

UNIVERSIDAD NACIONAL DE CAJAMARCA

ESCUELA DE POSGRADO



UNIDAD DE POSGRADO DE LA FACULTAD DE CIENCIAS AGRARIAS

PROGRAMA DE DOCTORADO EN CIENCIAS

TESIS:

**MODELO HÍBRIDO BASADO EN LÓGICA DIFUSA Y SPLINES
CÚBICOS PARA LA DOSIFICACIÓN ÓPTIMA DE SULFATO DE
ALUMINIO EN LA POTABILIZACIÓN DE AGUA: CASO DE ESTUDIO
EPS SEDACAJ -- PLANTA EL MILAGRO, CAJAMARCA**

Para optar el Grado Académico de:

DOCTOR EN CIENCIAS

MENCIÓN: GESTIÓN AMBIENTAL Y RECURSOS NATURALES

Presentado por:

Mg. EVER ROJAS HUAMÁN

Asesor:

Dr. GLICERIO EDUARDO TORRES CARRANZA

Cajamarca, Perú

2026

CONSTANCIA DE INFORME DE ORIGINALIDAD

1. Investigador:
Ever Rojas Huamán
DNI: 26694311
Escuela Profesional/Unidad de Posgrado de la Facultad de Ciencias Agrarias. Programa de Doctorado en Ciencias, Mención: Gestión Ambiental y Recursos Naturales
2. Asesor: Dr. Glicerio Eduardo Torres Carranza
3. Grado académico o título profesional
☐ Bachiller ☐ Título profesional ☐ Segunda especialidad
☐ Maestro ☒ Doctor
4. Tipo de Investigación:
☒ Tesis ☐ Trabajo de investigación ☐ Trabajo de suficiencia profesional
☐ Trabajo académico
5. Título de Trabajo de Investigación:
MODELO HÍBRIDO BASADO EN LÓGICA DIFUSA Y SPLINES CÚBICOS PARA LA DOSIFICACIÓN ÓPTIMA DE SULFATO DE ALUMINIO EN LA POTABILIZACIÓN DE AGUA: CASO DE ESTUDIO EPS SEDACAJ - PLANTA EL MILAGRO, CAJAMARCA
6. Fecha de evaluación: **18/02/2026**
7. Software antiplagio: ☒ TURNITIN ☐ URKUND (OURIGINAL) (*)
8. Porcentaje de Informe de Similitud: **17%**
9. Código Documento: **3117:558185539**
10. Resultado de la Evaluación de Similitud:
☒ APROBADO ☐ PARA LEVANTAMIENTO DE OBSERVACIONES O DESAPROBADO

Fecha Emisión: **18/02/2026**

*Firma y/o Sello
Emisor Constancia*



.....
Dr. Glicerio Eduardo Torres Carranza
DNI: 26602048

* En caso se realizó la evaluación hasta setiembre de 2023

COPYRIGHT © 2026 by
EVER ROJAS HUAMÁN
Todos los derechos reservados



Universidad Nacional de Cajamarca
LICENCIADA CON RESOLUCIÓN DE CONSEJO DIRECTIVO N° 080-2018-SUNEDUC/D
Escuela de Posgrado
CAJAMARCA - PERÚ



PROGRAMA DE DOCTORADO EN CIENCIAS

ACTA DE SUSTENTACIÓN DE TESIS

MENCIÓN: GESTIÓN AMBIENTAL Y RECURSOS NATURALES

Siendo las 16:10 horas, del día 12 de febrero del año dos mil veintiséis, reunidos en el Auditorio de la Escuela de Posgrado de la Universidad Nacional de Cajamarca, el Jurado Evaluador presidido por el Dr. MARCIAL HIDELSO MENDO VELÁSQUEZ, Dr. NILTON EDUARDO DEZA ARROYO, Dr. JIMY FRANK OBLITAS CRUZ y en calidad de Asesor, el Dr. GLICERIO EDUARDO TORRES CARRANZA Actuando de conformidad con el Reglamento Interno de la Escuela de Posgrado y el Reglamento del Programa de Doctorado de la Escuela de Posgrado de la Universidad Nacional de Cajamarca, se inició la SUSTENTACIÓN de la tesis titulada: **MODELO HÍBRIDO BASADO EN LÓGICA DIFUSA Y SPLINES CÚBICOS PARA LA DOSIFICACIÓN ÓPTIMA DE SULFATO DE ALUMINIO EN LA POTABILIZACIÓN DE AGUA: CASO DE ESTUDIO EPS SEDACAJ – PLANTA EL MILAGRO, CAJAMARCA**; presentada por el Maestro en Ciencias en Matemática Aplicada **EVER ROJAS HUAMÁN**

Realizada la exposición de la Tesis y absueltas las preguntas formuladas por el Jurado Evaluador, y luego de la deliberación, se acordó aprobar por unanimidad con la calificación de diez y ocho (18) - excelente la mencionada Tesis; en tal virtud, el Maestro en Ciencias en Matemática Aplicada **EVER ROJAS HUAMÁN**, está apto para recibir en ceremonia especial el Diploma que la acredita como **DOCTOR EN CIENCIAS**, de la Unidad de Posgrado de la Facultad de Ciencias Agrarias, Mención **GESTIÓN AMBIENTAL Y RECURSOS NATURALES**

Siendo las 17:30 horas del mismo día, se dio por concluido el acto.


.....
Dr. Glicerio Eduardo Torres Carranza
Asesor


.....
Dr. Marcial Hidelo Mendo Velásquez
Presidente-Jurado Evaluador


.....
Dr. Nilton Eduardo Deza Arroyo
Jurado Evaluador


.....
Dr. Jimmy Frank Oblitas Cruz
Jurado Evaluador

A:

Dios, por ser guía en mi vida y fuente de esperanza constante.

Mis padres, **César Octavio Rojas Zegarra** (†) e **Ina Huamán**, por su amor incondicional,
sacrificio y ejemplo de integridad.

Mis hermanos, por su compañía silenciosa en los momentos difíciles y su aliento en los días
de esfuerzo.

Mis hijos, **Jorge, Anggie Sofía y Nataly**, razón profunda de mi lucha diaria, esperanza viva
del futuro que sueño construir.

A todos ellos, ofrezco con humildad este logro académico

La educación es el arma más poderosa que puedes usar para cambiar el mundo.

