., Escobar, C., Duarte, J., & Ramírez, P. (2007). Aplicación del Modelo de Regresión Logística en la predicción del rendimiento estudiantil. *Estudios Pedagógicos XXXIII*, 33(2), 101-120. Obtenido de [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-07052007000200006](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-07052007000200006)
- Rubalcaba G., J. (2016). *Cosas que conviene saber al usar AIC, DIC y otros criterios de información*. Obtenido de Modelización y Estadística en Biología: <https://jgrubalcaba.wordpress.com/2016/02/21/cosas-que-conviene-saber-al-usar-aic-dic-y-otros-criterios-de-informacion/>
- Santana. (2014). *Validación cruzada en R*. Recuperado el 2017, de EcuRed: [https://www.ecured.cu/Validaci%C3%B3n\\_cruzada\\_en\\_R](https://www.ecured.cu/Validaci%C3%B3n_cruzada_en_R)
- Seoane, J. (s.f.). *Métodos para el estudio de Sistemas Ecológicos*. Recuperado el 2017, de Universidad Autónoma de Madrid: [https://www.uam.es/personal\\_pdi/ciencias/jspinill/documentos/METODOS/Metodos304011\\_Intro\\_y\\_GLM.pdf](https://www.uam.es/personal_pdi/ciencias/jspinill/documentos/METODOS/Metodos304011_Intro_y_GLM.pdf)
- Universidad de Sonora. (s.f.). *Capítulo 5: Variograma*, 15-19. México. Recuperado el 2017, de <http://tesis.uson.mx/digital/tesis/docs/2889/Capitulo5.pdf>
- Wikipedia Enciclopedia Libre*. (s.f.). Recuperado el 2017, de <https://es.wikipedia.org/wiki/Chironomidae>