



UNIVERSIDAD TÉCNICA DE COTOPAXI

FACULTAD DE CIENCIAS AGROPECUARIAS Y RECURSOS NATURALES

CARRERA DE INGENIERÍA AMBIENTAL

PROYECTO DE INVESTIGACIÓN

“EVALUACIÓN DE CONTAMINANTES ORGÁNICOS Y PARAMETROS FÍSICO QUÍMICOS, MEDIANTE UN MODELO DE PREDICCIÓN RANDOM FOREST CON VALIDACIÓN LOOCV, EN EL RÍO ISINCHE PARA AGUAS DE USO RECREATIVO,

.COTOPAXI 2025”

Proyecto de Investigación presentado previo a la obtención del Título de
Ingeniero Ambiental

Autor:
Martín Espartaco

Tutora:
Tapia Borja Alexandra Isabel

LATACUNGA – ECUADOR Julio 2025

DECLARACIÓN DE AUTORÍA

Cruz Duque Martín Espartaco, con cédula de ciudadanía No. 1718682345, declaro ser autor del presente Proyecto de Investigación: **“EVALUACIÓN DE CONTAMINANTES ORGÁNICOS Y PARAMETROS FÍSICO QUÍMICOS, MEDIANTE UN MODELO DE PREDICCIÓN RANDOM FOREST CON VALIDACIÓN LOOCV, EN EL RÍO ISINCHE PARA AGUAS DE USO RECREATIVO, COTOPAXI 2025”**, siendo la Ingeniera Mg. Alexandra Isabel Tapia Borja, Tutora del presente trabajo; y, eximo expresamente a la Universidad Técnica de Cotopaxi y a sus representantes legales de posibles reclamos o acciones legales.

Además, certifico que las ideas, conceptos, procedimientos y resultados vertidos en el presente trabajo investigativo, son de mi exclusiva responsabilidad.

Latacunga, 23 de julio del 2025



Martín Espartaco Cruz Duque
C.C: 1718682345
ESTUDIANTE

CONTRATO DE CESIÓN NO EXCLUSIVA DE DERECHOS DE AUTOR

Comparecen a la celebración del presente instrumento de cesión no exclusiva de obra, que celebran de una parte **CRUZ DUQUE MARTIN ESPARTACO**, identificado con cédula de ciudadanía 1718682345 de estado civil soltero, a quien en lo sucesivo se denominará **EL CEDENTE**; y, de otra parte, la Doctora Idalia Eleonora Pacheco Tigselema, en calidad de Rectora, y por tanto representante legal de la Universidad Técnica de Cotopaxi, con domicilio en la Av. Simón Rodríguez, Barrio El Ejido, Sector San Felipe, a quien en lo sucesivo se le denominará **LA CESIONARIA** en los términos contenidos en las cláusulas siguientes:

ANTECEDENTES: CLÁUSULA PRIMERA. - EL CEDENTE es una persona natural estudiante de la carrera de Ambiente, titular de los derechos patrimoniales y morales sobre el trabajo de grado **“EVALUACIÓN DE CONTAMINANTES ORGÁNICOS Y PARAMETROS FÍSICO QUÍMICOS, MEDIANTE UN MODELO DE PREDICCIÓN RANDOM FOREST CON VALIDACIÓN LOOCV, EN EL RÍO ISINCHE PARA AGUAS DE USO RECREATIVO, COTOPAXI 2025”**, la cual se encuentra elaborada según los requerimientos académicos propios de la Facultad; y, las características que a continuación se detallan:

ANTECEDENTES: CLÁUSULA PRIMERA. - EL CEDENTE es una persona natural estudiante de la carrera de Ambiente, titular de los derechos patrimoniales y morales sobre el trabajo de grado **“EVALUACIÓN DE CONTAMINANTES ORGÁNICOS Y PARAMETROS FÍSICO QUÍMICOS, MEDIANTE UN MODELO DE PREDICCIÓN RANDOM FOREST CON VALIDACIÓN LOOCV, EN EL RÍO ISINCHE PARA AGUAS DE USO RECREATIVO, COTOPAXI 2025”**, la cual se encuentra elaborada según los requerimientos académicos propios de la Facultad; y, las características que a continuación se detallan:

Historial Académico

Inicio de la carrera: octubre 2020 – marzo 2021

Finalización de la carrera: abril 2025 – agosto 2025

Tutor: Ing. Tapia Borja Alexandra Mg.

Tema: **“EVALUACIÓN DE CONTAMINANTES ORGÁNICOS Y PARAMETROS FÍSICO QUÍMICOS, MEDIANTE UN MODELO DE PREDICCIÓN RANDOM FOREST CON VALIDACIÓN LOOCV, EN EL RÍO ISINCHE PARA AGUAS DE USO RECREATIVO, COTOPAXI 2025”**

CLÁUSULA SEGUNDA. - LA CESIONARIA es una persona jurídica de derecho público creada por ley, cuya actividad principal está encaminada a la educación superior formando profesionales de tercer y cuarto nivel normada por la legislación ecuatoriana la misma que establece como requisito obligatorio para publicación de trabajos de investigación de grado en su repositorio institucional, hacerlo en formato digital de la presente investigación.

CLÁUSULA TERCERA. - Por el presente contrato, **EL CEDENTE** autoriza a **LA CESIONARIA** a explotar el trabajo de grado en forma exclusiva dentro del territorio de la República del Ecuador.

CLÁUSULA CUARTA. - OBJETO DEL CONTRATO: Por el presente contrato **EL CEDENTE**, transfiere definitivamente a **LA CESIONARIA** y en forma exclusiva los siguientes derechos patrimoniales; pudiendo a partir de la firma del contrato, realizar, autorizar o prohibir:

- a) La reproducción parcial del trabajo de grado por medio de su fijación en el soporte informático conocido como repositorio institucional que se ajuste a ese fin.

- b) La publicación del trabajo de grado.
- c) La traducción, adaptación, arreglo u otra transformación del trabajo de grado con fines académicos y de consulta.
- d) La importación al territorio nacional de copias del trabajo de grado hechas sin autorización del titular del derecho por cualquier medio incluyendo mediante transmisión.
- e) Cualquier otra forma de utilización del trabajo de grado que no está contemplada en la ley como excepción al derecho patrimonial.

CLÁUSULA QUINTA. - El presente contrato se lo realiza a título gratuito por lo que **LA CESIONARIA** no se halla obligada a reconocer pago alguno en igual sentido **EL CEDENTE** declara que no existe obligación pendiente a su favor.

CLÁUSULA SEXTA. - El presente contrato tendrá una duración indefinida, contados a partir de la firma del presente instrumento por ambas partes.

CLÁUSULA SÉPTIMA. - CLÁUSULA DE EXCLUSIVIDAD. - Por medio del presente contrato, se cede en favor de **LA CESIONARIA** el derecho a explotar la obra en forma exclusiva, dentro del marco establecido en la cláusula cuarta, lo que implica que ninguna otra persona incluyendo **EL CEDENTE** podrá utilizarla.

CLÁUSULA OCTAVA. - LICENCIA A FAVOR DE TERCEROS. - **LA CESIONARIA** podrá licenciar la investigación a terceras personas siempre que cuente con el consentimiento de **EL CEDENTE** en forma escrita.

CLÁUSULA NOVENA. - El incumplimiento de la obligación asumida por las partes en la cláusula cuarta, constituirá causal de resolución del presente contrato. En consecuencia, la resolución se producirá de pleno derecho cuando una de las partes comunique, por carta notarial, a la otra que quiere valerse de esta cláusula.

CLÁUSULA DÉCIMA. - En todo lo no previsto por las partes en el presente contrato, ambas se someten a lo establecido por la Ley de Propiedad Intelectual, Código Civil y demás del sistema jurídico que resulten aplicables.

CLÁUSULA UNDÉCIMA. - Las controversias que pudieran suscitarse en torno al presente contrato, serán sometidas a mediación, mediante el Centro de Mediación del Consejo de la Judicatura en la ciudad de Latacunga. La resolución adoptada será definitiva e inapelable, así como de obligatorio cumplimiento y ejecución para las partes y, en su caso, para la sociedad. El costo de tasas judiciales por tal concepto será cubierto por parte del estudiante que lo solicitare.

En señal de conformidad las partes suscriben este documento en dos ejemplares de igual valor y tenor en la ciudad de Latacunga, a los 23 días del mes de julio del 2025.



Martín Espartaco Cruz Duque

EL CEDENTE

Dra. Idalia Pacheco Tigselema, Ph.D.

LA CESIONARIA

AVAL DE LA TUTORA DEL PROYECTO DE INVESTIGACIÓN

En calidad de Tutora del Proyecto de Investigación con el título:

“EVALUACIÓN DE CONTAMINANTES ORGÁNICOS Y PARAMETROS FÍSICO QUÍMICOS, MEDIANTE UN MODELO DE PREDICCIÓN RANDOM FOREST CON VALIDACIÓN LOOCV, EN EL RÍO ISINCHE PARA AGUAS DE USO RECREATIVO, COTOPAXI 2025”, de Cruz Duque Martín Espartaco, de la carrera de Ingeniería Ambiental, considero que el presente trabajo investigativo es merecedor del Aval de aprobación al cumplir las normas, técnicas y formatos previstos, así como también ha incorporado las observaciones y recomendaciones propuestas en la pre-defensa.

Latacunga, 23 de julio del 2025



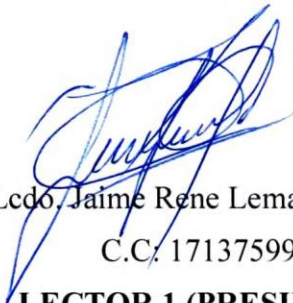
Ing. Alexandra Isabel Tapia Borja, Mg.
C.C: 0502661754
DOCENTE TUTORA

AVAL DE APROBACIÓN DEL TRIBUNAL DE TITULACIÓN

En calidad de Tribunal de Lectores, aprobamos el presente Informe de Investigación de acuerdo a las disposiciones reglamentarias emitidas por la Universidad Técnica de Cotopaxi; y, por la Facultad de Ciencias Agropecuarias y Recursos Naturales; por cuanto, el postulante: Cruz Duque Martín Espartaco, con el título del Proyecto de Investigación: **“EVALUACIÓN DE CONTAMINANTES ORGÁNICOS Y PARAMETROS FÍSICO QUÍMICOS, MEDIANTE UN MODELO DE PREDICCIÓN RANDOM FOREST CON VALIDACIÓN LOOCV, EN EL RÍO ISINCHE PARA AGUAS DE USO RECREATIVO, COTOPAXI 2025”**, ha considerado las recomendaciones emitidas oportunamente y reúne los méritos suficientes para ser sometido al acto de sustentación del trabajo de titulación.

Por lo antes expuesto, se autoriza grabar los archivos correspondientes en un CD, según la normativa institucional.

Latacunga, 23 de julio del 2025



Lcdo. Jaime Rene Lema Pillalaza Mg.
C.C. 1713759932
LECTOR 1 (PRESIDENTE)



Ing. Juan Pio Salazar Arias, Mg.
C.C: 0501389530
LECTOR 2 (MIEMBRO)



Ing. Kalina Marcela Fonseca Largo, PhD.
CC: 1723534457
LECTOR 3 (MIEMBRO)

AGRADECIMIENTO

La culminación de este trabajo, se lo debo a mis 2 padres, Catalina María Duque Rosales y Luis Ernesto Cruz Cedeño, quien, con su guía y apoyo incondicional, me permitieron seguir el camino que quise, siempre velando por mi bienestar a pesar de las adversidades. A mi tutora de tesis, Ing, Alexandra Isabel Tapia Borja, sin su ayuda y constante apoyo, y jalones de oreja, esta investigación no se habría llevado a cabo, pues siempre estuvo brindando inspiración cuando más lo necesitaba, y a toda aquella persona que me ayudo dar un paso en mi formación académica, les estaré eternamente agradecido

0DUWtQ(VSDUWDFR&UXJ'XTXH

DEDICATORIA

Este trabajo se lo dedico a mis guías y padres, Catalina y Luis, me han demostrado que todo es posible si uno se esfuerza lo suficiente, gracias por ser mi ejemplo a seguir, a mi madre por enseñarme que la bondad es parte de las cualidades de una persona, y a mi padre, por siempre ser un ejemplo de perseverancia y fortaleza, a mi tutora Alexandra Tapia, por ser una amiga en la que se puede confiar y una persona que brinda luz cuando se necesita, a aquellos amigos que me brindaron una parte de su vida e hicieron que la universidad fuera menos pesada, y a toda aquella persona que brindo un grano de arena apoyándome en este recorrido, todo esto fue gracias a ustedes

0DUWtQ(VSDUWDFR&UXJ'XTXH

UNIVERSIDAD TÉCNICA DE COTOPAXI

FACULTAD DE CIENCIAS AGROPECUARIAS Y RECURSOS NATURALES

TÍTULO: “EVALUACIÓN DE CONTAMINANTES ORGÁNICOS Y PARAMETROS FÍSICO QUÍMICOS, MEDIANTE UN MODELO DE PREDICCIÓN RANDOM FOREST CON VALIDACIÓN LOOCV, EN EL RÍO ISINCHE PARA AGUAS DE USO RECREATIVO, COTOPAXI 2025”.

Autor:

Cruz Duque Martín Espartaco

RESUMEN

El presente estudio se centró en evaluar la calidad del agua del río Isinche para uso recreativo y desarrollar un modelo predictivo de aprendizaje automático (Random Forest) para estimar los cambios en los parámetros analizados en un escenario con datos limitados. Se recolectaron muestras en cinco puntos de interés distribuidos desde la cuenca alta hasta la cuenca baja del río, durante tres fechas de muestreo realizadas en junio de 2025. Las muestras de agua fueron analizadas en laboratorio para determinar los parámetros de coliformes totales, coliformes fecales, E. coli, nitritos, nitratos, fosfatos, pH y oxígeno disuelto. Posteriormente, los resultados fueron comparados con la normativa TULSMA para determinar la calidad del agua. Para la construcción del modelo predictivo se utilizó el lenguaje de programación Python. Mediante una serie de procedimientos sistemáticos, se entrenó el modelo de predicción Random Forest para predecir los valores futuros de las muestras recolectadas. Adicionalmente, se implementó una validación cruzada LOOCV y se evaluó el desempeño del modelo mediante RMSE, generando mapas que reflejan las predicciones de los valores futuros de los parámetros estudiados. Los resultados obtenidos demostraron que, en promedio, parámetros como los coliformes totales y fecales presentan un aumento significativo en sus concentraciones durante los fines de semana. De manera general, el pH, el oxígeno disuelto y la relación nitrógenofósforo no cumplen con los límites permisibles máximos (LPM) establecidos por la normativa TULSMA, indicando que el agua del río Isinche no es apta para uso recreacional. Las predicciones realizadas por el modelo demostraron alta precisión para parámetros como nitritos, nitratos, fosfatos, pH y oxígeno disuelto, los cuales mostrarán un aumento en sus concentraciones en la cuenca media del río. En conclusión, tanto en el escenario actual como en el futuro, el agua del río Isinche presenta graves problemas de contaminación en todos los parámetros estudiados, lo que hace que sus aguas no sean aptas para aprovechamiento recreativo. Esta situación representa una problemática significativa donde tanto la salud pública como la ecosistémica se ven comprometidas.

Palabras clave: río Isinche, Python, Random Forest, predicción, calidad del agua

TECHNICAL UNIVERSITY OF COTOPAXI FACULTY OF AGRICULTURAL SCIENCE AND NATURAL RESOURCES

THEME: “EVALUATION OF ORGANIC CONTAMINANTS AND PHYSICOCHEMICAL PARAMETERS USING A RANDOM FOREST PREDICTION MODEL WITH LOOCV VALIDATION IN THE ISINCHE RIVER FOR RECREATIONAL WATERS, COTOPAXI 2025”

Author:
Cruz Duque Martín Espartaco

ABSTRACT

The present study focused on evaluating the water quality of the Isinche River for recreational use and developing a machine learning predictive model (Random Forest) to estimate changes in the analyzed parameters under a limited data scenario. Samples were collected at five points of interest distributed from the upper basin to the lower basin of the river during three sampling dates conducted in June 2025. Water samples were analyzed in the laboratory to determine parameters including total coliforms, fecal coliforms, E. coli, nitrites, nitrates, phosphates, pH, and dissolved oxygen. Subsequently, the results were compared with TULSMA regulations to determine water quality. Python programming language was used for the construction of the predictive model. Through a series of systematic procedures, the Random Forest prediction model was trained to predict future values of the collected samples. Additionally, Leave-OneOut Cross-Validation (LOOCV) was implemented and model performance was evaluated using RMSE, generating maps that reflect predictions of future values for the studied parameters. The obtained results demonstrated that, on average, parameters such as total and fecal coliforms show a significant increase in their concentrations during weekends. Generally, pH, dissolved oxygen, and the nitrogen-phosphorus ratio do not comply with the maximum permissible limits (MPL) established by TULSMA regulations, indicating that the water of the Isinche River is not suitable for recreational use. The predictions made by the model demonstrated high accuracy for parameters such as nitrites, nitrates, phosphates, pH, and dissolved oxygen, which will show an increase in their concentrations in the middle basin of the river. In conclusion, both in the current scenario and in the future, the water of the Isinche River presents serious contamination problems in all studied parameters, making its waters unsuitable for recreational exploitation. This situation represents a significant problem where both public and ecosystem health are compromised.