— Nelson Mandela

ÍNDICE GENERAL

Índice de tablas	xi
Índice de figuras	xii
Agradecimientos	xiv
Lista de abreviaciones y siglas	xv
Resumen	xx
Abstract	xxi
1 INTRODUCCIÓN	1
2 MARCO TEÓRICO	4
2.1 Antecedentes de la investigación	4
2.2 Bases teóricas	12
2.2.1 Fundamentos de la potabilización del agua	12
2.2.2 Proceso de coagulación - floculación	15
2.2.3 Dosificación de coagulantes	18
2.2.4 Tecnologías de tratamiento de agua	21
2.2.5 Pruebas de jarras y control de procesos	23
2.2.6 Modelamiento matemático en tratamiento de agua	28
2.2.7 Interpolación mediante splines cúbicos	31
2.2.8 Sistemas de lógica difusa	33
2.2.9 Fundamentos de sistemas híbridos	36
2.2.10 Modelos híbridos en dosificación	38
2.2.11 Integración de componentes	41
2.2.12 Metodologías de validación	44
2.2.13 Técnicas de optimización	46
2.2.14 Implementación práctica	49

2.3	Definición de términos básicos	53
3	MATERIALES Y MÉTODOS	55
3.1	Localización del estudio	55
3.1.1	Ubicación geográfica	55
3.1.2	Área de influencia	57
3.1.3	Características operativas del sistema	59
3.1.4	Condiciones climáticas	59
3.2	Materiales	60
3.2.1	Ubicación y características del sistema de tratamiento	60
3.2.2	Equipos e instrumentación de laboratorio	60
3.2.3	Herramientas computacionales	61
3.2.4	Fuentes de información	62
3.3	Metodología	63
3.3.1	Método para el diseño de puntos de control representativos	63
3.3.2	Método para desarrollo del componente de splines cúbicos	64
3.3.3	Método para construcción del sistema de lógica difusa	65
3.3.4	Método para integración del modelo híbrido	67
3.3.5	Método para evaluación del desempeño	69
3.3.6	Implementación del dashboard operativo	71
3.3.7	Validación de cumplimiento normativo	72
3.3.8	Control de calidad y trazabilidad experimental	72
4	RESULTADOS Y DISCUSIÓN	74
4.1	Diseño y validación de puntos de control representativos para dosificación . .	74
4.1.1	Caracterización estadística de los puntos de control experimentales .	75
4.1.2	Distribución espacial y representatividad de los puntos de control . .	76
4.1.3	Resultados de la validación experimental	77
4.1.4	Contraste con estudios internacionales	82

4.2	Desarrollo del componente de interpolación por splines cúbicos	84
4.2.1	Evolución del desempeño por fold de validación cruzada	85
4.2.2	Análisis de sensibilidad paramétrica del modelo spline	89
4.2.3	Superficie tridimensional de interpolación	91
4.2.4	Contraste del componente spline con enfoques internacionales	92
4.2.5	Síntesis de capacidades y limitaciones del componente spline	93
4.3	Construcción del sistema de lógica difusa	94
4.3.1	Configuración de funciones de pertenencia	95
4.3.2	Base de conocimiento y reglas de inferencia	96
4.3.3	Superficie de respuesta en dominio de calibración	98
4.3.4	Capacidad de extrapolación para condiciones extremas	99
4.3.5	Resultados de validación experimental del sistema difuso	100
4.3.6	Contraste del componente difuso con enfoques internacionales	101
4.3.7	Resiliencia ante perturbaciones operativas	101
4.3.8	Integración del modelo híbrido con criterio adaptativo	102
4.3.9	Resultados de validación experimental del modelo híbrido	103
4.3.10	Análisis de desempeño por regiones operativas	104
4.3.11	Comparación con enfoques híbridos internacionales	105
4.4	Evaluación de desempeño del modelo híbrido mediante validación experimental	107
4.4.1	Comparación experimental con métodos de referencia	109
4.4.2	Análisis diferencial entre métodos de interpolación	111
4.4.3	Evaluación de resiliencia ante condiciones operativas extremas	114
4.4.4	Análisis temporal y estabilidad operativa	116
4.4.5	Curvas de dosificación optimizada	118
4.4.6	Cumplimiento normativo de dosis óptimas experimentales	119
4.4.7	Análisis estadístico comparativo	120

ANEXOS	136
Anexo A: Tablas de Dosificación Desarrolladas	137
Anexo B: Instrumentación y Equipos Utilizados	141
Anexo C: Implementación Computacional del Modelo Híbrido	146

ÍNDICE DE TABLAS

1	Matriz de puntos de control experimentales para calibración del modelo híbrido	74
2	Estadísticos descriptivos de las variables experimentales (n=16 caudales)	75
3	Dosis óptimas obtenidas para condiciones operativas representativas	77
4	Resultados de validación cruzada k-fold para el modelo de interpolación spline cúbico natural	84
5	Universos de discurso y rangos operativos del sistema difuso	94
6	Configuración de funciones de pertenencia por variable	95
7	Muestra representativa de reglas difusas implementadas	96
8	Parámetros optimizados del modelo híbrido	102
9	Resultados de validación experimental del modelo híbrido	103
10	Desempeño del modelo híbrido por regiones operativas	104
11	Métricas de evaluación cuantitativa del modelo híbrido	107
12	Comparaciones post-hoc entre métodos de dosificación	120
A1	Relación Turbidez-Dosis para Caudal de 150 L s^{-1}	137
A2	Puntos de Control Establecidos para Calibración del Modelo Híbrido	139
B1	Especificaciones Técnicas del Equipo de Prueba de Jarras	141
B2	Instrumentos de Medición Utilizados en la Investigación	143
B3	Especificaciones Detalladas del Turbidímetro HACH 2100AN	143
C1	Funcionalidades del dashboard operativo	155
C2	Especificaciones técnicas del sistema	156

ÍNDICE DE FIGURAS

1	Ubicación geográfica de la planta de tratamiento de agua potable “El Milagro” - EPS SEDACAJ	56
2	Área de influencia de la Planta “El Milagro” en la ciudad de Cajamarca	58
1	Distribución espacial de turbidez por categorías operativas en función del caudal	76
2	Evolución de métricas de desempeño por fold en validación cruzada del modelo spline cúbico	85
3	Análisis de sensibilidad paramétrica del modelo spline cúbico ante perturbacio- nes en turbidez	89
4	Superficie tridimensional del modelo spline cúbico: relación caudal-turbidez- dosis	91
5	Superficie de respuesta del sistema difuso en el dominio de calibración (pH = 7.2)	98
6	Respuesta del sistema difuso ante turbideces extremas (3,500–25,000 UNT) . .	99
7	Evolución de métricas de desempeño del sistema difuso por fold (k=5)	100
8	Comparación de dosificación: método tradicional vs. modelo híbrido propuesto	109
9	Análisis diferencial entre interpolación lineal y splines cúbicos para caudal 180 L s^{-1}	111
10	Respuesta del modelo híbrido ante condiciones operativas extremas (turbidez = 15,000 UNT, pH = 5.8)	114
11	Serie temporal de turbidez de entrada y dosis híbrida estimada durante transi- ción estacional (abril–mayo 2024)	116
12	Curvas de dosificación optimizada generadas por el modelo híbrido para cau- dales operativos representativos ($160, 180, 200 \text{ L s}^{-1}$)	118
B1	Equipo de prueba de jarras utilizado para validación experimental de puntos de control.	142

B2	Material de laboratorio utilizado para preparación de muestras	144
C2	Encabezado institucional y formulario de Etapa 2	152
C3	Resultado de estimación y comparación de componentes en historial	153
C4	Panel de tendencias, exportación y pie institucional	154

AGRADECIMIENTOS

Agradezco profundamente a:

SEDACAJ S.A. por permitirme realizar las pruebas experimentales en sus instalaciones de la **Planta El Milagro**, y a sus operadores por su colaboración durante el desarrollo de la investigación.

Los docentes del **Doctorado en Gestión Ambiental** de la UNC por su excelencia académica.

Colegas, amigos y familiares por su apoyo incondicional.

Eternamente agradecido

LISTA DE ABREVIACIONES Y SIGLAS

AG	: Algoritmo Genético (<i>Genetic Algorithm</i> , AG)
ANFIS	: <i>Adaptive Neuro-Fuzzy Inference System</i> (Sistema de Inferencia Neuro-Difusa Adaptativo)
APHA	: <i>American Public Health Association</i> (Asociación Americana de Salud Pública)
AWWA	: <i>American Water Works Association</i> (Asociación Americana de Obras de Agua)
CAO	: Carbono Orgánico Total (<i>Total Organic Carbon</i> , TOC)
CPA	: Cloruro de Polialuminio (<i>Poly Aluminum Chloride</i> , PAC)
CPM	: Control Predictivo basado en Modelos (<i>Model Predictive Control</i> , MPC)
CPMDF	: Control Predictivo basado en Modelo Difuso (<i>Fuzzy Model Predictive Control</i> , FMPC)
CPMNL	: Control Predictivo de Modelo No Lineal (<i>Nonlinear Model Predictive Control</i> , NMPC)
D2FIS	: <i>Data-Driven Fuzzy Inference System</i> (Sistema de Inferencia Difusa Basado en Datos)
DBPs	: Subproductos de Desinfección (<i>Disinfection By-Products</i>)

DQO	: Demanda Química de Oxígeno (<i>Chemical Oxygen Demand</i> , COD)
EAM	: Error Absoluto Medio (<i>Mean Absolute Error</i> , MAE)
ECM	: Error Cuadrático Medio (<i>Mean Squared Error</i> , MSE)
ECI	: Expectativa Condicional Individual (<i>Individual Conditional Expectation</i> , ICE)
ENS	: Eficiencia de Nash-Sutcliffe (<i>Nash-Sutcliffe Efficiency</i> , NSE)
EPA	: <i>U.S. Environmental Protection Agency</i> (Agencia de Protección Ambiental de EE.UU.)
EPAM	: Error Porcentual Absoluto Medio (<i>Mean Absolute Percentage Error</i> , MAPE)
Ethernet/IP	: Protocolo Industrial Ethernet (<i>Ethernet Industrial Protocol</i>)
FCM	: <i>Fuzzy C-Means</i> (Agrupamiento Difuso de C-Medias)
HAAs	: Ácidos Haloacéticos (<i>Haloacetic Acids</i>)
IAE	: Inteligencia Artificial Explicable (<i>Explainable Artificial Intelligence</i> , XAI)
IARC	: Agencia Internacional para la Investigación del Cáncer (<i>International Agency for Research on Cancer</i>)
INACAL	: Instituto Nacional de Calidad (Perú)
INEI	: Instituto Nacional de Estadística e Informática (Perú)

IT2FLS	: <i>Interval Type-2 Fuzzy Logic System</i> (Sistema de Lógica Difusa de Intervalo Tipo 2)
IWA	: <i>International Water Association</i> (Asociación Internacional del Agua)
MEMS	: Múltiples Entradas y Múltiples Salidas (<i>Multiple Input Multiple Output</i> , MIMO)
MEUS	: Múltiples Entradas y Una Salida (<i>Multiple Input Single Output</i> , MISO)
MON	: Materia Orgánica Natural (<i>Natural Organic Matter</i> , NOM)
Modbus	: Protocolo de Comunicación Modbus (<i>Modbus Communication Protocol</i>)
MSR	: Metodología de Superficie de Respuesta (<i>Response Surface Methodology</i> , RSM)
NIST	: <i>National Institute of Standards and Technology</i> (Instituto Nacional de Estándares y Tecnología de EE.UU.)
OMS	: Organización Mundial de la Salud (<i>World Health Organization</i> , WHO)
OPC-UA	: <i>OPC Unified Architecture</i> (Arquitectura Unificada OPC)
OTASS	: Organismo Técnico de la Administración de los Servicios de Saneamiento
PMC	: Perceptrón Multicapa (<i>Multi-Layer Perceptron</i> , MLP)
PTAP	: Planta de Tratamiento de Agua Potable

RECM	: Raíz del Error Cuadrático Medio (<i>Root Mean Squared Error</i> , RMSE)
RECMN	: Raíz del Error Cuadrático Medio Normalizado (<i>Normalized Root Mean Square Error</i> , NRMSE)
RLM	: Regresión Lineal Múltiple (<i>Multiple Linear Regression</i> , MLR)
RNA	: Redes Neuronales Artificiales (<i>Artificial Neural Networks</i> , ANN)
SCADA	: <i>Supervisory Control and Data Acquisition</i> (Control Supervisor y Adquisición de Datos)
SDSA	: Sistema de Soporte a Decisiones Ambientales (<i>Environmental Decision Support System</i> , EDSS)
SEDACAJ	: Servicio de Agua Potable y Alcantarillado de Cajamarca
SENAMHI	: Servicio Nacional de Meteorología e Hidrología del Perú
SHAP	: <i>Shapley Additive Explanations</i> (Explicaciones Aditivas de Shapley)
SLDIT2	: Sistema de Lógica Difusa de Intervalo Tipo 2 (<i>Interval Type-2 Fuzzy Logic System</i> , IT2FLS)
SUB	: <i>Subtractive Clustering</i> (Agrupamiento Sustractivo)
SUNASS	: Superintendencia Nacional de Servicios de Saneamiento
SVM	: <i>Support Vector Machines</i> (Máquinas de Vectores de Soporte)

THM	: Trihalometanos (<i>Trihalomethanes</i>)
T-S	: Takagi-Sugeno (Modelo Difuso Takagi-Sugeno)
UNT	: Unidades Nefelométricas de Turbidez (<i>Nephelometric Turbidity Units</i> , NTU)
UV254	: Absorbancia Ultravioleta a 254 nm (<i>Ultraviolet Absorbance at 254 nm</i>)
WEF	: <i>Water Environment Federation</i> (Federación del Ambiente Acuático)
WRF	: <i>Water Research Foundation</i> (Fundación de Investigación del Agua)

RESUMEN

La presente investigación tuvo como propósito desarrollar un modelo híbrido basado en lógica difusa y splines cúbicos para optimizar la dosificación de sulfato de aluminio en la potabilización de agua, aplicado a la Planta El Milagro de EPS SEDACAJ – Cajamarca. La metodología se estructuró en cinco fases: identificación de 48 puntos de control validados experimentalmente, desarrollo del componente de interpolación por splines cúbicos, construcción de un sistema de lógica difusa tipo Mamdani con 19 reglas fundamentales, integración híbrida con ponderación adaptativa optimizada, y validación experimental mediante técnicas estadísticas. Los resultados demostraron desempeño destacable del modelo híbrido con coeficiente de determinación $R^2 = 0,9985$, error cuadrático medio RECM = $0,087 \text{ mg L}^{-1}$, error porcentual absoluto medio EPAM = $0,42 \%$ y eficiencia de Nash-Sutcliffe ENS = $0,9980$, mostrando mejoras del 48 a 68% respecto a benchmarks internacionales ($p < 0,001$) en comparación con el método tradicional. Se concluye que el modelo híbrido desarrollado combina exitosamente la precisión matemática de splines cúbicos ($R^2 = 0,99991$) con la capacidad de extrapolación de lógica difusa ($R^2 = 0,940$), logrando 100% de cumplimiento normativo en los escenarios evaluados y optimización efectiva de la dosificación bajo condiciones de alta variabilidad operativa.