Keywords: Isinche river, Python, Random Forest, prediction, wáter quality

ÍNDICE DE CONTENIDOS

1. INFORMACIÓN GENERAL	1
2. JUSTIFICACIÓN DEL PROYECTO	2
3. BENEFICIARIOS DEL PROYECTO	3
3.1. Beneficiarios Directos (Impacto inmediato)	3

3.2. Beneficiarios Indirectos (Impacto a mediano/largo plazo)	3
4. EL PROBLEMA DE INVESTIGACIÓN	4
5. OBJETIVOS	5
5.1. General	5
5.2. Específicos	5
6. ACTIVIDADES Y SISTEMA DE TAREAS EN RELACIÓN A LOS OBJETIVOS PLANTEADOS	6
7. FUNDAMENTACIÓN CIENTÍFICO TÉCNICA	8
7.1. Aguas recreativas	8
7.2. Aguas Superficiales	8
7.3. Contaminación de aguas superficiales	8
7.4. Fuentes de contaminación	9
7.4.1. Aguas residuales domésticas (ARD)	9
7.4.2. Aguas residuales agropecuarias (ARA)	9
7.5. Riesgos asociados a la contaminación por ARD y ARA	10
7.6. Coeficiente de correlación de Pearson	12
7.7. Random Forest	13
7.7.1. Script	13
7.7.2. Principio de Bagging	13
7.7.3. Validación LOOCV	14
7.8. Marco Legal	14
7.8.1. Constitución de la República del Ecuador	14
7.8.2. Código Orgánico del Ambiente (COA)	15
7.8.3. Acuerdo Ministerial 097 – A (Texto Unificado de Legislación Ambiental Secundaria (TULSMA))	16
8. VALIDACIÓN DE LAS PREGUNTAS CIENTÍFICAS O HIPÓTESIS	18
9. METODOLOGÍA	19
9.1. Enfoque de la investigación	19
9.1.1. Enfoque cuantitativo	19
9.2. Análisis espacial del área de estudio	19

9.1. Técnicas e instrumentos	19
9.1.1. Técnicas de campo	19
9.2. Recolección de muestras	21
9.2.1. Método	21
9.2.2. Materiales y equipos usados para la recolección de muestras	21
9.3. Cálculo de relación nitrógeno – fosforo	22
9.4. Métodos	22
9.4.1. Métodos de programación – construcción del modelo	22
10. ANÁLISIS Y DISCUSIÓN DE RESULTADOS	26
10.1. Puntos de muestreo	26
10.2. Análisis de las muestras de agua recolectadas	29
10.2.1. Análisis normativo de contaminantes orgánicos	29
10.2.2. Relación nitrógeno – fosforo	34
10.2.3. Análisis normativo de los parámetros físico – químicos	35
10.3. Construcción del modelo para la predicción de contaminantes y parámetros físico - químicos	37
10.3.1. Análisis estadísticos	42
10.4. Mapas de predicción	44
10.4.1. Nitritos y nitratos	44
10.4.2. Fosfatos	47
10.4.3. pH	48
10.4.4. OD	49
11. IMPACTOS (TÉCNICOS, SOCIALES, AMBIENTALES O ECONÓMICOS)	51
11.1. Impactos técnicos	51
11.2. Impactos ambientales	52
12. PRESUPUESTO DEL PROYECTO	53
13. CONCLUSIONES	54
14. RECOMENDACIONES	54

ÍNDICE DE TABLAS

Tabla 1. Beneficiarios directos	3
Tabla 2. Beneficiarios indirectos	3
Tabla 3. Sistematización de actividades	6
Tabla 4. Riesgos asociados a contaminación por ARD y ARA.	10
Tabla 5. Criterios de calidad de aguas para fines recreativos mediante contacto primario.	17
Tabla 6. Criterios de calidad de aguas para fines recreativos mediante contacto secundario.	17
Tabla 7. Librerías utilizadas	22
Tabla 8. Puntos de muestreo	27
Tabla 9. Valores de los contaminantes orgánicos obtenidos de las muestras de agua.	28
Tabla 10. Valores promedio de los contaminantes.	29
Tabla 11. Valores de los parámetros físico químicos obtenidos de las muestras de agua.	32
Tabla 12. Valores promedio de los parámetros físico – químicos.	33
Tabla 13. Resultados del cálculo de la relación N:P	34
Tabla 14. Valores promediados de parámetros físico – químicos	34
Tabla 15. Resultados de Error Cuadrático Medio	42
Tabla 16. Resultados de la correlación entre parámetros y variables meteorológicas	43

ÍNDICE DE FIGURAS

Figura 1. Tipos de correlación – gráficos de dispersión.	12
Figura 2. GPS utilizado para la georreferenciación	20
Figura 3. Multiparámetros utilizado in situ	21
Figura 4. Flujo de trabajo para el preprocesamiento	23
Figura 5. Flujo de trabajo para el entrenamiento del modelo	24
Figura 6. Flujo de trabajo para el proceso de predicción	25
Figura 7. Mapa de ruta - río Isinche	26
Figura 8. Sección 1: importaciones y configuración	37
Figura 9. Sección 2: rutas y estructura de carpetas.	37
Figura 10. Sección 3: funciones Auxiliares.	38
Figura 11. Sección 4: carga y preprocesamiento.	39
Figura 12. Sección 5: modelado y predicción	39
Figura 13. Sección 6: análisis estadístico.	40
Figura 14. Sección 7: creación de mapas	40
Figura 15. Sección 8: salida de resultados.	41
Figura 16. Mapa de predicción nitritos.	45
Figura 17. Mapa de predicción Nitratos	46
Figura 18. Mapa de predicción fosfatos	47
Figura 19. Mapa predicción pH	48
Figura 20. Mapa de predicción OD	49

ÍNDICE DE GRÁFICOS

¿Gráfico 1. Comparación de los valores promediados de C.T. vs TULSMA	30
Gráfico 2. Comparación de los valores promedio de C.F. vs TULSMA	31
Gráfico 3. Comparación de los valores promedio de pH vs TULSMA	35
Gráfico 4. Comparación de valores OD vs TULSMA.	36

LISTA DE ACRÓNIMOS

ARA	Aguas Residuales Agropecuarias
ARD	Aguas Residuales Domésticas
CF	Coliformes Fecales
CT	Coliformes Totales
LOOCV	Leave One Out Cross Validation
NPM	Numero Mas Probable
OD	Oxígeno Disuelto
TULSMA	Texto Unificado de Legislación Secundaria del Ministerio del Ambiente
UFC	Unidades Formadoras de Colonias

1. INFORMACIÓN GENERAL

Título del Proyecto: Evaluación de contaminantes orgánicos y parámetros físico químicos, mediante un modelo de predicción Random Forest con validación LOOCSV, en el río Isinche para aguas de uso recreativo, Cotopaxi 2025.

Fecha de inicio: Abril del 2025 **Fecha de finalización:** Julio del 2025

Lugar de ejecución: Cantón Latacunga, Provincia de Cotopaxi.

Facultad que auspicia: Facultad de Ciencias Agropecuarias y Recursos Naturales

Carrera que auspicia: Ingeniería Ambiental **Equipo de Trabajo:**

Tutora: Ing. Alexandra Tapia Borja Msc. **Coordinador**

del Proyecto:

Nombre/s: Martín Espartaco Cruz Duque

Teléfonos: 0979303733

Correo electrónico: martin.cruz2345artin.cruz2345@utc.edu.ec

Área de Conocimiento: Ciencia Naturales. Medio Ambiente, Ciencias Ambientales. **Línea de investigación:** Energías Alternativas y Renovables, eficiencia energética y protección ambiental.

Línea de vinculación de la carrera: Manejo y Conservación del Recurso Hídrico.

Línea de Vinculación de la Facultad: Gestión de recursos naturales, biodiversidad, biotecnología y genética, para el desarrollo humano y social

2. JUSTIFICACIÓN DEL PROYECTO

La carencia de sistemas de monitoreo continuo en ríos de bajo orden, como el río Isinche (afluente de la cuenca del Cutuchi), ha impedido una evaluación sistemática de su deterioro ambiental, situación agravada por la insuficiente inversión en conservación y control de calidad hídrica. Estudios recientes demuestran que este cuerpo de agua presenta niveles de contaminación biológica que exceden significativamente los límites establecidos en el TULSMA (Texto Unificado de Legislación Ambiental Secundaria). Particularmente preocupante resulta la concentración de nitritos de origen fecal, que alcanzan valores de 0.428 mg/L (Hidalgo Pio & Veintimilla Tipanguano, 2024). superando en 114% el límite permisible recomendado para la preservación de la vida acuática (0.2 mg/L) (TULSMA, 2015; EPA, 2022).

Paralelamente, investigaciones microbiológicas han identificado contaminación por *Escherichia coli* con concentraciones entre 1,833 - 3,333 UFC/100 mL (Bungacho Michilena & Quezada Suárez, 2021), valores que exceden en 36 a 66 veces el estándar para aguas recreativas de contacto primario (100 UFC/100 mL) (Organización Mundial de la Salud, 2021). Estas cargas microbianas se originan principalmente en asentamientos rurales con disposición directa de excretas humanas y animales al cauce.

Ante este escenario, el contacto directo con las aguas del Isinche representa un riesgo sanitario crítico para las comunidades aledañas. Estas no solo dependen del río para la producción agrícola, sino que grupos vulnerables –especialmente niños– realizan actividades recreativas en sus aguas. Esta exposición constante incrementa la probabilidad de enfermedades gastrointestinales graves, un impacto alarmante considerando las limitaciones de acceso a servicios de salud pública en estas poblaciones.

Frente a esta problemática, la implementación de un modelo predictivo Random Forest con validación LOOCV (Leave-One-Out Cross-Validation) ofrece ventajas técnicas clave: capacidad para procesar conjuntos de datos fragmentarios mediante imputación multivariada (Nong et al., 2025), reducción de ruido estadístico (minimizando errores de generalización en predicciones de calidad hídrica), e identificación de patrones no lineales entre variables ambientales y concentración de patógenos (Singh et al., 2014).

3. BENEFICIARIOS DEL PROYECTO

3.1. Beneficiarios Directos (Impacto inmediato)

Tabla 1. Beneficiarios directos

Grupo	Cantidad Aprox.	Beneficio
Comunidades rurales aledañas (Isinche, Pujilí, Salache)	~500–1,000 familias	Acceso a datos científicos para gestionar agua segura en recreación y agricultura.
GADs Parroquiales/Cantonaes (Pujilí, Salcedo)	3–5 gobiernos locales	Bases técnicas para políticas públicas y proyectos de remediación.
Turistas y operadores turísticos	~200–500 personas/año	Garantía de seguridad en actividades recreativas (balnearios, ecoturismo).

3.2. Beneficiarios Indirectos (Impacto a mediano/largo plazo)

Tabla 2. Beneficiarios indirectos

Grupo	Cantidad Aprox.	Beneficio
Universidades y centros de investigación	5–10 instituciones	Datos de referencia para futuros estudios en cuencas andinas.
MAATE y SENAGUA	Entes nacionales	Validación de normativas y escalamiento de estrategias de mitigación.
Población aguas abajo (río Cutuchi-Patate)	~2,000–5,000 hab.	Reducción de contaminación en la cuenca hidrográfica.

4. EL PROBLEMA DE INVESTIGACIÓN

El agua dulce disponible para el consumo humano y el mantenimiento de los ecosistemas representa apenas el 1% del total de agua en el planeta (González, 2018). Sin embargo, este recurso limitado se ha visto gravemente afectado en las últimas décadas debido a factores como el crecimiento poblacional, la descarga de efluentes domésticos e industriales, las actividades agropecuarias y el arrastre de contaminantes por escorrentías (Castro, 2020). Entre las principales fuentes de contaminación hídrica, aquellas de origen biológico representan un riesgo significativo para la salud pública, ya que facilitan la propagación de microorganismos patógenos, incluyendo bacterias, virus, hongos y parásitos, los cuales pueden dispersarse a través de los cuerpos de agua (Pauta et al., 2020). La exposición a estos agentes puede desencadenar enfermedades como diarrea, cólera, disentería, hepatitis A y fiebre tifoidea, entre otras. De hecho, se estima que alrededor de 2,000 millones de personas en el mundo consumen agua contaminada con materia fecal, lo que provoca aproximadamente 502,000 muertes anuales por enfermedades diarreicas (Organización Mundial de la Salud, 2023).

En el contexto ecuatoriano, la gestión del agua ha priorizado el abastecimiento sobre la conservación y el control de calidad, agravando la problemática debido a la insuficiente inversión en infraestructura y la falta de criterios técnicos en su manejo (Zambrano Mero et al., 2022). Esta situación se ve exacerbada por la descarga directa de aguas residuales, tanto en zonas urbanas —a través de sistemas de alcantarillado— como en áreas rurales, donde las excretas de animales y humanos son transportadas hacia los ríos mediante escorrentías (Campaña et al., 2017; Cedeño-Muñoz, 2020).

El presente estudio se centra en el río Isinche, el cual nace en las estribaciones occidentales de la cordillera de los Andes, cerca de Latacunga, y fluye hacia el suroeste, atravesando zonas rurales y agrícolas. Además, este cuerpo de agua sirve como límite natural entre los cantones Salcedo y Latacunga (GAD Parroquial de Salcedo, 2014) y entre Belisario Quevedo y Latacunga (GAD Parroquial de Belisario Quevedo, 2011). Dada su trayectoria, el río Isinche está expuesto a una creciente concentración de contaminantes biológicos, lo que representa un riesgo potencial para las comunidades y ecosistemas asociados. Esta preocupación se acentúa al considerar que el Isinche es un afluente del río Cutuchi, actualmente catalogado como uno de los ríos más contaminados del país debido a las descargas industriales, domésticas y de camales municipales (Mena Maisincho, 2020). De no implementarse medidas correctivas, el río Isinche podría enfrentar un deterioro similar en un futuro próximo.

5. OBJETIVOS

5.1. General

- Evaluar contaminantes orgánicos y parámetros físico químicos, mediante un modelo de predicción Random Forest con validación LOOCSV, en el río Isinche para aguas de uso recreativo, Cotopaxi 2025.

5.2. Específicos

- Reconocer el área de estudio mediante análisis espacial.
- Evaluar los parámetros físico – químicos y microbiológicos del agua en relación con los límites establecidos por la normativa nacional para actividades recreativas.
- Desarrollar un modelo de predicción basado en Random Forest con validación LOOCV, capaz de estimar concentraciones de contaminantes críticos en escenarios de datos limitado

6. ACTIVIDADES Y SISTEMA DE TAREAS EN RELACIÓN A LOS OBJETIVOS PLANTEADOS

Tabla 3. Sistematización de actividades

OBJETIVO ESPECÍFICO	ACTIVIDAD	METODOLOGÍA	RESULTADOS
Reconocer el área de estudio mediante análisis espacial	<ul style="list-style-type: none"> - Reconocimiento del área de estudio por medio de imágenes satelitales - Identificación de fuentes de contaminación - Georreferenciación de puntos críticos 	<ul style="list-style-type: none"> - GIS (Sistemas de Información Geográfica) - Observación directa 	<ul style="list-style-type: none"> - Puntos de muestreo determinados mediante su accesibilidad y cercanía con fuentes de contaminación.
Evaluar los parámetros físico – químicos y microbiológicos del agua en relación con los límites establecidos por la normativa nacional e internacional para actividades recreativas.	<ul style="list-style-type: none"> - Muestreo cada 5 días en 5 puntos críticos (15 muestras) - Análisis en laboratorio acreditado - Comparación con estándares 	<ul style="list-style-type: none"> - la normativa NTE INEN 2169 - AGUA, Calidad del Agua. Muestreo. Manejo y conservación de muestras. - Análisis estadístico descriptivo 	<ul style="list-style-type: none"> - Coliformes totales, Coliformes fecales, E. coli, Nitritos, Nitratos, Fosfatos, pH y Oxígeno Disuelto - Cumplimiento normativo con TULSMA
Desarrollar un modelo de predicción basado en Random Forest con validación LOOCV, capaz de estimar concentraciones de contaminantes críticos en escenarios de datos limitados.	<ul style="list-style-type: none"> - Preprocesamiento de datos (imputación de valores faltantes) - Entrenamiento del modelo - Validación con LOOCV y métricas. 	<ul style="list-style-type: none"> - Algoritmo Random Forest (Python) - Métricas: RMSE - Correlación con variables meteorológicas. 	<ul style="list-style-type: none"> - Predicción con 15 muestras: - Valores de las muestras recogidas - Mapas de representación espacial de los valores de predicción - Métricas de validación

7. FUNDAMENTACIÓN CIENTÍFICO TÉCNICA

7.1. Aguas recreativas

Las aguas recreativas son cuerpos de agua —naturales (ríos, lagos, zonas costeras) o artificiales (piscinas, balnearios)— destinados al contacto humano directo (natación, surf, inmersión, inhalación de aerosoles). Su calidad se evalúa mediante un monitoreo integral entre contaminación orgánica y parámetros físico y químico, con el objetivo científico-técnico de reducir riesgos sanitarios y ecológicos asociados al uso recreativo (Sacoto Acaro et al., 2020).

7.2. Aguas Superficiales

Se entiende por aguas superficiales a todo aquel cuerpo de agua que se encuentre en contacto directo con la atmósfera, es decir, localizado en el manto terrestre, tal como los manantiales, ríos, lagos, lagunas, embalses y canales, entre otros (Pavan et al., 2022). Este recurso hídrico posee un alto valor estratégico, ya que se utiliza para satisfacer diversas necesidades humanas, como abastecimiento de agua potable, descargas de efluentes, pesca, irrigación agrícola, generación de energía, turismo, conservación de hábitats naturales y actividades recreativas (Ibana-Lopez et al., 2021).

7.3. Contaminación de aguas superficiales

Debido a su exposición, las aguas superficiales son susceptibles a dos tipos de contaminación: la contaminación natural y la contaminación antrópica. La primera asociada a variaciones en la disponibilidad del recurso debido a factores como precipitaciones, cambios estacionales y fluctuaciones de temperatura, condiciones que favorecen la proliferación de algas y microorganismos (Tian et al., 2019). La segunda, en cambio, deriva de actividades humanas, como la agricultura y ganadería, que incrementan la presencia de microorganismos nocivos a través de la escorrentía de desechos hacia los cuerpos de agua, generando un exceso de nutrientes (eutrofización) y un aumento de patógenos (Federigi et al., 2017).

Asimismo, el crecimiento poblacional contribuye a esta problemática al incrementar el vertido de aguas residuales —tratadas o no— en los sistemas fluviales (Wang et al., 2019).

7.4. Fuentes de contaminación

7.4.1. Aguas residuales domésticas (ARD)

Según (Huamán et al., 2022; Osorio Rivera et al., 2021) las ARD están compuestas principalmente por aguas derivadas de actividades higiénicas y —en un 80% a escala global— descargadas sin tratamiento previo. Esta composición les confiere propiedades variables, entre las que destacan:

- Olor desagradable, debido a la descomposición de materia orgánica.
- Efectos tóxicos, por la presencia de minerales y contaminantes químicos.
- Potencial infeccioso, especialmente en regiones tropicales, donde su uso está asociado a enfermedades como amebiasis, cólera y parasitosis intestinales (Organización Mundial de la Salud, 2021; Vargas et al., 2020).

Estos riesgos se atribuyen a la alta concentración de patógenos (bacterias, virus, helmintos), materia orgánica en descomposición y metales pesados, que amenazan a poblaciones vulnerables ubicadas aguas abajo de los puntos de descarga (Grundmann & Maaß, 2017).

Las aguas residuales domésticas (ARD), representan aproximadamente el 10% del agua utilizada a nivel mundial para irrigación agrícola, una práctica que se popularizó durante el periodo de la posguerra del siglo XX, como estrategia de optimización en el uso de recursos hídricos en países como Alemania, Estados Unidos, México, Perú, Francia, Italia y Sudáfrica, entre otros (Silva et al., 2008). Aunque su empleo en agricultura es limitado, su manejo inadecuado ha generado brotes epidémicos, como los documentados en zonas de regadío de Perú y México, donde se reportaron casos de amebiasis vinculados al riego con ARD no tratadas (Prado et al., 2023).

7.4.2. Aguas residuales agropecuarias (ARA)

Las aguas residuales agropecuarias (ARA) son el resultado de las actividades agrícolas y ganaderas, donde los principales contaminantes provienen del uso de fertilizantes nitrogenados y fosforados, así como de las excretas animales (Mancilla-Villa et al., 2022). Clasificadas como fuentes de contaminación difusa (no puntuales), estas aguas se generan principalmente por escorrentías agrícolas, escorrentías de pastizales, lixiviados de sistemas sépticos, drenajes mineros y descargas de zonas rurales no conectadas a alcantarillado, las cuales transportan los contaminantes hacia cuerpos de agua superficiales (Pérez-Sane et al., 2024).

Un estudio en China determinó que más del 50% de la carga de nitrógeno (N) y fósforo (P) en cuerpos de agua proviene de escorrentías agropecuarias, acelerando los procesos de eutrofización y reduciendo el tiempo de desarrollo de floraciones algales nocivas (Hou et al., 2022). Asimismo, (Zhao et al., 2022) destacan que los aportes de contaminantes orgánicos — como pesticidas y materia fecal— desde actividades agropecuarias superan incluso a los provenientes de desechos industriales y ARD.