Palabras clave: modelo híbrido, lógica difusa, splines cúbicos, dosificación óptima, sulfato de aluminio, tratamiento de agua potable.

ABSTRACT

This research aimed to develop a hybrid model based on fuzzy logic and cubic splines for optimal aluminum sulfate dosing in water treatment, applied to El Milagro Plant of EPS SEDACAJ – Cajamarca. The methodology was structured in five phases: identification of 48 experimentally validated control points, development of cubic spline interpolation component, construction of a Mamdani-type fuzzy logic system with 19 fundamental rules, hybrid integration with optimized adaptive weighting, and experimental validation through statistical techniques. Results demonstrated notable performance of the hybrid model with coefficient of determination $R^2 = 0,9985$, root mean square error $RMSE = 0,087 \text{ mg L}^{-1}$, mean absolute percentage error $MAPE = 0,42 \%$, and Nash-Sutcliffe efficiency $ENS = 0,9980$, showing improvements of 48 a 68 % compared to international benchmarks ($p < 0,001$) relative to traditional methods. It is concluded that the developed hybrid model successfully combines the mathematical precision of cubic splines ($R^2 = 0,99991$) with the extrapolation capability of fuzzy logic ($R^2 = 0,940$), achieving 100 % regulatory compliance in evaluated scenarios and effective dosing optimization under high operational variability conditions.

Keywords: hybrid model, fuzzy logic, cubic splines, optimal dosing, aluminum sulfate, water treatment.

CAPÍTULO 1

INTRODUCCIÓN

El tratamiento eficiente del agua potable constituye una problemática central para las empresas prestadoras de servicios de saneamiento, particularmente en contextos urbanos de crecimiento acelerado y condiciones hidroclimáticas variables. En este contexto, la dosificación óptima de coagulantes desempeña un rol fundamental en los procesos de potabilización, especialmente en la etapa de coagulación-floculación, donde la cantidad y oportunidad del reactivo condicionan la calidad del agua tratada, la estabilidad del proceso e incluso los costos operativos.

La planta de tratamiento de agua potable (PTAP) El Milagro, operada por la EPS SEDACAJ S.A. en la ciudad de Cajamarca, presenta dificultades significativas relacionadas con la variabilidad en la turbidez del agua cruda proveniente de fuentes superficiales, además de limitaciones en la estandarización de criterios para la selección de dosis de coagulante. Actualmente, dicha selección se basa fundamentalmente en la experiencia del operador y en la aplicación periódica de pruebas de jarras, lo que, aunque proporciona una respuesta basada en la experiencia, muestra limitaciones ante escenarios de alta incertidumbre, situaciones atípicas o condiciones no evaluadas previamente.

En este sentido, la presente investigación propone el desarrollo de un *modelo híbrido* que integre técnicas de **lógica difusa** y **splines cúbicos**, con el objetivo de optimizar la dosificación de sulfato de aluminio en función de variables operativas fundamentales (como la turbidez del agua cruda, el pH y el caudal de entrada). Esta combinación metodológica busca aprovechar la capacidad interpolativa que poseen los splines en escenarios bien documentados, junto con la capacidad de generalización y manejo de incertidumbre que ofrecen los sistemas difusos, particularmente en escenarios no registrados previamente o poco representados en

los datos históricos.

Este enfoque busca proporcionar a la PTAP El Milagro una herramienta computacional confiable y adaptable, que permita una toma de decisiones más precisa y estable en la selección de dosis coagulante. Para este fin se utilizaron más de 2,500 registros operativos recolectados entre los años 2018 y 2024, acompañados de datos experimentales obtenidos mediante pruebas de jarras bajo condiciones controladas, asegurando un fundamento experimental consistente para la calibración y validación del modelo propuesto.

Esta investigación tiene como objetivo general desarrollar un modelo híbrido basado en lógica difusa y splines cúbicos para optimizar la dosificación de sulfato de aluminio durante el proceso de potabilización de agua, específicamente en la Planta El Milagro de EPS SEDACAJ - Cajamarca. Para lograr esto, se establecieron los siguientes objetivos específicos: identificar y validar experimentalmente puntos de control representativos para diferentes condiciones operativas; desarrollar un componente de interpolación mediante splines cúbicos naturales que asegure precisión matemática en rangos calibrados; implementar un sistema de inferencia difusa tipo Mamdani para permitir extrapolación consistente hacia condiciones extremas; integrar de manera inteligente ambos enfoques mediante ponderación adaptativa; y validar experimentalmente el modelo híbrido mediante métricas estadísticas apropiadas y análisis comparativo con métodos tradicionales.

Este trabajo se enmarca dentro del programa de Doctorado en Ciencias con mención en Gestión Ambiental y Recursos Naturales, específicamente en la línea de investigación de tratamiento de aguas, residuos sólidos y gaseosos. La investigación busca aportar, desde un enfoque matemático-computacional, alternativas efectivas para la optimización de procesos de potabilización, contribuyendo al desarrollo de tecnologías ambientalmente sostenibles que mejoren la eficiencia operativa en el tratamiento de agua potable. El modelo híbrido desarrollado constituye una herramienta de gestión ambiental que permite la adaptación automática de procesos de coagulación-floculación frente a la variabilidad natural de la calidad del agua cruda, favoreciendo un mejor aprovechamiento de recursos químicos y reduciendo el impacto

ambiental asociado a la generación de lodos de tratamiento.

CAPÍTULO 2

MARCO TEÓRICO

2.1 Antecedentes de la investigación

Park et al. (2024) en su investigación realizada en la Corporación de Recursos Hídricos de Corea (K-water), titulada «Explicación y Optimización del Aprendizaje Automático para la Dosificación de Coagulantes en Plantas de Tratamiento de Agua mediante IAE (IA Explicable)», propusieron una metodología para optimizar y clarificar la dosificación de coagulantes en plantas de tratamiento de agua potable utilizando Inteligencia Artificial Explicable (IAE). La investigación buscó resolver los desafíos en la dosificación de coagulantes causados por la contaminación industrial y el cambio climático. Estos han vuelto insuficientes los sistemas convencionales de tratamiento automatizado. Para esto utilizaron técnicas específicas de IAE. Particularmente SHAP (*Shapley Additive exPlanations*) y ECI (*Individual Conditional Expectation*). Estas permitieron comprender las contribuciones de las variables y las predicciones individuales en el proceso de dosificación. Los resultados demostraron que el enfoque basado en IAE no solo mejoró la precisión en la dosificación de coagulantes. También proporcionó información valiosa a los operadores. Esto facilitó una toma de decisiones más informada. El modelo optimizado alcanzó un R-cuadrado de 0,95, superando al modelo base que tenía 0,92. Fue particularmente efectivo en la respuesta a alta turbidez durante los meses de verano. La integración de *machine learning* y IAE en el tratamiento de agua resultó efectiva para lograr operaciones más eficientes, transparentes y responsables.

El-Marmar et al. (2024) en un estudio realizado en Marruecos, titulado "Modelado y Optimización de la Coagulación-Floculación para el Tratamiento de Lixiviados", propusieron un modelo neuro-difuso Takagi-Sugeno (*ANFIS*) para predecir y optimizar la dosis de coagulantes y floculantes durante el proceso de clarificación en una planta de tratamiento de

lixiviados. El estudio pretende resolver las limitaciones de las pruebas de jarras tradicionales. Estas son lentas y no permiten un control automático del proceso de coagulación-floculación. Para esto utilizaron datos históricos de la planta para calibrar y validar el modelo *ANFIS*. Se emplearon variables como pH, turbidez y características del lixiviado. Los resultados obtenidos fueron comparados con los de una red neuronal de perceptrón multicapa (*PMC*) y un modelo de regresión lineal múltiple (*RLM*). Durante la fase de validación, el modelo *ANFIS* mostró un coeficiente de determinación ($R^2 = 0,92$). Superó ampliamente a la *PMC* ($R^2 = 0,65$) y a la *RLM* ($R^2 = 0,65$). El enfoque neuro-difuso es significativamente más preciso para predecir la dosis óptima de químicos. Esto reduce costos y mejora la eficiencia operativa de la planta. Los autores concluyen que el estudio destaca el potencial del modelo *ANFIS* para gestionar de manera efectiva las operaciones de tratamiento de lixiviados. Proporciona una herramienta avanzada para optimizar procesos y garantizar la calidad de salida del agua tratada. Estos resultados son relevantes para la gestión de instalaciones que buscan combinar precisión, eficiencia y sostenibilidad en sus operaciones.

Ferrández-Galceran et al. (2024) en Cataluña, España, realizaron el estudio «Desarrollo de un Modelo de Lógica Difusa Predictiva para el Control de la Tasa de Dosificación de Ozono en una Planta de Tratamiento de Agua Potable», orientado al desarrollo de un modelo predictivo basado en lógica difusa para optimizar la dosificación de ozono en la planta de tratamiento de agua potable de Figueres. El propósito fue disminuir los costos operativos y mejorar la sostenibilidad ambiental del proceso de tratamiento. Para esto seleccionaron dos parámetros fundamentales. La absorbancia *UV254* y la temperatura del agua fueron definidos en colaboración con los gerentes de la planta. A través de funciones de membresía y reglas de inferencia difusa, el modelo predijo la tasa óptima de dosificación de ozono. Los autores analizaron los datos en *MATLAB*. Las variables de entrada y salida se procesaron mediante las etapas de *fuzzificación*, inferencia y *desfuzzificación* para generar decisiones accionables. El modelo logró una reducción del 30 % en la tasa de dosificación de ozono. Esto generó una reducción del 14 % en el consumo de energía de la planta. También se redujo en 18.5 % la

formación de trihalometanos (*THM*). Actualmente el modelo está siendo validado continuamente. Está proyectada su incorporación en un sistema de soporte a decisiones ambientales (*SDSA*) para automatizar la toma de decisiones en la planta. En conclusión, la investigación evidencia que los modelos de lógica difusa pueden mejorar significativamente la eficiencia operativa. Reducen costos y minimizan el impacto ambiental en los procesos de tratamiento de agua potable.

Zanil y Ramzan (2023) desarrollaron en Malasia una investigación publicada en la *Revista de la ASEAN sobre Control de Procesos* titulada "Optimización de los Procesos del Tanque de Coagulación mediante un Sistema de Lógica Difusa Tipo-2 por Intervalos: Estudio de la Reducción de la Turbidez". El propósito fue optimizar parámetros fundamentales como la dosificación de coagulantes, la velocidad y tiempo de mezcla, el pH y la temperatura. Así se logra un tratamiento más eficiente y sostenible. Los autores compararon los resultados de los modelos *IT2FLS* y *AG* con las pruebas de jarras tradicionales bajo condiciones controladas. La precisión predictiva de los modelos se evaluó utilizando métricas como el error absoluto medio (*EAM*), el error cuadrático medio (*ECM*) y el error cuadrático medio raíz (*RECM*). Los resultados mostraron que el *AG* produjo valores de turbidez final ligeramente superiores. Sin embargo, el *IT2FLS* demostró una notable precisión al replicar las tendencias de las pruebas de jarras. El estudio concluyó que el *IT2FLS* tiene un gran potencial para predecir y optimizar los parámetros de coagulación con alta precisión. Esto contribuye a la eficiencia del tratamiento de aguas residuales, la sostenibilidad ambiental y la reducción de costos operativos. Los autores sugieren validar el modelo con conjuntos de datos más amplios y variados para garantizar su aplicabilidad en escenarios del mundo real. El estudio sobresale por su enfoque novedoso en la integración de técnicas de optimización y modelos difusos avanzados. Muestra potencial como alternativa efectiva para mejorar los procesos de tratamiento de agua.

Bressane et al. (2023) en un estudio desarrollado en São Paulo, Brasil, publicaron la investigación titulada "Sistema de Inferencia Difusa Basado Exclusivamente en Datos (No Híbrido) para la Dosificación de Coagulantes en una Planta de Tratamiento de Agua Potable:

Aprendizaje Automático para la Predicción Precisa en Tiempo Real”. Los autores buscaron desarrollar y validar un sistema de inferencia difusa basado exclusivamente en datos (*D2FIS*). Este debía ser capaz de predecir en tiempo real la dosis óptima de coagulante en una planta de tratamiento de agua potable. Se enfrentaron a las deficiencias de las pruebas de jarras tradicionales. Estas presentan escasa capacidad de adaptación dinámica y alta dependencia del criterio del operador. Para esto desarrollaron y calibraron un *D2FIS* utilizando un año de datos históricos de calidad del agua (2021). Se consideraron parámetros como pH, turbidez, color, flúor y cloruros. El coagulante estudiado fue cloruro de poli aluminio (*CPA*). Se comparó la precisión del *D2FIS* con diversos algoritmos de *machine learning*. Estos incluyeron redes neuronales artificiales, sistemas neuro-difusos adaptativos (*ANFIS*), máquinas de soporte vectorial y modelos basados en árboles. Se utilizó el error cuadrático medio (*RECM*) como métrica de evaluación. Se encontró que el *D2FIS* superó a todos los otros métodos evaluados. Alcanzó un *RECM* de $0,69 \text{ mg L}^{-1}$, frente a 1.09 del *ANFIS* y mayores errores en otros algoritmos. Esto evidencia la capacidad del *D2FIS* para manejar incertidumbres inherentes al proceso de coagulación. También realiza predicciones altamente precisas en tiempo real. Además, el *D2FIS* mostró una significativa mejora en la generalización tras ajustes adicionales en los parámetros de entrada. Los autores concluyen que el *D2FIS* se presenta como una herramienta útil para la dosificación automatizada de coagulantes. Permite reducir costos operativos, minimizar errores y optimizar la eficacia del tratamiento de agua potable. Todo ello con un enfoque práctico y adaptable a variaciones rápidas en la calidad del agua. Bouach y Benmamar (2023) desarrollaron la investigación «Predicción de la Dosificación de Coagulantes para el Tratamiento de Agua Potable» en la planta de tratamiento de agua potable Boudouaou, Argelia. Los autores buscaron desarrollar un modelo predictivo que permitiera optimizar la dosificación de coagulantes mediante redes neuronales artificiales (*RNA*), como una alternativa más precisa, rápida y eficiente frente al método tradicional de pruebas de jarras (*Jar-Test*). Para esto, diseñaron un modelo *RNA* de tres capas con dos neuronas en la capa oculta, utilizando como *inputs* los parámetros de calidad del agua cruda: temperatura,

pH, conductividad eléctrica, turbidez y alcalinidad total. El modelo fue entrenado con el 83 % de la base de datos histórica de la planta y validado con el 17 % restante. El modelo mostró un desempeño aceptable, alcanzando un coeficiente de correlación (R) de 0,7764, un error porcentual absoluto medio ($EPAM$) de 103.6727, un error cuadrático medio ($RECM$) de 0,2709 y un error absoluto medio (EAM) de 0,3381, evidenciando que es posible usar técnicas de inteligencia artificial para mejorar el control y gestión de los sistemas de tratamiento de agua potable.