En Ecuador, este problema es particularmente relevante en zonas de alta producción agrícola, como la cuenca del río Guayas, donde estudios han identificado altas concentraciones de nitratos y coliformes fecales debido al uso intensivo de fertilizantes y la ganadería (Cedeño et al., 2021).

7.5. Riesgos asociados a la contaminación por ARD y ARA.

Un ejemplo histórico de los riesgos asociados a la contaminación por escorrentías de ARA ocurrió en 1993 en Wisconsin (EE.UU.), donde un brote de criptosporidiosis —enfermedad diarreica grave— afectó a más de 400 000 personas y causó la muerte de 104 individuos. Este episodio se originó por la acción de la escorrentía agropecuaria, la cual arrastró heces de ganado contaminadas con *Cryptosporidium spp.* hacia el río Milwaukee, fuente principal de agua potable para la región. A pesar de que el agua recibió tratamiento convencional con cloro, el parásito sobrevivió debido a su resistencia a los desinfectantes clorados, exponiendo las limitaciones de los sistemas de purificación básicos (Rock & Rivera, 2014).

Así, los riesgos presentes por la contaminación proveniente de las ARD y ARA, se pueden asociar al poco o nulo tratamiento que se brinda a las aguas residuales antes de disponerlas a los cuerpos de agua superficial. Por ejemplo, en el Ecuador, el 92% de los municipios en el país no cuentan con sistemas de tratamiento para basura y aguas residuales (Zambrano Mero et al., 2022).

A continuación, la tabla 4 presenta las posibles afecciones de la contaminación por ARD y ARA hacia cuerpos de agua superficiales.

Tabla 4. Riesgos asociados a contaminación por ARD y ARA.

AREA	TIPO DE CONTAMINANTE	IMPACTO/AFECCIÓN
Salud pública	Biológica (<i>E. coli</i> , <i>Salmonella</i> , <i>enterococos</i>)	<ul style="list-style-type: none"> - Enfermedades como: gastroenteritis, conjuntivitis, dermatitis, daño celular en intestino delgado, colitis hemorrágica, insuficiencia renal aguda, úlceras intestinales y muerte. - Incidencia en la producción, calidad e inocuidad alimentaria.
	Metales pesados (Pb, As, Hg)	<ul style="list-style-type: none"> - Toxicidad crónica, daño neurológico, afecciones al sistema nervioso y diversos tipos de cáncer (Guadie et al., 2021)
	Compuestos farmacéuticos	<ul style="list-style-type: none"> - Desarrollo de resistencia antibiótica en determinadas cepas, alteración endocrina (Wilkinson et al., 2022)
Medio Ambientes	Nutrientes (N, P)	<ul style="list-style-type: none"> - Aumentan los porcentajes de eutrofización: zonas con hipoxia y pérdida de biodiversidad.
	Biológica (<i>E. coli</i> , <i>Salmonella</i> , <i>enterococos</i>)	<ul style="list-style-type: none"> - Ictiofauna: Septicemia general, ulceraciones cutáneas, necrosis de

- hígado, congestiones
branquiales
- Herpetofauna: Lesiones cutáneas, muerte por septicemia.
 - En manglares, junto a la presencia de residuos sólidos y metales pesados, ocasiona tumoraciones en especies de cangrejos (Pernía et al., 2019)

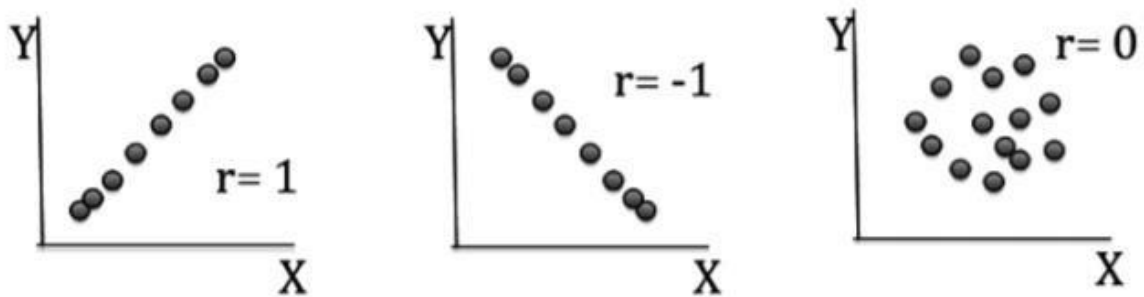
Nota Las afecciones recopiladas en la presente tabla se obtuvieron de las siguientes fuentes (Mancilla-Villa et al., 2022; Pérez-Sane et al., 2024; Zambrano Mero et al., 2022b).

En Ecuador, problemas similares han sido documentados. Por ejemplo, en 2017, un estudio en Quito reveló que el 30% de las muestras de aguas superficiales en zonas periurbanas presentaban contaminación por *Cryptosporidium* y *Giardia*, asociada a descargas de aguas residuales domésticas y agropecuarias sin tratamiento (Ortega-Paredes et al., 2019). Adicionalmente, en la Amazonía ecuatoriana, comunidades que dependen de ríos para consumo humano reportaron casos recurrentes de diarrea por patógenos transmitidos por agua, vinculados a la falta de saneamiento básico (Ortiz-Prado et al., 2022). Estos casos subrayan la urgente necesidad de contar con un sistema de monitoreo de aguas residuales adecuado especialmente en regiones con alta vulnerabilidad climática y sanitaria.

7.6. Coeficiente de correlación de Pearson

Este método mide el grado de relación entre dos o más variables, mayormente cuantitativas, de manera ordinal, cabe recalcar que este coeficiente, no analiza relaciones causa - efecto, ya que no se trata a las variables como dependiente – independiente. Las relaciones establecidas entre variables pueden ser directa, indirecta o nula, las cuales son representadas mediante gráficos de dispersión (Sucasaire Pilco & Ticona Vilcapaza, 2023), como se puede observar en la figura 1.

Figura 1. Tipos de correlación – gráficos de dispersión.



Fuente (Danigno, 2015).

En la figura 1 se puede observar los distintos tipos de correlación, donde $r = 1$ (directa), significa una mayor relación o fuerza de asociación entre X e Y, es decir, X e Y son directamente proporcionales, en contraparte, $r = -1$ (indirecta), significa que X e Y son inversamente proporcionales, mientras que $r = 0$ (nula), representa la falta de asociación entre X e Y

7.7. Random Forest

El modelo de aprendizaje automático, o machine learning en inglés, Random Forest, es definido como un algoritmo de aprendizaje automático supervisado. Basándose en la construcción de múltiples árboles de decisión, la combinación con el principio de bagging y la aleatorización de las características de los árboles, teniendo la capacidad de realizar procesos de clasificación y regresión (Alomani et al., 2022).

7.7.1. Script

Se conoce a un script como el conjunto de instrucciones escrita en un lenguaje de scripting (interpretado) específico, el cual automatiza operaciones específicas dentro de un sistema, aplicación, o flujos de trabajo. Se diferencia de los programa de compilación por: ejecutar tareas manuales de forma programada, ser flexible e interplataforma, es decir, los scripts creados pueden ser usados entre un programa de interpretación otro, ser interoperable por conectarse a varias sistemas a la vez y ser eficiente a la hora de reducir los tiempos de ejecución y errores humanos en conjunto, pues al ejecutarse línea por línea en tiempo real, se simplifican operaciones repetitivas, integran componentes y manipulan datos de forma ágil (Lutz, 2013; Sebesta, 2015).

7.7.2. Principio de Bagging

Se refiere a una técnica ensemble learning, es decir, que combinan múltiples modelos estadísticos individuales, con la finalidad de mejorar la precisión de los algoritmos de

aprendizaje automáticos, minimizar la varianza y el sobreajuste dentro de modelos que combinan varios subconjuntos aleatorios de datos (Fernández, 2025).

7.7.3. Validación LOOCV

“Leave-One-Out Cross-Validation” LOOCV, también conocida como validación cruzada, es un método de muestreo de datos que, consiste en entrenar un modelo de aprendizaje automático basado en el número de datos, es decir muestras, que se posee, haciendo que cada dato funcione como un submodelo independiente, ejecutando procesos de regresión o clasificación, dejando un conjunto de datos fuera los procesos para evaluar su funcionamiento, con la finalidad de evaluar la capacidad de generalización del modelo y evitar el sobreajuste (Cerdeira Corrotea, 2020).

7.8. Marco Legal

7.8.1. Constitución de la República del Ecuador

TÍTULO I ELEMENTOS CONSTITUTIVOS DEL ESTADO

Capítulo primero Principios Fundamentales

Art. 3.- Son deberes primordiales del Estado:

1. Garantizar sin discriminación alguna el efectivo goce de los derechos establecidos en la Constitución y en los instrumentos internacionales, en particular la educación, la salud, la alimentación, la seguridad social y el agua para sus habitantes.

TÍTULO II DERECHOS

Capítulo segundo Sección primera Agua y alimentación

Art. 12.- El derecho humano al agua es fundamental e irrenunciable. El agua constituye patrimonio nacional estratégico de uso público, inalienable, imprescriptible, inembargable y esencial para la vida.

Sección segunda Ambiente sano

Art. 14.- Se reconoce el derecho de la población a vivir en un ambiente sano y ecológicamente equilibrado, que garantice la sostenibilidad y el buen vivir, *sumak kawsay*.

Se declara de interés público la preservación del ambiente, la conservación de los ecosistemas, la biodiversidad y la integridad del patrimonio genético del país, la prevención del daño ambiental y la recuperación de los espacios naturales degradados.

Art. 15.- El Estado promoverá, en el sector público y privado, el uso de tecnologías ambientalmente limpias y de energías alternativas no contaminantes y de bajo impacto. La

soberanía energética no se alcanzará en detrimento de la soberanía alimentaria, ni afectará el derecho al agua. **Sección séptima**

Salud

Art. 32.- La salud es un derecho que garantiza el Estado, cuya realización se vincula al ejercicio de otros derechos, entre ellos el derecho al agua, la alimentación, la educación, la cultura física, el trabajo, la seguridad social, los ambientes sanos y otros que sustentan el buen vivir.

Capítulo sexto Derechos de libertad

Art. 66.- Se reconocer y garantiza a las personas:

1. El derecho a una vida digna, que asegure la salud, alimentación y nutrición, agua potable, vivienda, saneamiento ambiental, educación, trabajo, empleo, descanso y ocio, cultura física, vestido, seguridad social y otros servicios sociales necesarios.

TÍTULO VI REGIMEN DE DESARROLLO

Capítulo primero Principios generales

Art. 276.- El régimen de desarrollo tendrá los siguientes objetivos:

4. Recuperar y conservar la naturaleza y mantener un ambiente sano y sustentable que garanticen a las personas y colectividades el acceso equitativo, permanente y de calidad al agua, aire y suelo, y a los beneficios de los recursos del subsuelo y del patrimonio natural.

TÍTULO VII REGIMEN DEL BUEN VIVIR

Capítulo segundo Biodiversidad y recursos naturales Sección sexta Agua

Art. 411.- El Estado garantizará la conservación, recuperación y manejo integral de los recursos hídricos, cuencas hidrográficas y caudales ecológicos asociados al ciclo hidrológico. Se regulará toda actividad que pueda afectar la calidad y cantidad de agua, y el equilibrio de los ecosistemas, en especial en las fuentes y zonas de recarga de agua.

La sustentabilidad de los ecosistemas y el consumo humano serán prioritarios en el uso y aprovechamiento del agua.

7.8.2. Código Orgánico del Ambiente (COA)

TÍTULO II

INSTITUCIONALIDAD Y ARTICULACIÓN DE LOS NIVELES DE GOBIERNO EN EL SISTEMA NACIONAL DESCENTRALIZADO DE GESTIÓN AMBIENTAL

CAPÍTULO II

DE LAS FACULTADES AMBIENTALES DE LOS GOBIERNOS AUTONOMOS DESCENTRALIZADOS

Art. 26.- Facultades de los Gobiernos Autónomos Descentralizados Provinciales en materia ambiental.

8. Controlar el cumplimiento de los parámetros ambientales y la aplicación de normas técnicas de los componentes agua, suelo, aire y ruido;

TITULO II

SISTEMA UNICO DE MANEJO AMBIENTAL

CAPITULO V

CALIDAD DE LOS COMPONENTES ABIOTICOS Y ESTADO DE LOS COMPONENTES BIOTICOS

Art. 191.- Del monitoreo de la calidad del aire, agua y suelo. La Autoridad Ambiental Nacional o el Gobierno Autónomo Descentralizado competente, en coordinación con las demás autoridades competentes, según corresponda, realizarán el monitoreo y seguimiento de la calidad del aire, agua y suelo, de conformidad con las normas reglamentarias y técnicas que se expidan para el efecto.

Art. 196.- Tratamiento de aguas residuales urbanas y rurales. Los Gobiernos Autónomos Descentralizados Municipales deberán contar con la infraestructura técnica para la instalación de sistemas de alcantarillado y tratamiento de aguas residuales urbanas y rurales, de conformidad con la ley y la normativa técnica expedida para el efecto. Asimismo, deberán fomentar el tratamiento de aguas residuales con fines de reutilización, siempre y cuando estas recuperen los niveles cualitativos y cuantitativos que exija la autoridad competente y no se afecte la salubridad pública. Cuando las aguas residuales no puedan llevarse al sistema de alcantarillado, su tratamiento deberá hacerse de modo que no perjudique las fuentes receptoras, los suelos o la vida silvestre. Las obras deberán ser previamente aprobadas a través de las autorizaciones respectivas emitidas por las autoridades competentes en la materia.

7.8.3. Acuerdo Ministerial 097 – A (Texto Unificado de Legislación Ambiental Secundaria (TULSMA))

LIBRO I

DE LA AUTORIDAD AMBIENTAL

TITULO I

De la Misión, Visión y Objetivos del Ministerio del Ambiente

Art. 3.- OBJETIVOS DEL MINISTERIO DEL AMBIENTE: Los objetivos estratégicos institucionales son los siguiente:

3. Mantener y mejorar la cantidad y calidad del agua, manejando sustentablemente las cuencas hidrográficas.

ANEXO 1 DEL LIBRO VI

NORMA DE CALIDAD AMBIENTAL Y DE DESCARGA DE EFLUENTES AL RECURSO AGUA

Para determinar los criterios en calidad de agua, con respectos a contaminantes orgánicos presentes en el río Isinche, se hizo uso de los parámetros establecidos en el libro VI de TULSMA, en específico la tabla 6 y la tabla 7, las cuales pueden ser visualizadas a continuación:

Tabla 5. Criterios de calidad de aguas para fines recreativos mediante contacto primario.

PARAMETRO	EXPRESADO COMO	UNIDAD	CRITERIO DE CALIDAD
Parásitos Nemátodos intestinales	-	-	Ausencia
Coliformes Fecales	NPM	NPM/100 ml	200
Coliformes Totales	NPM	NPM/100 ml	2000
Compuestos fenólicos	Fenol	mg/l	0,002
Grasas y aceites	Película visible	-	Ausencia
Material flotante	Visible	-	Ausencia
Oxígeno Disuelto	OD	% de saturación	>80
pH	pH		6,5 – 8,3
Relación Nitrógeno Fosforo Total	-	-	15:1
Tensoactivos	Sustancias activas al azul de metileno	mg/l	0,5

Fuente libro VI de TULSMA, 2017

Tabla 6. Criterios de calidad de aguas para fines recreativos mediante contacto secundario.

PARAMETRO	EXPRESADO COMO	UNIDAD	CRITERIO DE CALIDAD
Parásitos Nemátodos intestinales	-	-	Ausencia
Coliformes Totales	NPM	NPM/100 ml	2000
Grasas y aceites	Película visible	-	Ausencia
Material flotante	Visible	-	Ausencia
Oxígeno Disuelto	OD	% de saturación	>80
pH	pH	-	6,5 – 8,3
Relación Nitrógeno Fosforo Total	-	-	15:1
	Sustancias		
Tensoactivos	activas al azul de metileno	mg/l	0,5

Fuente libro VI de TULSMA, 2017

8. VALIDACIÓN DE LAS PREGUNTAS CIENTÍFICAS O HIPÓTESIS

¿Qué tan preciso fue el modelo predictivo Random Forest con validación LOOCV para estimar la variación en los valores de contaminantes orgánicos y parámetros físico químicos para ser empleado como herramienta en evaluaciones de calidad hídrica?

El modelo predictivo Random Forest con validación LOOCV demuestra precisión suficiente para estimar contaminantes físico-químicos en el río Isinche —nitritos (RMSE=0.0359), nitratos (RMSE=0.3232), fosfatos (RMSE=0.0960) y pH (RMSE=0.0775)— validado mediante mapas espaciales coherentes con fuentes de contaminación identificadas, lo que lo posiciona como una herramienta confiable para optimizar evaluaciones de calidad hídrica en escenarios de datos limitados y reducir costos de monitoreo; sin embargo, presenta limitaciones significativas en la predicción de coliformes fecales (RMSE=93.19) y totales (RMSE=297.35) debido a su alta dispersión por escorrentías y la insuficiencia de datos, lo que restringe su

aplicabilidad para contaminantes microbiológicos y exige complementarse con muestreos in situ.

9. METODOLOGÍA

9.1. Enfoque de la investigación

9.1.1. Enfoque cuantitativo

Basado en la recolección y análisis de datos numéricos, mediante métodos de regresión, pruebas de hipótesis o correlaciones. Tiene como objetivo establecer relaciones entre variables, obteniendo resultados con la capacidad de ser replicados (Romero et al., 2023). En el desarrollo de la investigación, permitió cuantificar los niveles de contaminación presentes en los cuerpos de agua analizados, para poder procesarlos, y de esta manera, crear relaciones estadísticas entre las variables ambientales y la concentración de contaminantes para analizar su comportamiento. Además, la cuantificación de los contaminantes presentes en el río Isinche, permitió su comparación de acuerdo a la normativa – TULSMA – para determinar si el agua del río Isinche se encuentra dentro de los rangos permisibles para usarla como agua con fines recreativos.

9.2. Análisis espacial del área de estudio

El área de estudio se encuentra ubicada desde la cuenca alta del río Isinche, ubicada a 3018 msnm hasta su cuenca baja a 2663 msnm, donde confluye con el río Cutuchi. Se hizo uso del software Google Earth, mismo que permitió una visualización del recorrido del río y proporciono una visión general de las actividades que se llevan a cabo en los alrededores del mismo. Posteriormente se realizó un recorrido para determinar puntos de interés para la recolección de muestras basado en el análisis previo realizados.

9.1. Técnicas e instrumentos

9.1.1. Técnicas de campo

Realizadas in situ, permiten obtener una perspectiva más completa del área de estudio, pues al estar el investigador en contacto directo con el área y sus componentes, se pueden tomar decisiones de manera más precisa para el correcto cumplimiento de los objetivos establecidos en el estudio.

9.1.1.1. Observación directa

Definida como una técnica de recolección sistemática de datos, similar al análisis de contenido, la observación puede darse de manera directa o indirecta, donde el observador interactúa con

los sujetos de observación o no interactúa, además, se la considera una manera de recolección inicial de datos, que emerge de la necesidad de exploración en un contexto de investigación (Hernández Sampieri et al., 2014).