Langat et al. (2022) en Kenia, durante la Conferencia de Investigación e Innovación Sostenible, presentaron el estudio «Estrategias de Control para la Remoción de Turbidez en Plantas de Tratamiento de Agua Potable: Una Revisión». El propósito fue evaluar las limitaciones de las técnicas convencionales para la remoción de turbidez en plantas de tratamiento de agua potable. También propusieron estrategias de control más avanzadas que permitan mejorar la eficiencia operativa y la calidad del agua tratada. Para esto realizaron una revisión exhaustiva de técnicas de control clasificadas en estrategias de *feedforward*, *feedback* y combinaciones *feedforward-feedback*. Los autores analizaron métodos como modelos predictivos basados en inteligencia artificial. Estos incluyen redes neuronales y sistemas de inferencia neuro-difusa (*ANFIS*). Se evaluaron sus ventajas y limitaciones en comparación con los métodos tradicionales como las pruebas de jarras. Se encontró que las estrategias avanzadas ofrecen una mayor precisión en la dosificación de coagulantes. El control predictivo basado en modelos (*CPM*) y los sistemas combinados son ejemplos. Estos reducen errores humanos y costos operativos. No obstante, encontraron limitaciones como la dependencia de datos precisos y la necesidad de modelos dinámicos más estables para adaptarse a la variabilidad en la calidad del agua cruda. Concluyen que el desarrollo de estrategias de control avanzadas es esencial para cumplir con los estándares de calidad del agua. También asegura el funcionamiento continuo de las plantas de tratamiento.

Guo et al. (2022) en un estudio desarrollado en China y publicado en la Revista de Ingeniería de Procesos del Agua, realizaron la investigación «Sistema de Control Difuso Adaptati-

vo para la Optimización de la Dosificación de Coagulantes». Los autores buscaron desarrollar un sistema de control inteligente que permita optimizar dinámicamente la dosificación de coagulantes en plantas de tratamiento de agua potable. El objetivo fue superar las deficiencias de los métodos tradicionales en contextos de alta variabilidad en la calidad del agua cruda. Utilizaron un sistema de inferencia neuro-difuso adaptativo (*ANFIS*) mejorado mediante algoritmos genéticos (*AG*) y *clustering* difuso (*FCM*) dentro de un marco de control predictivo (*CPM*). Este modelo multiobjetivo integra optimización y control para determinar los puntos de ajuste de la dosis de coagulante y las leyes de control correspondientes. Ejecutaron simulaciones en *MATLAB/Simulink*. Luego realizaron implementaciones prácticas en plantas de tratamiento de agua para validar su desempeño. Los resultados evidenciaron beneficios significativos del modelo propuesto. Se incluyen una reducción de costos operativos y una mayor estabilidad en la calidad del agua tratada. El enfoque confirmó su capacidad para adaptarse a condiciones dinámicas del agua cruda. Se establece como una alternativa efectiva para el control de la dosificación de coagulantes en instalaciones de tratamiento de agua. Los autores concluyen que la integración de técnicas avanzadas como *AG*, *FCM* y *ANFIS* en un marco de control predictivo es efectiva para optimizar procesos de tratamiento de agua potable. Mejora tanto la eficiencia operativa como la calidad del producto final.

Shahsavari et al. (2021) en una investigación realizada en Irán, publicaron el estudio titulado "Predicción de la Dosis Óptima de Coagulantes en Plantas de Tratamiento de Agua mediante el Desarrollo de Modelos Basados en Sistemas de Inferencia Difusa con Redes Neuronales Artificiales (*ANFIS*)". Buscaron predecir con alta precisión la dosis óptima de coagulantes en plantas de tratamiento de agua potable. El objetivo fue superar las limitaciones operativas y metodológicas de las pruebas de jarras tradicionales. Estas suelen ser lentas, subjetivas y sensibles a las variaciones en la calidad del agua cruda. Los autores aplicaron el método de *clustering* sustitutivo (*SUB*) para reducir el número de reglas y conexiones dentro del modelo *ANFIS*. Esto mejoró su capacidad de generalización y reconocimiento. Los parámetros utilizados como entrada incluyeron el pH, la turbidez del agua cruda, la alcalinidad,

la temperatura y la conductividad eléctrica. Los coagulantes evaluados fueron sulfato de aluminio (*alum*) y cloruro de poli aluminio (*CPA*). Los resultados mostraron que *ANFIS* logró coeficientes de correlación de 0,85 y 0,84, con errores cuadráticos medios (*RECM*) de 1,32 y 1,83 para las dosis de *alum* y *CPA*, respectivamente. Esto demuestra la efectividad del modelo para determinar la dosis óptima de coagulantes. Ofrece una solución confiable que aborda la complejidad inherente al proceso de coagulación. Los autores concluyen que el uso de *ANFIS* con *clustering* sustitutivo es una alternativa precisa y eficiente para la dosificación en plantas de tratamiento de agua. Supera los métodos tradicionales y se adapta a las variaciones en la calidad del agua cruda.

Tahraoui et al. (2021) realizaron en Argelia la investigación titulada "Optimización y Predicción de la Dosis de Coagulante para la Eliminación de Micropolutantes Orgánicos Basada en la Turbidez". Esta se orientó hacia la optimización y predicción de la dosis de coagulante necesaria para eliminar micropolutantes orgánicos y turbidez en el tratamiento de agua potable. Los autores evaluaron cuatro modelos matemáticos: metodología de superficie de respuesta (*MSR*), redes neuronales artificiales (*RNA*), máquinas de soporte vectorial (*SVM*) y el sistema neuro-difuso adaptativo (*ANFIS*). El objetivo fue identificar el método más eficaz y eficiente para predecir la dosificación de coagulantes basada en parámetros fisicoquímicos del agua. Para esto recolectaron muestras de agua cruda de diferentes localidades en la región de Médéa. Estas fueron ajustadas con ácido húmico para simular condiciones reales. Los parámetros analizados incluyeron pH, temperatura, conductividad, turbidez y alcalinidad total. Los autores utilizaron estos datos para entrenar y validar los modelos. Se emplearon métricas como el coeficiente de correlación y el error cuadrático medio (*RECM*) para evaluar el desempeño de cada modelo. El análisis mostró que el modelo *RNA* fue el más preciso. Alcanzó un coeficiente de determinación ajustado (R^2_{adj}) de 0,99400 y un *RECM* de 1,79535. Le siguió *SVM* ($R^2_{adj} = 0,99204$, *RECM* = 2,08927). Aunque el modelo *ANFIS* presentó resultados satisfactorios ($R^2_{adj} = 0,93227$, *RECM* = 6,14477), su alta complejidad resultó menos práctica. Tenía 309 parámetros. Por su parte, *MSR* ofreció un balance adecuado entre

precisión ($R^2_{adj} = 0,96878$, $RECM = 3,82890$) y simplicidad. Solo requirió 13 parámetros. En conclusión, *RNA* y *SVM* resultaron las mejores alternativas para la predicción precisa de la dosis de coagulante. Por otro lado, *MSR* es una alternativa económica y práctica para reducir costos operativos en plantas de tratamiento de agua. *ANFIS* presentó limitaciones debido a su alta complejidad. El trabajo favorece el desarrollo de herramientas avanzadas para optimizar la eficiencia en el tratamiento de agua potable.