En el contexto de la investigación, se hizo uso de dicha técnica al recorrer el río desde su cuenca alta hasta su unión con el río Cutuchi – cuenca baja – con la finalidad de explorar los alrededores del río en busca de fuentes que aporten contaminantes al río, determinar los puntos de interés – puntos de muestreo – con respecto a su accesibilidad y su cercanía a puntos de interés que puedan aportar contaminantes al río Isinche. Los recorridos se realizaron en cada muestreo realizado para identificar cualquier cambio en las actividades que puedan influir en los parámetros estudiados.

9.1.1.2. Georreferenciación

Definida por CanquiLlusco, 2012 como el establecimiento de relaciones de información existente dentro del espacio físico – ubicaciones geográficas – e información – documentos, mapas, imágenes, etc. – a través de códigos postales o referencias geoespaciales, conocidas también como sistemas de coordenadas. Se realizó mediante el uso de un GPSMap64s para poder obtener las coordenadas con exactitud de los puntos de interés para la toma de muestras en los puntos de interés establecidos posterior a los recorridos realizados para reconocer el área de estudio.

Figura 2. GPS utilizado para la georreferenciación



Nota el equipo visualizado en la figura 3 utiliza el sistema de coordenadas planas con proyección UTM zona 17

9.2. Recolección de muestras

Se aplicó la normativa NTE INEN 2169 - AGUA, Calidad del Agua. Muestreo. Manejo y conservación de muestras, 2013. Debido a que dicha norma prescribe las técnicas y precauciones generales necesarias para la preservación y transporte de diversas muestras de agua. La finalidad de seguir esta normativa fue prevenir alteraciones en los análisis y asegurar la integridad de las muestras durante el proceso.

9.2.1. Método

Fundamentado en el método analítico – experimental, la recolección sistemática de datos empíricos mediante técnicas estandarizadas, seguida de su análisis en condiciones controladas para validar hipótesis científicas. En la presente investigación, dicho método fue aplicado mediante: un muestreo sistemático, recolectando las muestras de agua en puntos de interés a lo largo del recorrido del río y el análisis en laboratorio de las muestras recolectadas tomando en consideración la normativa y estándares nacionales para su comparación y la determinación de su calidad para su uso recreativo.

9.2.2. Materiales y equipos usados para la recolección de muestras

9.2.2.1. Multiparámetros Hanna HI98194

Este dispositivo fue utilizado para medir, in situ, los valores tales como pH, conductividad eléctrica y oxígeno disuelto en el río, con la finalidad de obtener datos más precisos, se procedió a colocar la sonda del multiparámetro en la parte media de cada punto de muestreo, dejándolo sumergido alrededor de 15 segundos antes de realizar el registro de los datos.

Figura 3. Multiparámetros utilizado in situ



Nota el equipo de la figura 4 fue proporcionado por el laboratorio de ambiente de la Universidad Técnica de Cotopaxi

Posterior al uso del multiparámetro para el levantamiento de información entre cada punto, se utilizó agua destilada para limpiar las sondas del equipo y evitar una contaminación en la toma de muestras del siguiente punto muestreado.

9.2.2.2. Frascos de muestras

Se hizo uso de frascos para la recolección de orina por sus características de esterilización microbiana, los cuales, fueron primeramente lavados en el agua del río 3 veces para remover cualquier tipo de contaminante hacia las muestras, para luego llenarlo completamente mientras está sumergido en el agua del río, y cerrarlo del mismo modo.

9.2.2.3. Cooler

Una vez obtenidas las muestras fueron transportadas haciendo uso de un cooler con hielo para asegurar la inocuidad de las muestras en un ambiente propicio, para ser trasladadas al laboratorio donde se realizará su análisis.

9.3. Cálculo de relación nitrógeno – fosforo

Para realizar el cálculo de la relación nitrógeno – fosforo, por medio de los datos obtenidos en laboratorio, se realizó un proceso de cálculo, en primer lugar, del nitrógeno total por medio de la suma de los nitritos y nitratos, y luego del fosforo total, tomando directamente el valor de los fosfatos como indicador del fosforo disponible. Seguidamente, se realizó la relación nitrógeno – fosforo por medio de la siguiente fórmula:

$$N: P = N/P$$

9.4. Métodos

9.4.1. Métodos de programación – construcción del modelo

Se utilizo el lenguaje de programación Python 3.14 para la elaboración del script, así como la implementación de librerías que ayudaran a ejecutar de una manera más precisa las funciones destinadas al modelo. En la tabla 7 se detalla una lista de las librerías que se usaron.

Tabla 7. Librerías utilizadas

Categoría	Librerías	Función principal
------------------	------------------	--------------------------

Procesamiento de datos	Pandas, Numpy	Manipulación de tablas y operaciones matriciales
Machine Learning	Scikit – Learn	Modelos predictivos y validación
Geoespacial	Rasterio, SciPy.interpolate	Generación de mapas raster
Estadística	SciPy.stats, Seaborn	Análisis de correlaciones y visualización
Utilidades	OS, Sys, Datatime	Manejo de sistema y fechas

Nota las librerías mostradas en la tabla 7 fueron obtenidas directamente desde el sistema.

9.4.1.1. Preprocesamiento

Según Vidal & Fernando, 2024 el preprocesamiento es una de las etapas de mayor relevancia en lo que análisis de datos y aprendizaje automático se refiere. En esta etapa se limpió los datos brutos obtenidos de las muestras de aguas recogidas y los datos meteorológicos obtenidos de la estación M0004 Rumipamba – Salcedo, esta etapa consistió en la eliminación del ruido estadístico generado entre los días de diferencia en la toma de muestras (5 días entre cada muestreo), a continuación, la siguiente figura se representa el flujo de trabajo que siguieron las distintas etapas del preprocesamiento.

Figura 4. Flujo de trabajo para el preprocesamiento



Las etapas del preprocesamiento constaron en:

1. Carga de datos: los datos fueron cargados en formato CSV UFT – 8 (delimitado por comas), para evitar problemas de formato al ejecutar el script.
2. Limpieza y validación: mediante un formato CSV se detectó la categoría de los datos ingresados, desde la fecha hasta el tipo de contaminante estudiado.
3. Imputación de valores faltantes: se hizo uso de estrategias específicas para cada variable -
 - Precipitación: se tomaron los valores de 0 como ausencia de lluvia.
 - Temperatura: interpolación temporal.

- Humedad / Velocidad del viento: media móvil de 3 días.
 - Contaminantes: imputación con KNN (k = 3 vecinos más cercanos).
4. Integración de fuentes: se combinaron las muestras de datos meteorológicas mediante merge, es decir, mediante distintas ramas de código, basándose en fechas progresivas.
 5. Verificación de calidad: se realizó la creación de columnas esenciales, basadas en las 15 muestras de agua recolectadas, detección de outliers mediante análisis de distribución y control de rangos numéricos plausibles.

9.4.1.2. Entrenamiento Random Forest y Validación LOOCV

La etapa de entrenamiento del modelo, basado en metodología Random Forest, contando con una validación LOOCV, consistió en un modelado individualizado que consta de 8 modelos Random Forest, 1 por cada variable estudiada.

Figura 5. Flujo de trabajo para el entrenamiento del modelo



1. Optimización de hiperparámetros: se hizo uso de distintas funciones de Python para estimar el número adecuado acorde a los datos que se poseen.
2. Validación LOOCV: se crearon 15 iteraciones, cada una constando de 14 muestras para el entrenamiento del modelo y 1 para la validación de ellas.
3. Evaluación de métricas: se hizo uso conjunto de la validación LOOCV y métodos como el error cuadrático medio y el coeficiente de Pearson, con la finalidad de determinar la precisión del modelo cuantificando la diferencia entre los valores que el modelo prediga y los valores obtenidos en la recolección de muestras.

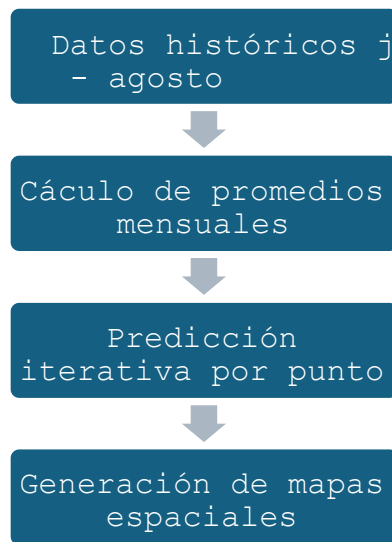
9.4.1.3. Predicción

Debido a la limitada cantidad de datos que se posee del área de estudio, las predicciones realizadas sobre las variables de estudio tienen un horizonte de confiabilidad de 60 días. Para poder obtener resultados con mayor precisión se trabajó bajo 2 segmentos. El primero consiste en: los valores de las variables de estudio (contaminantes y parámetros físico químicos) de las

muestras recolectadas y los valores de las variables meteorológicas desde el primer muestreo (09 de junio) hasta el último muestreo (19 de junio). El segundo segmento consta de los datos obtenidos con el procesamiento de los valores obtenidos del primer segmento y el uso de datos históricos del periodo desde el 16 julio – 11 de agosto año 2024, para que el modelo tenga una base de datos meteorológicos sobre el cual trabajar en ese periodo de estudio.

A continuación, el siguiente gráfico tiene como finalidad explicar de manera más sencilla el flujo de trabajo que se usó para el segundo segmento, el cual es el encargado de realizar las predicciones.

Figura 6. Flujo de trabajo para el proceso de predicción



1. Datos históricos julio – agosto: se prepararon los datos históricos recolectados para que sean tratadas como una variable de influencia en el comportamiento de las variables estudiadas.
2. Cálculo de promedios mensuales: Se hizo uso de los promedios mensuales de las variables meteorológicas (datos históricos) estudiadas para usarlo como datos futuros para la predicción.
3. Predicción iterativa por punto: se realizó un proceso individual de predicción espaciotemporal por cada punto de muestreo, mismo que consto de la construcción de un vector de caracterización individual, escalado de cada variable con el mismo scaler del entrenamiento y la predicción por medio del modelo Random Forest.
4. Generación de mapas espaciales: para generar archivos compatibles con el software QGIS se hizo uso de la técnica Interpolación IDW (Inverse Distance Weighting), a la

cual se le fijo una resolución de 100 x 100 puntos por mapa. Se utilizo el formato GeoTIFF para poder ser representado en mapas.

9.4.1.4. Análisis ANOVA

El análisis descriptivo, según Molano & Katerine, 2020 se define como una fase fundamental de la estadística en la cual se organizan los conjuntos de datos, identifican tendencias y anomalías de varianzas y facilita análisis posteriores, sean de carácter inferencial o predictivo. Dentro del proyecto de investigación se utilizó este método para entender la distribución de los contaminantes orgánicos estudiados con respecto a las variables meteorológicas, como método antes de realizar el modelamiento.

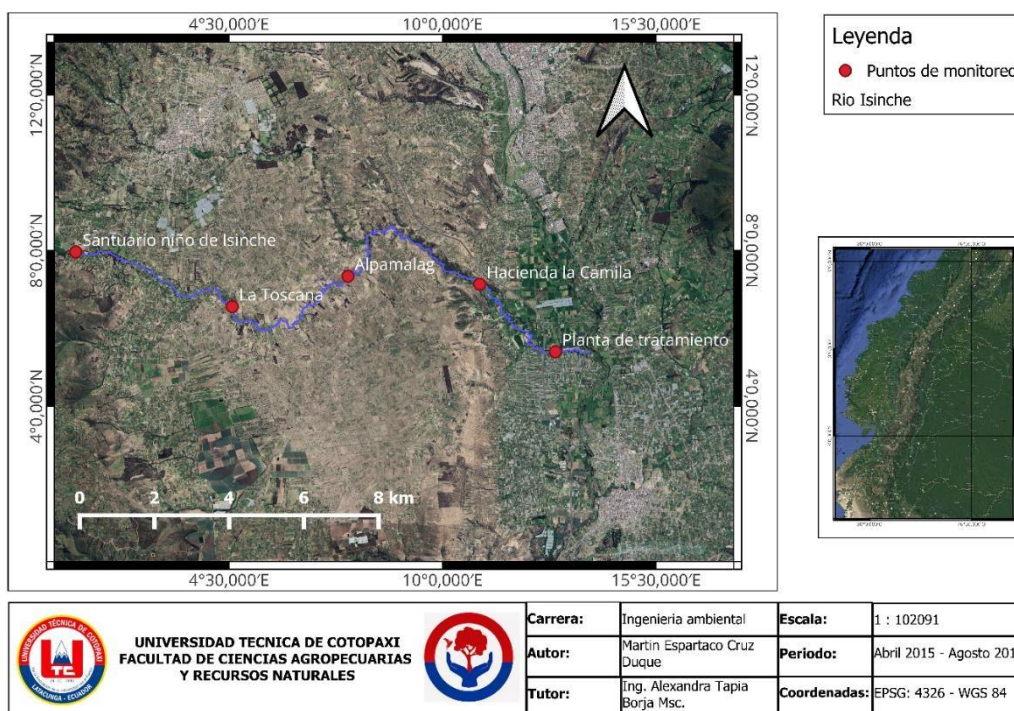
10. ANÁLISIS Y DISCUSIÓN DE RESULTADOS

10.1. Puntos de muestreo

Los puntos de muestreo se definieron mediante reconocimientos in situ del área de estudio, seleccionándose por su accesibilidad operativa y proximidad a focos de contaminación. Con base en diagnósticos preliminares, se establecieron tres campañas de muestreo: M1 (09 junio), M2 (14 junio) y M3 (19 junio).

Figura 7. Mapa de ruta - río Isinche

MAPA DE UBICACION DE PUNTOS DE MUESTREO



En el análisis realizado en el software Google Earth, se pudo identificar que el recorrido total del río, desde el primer punto de muestreo (Santuario niño de Isinche), hasta su unión con el río Cutuchi consta de 22.6 km, además, se pudo identificar que el área a los alrededores del río Isinche, en su mayoría, son rurales, mismas en las que predomina una producción agropecuaria como principal actividad económica.

Tabla 8. Puntos de muestreo

Punto	Ubicación	Coordenadas		Altitud (msnm)
		Lat.	Lon.	
P1	Santuario divino niño de Isinche	-0,983596	-78,722228	3018
P2	La Toscana	-0,997414	-78,684098	2914
P3	Apamalag de Acurios	-0,990622	-78,6568369	2870
P4	Hacienda la Camila	-0,988597	-78,629108	2735

P5	Planta de tratamiento Latacunga	-1,007678	-78,605864	2681
----	---------------------------------------	-----------	------------	------

P1, Santuario católico divino niño de Isinche: se seleccionó debido a su cercanía con la cuenca alta del río, es decir, aquella que se encuentre con la menor carga de contaminantes posible, además de su accesibilidad, ya que más allá del santuario viven comunidades ganaderas y agrícolas, que, por motivos ajenos, dificultan el paso aguas arriba. En los alrededores del punto se puede observar una significativa concentración de áreas dedicadas a la producción agrícola y ganadera. Las actividades religiosas en este lugar se desarrollan los fines de semana, por lo cual, se pudo observar una gran aglomeración de personas en el santuario y sus alrededores, desarrollando diferentes actividades.

P2, La Toscana: seleccionada debido a que sus actividades se desarrollan en el sector de la producción de alimentos, por lo cual el río Isinche representan un recurso estratégico clave para el desarrollo de sus actividades, se desconoce si dicha fabrica posee o no planta de tratamiento para las aguas residuales producto de sus actividades. De igual forma, al intentar obtener información en línea no se obtuvieron datos, así como al intentar obtener información de manera directa al consultar en las instalaciones.

P3, Alpamalag de Acurios: comunidad ubicada en la cuenca media del río, se encuentra en la región noroccidental del cantón Salcedo, conectando con la parte sureste de Pujilí, se desarrollan actividades agrícolas y ganaderas, por lo cual el río es aprovechado como fuente de descargas y usado para el agua de riego dentro de la comunidad, motivo por el cual se lo considero punto de interés para la toma de muestras.

P4, Hacienda la Camila: Ubicada paralelo a la vía Latacunga – Ambato, Barrio Salache, desarrolla sus actividades como restaurante, además de poseer un acceso directo al río, su ubicación la posiciona junto con otras haciendas y lugares destinados a la producción agrícola y ganadera, motivos para su selección.

P5, Planta de tratamiento – Latacunga: la Planta de tratamiento de agua, se encuentra ubicada en la Provincia de Cotopaxi, a uno metro antes de llegar al río Cutuchi, la planta realiza sus descargas hacia el río Isinche. Por su importancia en la calidad del agua se escogió como último punto de interés para realizar los muestreos de agua.

10.2. Análisis de las muestras de agua recolectadas

10.2.1. Análisis normativo de contaminantes orgánicos

Tabla 9. Valores de los contaminantes orgánicos obtenidos de las muestras de agua.

PARAMETRO	UNIDAD	Muestra	P1	P2	P3	P4	P5	LMP
Coliformes Totales	UFC/100 ml	1	1000	1120	1800	1000	1100	2000
		2	1300	1400	2100	2200	2500	
		3	1100	1180	1550	1300	1400	
Coliformes Fecales	UFC/100 ml	1	100	160	220	50	100	200
		2	150	210	150	540	620	
		3	115	132	167	153	160	
<i>E. Coli</i>	UFC/ml	1, 2, 3			Presencia			Ausencia

En la tabla 9, se pueden observar los valores obtenidos para los contaminantes de estudio, cabe destacar que el parámetro *E. Coli* representa a la categoría “Parásitos nemátodos intestinales” indicado en la normativa TULSMA, es un indicador frecuentemente usado para representar la presencia o ausencia de los mismos, sus valores se encuentran representados por la palabra “Presencia”, debido a que, de acuerdo a la normativa TULSMA, no debe existir presencia alguna de estos parásitos.

Llama la atención que la fecha de muestreo M2 presenta valores por encima de LMP establecido para los parámetros de coliformes totales y fecales, a partir de los punto P3, P4 y P5, con valores 2100, 2200 y 2500 UFC/100 ml respectivamente, esto puede estar directamente relacionado con la aglomeración de los poblados adyacentes al río y el desarrollo de sus actividades recreativas y religiosas en los fines de semana, pues al realizar el recorrido del río se observó que muchas comunidades los fines de semana organizan actividades en conjunto, lo que puede tener una relación directa con los valores reflejados en los análisis.

Altas concentraciones de Coliformes Totales y Fecales se presentan como un riesgo para la salud ecosistémica, pues se ha documentado que las coliformes en especies acuáticas pueden ocasionar problemas como lesiones cutáneas, necrosis de hígado, septicemia general y congestiones branquiales (Mancilla-Villa et al., 2022; Pernía et al., 2019). En cuanto a la salud pública, la contaminación por coliformes fecales en aguas usadas para el aprovechamiento

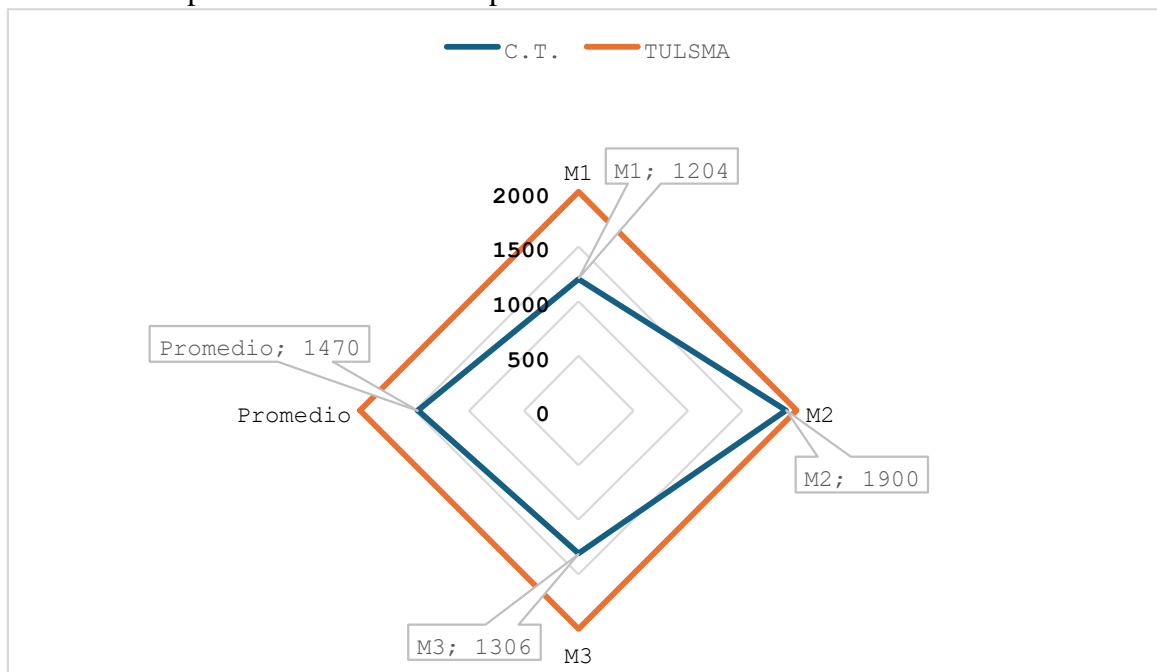
humano ocasionan enfermedades tales como: gastroenteritis, conjuntivitis, dermatitis, daño celular en intestino delgado, colitis hemorrágica, insuficiencia renal aguda, úlceras intestinales y muerte (Pérez-Sane et al., 2024; Zambrano Mero et al., 2022).

Tabla 10. Valores promedio de los contaminantes.

Promedio	PARAMETRO	UNIDAD	por	Promedio	LMP muestreo
			M1 (1204)		
			M2 (1900)	1470	2000
			M3 (1306)		
			M1 (126)		
			M2 (334)	UFC/100 ml 201,80	200
			M3 (145.40)		

El uso de los valores promediados entre cada fecha de muestreo y su promedio general, como se observa en la tabla anterior, es debido a que se desea transmitir los datos con mayor facilidad, además de que su manejo por promedio brinda una visión más generalizada de la problemática a escala general.

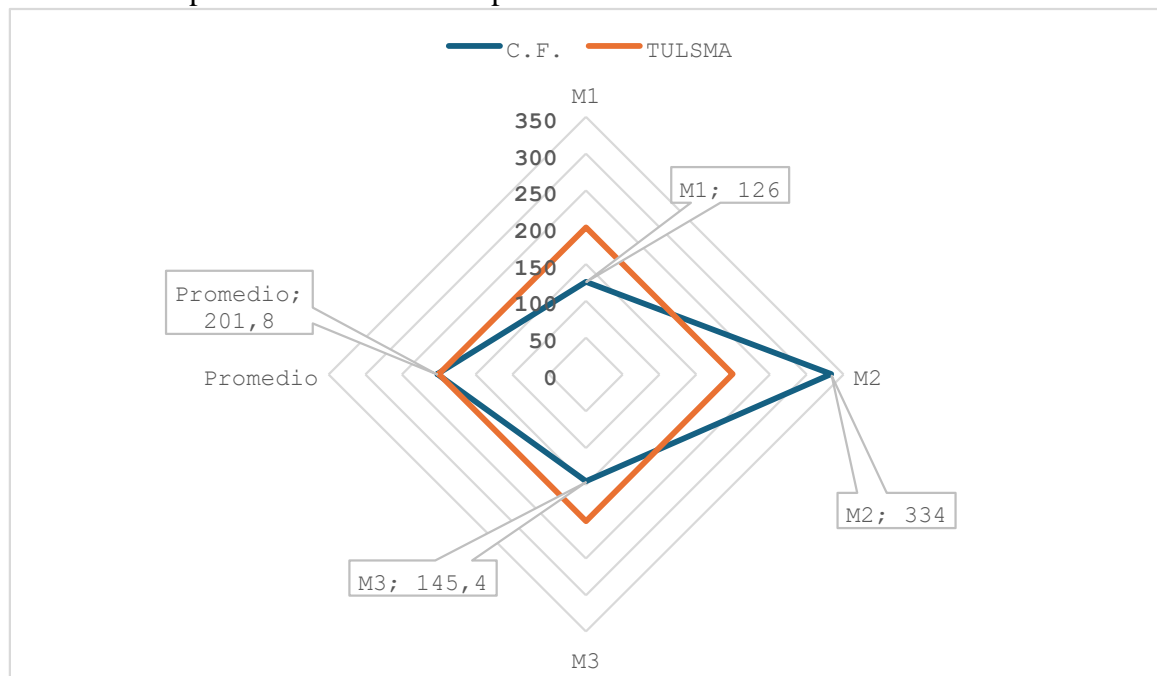
Gráfico 1. Comparación de los valores promediados de C.T. vs TULSMA



Nota los valores de los C.T. se encuentran en unidades de UFC/100 ml.

De acuerdo a los valores promediados para las coliformes totales, se observa que, en general todos los valores se encuentran dentro del LMP establecido, sin embargo, su fecha de muestreo M2 presenta un valor de 1900 UFC/ml, aproximándose al 95% del LMP, aunque no representa un riesgo agudo, indican una alta presencia de fuentes de contaminación orgánica, ya sean aguas residuales provenientes de las descargas de hogares o el material fecal animal arrastrada por la escorrentía agrícola, escenario que podría favorecer el crecimiento de especies tolerantes, lo que implica una alteración directa a la cadena trófica (Hounsinou et al., 2015).

Gráfico 2. Comparación de los valores promedio de C.F. vs TULSMA



Nota los valores de los C.T. se encuentran en unidades de UFC/100 ml.

Los coliformes fecales reflejan, para sus fechas de muestreo M1 y M3, valores dentro del LMP establecido, con 126 UFC/ml y 145.40 UFC/ml respectivamente. Sin embargo, sus fechas de muestreo M2 y su promedio cuentan con valores que superan el LMP, con valores de 334 y 201.80 UFC/ml respectivamente, lo cual representa un alto riesgo para la salud humana, por la presencia de bacterias patógenas como *E. coli* (comprobada su presencia en los análisis de laboratorio), *Salmonella* o *Campylobacter*, mismas que se asocian a enfermedades como el síndrome hemolítico agudo o la disentería (Drinking Water Program, 2023). Dichas

concentraciones pueden deberse a las escorrentías, que arrastran heces animales, y las descargas directas de agua residuales domésticas hacia el cuerpo de agua, presentándose así un escenario donde la salud pública, en comunidades rurales como Isinche o Alpamalag de Acurios, se ve vulnerada, pues al ser comunidades rurales, que usan el agua del río como aguas de aprovechamiento humano, el riesgo en contraer enfermedades por contaminación fecal aumenta.

Tabla 11. Valores de los parámetros físico químicos obtenidos de las muestras de agua.

PARAMETRO	UNIDAD	Muestreo	P1	P2	P3	P4	P5	Promedio	LMP
Nitritos	mg/L	1	0.11	0.21	0.11	0.02	0.02	0.11	-
		2	0.14	0.18	0.11	0.11	0.12		
		3	0.34	0.17	0.12	0.12	0.03		
Nitratos	mg/L	1	2.1	1.98	2.45	3.10	2.98	2.70	-
		2	3.01	2.41	3.12	2.94	3.24		
		3	2.22	2.12	2.86	2.65	3.12		
Fosfatos	mg/L	1	0.12	0.28	0.61	0.39	0.45	0.37	-
		2	0.28	0.32	0.42	0.22	0.35		
		3	0.31	0.27	0.78	0.43	0.43		
pH	pH	1	8.48	8.43	8.39	8.48	8.37	8.47	8.3
		2	8.68	8.48	8.55	8.39	8.43		
		3	8.57	8.48	8.37	8.55	8.48		

Oxígeno Disuelto	% de saturación	1	1.8	1.5	4.58	0.82	1.80		
		2	1.9	0.19	0.15	0.87	2.15	1.86	>80%
		3	1.7	1.52	2.35	11.22	1.97		

Los valores reflejados en la tabla 11, para nitritos, nitratos y fosfatos no presentan un LMP establecido, ya que la relación N:P establecida por la normativa TULSMA se aplica en sus formas de nitrógeno y fósforo total, mismas que se verán reflejadas en resultados posteriores para una mejor comprensión y manejo de los datos. Para los parámetros pH y oxígeno disuelto, se observa que en todas sus fechas y puntos de muestreo se superan los LMP establecidos, esto puede deberse a un mayor incremento de precipitaciones y arrastre de contaminantes orgánicos y nutrientes por la acción de las escorrentías, así como la posible presencia de metales pesados que aumenten los niveles de pH.

Tabla 12. Valores promedio de los parámetros físico – químicos.

Promedio	PARAMETRO	UNIDAD	por	Promedio	LMP
		muestreo			
			M1 (0.094)		
Nitritos	mg/l	M2 (0.13)	0.12	-	
		M3 (0.12)			
		M1 (2.52)			
Nitratos	mg/l	M2 (2.94)	2.70	-	
		M3 (2.70)			
		M1 (0.37)			
Fosfatos	mg/l	M2 (0.32)	0.37	-	
		M3 (0.47)			
		M1 (8.43)			
pH	pH	M2 (8.50)	8.48	6.5 – 8.3	
		M3 (8.49)			
		M1 (2.1)			
OD	% de saturación	M2 (1.05)	1.87	2000	
		M3 (5.44)			

10.2.2. Relación nitrógeno – fosforo

Tabla 13. Resultados del cálculo de la relación N:P

Parámetro	Valor	Valor ideal (15:1)	Déficit/Exceso
Nitrógeno	2.82	5.570	-49.5%
Fosforo	0.38	0.371	Equilibrado

Los datos obtenidos de la relación N:P muestran una clara limitación por nitrógeno, con una relación de 7.58:1, indicando que el N es el nutriente limitante en el sistema, señalando que, aunque existe fosforo como nutriente suficiente para el crecimiento de algas / plantas acuáticas, la escasez de nitrógeno restringirá el desarrollo de las mismas. Al existir una clara escasez de

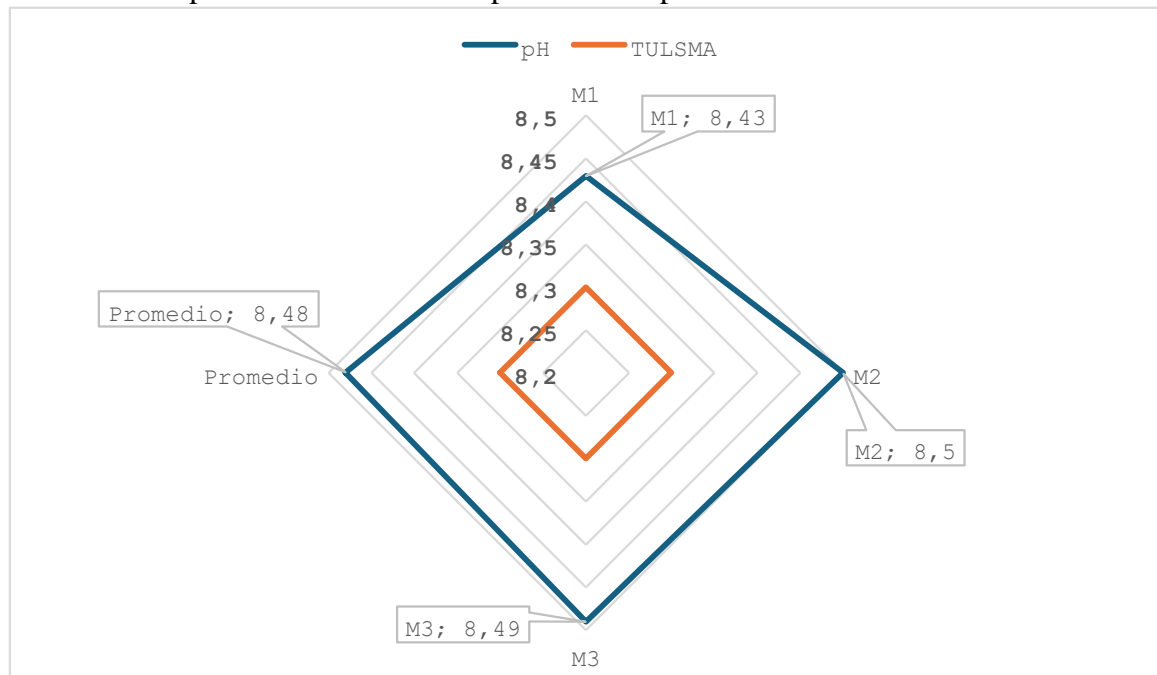
N como nutriente para el desarrollo de plantas acuáticas, bacterias como *Anabaena* o *Microcystis*, por su capacidad de fijación de N atmosférico, pueden llegar a dominar el ecosistema, ya que no dependen puramente del N disuelto (Ramirez, 2023).

10.2.3. Análisis normativo de los parámetros físico – químicos.

Tabla 14. Valores promediados de parámetros físico – químicos

PARAMETRO	UNIDAD	Promedio	LMP
Relación N:P	-	7.58:1	15:1
pH	pH	8.47	6.5 – 8.3
Oxígeno Disuelto	% de saturación	1.86	> 80%

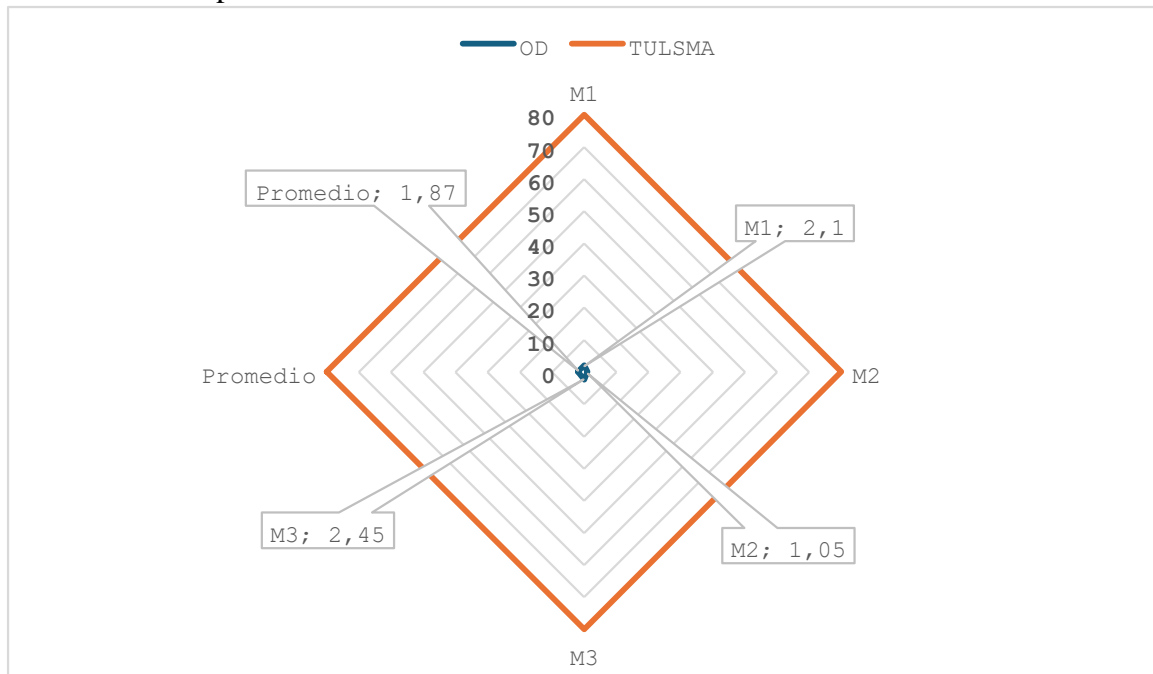
Gráfico 3. Comparación de los valores promedio de pH vs TULSMA



Se pueden observar que los valores, en todas sus fechas de muestreo, superan de manera crítica el LMP establecido, lo que refleja una alta alcalinidad del agua siendo su valor más elevado de 8.50, esto puede estar relacionado con la meteorización de rocas calcáreas y fotosíntesis agresiva por parte de organismos algales o microbiológicos que aprovechan la presencia de

coliformes y nutrientes para su crecimiento. Dichos valores hacen que el agua tenga la capacidad para ocasionar daños tales como: daños en infraestructuras como tuberías, lo que ocasiona la liberación de plomo (Pb) o cobre (Cu) en agua de consumo, por consiguiente, se pueden derivar problemas de salud pública como trastornos o deterioro del sistema nervioso o pérdida de funciones orgánicas, además, debido a su alta alcalinidad, se puede liberar amonio tóxico (NH_3) mismo que puede ocasionar afecciones como confusión, desorientación, letargo, y en casos más graves, alucinaciones, psicosis, convulsiones, coma severo y muerte (Bonilla, 2015; Pabón et al., 2020).

Gráfico 4. Comparación de valores OD vs TULSMA.



Nota los valores representados en el gráfico se encuentran en % de saturación.

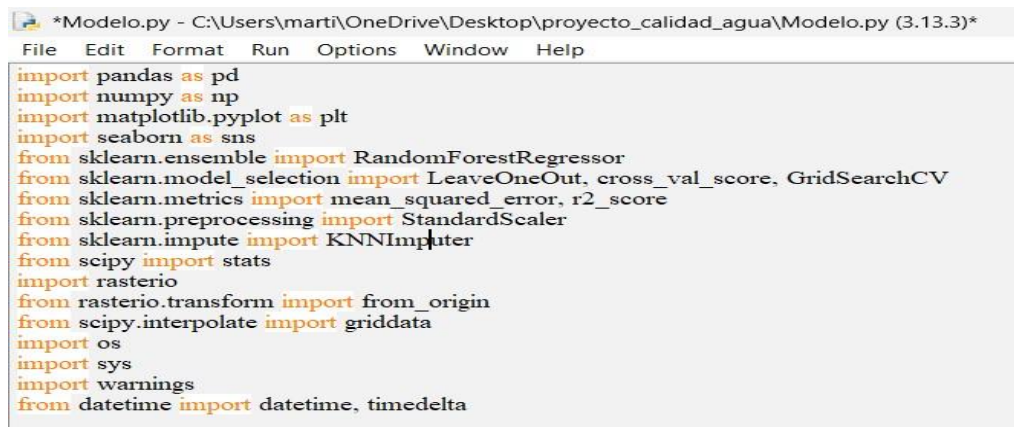
En cuanto al OD, se puede observar valores por debajo del LMP que establece un % de saturación $> 80\%$, es decir, el agua del río Isinche presenta condiciones de hipoxia críticas en todos los puntos y fechas de muestreo, lo que hace que no sea apto para sustentar vida acuática, representando así un riesgo para la biodiversidad y el equilibrio de los ecosistemas que usan el río como medio de subsistencia. Estas condiciones pueden darse debido a las escorrentías que

arrastran nutrientes y contaminantes orgánicos, lo cual aumenta la eutrofización en el río, y por consecuencia, disminuye el porcentaje de oxígeno disponible para otros organismos del río. El único valor que supera el 10% se da en el punto de muestreo P4 en su fecha M3, valor del 11.22%, esto puede estar relacionado a la falta de precipitaciones fuertes en el sector.