Bello et al. (2020) realizaron en la Universidad Tecnológica de Tshwane, Sudáfrica, el estudio "Modelado Dinámico Difuso y Control Predictivo de una Unidad de Dosificación de Químicos de Coagulación para Plantas de Tratamiento de Agua". El objetivo fue implementar una estrategia de control predictivo basado en lógica difusa (*FMPC*) para regular variables de salida en unidades de dosificación química durante la coagulación. Aplicaron un enfoque basado en un modelo de proceso de múltiples entradas y salidas (*MEMS*) en forma de modelo difuso Takagi-Sugeno (*T-S*) linealizado. Este se obtuvo a través de agrupamiento sustractivo de datos de planta. Los autores describieron el modelo *MEMS* mediante modelos acoplados de múltiples entradas y una salida (*MEUS*). Se encontró un desempeño superior del controlador *FMPC* comparado con el control predictivo de modelo no lineal (*NMPC*). Las métricas de rendimiento fueron 21.25 y 26.42 respectivamente. El modelo mostró capacidad suficiente para rechazar perturbaciones. Este es un aspecto fundamental para el control de coagulación y optimización del tratamiento de agua. Los autores concluyeron que el *FMPC* resulta efectivo y práctico para procesos de tratamiento de agua. Es especialmente útil ante variaciones frecuentes en condiciones operativas.

2.2 Bases teóricas

2.2.1 Fundamentos de la potabilización del agua

Parámetros fisicoquímicos de calidad. Las directrices de la Organización Mundial de la Salud (OMS, 2022) establecen criterios bien definidos para caracterizar y monitorear los parámetros fisicoquímicos del agua potable. La turbidez indica presencia de partículas suspendidas. Pero también puede comprometer la desinfección si supera 5 UNT. Las plantas de tratamiento trabajan con valores menores a 1 UNT (medidos en Unidades Nefelométricas de Turbidez) para evitar este riesgo sanitario. La dureza depende de las concentraciones de calcio y magnesio, aunque no tiene límites sanitarios específicos. El problema aparece cuando supera 500 mg L^{-1} : deteriora la calidad organoléptica más que la seguridad. El cloro residual debe mantenerse en un rango estrecho ($0,2\text{-}0,5 \text{ mg L}^{-1}$), ya que ese equilibrio garantiza desinfección efectiva sin generar subproductos indeseables.

La Asociación Americana de Obras de Agua (AWWA, 2023b) establece límites adicionales a estos criterios. El color verdadero (medido en unidades Pt-Co) es uno de ellos. Si supera 15 unidades, hay metales disueltos o compuestos orgánicos que el tratamiento no eliminó. El oxígeno disuelto también es relevante. Cuando cae por debajo de 6 mg L^{-1} se crean zonas anaeróbicas en las tuberías, lo cual desestabiliza el agua químicamente y corroe la infraestructura.

Respecto a los rangos operativos específicos, la Agencia de Protección Ambiental, EE.UU. (EPA, 2023) define parámetros fundamentales para el control fisicoquímico del agua potable. La conductividad eléctrica debe estar entre $50\text{--}1500 \mu\text{S cm}^{-1}$. Este parámetro indica cuántas sales disueltas hay en el agua y su capacidad de conducir corriente. Para el Carbono Orgánico Total (CAO) se necesitan remociones del 15 al 35 % según las concentraciones iniciales. El agua cruda varía mucho dependiendo de su origen.

Según la Asociación Americana de Salud Pública et al. (APHA, 2017) el pH es el factor más importante en el tratamiento de agua potable porque define la eficacia de la coagu-

lación, la desinfección y también la prevención de corrosión en el sistema. Un pH entre 6.5 y 8.5 permite que estos procesos trabajen adecuadamente.

Por su parte, Davis (2023) caracteriza la alcalinidad (expresada como $\text{mg L}^{-1} \text{CaCO}_3$) como el principal sistema buffer que protege al agua de cambios bruscos de pH cuando se añaden ácidos. Con valores entre 30 y 400 mg L^{-1} los sistemas de tratamiento funcionan de forma eficaz y estable.

Crittenden et al. (2021) determina normas técnicas para los parámetros de contaminación orgánica en agua potable tratada. La Demanda Química de Oxígeno (DQO) debe estar por debajo de 10 $\text{mg L}^{-1} \text{O}_2$. Este valor indica cuánta materia orgánica puede oxidarse químicamente en el sistema. La importancia de estos parámetros depende de varios factores: la fuente del agua cruda, la tecnología de tratamiento, las normas locales y los objetivos de calidad. Por eso se requiere monitoreo constante. Los muestreos deben ser frecuentes, los métodos analíticos estandarizados y el control de calidad riguroso

Procesos unitarios. El tratamiento convencional de agua potable combina procesos químicos y físicos, según Crittenden et al. (2021). Los procesos físicos incluyen filtración con membranas o medios granulados, sedimentación por gravedad para separar partículas y cribado para eliminar sólidos gruesos. Estos permiten remover contaminantes de diferentes tamaños del agua cruda.

En cuanto a los procesos químicos, OMS (2022) señala tres esenciales. La coagulación desestabiliza coloides. La floculación agrupa partículas. Y la desinfección elimina patógenos. Trabajan junto con los físicos para remover contaminantes químicos y microbiológicos, lo que garantiza la calidad sanitaria del agua tratada.

Ahora bien, la Fundación de Investigación del Agua (WRF, 2024f) señala que la eficiencia de cada proceso unitario depende mucho de su diseño y operación. En la práctica estos procesos están interconectados. No trabajan aislados. Su optimización requiere ver el sistema completo, ya que el desempeño de cada etapa afecta las siguientes en el tren de tratamiento.

Marco regulatorio internacional y nacional. La OMS (2022) define las directrices globales para la calidad del agua potable. Así mismo, actualiza regularmente estas guías con nuevos parámetros y límites basados en la evidencia científica más reciente. Esto permite que los países desarrollen sus propias normativas adaptadas a sus condiciones específicas.

En el contexto nacional peruano, el Ministerio de Salud (2010) define el marco regulatorio principal a través del Reglamento de la Calidad del Agua para Consumo Humano (DS N° 031-2010-SA). Este establece los límites máximos permisibles para diversos parámetros fisicoquímicos y microbiológicos. El reglamento constituye la base normativa que deben cumplir todas las empresas prestadoras de servicios de saneamiento en el territorio nacional. Garantiza estándares uniformes de calidad del agua potable.