10.3. Construcción del modelo para la predicción de contaminantes y parámetros físico - químicos

La construcción del modelo predictivo consta en un total de 835 líneas de código, entre las cuales se incluyen comentarios y espacio entre líneas de código, distribuyéndose en un total de 8 secciones clave para el proceso de modelamiento, las cuales serán presentadas a continuación mediante las siguientes figuras, cabe destacar que, con la finalidad de optimizar espacio dentro del documento, las figuras son visualizaciones recortadas de cada sección:

Figura 8. Sección 1: importaciones y configuración



```

*Modelo.py - C:\Users\marti\OneDrive\Desktop\proyecto_calidad_agua\Modelo.py (3.13.3)*
File Edit Format Run Options Window Help
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import LeaveOneOut, cross_val_score, GridSearchCV
from sklearn.metrics import mean_squared_error, r2_score
from sklearn.preprocessing import StandardScaler
from sklearn.impute import KNNImputer
from scipy import stats
import rasterio
from rasterio.transform import from_origin
from scipy.interpolate import griddata
import os
import sys
import warnings
from datetime import datetime, timedelta

```

Consta de 17 líneas de código, en las cuales se importaron y cargaron las librerías y funciones necesarias para llevar a cabo los cálculos propuestos.

Figura 9. Sección 2: rutas y estructura de carpetas.

```

# Configuración inicial
warnings.filterwarnings('ignore')
pd.set_option('display.max_columns', None)
pd.set_option('display.width', 1000)
pd.set_option('display.max_colwidth', 50)
np.random.seed(42)

# -----
# 1. CONFIGURACIÓN DE RUTAS
# -----

# Establecimiento de ruta base
carpeta_base = r'C:\Users\marti\OneDrive\Desktop\proyecto_calidad_agua'
carpeta_datos = os.path.join(carpeta_base, 'datos')
carpeta_resultados = os.path.join(carpeta_base, 'resultados')

# Crear carpetas si no existen
os.makedirs(carpeta_datos, exist_ok=True)
os.makedirs(carpeta_resultados, exist_ok=True)
os.makedirs(os.path.join(carpeta_resultados, 'mapas'), exist_ok=True)
os.makedirs(os.path.join(carpeta_resultados, 'graficos'), exist_ok=True)

# Definir rutas de archivos
rutas = {
    'muestras': os.path.join(carpeta_datos, 'muestras.csv'),
    'meteorologia': os.path.join(carpeta_datos, 'meteorologia.csv'),
    'meteorologia_historica': os.path.join(carpeta_datos, 'meteorologia_historica.csv')
}

```

Consta de 27 líneas de código, las cuales tienen como finalidad escoger de manera lógica la ruta de acceso a los datos necesarios para que el modelo pueda trabajar.

Figura 10. Sección 3: funciones Auxiliares.

```

*Modelo.py - C:\Users\marti\OneDrive\Desktop\proyecto_calidad_agua\Modelo.py (3.13.3)*
File Edit Format Run Options Window Help
}

# -----
# 2. FUNCIONES AUXILIARES
# -----

def cargar_datos(ruta_archivo):
    """Carga un archivo CSV con manejo robusto de fechas y formatos numéricos"""
    print(f"\nCargando {os.path.basename(ruta_archivo)}...")

    try:
        # 1. Leer archivo conservando todos los datos
        df = pd.read_csv(ruta_archivo, sep=',', dtype=str, keep_default_na=False)
        print(f"Registros crudos: {len(df)}")

        # Limpiar espacios en nombres de columnas
        df.columns = df.columns.str.strip()
        print(f"Columnas encontradas: {list(df.columns)}")

        # Renombrar OD a Oxígeno_Disuelto si existe
        if 'OD' in df.columns:
            df = df.rename(columns={'OD': 'Oxígeno_Disuelto'})
            print("Columna 'OD' renombrada a 'Oxígeno_Disuelto'")

        # 2. Manejo especial de fechas - Múltiples intentos de parseo
        posibles_fechas = [col for col in df.columns if 'fecha' in col.lower() or 'date' in col.lower()]

```

Consta de 284 líneas de código, las cuales cumplen las siguientes funciones:

1. Carga de datos:

- Propósito: carga de archivos CSV con manejo robusto de formatos □
- Características:

Conversión automática de fechas
 Renombrado de columnas de ser el caso
 Manejo de decimales (comas a puntos) Diagnóstico detallado de errores.

2. Manejo de datos faltantes □ Estrategias:

Rellena Precipitación con 0.
 Interpola Temperatura temporalmente.
 Usa media móvil para Humedad y Velocidad Viento.
 Aplica KNN - Imputer para datos numéricos restantes.

3. Creación de mapas

- Genera mapas raster usando interpolación IDW (Inverse Distance Weighting).
- Verifica mínimos de puntos (≥ 4) y maneja errores de georreferenciación.

4. Cálculo de correlaciones

- Calcula correlaciones de Pearson con diagnóstico de significancia estadística (pvalue).

A continuación, se muestra el flujo principal de trabajo que consta de 411 línea de código, distribuidas en las siguientes secciones.

Figura 11. Sección 4: carga y preprocesamiento.



```
*Modelo.py - C:\Users\marti\OneDrive\Desktop\proyecto_calidad_agua\Modelo.py (3.13.3)*
File Edit Format Run Options Window Help
    reporte += f'Correlación con {var}: {corr:.4f} {significativo} | p-value={p_val:.4f} | Obs: {n_obs}\n'
except Exception as e:
    reporte += f' ✘ Error calculando correlación: {str(e)}\n'
return reporte

# -----
# 3. CARGA Y PREPROCESAMIENTO - CON VERIFICACIÓN DE DATOS
# -----

print("\n" + "="*80)
print(" INICIANDO CARGA Y PREPROCESAMIENTO DE DATOS ".center(80, '='))
print("="*80)

# Cargar datos con funciones mejoradas
df_muestras = cargar_datos(rutas['muestras'])
df_meteo_actual = cargar_datos(rutas['meteorologia'])
df_meteo_historica = cargar_datos(rutas['meteorologia_historica'])

# Verificar carga exitosa y cantidad mínima de datos
if df_muestras is None or df_meteo_actual is None or df_meteo_historica is None:
    print("\n ✘ ERROR: No se pudieron cargar uno o más archivos. Verifica los errores anteriores.")
    sys.exit(1)
```

Se realizó la carga de los datos de los parámetros de las muestras de agua, los datos meteorológicos actuales y los datos meteorológicos del anterior año, en formato CSV para poder ser procesados por el programa, tomando en consideración las fechas y realizando una imputación de datos faltantes, obteniendo como resultado un archivo CSV de diagnóstico con los datos combinados entre: parámetros de muestras de agua + meteorología actual + meteorología anterior. Generando una base de datos primaria para el análisis espacial – temporal.

Figura 12. Sección 5: modelado y predicción

```
*Modelo.py - C:\Users\marti\OneDrive\Desktop\proyecto_calidad_agua\Modelo.py (3.13.3)*
File Edit Format Run Options Window Help

# -----
# 4. MODELADO Y PREDICCIÓN - CON VERIFICACIÓN DE DATOS
# -----

print("\n" + "="*80)
print(" INICIANDO MODELADO PREDICTIVO ".center(80, '='))
print("="*80)

# Configuración
targets = ['Nitritos', 'Nitratos', 'Fosfatos', 'Coliformes_Totales',
           'Coliformes_Fecales', 'pH', 'Conductividad', 'Oxigeno_Disuelto']
features = ['Temperatura', 'Precipitacion', 'Humedad', 'Velocidad_Viento', 'Lat', 'Lon']
meteo_vars = ['Temperatura', 'Precipitacion', 'Humedad', 'Velocidad_Viento']

# Resultados
modelos = {}
resultados = {}
predicciones_futuro = pd.DataFrame()
```

Con los datos cargados y preprocesados de manera previa, se realizó el siguiente proceso:

1. Algoritmo: se utilizó el modelo Random Forest Regressor, realizando un loop por cada parámetro
2. Variables objetivo: se trabajó con los 7 parámetros de estudio.
3. Aplicación de la validación LOOCV: luego del modelamiento con el algoritmo Random Forest se obtuvo que los datos cumplen con los requisitos necesarios para poder aplicar este tipo de validación (obtención de > 10 muestras validas)
4. Predicción a futuro: Uso de promedio históricos de datos meteorológicos (16 de julio – 11 de agosto 2024).

Figura 13. Sección 6: análisis estadístico.

```
*Modelo.py - C:\Users\marti\OneDrive\Desktop\proyecto_calidad_agua\Modelo.py (3.13.3)*
File Edit Format Run Options Window Help

    'Valor': [pred]
    })]

print("\n  Modelado predictivo completado exitosamente")

# -----
# 5. ANÁLISIS ESTADÍSTICO - MEJORADO
# -----

print("\n" + "="*80)
print(" REALIZANDO ANÁLISIS ESTADÍSTICO ".center(80, '='))
print("="*80)

# Reporte de correlaciones
print("Calculando correlaciones de Pearson...")
reporte_pearson = calcular_correlaciones(df, targets, meteo_vars)
print(reporte_pearson) # Mostrar en consola para diagnóstico inmediato

# Guardar reportes
ruta_pearson = os.path.join(carpeta_resultados, 'reporte_pearson.txt')
with open(ruta_pearson, 'w', encoding='utf-8') as f:
    f.write(reporte_pearson)
print(f'  Reporte de correlaciones guardado: {ruta_pearson}')
```

Un total de 63 líneas de código destinadas a llevar a cabo la validación sobre la precisión del modelo por medio de gráficos de dispersión y realizar un reporte de correlaciones., cuyos resultados pueden visualizarse en las tablas 15 y 16.

Figura 14. Sección 7: creación de mapas

```

*Modelo.py - C:\Users\marti\OneDrive\Desktop\proyecto_calidad_agua\Modelo.py (3.13.3)*
File Edit Format Run Options Window Help
print("✅ Gráficos de dispersión guardados")

# -----
# 6. GENERACIÓN DE MAPAS - CON VERIFICACIÓN
# -----

print("\n" + "="*80)
print(" GENERANDO MAPAS PARA QGIS ".center(80, '='))
print("="*80)

# Mapa actual (última fecha de muestreo)
if 'Fecha' in df.columns:
    ultima_fecha = df['Fecha'].max()
    df_actual = df[df['Fecha'] == ultima_fecha]

print(f"Generando mapas actuales (fecha: {ultima_fecha.date()})")
for target in targets:
    if target in df_actual.columns:
        crear_mapa_geotiff(
            df_actual,
            target,
            os.path.join(carpeta_resultados, 'mapas', f'actual_{target}.tif')
        )

```

El modelo realizó dos procesos dentro de la generación de datos, en los cuales genero, en código, mapas de distribución actuales con los valores recogidos en cada fecha y punto de muestreo, y mapas de distribución futura, tomando en cuenta la última fecha de muestreo y los promedios históricos de meteorología para realizar la predicción.

Figura 15. Sección 8: salida de resultados.

```

*Modelo.py - C:\Users\marti\OneDrive\Desktop\proyecto_calidad_agua\Modelo.py (3.13.3)*
File Edit Format Run Options Window Help

# -----
# 7. GUARDAR RESULTADOS - DETALLADO
# -----

print("\n" + "="*80)
print(" GUARDANDO RESULTADOS FINALES ".center(80, '='))
print("="*80)

# Métricas de modelos
if resultados:
    df_mtricas = pd.DataFrame(resultados).T
    df_mtricas.to_csv(os.path.join(carpeta_resultados, 'metricas_modelos.csv'))
    print("✅ Métricas de modelos guardadas")
else:
    print("⚠ No se generaron métricas de modelos")

```

Un total de 55 líneas de código destinadas a generar y guardar los resultados del modelo predictivo, obteniendo de esta manera:

1. Métricas modelo
2. Predicciones futuras
3. Mapas GeoTIFF de predicción.

10.3.1. Análisis estadísticos

10.3.1.1. Error Cuadrático Medio

Tabla 15. Resultados de Error Cuadrático Medio

Variable	RMSE	Muestras	Media	Desv. Estándar	Mínimo	Máximo
Nitritos	0.0359	15	0.116	0.0569	0.02	0.21
Nitratos	0.3232	15	2.699	0.4234	1.98	3.24
Fosfatos	0.0960	15	0.371	0.1645	0.12	0.78
Coliformes Totales	297.3467	15	1470.0	470.2431	1000	2500
Coliformes Fecales	93.1853	15	201.8	159.9773	50	620
pH	0.0775	15	8.475	0.0858	8.37	8.68
Oxígeno Disuelto	1.2211	15	1.866	1.2003	0.15	4.58

Se puede observar en la columna los valores de RSME, los cuales indican el error cuadrático medio, es decir, la diferencia entre los valores de las muestras recolectadas y los valores predichos por el modelo, entre los cuales destacan los Nitritos, Nitratos, Fosfatos y pH, con 0.0359, 0.3232, 0.0960 y 0.0775 respectivamente, indicando una alta precisión al momento de predecir estos parámetros, por otro lado, el OD presenta un valor RSME = 1.2211, lo que representa una precisión aceptable al momento de predecir este parámetro. En contraparte los valores de Coliformes Totales y Coliformes Fecales presentan un RSME de 297.3467 y 93.1853

respectivamente, lo que sugiere la existencia de errores significativos al momento de predecir estos contaminantes, esto puede deberse a la gran capacidad de dispersión que poseen o la falta de datos con los que entrenar al modelo para que pueda entender cómo se compartan estos contaminantes.

Por otro lado, las columnas: media y desviación estándar, explican los valores RSME obtenidos de los análisis del modelo, dónde se puede observar valores <0 en el caso de los nitritos, nitratos, fosfatos y pH, lo que indica un alto índice de confiabilidad con respecto a los resultados proporcionados por el modelo y coherencia entre los valores de las muestras recogidas, los valores del modelo predictivo y los valores del RSME. Por otro lado, se observa una Desviación estándar de 470.2431 y 159.9773 Coliformes Totales y Coliformes Fecales respectivamente, indicando el porqué de los altos valores obtenidos en RSME, lo que significa que los datos de las muestras no son suficientes para que el modelo pueda predecir de manera correcta estos contaminantes. En el caso del OD su variación de valor mínimo es de 0.15 y su valor máximo de 4.58 haciendo que su desviación estándar, de 1.2003, sea aceptable.

10.3.1.1. Coeficiente de correlación de Pearson

Tabla 16. Resultados de la correlación entre parámetros y variables meteorológicas

Variable	Factor	Correlación (r)	pvalue	Significancia ($\alpha=0.05$)
Coliformes Totales	Humedad	-0.5535	0.0323	Sí
	Velocidad Viento	-0.6160	0.0145	Sí
Coliformes Fecales	Humedad	-0.5146	0.0497	Sí
	Velocidad Viento	-0.5657	0.0280	Sí
Oxígeno Disuelto	Humedad	0.4972	0.0593	No (<i>cercano a 0.05</i>)
	Velocidad Viento	0.5105	0.0518	No (<i>cercano a 0.05</i>)

Se puede observar que, para los Coliformes Totales y Fecales se determinó que existe una correlación significativa entre dicho parámetros y las variables meteorológicas humedad y velocidad del viento, contando con valores $r = -0.5535$ y -0.6160 respectivamente para C.T. y valores $r = -0.5146$ y -0.5657 respectivamente para C.F., estos valores reflejan una correlación inversa entre los parámetros C.T. y C.F., es decir, mientras mayor condiciones de humedad y

velocidad del viento exista en el área, disminuirán las concentraciones de estos contaminantes, esto puede estar relacionado con la facilidad de dispersión de estos contaminantes al estar en contacto con las condiciones del ambiente. Sin Embargo, debido a los valores obtenidos en la métrica RSME, desviación estándar y media, no se tomará en cuenta a los C.T. y C.F. para la elaboración de gráficos de dispersión en relación a las variables meteorológicas.

En cuanto a los parámetros físico químicos, el OD presenta valores no significantes entre los parámetros estudiados y variables meteorológicas, es decir, la relación entre ambas variables no es representativa estadísticamente para realizar las predicciones, al igual que los parámetros nitritos, nitratos, fosfatos y pH no se encontró relación significativa entre estos parámetros y las variables meteorológicas, esta falta de correlación se puede deber a que, en lo que a estos parámetros se refiere, las precipitaciones son la variable de mayor influencia sobre estos los mismos, y ya que el modelo utilizó los datos meteorológicos del anterior año para realizar las predicciones del periodo jul 16 - agosto 11, fechas en las que las precipitaciones fueron casi nulas, no se tomó en consideración como una variable de significancia para el proceso de predicción.

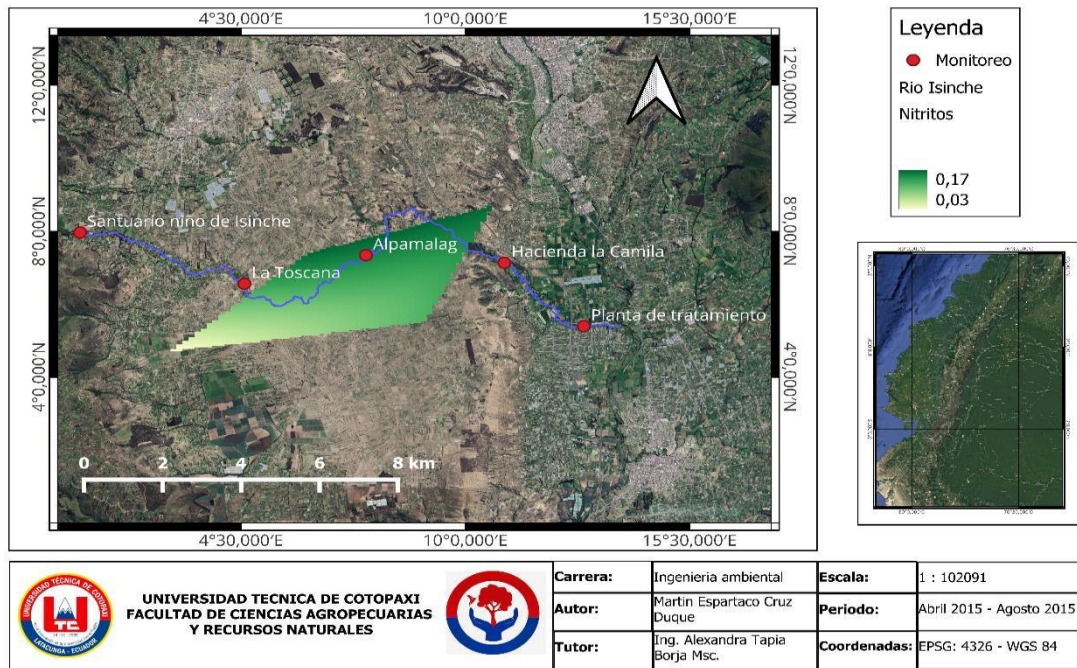
10.4. Mapas de predicción

Si bien los datos generados en el análisis estadístico, en cuanto a coliformes, tanto fecales y totales, reflejan una baja precisión para la predicción en el comportamiento y concentración de estos contaminantes, debido a que la base de datos para los mismo no es suficiente, para los parámetros estudiados tales como nitritos, nitratos, fosfatos, pH y OD se pudo predecir sus concentraciones a futuro, sin tomar en consideración las variables meteorológicas para dichas predicciones.

10.4.1. Nitritos y nitratos

Estudios como el de Xu et al., (2021) aplicaron el modelo Random Forest para predecir nitrógeno total (NT), temperatura y salinidad en aguas interiores y costeras. En aguas interiores, para NT, reportaron un RMSE = 0.509, lo que evidencia alta precisión. En contraste, el presente estudio calculó nitritos y nitratos con RMSE = 0.0359 y 0.03232, respectivamente, superando los resultados de Xu et al. Esta diferencia podría atribuirse al volumen de datos utilizado: mientras Xu et al. emplearon 1,917 muestras para aguas interiores, este trabajo utilizó solo 15 muestras. Pese a la disparidad en tamaño muestral, los resultados del estudio referenciado validan metodológicamente los hallazgos aquí presentados.

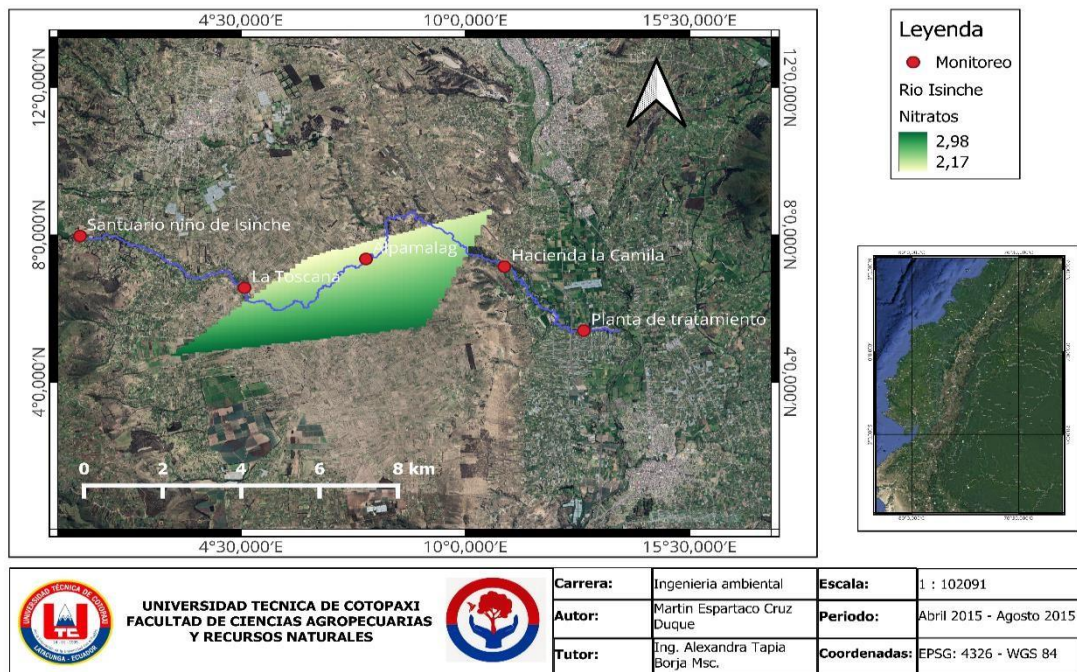
Figura
16. Mapa de predicción nitritos.



En la figura 16, se puede observar una clara tendencia de aumento en la concentración de nitritos dentro de la cuenca media del río Isinche, tendencia que puede ser explicada por la zona en la que se encuentra, entre los puntos de muestreo P2 y P3, pues en los recorridos realizados se observó una gran cantidad de invernaderos destinados a la producción agrícola de productos alimenticios y rosas, situación que se observó en menor medida dentro del área de los puntos de muestreo P1, P4 y P5. Sin embargo, se puede observar que el aumento de los nitritos se mantiene en valores bajos, lo que representa un factor a favor del cumplimiento de la relación N:P – 15:1 establecida por la normativa TULSMA.

Figura

17. Mapa de predicción Nitratos



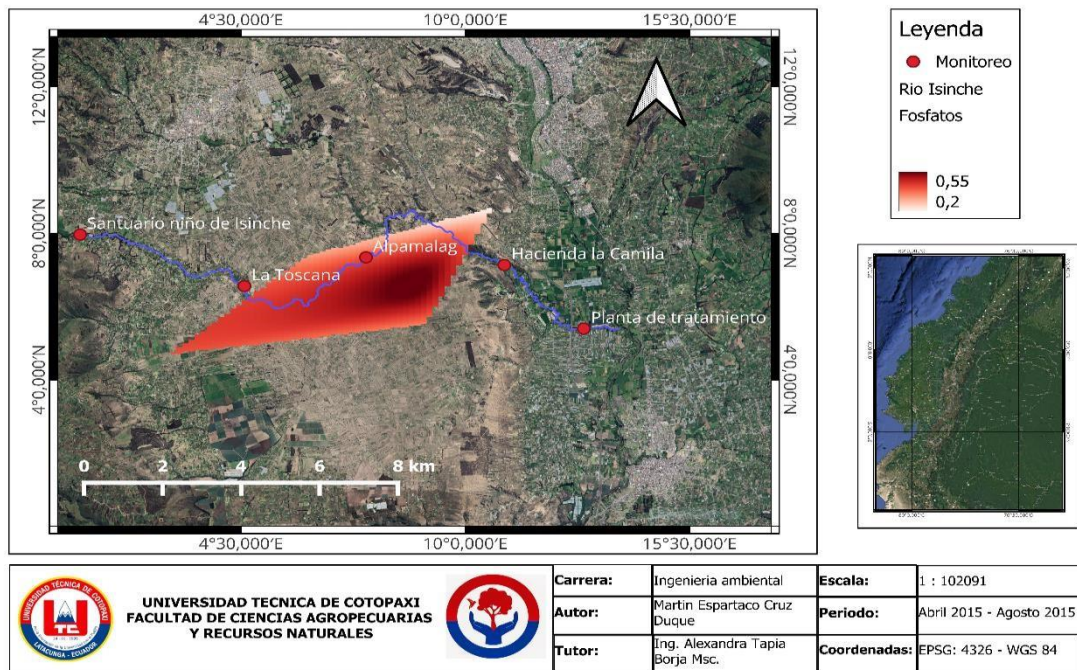
En cuanto a nitratos, se puede visualizar un claro aumento en los valores de los nitratos entre los puntos P2 y P3, es decir, el aumento se da desde el punto intermedio entre la cuenca alta y media del río, hacia la cuenca media del mismo, se puede visualizar un arrastre de nitratos sentido norte – sur, es decir viene desde los límites de Pujilí hasta concentrarse en los poblados aledaños al río Isinche, como es el caso de Alpamalag de Acurios.

10.4.2. Fosfatos

Respecto a la modelización de fosfatos, el estudio de Bhattarai et al., (2021) aplicó Random Forest en cinco cuencas, reportando para fósforo en zonas rurales un RMSE de 0.0539. Al comparar este resultado con el RMSE de 0.0960 obtenido en la cuenca rural del río Isinche (presente estudio), se observa que nuestro modelo presenta un rendimiento inferior en precisión predictiva. Pese a esta diferencia, nuestros valores de error se mantienen dentro de un rango aceptable para proyectar escenarios futuros de fosfatos en la cuenca estudiada.

Figura

18. Mapa de predicción fosfatos



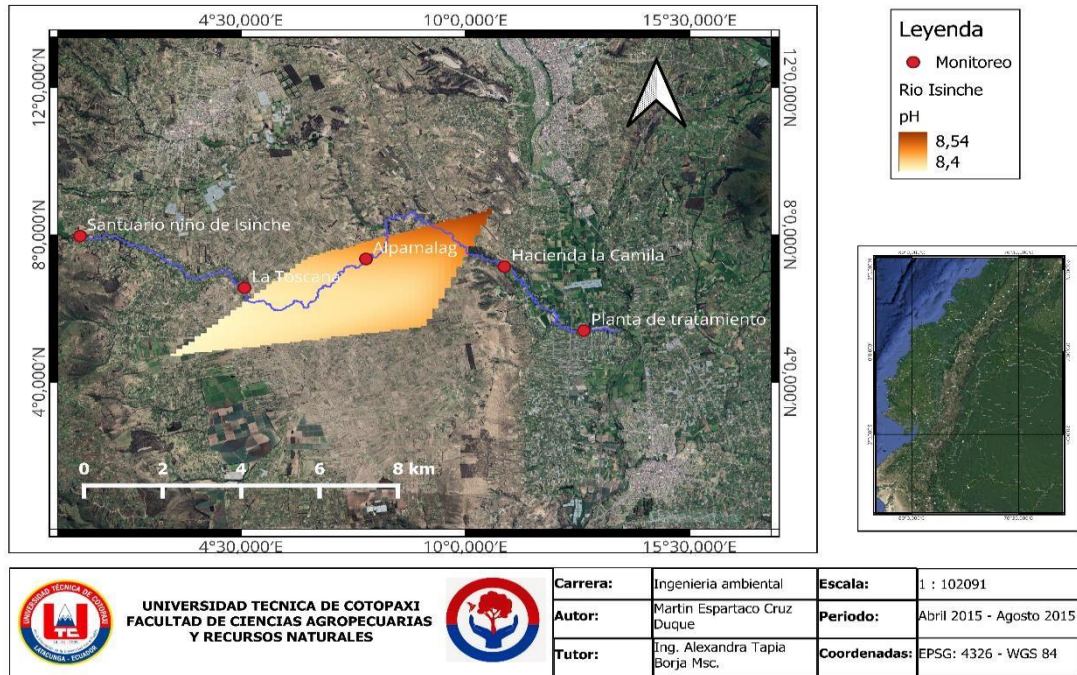
Se observa, en la figura 18, un aumento en la concentración de fosfatos en la cuenca media del río Isinche, que va acorde a los resultados obtenidos en los 3 muestreos para el punto de muestreo P3, cabe destacar que, para la fecha 11 de agosto los valores predichos superaron a los valores obtenidos en la recolección de muestras (Promedio 0.37 mg/L < Predicho 0.55 mg/L), lo que estaría desbalanceando la relación N:P de 15:1 establecidos por normativa TULSMA. En contraparte, se refleja una disminución de los fosfatos a partir de la ubicación intermedia entre el punto P3 y P5. Este aumento en las concentraciones de fosfatos, puede derivar en problemas de crecimientos algales, por consecuencia, reduciendo el OD disponible para los ecosistemas acuáticos.

10.4.3. pH

Para la predicción de pH, nuestro modelo de Random Forest registró RMSE = 0.0775, superando significativamente al estudio de Jena et al., (2022) (RMSE = 0.14). Esta diferencia – un 44.6% menos de error – podría atribuirse a que su enfoque incluyó metales como predictores, aumentando la complejidad del modelo. Pese a su precisión global del 90%, nuestros resultados demuestran mayor eficacia en la estimación de concentraciones.

Figura

19. Mapa predicción pH



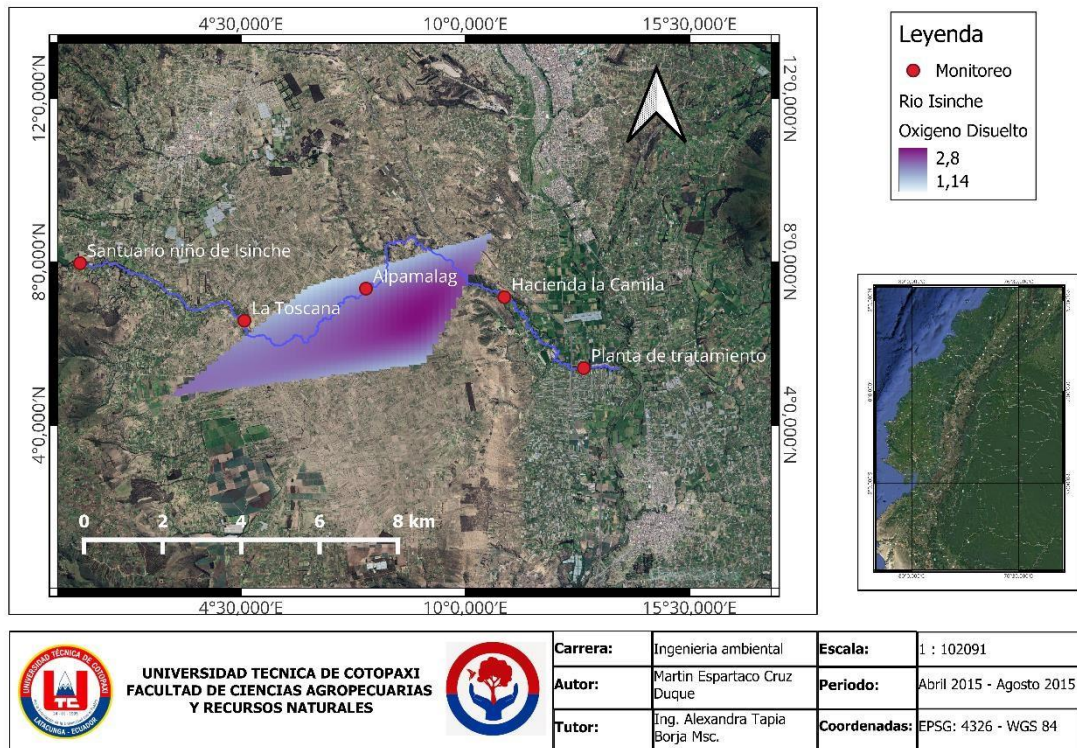
Los valores de pH para el futuro cercano (11 de agosto), presentan un aumento significativo con respecto a su valor promedio (Promedio 8.47 < Predicho 8.54), por otro lado, se observa un decaimiento en los valores de pH mientras más cerca se encuentra de la cuenca alta. Esta diferencia entre las concentraciones de pH entre el punto P1 y P3, P4, puede deberse a la ausencia de lluvias, que reduce el caudal del río y por consecuencia, factores como la meteorización de rocas calcáreas van a contribuir en menos medidas al aumento del pH.

10.4.4. OD

El estudio de Zavareh et al., (2024) evaluó la capacidad predictiva de Random Forest para oxígeno disuelto (OD) en 11 cuencas de EE. UU., integrando variables hidrometeorológicas, fisiográficas y de cobertura terrestre. Reportaron un RMSE = 0.10 en su escenario óptimo, que incluyó estos factores. En contraste, nuestro estudio obtuvo un RMSE = 1.2211, valor significativamente mayor. Esta diferencia puede atribuirse a la omisión en nuestro modelo de variables clave como las características fisiográficas del terreno y la cobertura terrestre, identificadas por Zavareh como determinantes para la precisión en OD. Pese a esta divergencia,

Figura

los resultados obtenidos en este estudio se mantienen dentro de rangos aceptables para aplicaciones prácticas.

20. Mapa de predicción OD

En la figura 20 se puede observar una mejora en los porcentajes de saturación del oxígeno disuelto en la cuenca media del río, de igual forma, se aprecia un decaimiento mientras más se acerca el modelo a la cuenca alta. Factores ya reportados como la ausencia de lluvias, y por ende, un menor porcentaje de eutrofización debido a un menor número de esorrentías, arrastre de contaminantes y nutrientes, hacen que esta época sea clave para la toma de decisiones, pues al haber un mayor % de saturación en las épocas secas, como es el mes de agosto, hace que los esfuerzos en la implementación de alternativas para mitigar el impacto de las fuentes de contaminación puedan ser dirigidos con un mayor tiempo de reacción y poder enfocarse en otras alternativas.

Figura**11. IMPACTOS (TÉCNICOS, SOCIALES, AMBIENTALES O ECONÓMICOS)****11.1. Impactos técnicos**

Los datos recolectados en el proceso de toma de muestras, y aquellos generados en la predicción de los contaminantes orgánicos y parámetros físico – químicos, proporcionan una perspectiva general del estado del río tomando en consideración factores externos como las actividades

realizadas por las comunidades aledañas al río o la implementación de variables meteorológicas para su evaluación, además, esta investigación puede ser aprovechada como base para modelos de predicción futuros, fundamentando el desarrollo de nuevos modelos, especialmente para aquellos parámetros con desafíos de precisión en este estudio, sírvase como ejemplo los coliformes, tanto totales como fecales. Además, puede servir como herramienta de monitoreo histórico, mismo que amplía las bases de datos del río, al analizar cambios temporales en la calidad del agua, verificar el cumplimiento normativo y establecer comparativas con estudios similares. Finalmente, se identifican vacíos científicos, que pueden servir como campo de estudio, en áreas donde el modelo desarrollado en esta investigación presenta limitaciones.

11.2. Impactos ambientales

Esta investigación proporciona una evaluación integral de los impactos potenciales de la contaminación orgánica en la salud ecosistémica y pública del río Isinche, donde el modelo predictivo desarrollado identifica con precisión espacial áreas críticas de alto riesgo, permitiendo llevar a cabo una gestión preventiva mediante predicción temprana de zonas con altas concentraciones de nitratos (2.70 mg/l), zonas de eutrofización, que afectan fuentes de agua para consumo humano y representan una amenaza para la biodiversidad acuática, además, en base a la evaluaciones realizadas se pueden determinar zonas de protección de ecosistemas estratégicos a través de monitoreo predictivo de oxígeno disuelto (<4 mg/L) para prevenir mortalidad de especies en áreas ribereñas vulnerables a contaminación ganadera ($r=0.78$) y mitigación de riesgos para la salud pública con alertas tempranas cuando los coliformes totales superan 1000 UFC/100ml en comunidades ribereñas, junto con modelización de propagación de contaminantes durante épocas lluviosas (+30% precipitación).

12. PRESUPUESTO DEL PROYECTO

Recurso	Cantidad	Valor Unitario (\$)	Valor Total (\$)
Cooler	1	9.50	9.50
Frascos muestra de orina	15	0.50	7.50
Hielo	6	1	6
Análisis de laboratorio	3	200	600
Transporte			45
Total			668

13. CONCLUSIONES

- La selección de puntos de muestreo por observación directa permitió identificar los factores que influyen en la calidad del río Isinche y los usos predominantes de su agua.
- Los análisis de laboratorio revelaron que el agua del río no cumple los límites máximos permisibles (LMP) del Acuerdo 097-A (TULSMA) para la relación nitrógeno-fósforo, oxígeno disuelto y pH, y presenta acumulación creciente de coliformes totales y fecales, lo que implica riesgos sanitarios y ecológicos si no se implementan medidas correctivas.
- El modelo predictivo físico-químico ofrece estimaciones fiables en un horizonte de 60 días, reduciendo costes de muestreo, pero no es apto para predecir coliformes debido a la escasez de datos y sus altos márgenes de error.

14. RECOMENDACIONES

- Realizar estudios específicos sobre las actividades antropogénicas en las riberas del río, cuantificando su impacto en los parámetros de calidad del agua.
- Incorporar el análisis de tensoactivos y grasas y aceites para evaluar el cumplimiento normativo integral.
- Utilizar hardware y software con una mayor capacidad de procesamiento de datos y una base de datos más amplia para entrenar un modelo que presente mayores índices de precisión a la hora de realizar predicciones.

15. BIBLIOGRAFÍA

- Alomani, S. M., Alhawiti, N. I., & Alhakamy, A. (2022). Prediction of Quality of Water According to a Random Forest Classifier. *International Journal of Advanced Computer Science and Applications (Ijacs)*, 13(6), Article 6. <https://doi.org/10.14569/IJACSA.2022.01306105>
- Bhattarai, A., Dhakal, S., Gautam, Y., & Bhattarai, R. (2021). Prediction of Nitrate and Phosphorus Concentrations Using Machine Learning Algorithms in Watersheds with Different Landuse. *Water*, 13(21), Article 21. <https://doi.org/10.3390/w13213096>
- Bonilla, A. (2015, septiembre 5). Amoniac (NH₃) y amonio (NH₄⁺). *Revista Galenus*. <https://www.galenusrevista.com/amoniac-nh3-y-amonio-nh4/>
- Bungacho Michilena, M. L., & Quezada Suárez, P. A. (2021). *Ocurrencia y caracterización de escherichia coli resistente a cefalosporina de tercera generación mediado por la familia génica blactx-m en el curso del río Isinche Latacunga*. <http://repositorio.utc.edu.ec/handle/27000/10863>
- Campaña, A., Gualoto, E., & Chiluisa-Utreras. (2017). Evaluación físico-química y microbiológica de la calidad del agua de los ríos Machángara y Monjas de la red hídrica del distrito metropolitano de Quito. *Revista Bionatura*. <http://dx.doi.org/10.21931/RB/2017.02.02.6>
- CanquiLlusco, J. E. (2012). Georreferenciación. *Revista de Información, Tecnología y Sociedad*, 22.

- Castro, D. (2020). *Evaluación de la oxidación avanzada para la eliminación de coliformes fecales en la remediación del río Puente Lucía, cantón Guayaquil*.
[https://cia.uagraria.edu.ec/Archivos/CASTRO%20CHILA%20DENISSE%20YOSSA RY.pdf](https://cia.uagraria.edu.ec/Archivos/CASTRO%20CHILA%20DENISSE%20YOSSA%20RY.pdf)
- Cedeño-Muñoz, H. A. (2020). Calidad de la descarga de aguas servidas de las lagunas de tratamiento de aguas residuales al río carrizal. *Polo del Conocimiento*, 5(01), Article 01.
<https://doi.org/10.23857/pc.v5i01.1222>
- Cerda Corrotea, C. E. (2020). *Efectos de forzantes hidroclimáticas sobre la calidad de agua superficial en cuencas ubicadas entre las regiones de Coquimbo y Los Ríos: Análisis histórico y proyecciones climáticas para el periodo (2030-2060)*.
<https://repositorio.uchile.cl/handle/2250/176986>
- Danigno, J. (2015, diciembre 7). Correlación. *Revista Chilena de Anestesia*.
<https://revistachilenadeanestesia.cl/correlacion/>
- Drinking Water Program. (2023). *Coliform Bacteria in Drinking Water Supplies*.
[https://wwwhealth-nygov.translate.goog/environmental/water/drinking/coliform_bacteria.htm?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc](https://www.health.ny.gov/environmental/water/drinking/coliform_bacteria.htm?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc)
- EPA. (2022). *Aquatic Life Criteria and Methods for Toxics*.
<https://www.epa.gov/wqc/aquaticlife-criteria-and-methods-toxics>

- Federigi, I., Verani, M., & Carducci, A. (2017). Sources of bathing water pollution in northern Tuscany (Italy): Effects of meteorological variables. *Marine Pollution Bulletin*, 114(2), 843-848. <https://doi.org/10.1016/j.marpolbul.2016.11.017>
- Fernández, A. (2025, febrero 10). *MOOC Machine Learning y Big Data para la Bioinformática*. 5ª Edición. Universidad de Granada - Facultad De. <https://fciencias.ugr.es/en/34noticias/4666-mooc-machine-learning-y-big-data-para-la-bioinformatica-5-edicion>
- GAD Parroquial de Belisario Quevedo. (2011, agosto 22). Datos Generales. *GAD Parroquial de Belisario Quevedo*. <https://belisarioquevedo.gob.ec/cotopaxi/datos-generales/>
- GAD Parroquial de Salcedo. (2014). *El Cantón Salcedo*. <https://www.salcedo.gob.ec/?p=3569>
- González, B. (2018). Agua y ciudad: Análisis y perspectivas del consumo de agua en el municipio de Guatemala. *Revista Análisis de la Realidad Nacional*. https://www.academia.edu/40988424/Agua_y_ciudad_an%C3%A1lisis_y_perspectivas_del_consumo_de_agua_en_el_municipio_de_Guatemala
- Grundmann, P., & Maaß, O. (2017). Chapter 3.2.1—Wastewater Reuse to Cope With Water and Nutrient Scarcity in Agriculture—A Case Study for Braunschweig in Germany. En J. R. Ziolkowska & J. M. Peterson (Eds.), *Competition for Water Resources* (pp. 352-365). Elsevier. <https://doi.org/10.1016/B978-0-12-803237-4.00020-3>
- Guadie, A., Yesigat, A., Gatew, S., Worku, A., Liu, W., Ajibade, F. O., & Wang, A. (2021). Evaluating the health risks of heavy metals from vegetables grown on soil irrigated with untreated and treated wastewater in Arba Minch, Ethiopia. *Science of The Total Environment*, 761, 143302. <https://doi.org/10.1016/j.scitotenv.2020.143302>

- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación*. McGraw Hill España. <https://dialnet.unirioja.es/servlet/libro?codigo=775008>
- Hidalgo Pio, D. A., & Veintimilla Tipanguano, C. I. (2024). *Caracterización de la calidad de agua el río Isinche para el análisis de la descarga en la planta de tratamiento del campus CEASA*. <http://repositorio.utc.edu.ec/handle/27000/11732>
- Hou, L., Zhou, Z., Wang, R., Li, J., Dong, F., & Liu, J. (2022). Research on the Non-Point Source Pollution Characteristics of Important Drinking Water Sources. *Water*, 14(2), Article 2. <https://doi.org/10.3390/w14020211>
- Hounsino, P. S., Mama, D., Tchibozo, M. A. D., Boukari, M., & Sohounhloue, D. (2015). Microbiological Pollution Indication as Tracer for the Pollution of Well Water: The Example of the District of Abomey-Calavi (Benin). *Journal of Environmental Protection*, 6(4), Article 4. <https://doi.org/10.4236/jep.2015.64029>
- Huamán, W. S., Pastrana, P. A. P., Victoria, H. M. D., & Córdova, L. A. T. (2022). Aguas residuales en la calidad de agua del río. *GnosisWisdom*, 2(3), Article 3. <https://doi.org/10.54556/gnosiswisdom.v2i3.43>
- Ibana-Lopez, K., Sihuay-Perales, M., Garate-Quispe, J., Araújo-Flores, J., Herrera-Machaca, M., Alarcón Aguirre, G., Rodriguez-Achata, L., Ibana-Lopez, K., Sihuay-Perales, M., Garate-Quispe, J., Araújo-Flores, J., Herrera-Machaca, M., Alarcón Aguirre, G., & Rodriguez-Achata, L. (2021). Contaminación de agua superficial de la periferia urbana de Puerto Maldonado, al sureste de la amazonia peruana. *Revista de Investigaciones Veterinarias del Perú*, 32(6). <https://doi.org/10.15381/rivep.v32i6.20365>

- Jena, P. K., Rahaman, S. M., Das Mohapatra, P. K., Barik, D. P., & Patra, D. S. (2022). Surface water quality assessment by Random Forest. *Water Practice and Technology*, 18(1), 201-214. <https://doi.org/10.2166/wpt.2022.156>
- Lutz, M. (2013). *Learning Python, 5th Edition*[Book]. <https://www.oreilly.com/library/view/learning-python-5th/9781449355722/>
- Mancilla-Villa, O. R., Gómez-Villaseñor, L., Olguín-Lopez, J. L., Guevara-Gutiérrez, R. D., Hernández-Vargas, O., Ortega-Escobar, H. M., Flores-Magdaleno, H., Can-Chulim, Á., Sánchez-Bernal, E. I., Cruz-Crespo, E., Palomera-García, C., Mancilla-Villa, O. R., Gómez-Villaseñor, L., Olguín-Lopez, J. L., Guevara-Gutiérrez, R. D., HernándezVargas, O., Ortega-Escobar, H. M., Flores-Magdaleno, H., Can-Chulim, Á., ...
- Palomera-García, C. (2022). Contaminación orgánica por coliformes, Nitrógeno y Fósforo en los ecosistemas acuáticos de la cuenca Ayuquila-Armería, Jalisco, México. *Biotecnia*, 24(1), 5-14. <https://doi.org/10.18633/biotecnia.v24i1.1283>
- Mena Maisincho, E. G. (2020). *Evaluación de las tendencias de la contaminación del recurso hídrico de la parte baja de la Microcuenca del Río Cutuchi, en la Provincia de Cotopaxi, periodo 2019-2020*. <http://repositorio.utc.edu.ec/handle/27000/7099>
- Molano, Z., & Katerine, A. (2020). *Análisis descriptivo de la contaminación por residuos sólidos en el municipio de Solano-Caquetá*. <http://repository.unad.edu.co/handle/10596/36701>
- Nong, X., He, Y., Chen, L., & Wei, J. (2025). Machine learning-based evolution of water quality prediction model: An integrated robust framework for comparative application

on periodic return and jitter data. *Environmental Pollution*, 369, 125834.

<https://doi.org/10.1016/j.envpol.2025.125834>

Organización Mundial de la Salud. (2021). *Guidelines on Recreational Water Quality. Volume 1: Coastal and Fresh Waters* (1st ed). World Health Organization.

Organización Mundial de la Salud. (2023). *Agua para consumo humano*.

<https://www.who.int/es/news-room/fact-sheets/detail/drinking-water>

Ortega-Paredes, D., Barba, P., Mena-López, S., Espinel, N., Crespo, V., & Zurita, J. (2019).

High quantities of multidrug-resistant *Escherichia coli* are present in the Machángara urban river in Quito, Ecuador. *Journal of Water and Health*, 18(1), 67-76.

<https://doi.org/10.2166/wh.2019.195>

Ortiz-Prado, E., Simbaña-Rivera, K., Cevallos, G., Gómez-Barreno, L., Cevallos, D., Lister,

A., Fernandez-Naranjo, R., Ríos-Touma, B., Vásconez-González, J., & Izquierdo-

Condoy, J. S. (2022). Waterborne diseases and ethnic-related disparities: A 10 years nationwide mortality and burden of disease analysis from Ecuador. *Frontiers in Public*

Health, 10,

1029375. <https://doi.org/10.3389/fpubh.2022.1029375>

Osorio Rivera, M. A., Carrillo Barahona, W. E., Negrete Costales, J. H., Llor Lalvay, X. A., &

Riera Guachichullca, E. J. (2021). La calidad de las aguas residuales domésticas. *Polo*

del Conocimiento: Revista científico - profesional, 6(3), 228-245.

Pabón, S. E., Benítez, R., Sarria, R. A., Gallo, J. A., Pabón, S. E., Benítez, R., Sarria, R. A., &

Gallo, J. A. (2020). Contaminación del agua por metales pesados, métodos de análisis

y tecnologías de remoción. Una revisión. *Entre Ciencia e Ingeniería*, 14(27), 9-18.

<https://doi.org/10.31908/19098367.0001>

- Pauta, G., Vázquez, G., Abril, A., Torres Inga, C. S., Loja Sari, M., & Palta Vera, A. (2020). Indicadores bacteriológicos de contaminación fecal en los ríos de Cuenca, Ecuador. *Maskana*, 11(2), 46-57.
- Pavan, J. V., Masachessi, G., Prez, V. E., Di Cola, G., Re, V. E., & Nates, S. V. (2022). Evaluación de la calidad de aguas superficiales en espacios recreacionales, una propuesta integradora de marcadores químicos y microbiológicos. *Revista de la Facultad de Ciencias Médicas*, 79(2), 210-214. <https://doi.org/10.31053/1853.0605.v79.n2.33403>
- Pérez-Sane, M., Torres, C., & Lara-Borrero, J. (2024). Estimación, Manejo y Control de la Contaminación No Puntual por Escorrentía Superficial Ganadera: Una Revisión de literatura. *Revista Politécnica*, 53(1), 111-123. <https://doi.org/10.33333/rp.vol53n1.10>
- Pernía, B., Mero, M., Cornejo, X., & Zambrano, J. (2019). *IMPACTOS DE LA CONTAMINACIÓN SOBRE LOS MANGLARES DE ECUADOR*. 375-419.
- Prado, T., Rey-Benito, G., Miagostovich, M. P., Sato, M. I. Z., Rajal, V. B., Filho, C. R. M., Pereira, A. D., Barbosa, M. R. F., Mannarino, C. F., & da Silva, A. S. (2023). Wastewater-based epidemiology for preventing outbreaks and epidemics in Latin America – Lessons from the past and a look to the future. *Science of The Total Environment*, 865, 161210. <https://doi.org/10.1016/j.scitotenv.2022.161210>
- Ramirez, P. (2023). *Repositorio de Tesis DGBSDI: Evaluación in vitro de la relación mutuamente excluyente entre especies de cianobacterias tóxicas pertenecientes a los géneros Microcystis sp. Y Anabaena sp. Bajo tres condiciones de irradiancia luminosa*. <https://ru.dgb.unam.mx/handle/20.500.14330/TES01000836772?mode=full>
- Rock, C., & Rivera, B. (2014). *La Calidad deL agua, E. coli y su saLud*.

- Romero, M. Á. M., Tiza, D. R. H., Murillo, J. P. M., Cervantez, D. O. O., & Ordóñez, G. I. (2023). Método mixto de investigación: Cuantitativo y cualitativo. En *Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú*. Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú. <https://doi.org/10.35622/inudi.b.105>
- Sacoto Acaro, D., Andueza Leal, F. D., Sacoto Acaro, D., & Andueza Leal, F. D. (2020). Microbiología del agua termal del balneario Ilaló. Pichincha, Ecuador. *FIGEMPA: Investigación y Desarrollo*, 9(1), 18-25. <https://doi.org/10.29166/revfig.v1i1.2178>
- Sebesta, R. W. (2015). *Concepts of Programming Languages*. Pearson Education.
- Silva, J., Torres, P., & Madera, C. (2008). Reuso de aguas residuales domésticas en agricultura. Una revisión. *Agronomía Colombiana*, 26(2), 347-359.
- Singh, G., Panda, R. K., & Lamers, M. (2014). Modeling of daily runoff from a small agricultural watershed using artificial neural network with resampling techniques. *Journal of Hydroinformatics*, 17(1), 56-74. <https://doi.org/10.2166/hydro.2014.106>
- Sucasaire Pilco, J., & Ticona Vilcapaza, R. (2023). *Métodos estadísticos: Guía básica para el uso de la estadística inferencial en investigación*. Jorge Sucasaire Pilco. <https://hdl.handle.net/20.500.12390/3219>
- Tian, H., Yu, G.-A., Tong, L., Li, R., Huang, H. Q., Bridhikitti, A., & Prabamroong, T. (2019). Water Quality of the Mun River in Thailand—Spatiotemporal Variations and Potential Causes. *International Journal of Environmental Research and Public Health*, 16(20), Article 20. <https://doi.org/10.3390/ijerph16203906>
- Vargas, A. K. N., Calderón, J., Velásquez, D., Castro, M., & Núñez, D. A. (2020). Análisis de los principales sistemas biológicos de tratamiento de aguas residuales domésticas en

- Colombia. *Ingeniare. Revista chilena de ingeniería*, 28(2), 315-322.
<https://doi.org/10.4067/S0718-33052020000200315>
- Vidal, G., & Fernando, J. (2024). *ProspectAI - Modelos de Machine Learning y Preprocesamiento de Lenguaje Natural para la Clasificación Efectiva de Clientes*.
<http://repository.unad.edu.co/handle/10596/62821>
- Wang, P., Zhao, J., Xiao, H., Yang, W., & Yu, X. (2019). Bacterial community composition shaped by water chemistry and geographic distance in an anthropogenically disturbed river. *Science of The Total Environment*, 655, 61-69.
<https://doi.org/10.1016/j.scitotenv.2018.11.234>
- Wilkinson, J. L., Boxall, A. B. A., Kolpin, D. W., Leung, K. M. Y., Lai, R. W. S., GalbánMalagón, C., Adell, A. D., Mondon, J., Metian, M., Marchant, R. A., Bouzas-Monroy, A., Cuni-Sanchez, A., Coors, A., Carriquiriborde, P., Rojo, M., Gordon, C., Cara, M., Moermond, M., Luarte, T., ... Teta, C. (2022). Pharmaceutical pollution of the world's rivers. *Proceedings of the National Academy of Sciences of the United States of America*, 119(8), e2113947119. <https://doi.org/10.1073/pnas.2113947119>
- Xu, J., Xu, Z., Kuang, J., Lin, C., Xiao, L., Huang, X., & Zhang, Y. (2021). An Alternative to Laboratory Testing: Random Forest-Based Water Quality Prediction Framework for Inland and Nearshore Water Bodies. *Water*, 13(22), Article 22.
<https://doi.org/10.3390/w13223262>
- Zambrano Mero, J. D., Delgado Párraga, A. G., Zambrano Mero, E. T., Peñaherrera Villafuerte, S. L., Zambrano Mero, J. D., Delgado Párraga, A. G., Zambrano Mero, E. T., & Peñaherrera Villafuerte, S. L. (2022a). Contaminantes biológicos presentes en fuentes

de agua del centro-sur de la provincia de Manabí, Ecuador. *Siembra*, 9(2).

<https://doi.org/10.29166/siembra.v9i2.4011>

Zambrano Mero, J. D., Delgado Párraga, A. G., Zambrano Mero, E. T., Peñaherrera Villafuerte,

S. L., Zambrano Mero, J. D., Delgado Párraga, A. G., Zambrano Mero, E. T., &

Peñaherrera Villafuerte, S. L. (2022b). Contaminantes biológicos presentes en fuentes

de agua del centro-sur de la provincia de Manabí, Ecuador. *Siembra*, 9(2).

<https://doi.org/10.29166/siembra.v9i2.4011>

Zavareh, M., Maggioni, V., & Zhang, X. (2024). Assessing the efficiency of a random forest

regression model for estimating water quality indicators. *Meteorology Hydrology and*

Water Management, 11(2), 52-69. <https://doi.org/10.26491/mhwm/183734>

Zhao, C., Li, M., Wang, X., Liu, B., Pan, X., & Fang, H. (2022). Improving the accuracy of

nonpoint-source pollution estimates in inland waters with coupled satellite-UAV data.

Water Research, 225, 119208. <https://doi.org/10.1016/j.watres.2022.119208>