Por su parte, la Superintendencia Nacional de Servicios de Saneamiento (SUNASS, 2023) supervisa que las empresas prestadoras cumplan estos estándares. Esta entidad reguladora realiza inspecciones y evaluaciones constantes. Verifica el cumplimiento de los parámetros establecidos. Así controla la calidad del agua potable en todo el país.

Tendencias en automatización. La Asociación Internacional del Agua (IWA, 2024a) señala que la automatización en plantas de agua potable ha crecido rápidamente gracias a los avances en tecnologías de control y monitoreo. Los sistemas de control avanzado ayudan a optimizar las operaciones de las plantas. El control predictivo y los sistemas adaptativos son ejemplos. Estos algoritmos se ajustan automáticamente cuando cambia la calidad del agua cruda o las condiciones de operación.

Sobre este punto, WRF (2024f) indica que la integración de sensores en línea y sistemas SCADA modernos facilita el monitoreo continuo de parámetros críticos. También permite tomar decisiones en tiempo real. La digitalización ahora incluye gemelos digitales y análisis de big data. Esto permite una gestión más eficiente del tratamiento mediante la simulación virtual de procesos y el análisis predictivo de tendencias operativas.

Ahora bien, AWWA (2023b) destaca que la ciberseguridad se ha vuelto fundamental en la automatización de plantas de tratamiento. Proteger los sistemas SCADA es esencial.

También lo es asegurar los datos operativos y crear protocolos de respuesta ante incidentes. Sin estas medidas los ataques pueden comprometer seriamente el suministro de agua.

2.2.2 *Proceso de coagulación - floculación*

Teoría de la coagulación. Crittenden et al. (2021) afirman que la coagulación es un proceso fundamental en el tratamiento de agua potable. Consiste en desestabilizar partículas coloidales mediante coagulantes químicos. El proceso neutraliza las cargas superficiales de las partículas. Esto reduce las fuerzas repulsivas entre ellas y permite su agregación posterior. La efectividad de la coagulación depende de múltiples factores: el pH del agua, la dosis de coagulante, la intensidad de mezclado y las características fisicoquímicas de las partículas coloidales en el agua cruda.

Respecto a la teoría, Gregory y Duan (2024) indican que la teoría DLVO sigue siendo el marco principal para explicar la estabilidad coloidal y los mecanismos de coagulación. Esta teoría fue desarrollada por Derjaguin, Landau, Verwey y Overbeek. Describe la interacción entre partículas coloidales como el balance entre fuerzas de repulsión electrostática y atracción de Van der Waals. Esto proporciona la base científica para comprender los fenómenos de desestabilización durante la coagulación. También permite diseñar estrategias racionales de dosificación de coagulantes.

Mecanismos de desestabilización. Sillanpää y Matilainen (2023) señalan que los principales mecanismos de desestabilización incluyen cuatro procesos fundamentales. Estos actúan de manera individual o combinada según las condiciones específicas del tratamiento. La compresión de la doble capa eléctrica ocurre mediante el incremento de la fuerza iónica del medio. Esto reduce el espesor de la capa difusa alrededor de las partículas. Es particularmente efectivo en aguas con alta concentración de electrolitos. La neutralización de cargas implica la adsorción de iones o polímeros de carga opuesta sobre la superficie de las partículas coloidales. Esto anula las fuerzas repulsivas que mantienen estables las suspensiones coloidales. Es el mecanismo predominante en la mayoría de aplicaciones de coagulación con

sales metálicas. El atrapamiento en precipitados (también conocido como sweep flocculation) es un mecanismo donde las partículas son incorporadas en precipitados voluminosos del coagulante. Resulta efectivo para remover partículas finas y material coloidal que no pueden ser desestabilizadas eficientemente por otros mecanismos. Por último, el puente entre partículas es característico de polímeros de alto peso molecular que pueden unir múltiples partículas simultáneamente. Es especialmente relevante cuando se emplean ayudantes de coagulación poliméricos o coagulantes orgánicos.

Cinética de la coagulación. Zhang et al. (2024) indican que la cinética de coagulación sigue típicamente un modelo de segundo orden. La velocidad de agregación depende de la frecuencia de colisiones efectivas entre partículas. El proceso se puede dividir en dos regímenes distintos según el mecanismo dominante de transporte. La coagulación pericinéctica está dominada por el movimiento browniano. Predomina para partículas submicrónicas. Las fuerzas térmicas proporcionan la energía necesaria para el contacto entre partículas. Por otro lado, la coagulación ortocinéctica está controlada por gradientes de velocidad del fluido. Es el mecanismo principal para partículas de mayor tamaño, donde la agitación mecánica facilita los encuentros interparticulares.

Sobre este punto, Wang y Yang (2023) indican que la eficiencia del proceso está determinada por el factor de efectividad de colisiones (α). Este representa la fracción de colisiones que resultan en agregación permanente. El factor es adimensional y varía entre 0 y 1. Valores cercanos a la unidad indican condiciones óptimas de coagulación. Así se puede medir cuantitativamente la probabilidad de que los encuentros entre partículas resulten en la formación de agregados estables durante el tratamiento.

Variables críticas del proceso. Crittenden et al. (2021) señalan que la eficiencia del proceso de coagulación está gobernada por diversas variables críticas que interactúan de manera compleja. El pH es uno de los parámetros más significativos. Influye directamente en la especiación del coagulante, la carga superficial de las partículas coloidales y la solubilidad

de los productos de hidrólisis. Por eso su control preciso es fundamental para optimizar la desestabilización de partículas.

En cuanto a la dosis de coagulante, AWWA (2023b) señala que debe ajustarse según las características del agua cruda. Una dosis insuficiente no logra desestabilizar todas las partículas. Una dosis excesiva puede reestabilizar los coloides por inversión de carga. También aumenta costos operativos y genera mayor producción de lodos. El punto óptimo de dosificación se determina mediante pruebas de jarras en laboratorio.

Por su parte, Sillanpää y Matilainen (2023) indican que la temperatura del agua afecta significativamente el rendimiento del proceso. Modifica la velocidad de las reacciones químicas involucradas en la coagulación. También altera la viscosidad del medio acuoso. Y afecta la cinética de formación de flóculos. Las variaciones estacionales de temperatura requieren ajustes operativos para mantener la eficiencia del tratamiento bajo diferentes condiciones climáticas.

Por otra parte, Wang y Yang (2023) afirman que determinar la dosis óptima de coagulante es un aspecto crítico. Depende de múltiples factores interrelacionados. Las características del agua cruda incluyen turbiedad, alcalinidad y contenido de materia orgánica. También influyen el tipo específico de coagulante empleado y el mecanismo de coagulación predominante. Todos estos factores determinan los requerimientos de dosificación. Optimizar este parámetro garantiza tanto la eficiencia del proceso como la economía del tratamiento.

Respecto al mezclado, Zhang et al. (2024) señalan que los parámetros de intensidad y duración juegan un papel fundamental en dos etapas distintas. La mezcla rápida inicial asegura una dispersión uniforme del coagulante en toda la masa de agua. La mezcla lenta posterior promueve el contacto entre partículas desestabilizadas. Así se forman flóculos de tamaño adecuado. Optimizar estos parámetros maximiza la eficiencia del proceso de coagulación-floculación.

2.2.3 *Dosificación de coagulantes*

Tipos y características de coagulantes. Sillanpää y Matilainen (2023) establecen que los coagulantes utilizados en el tratamiento de agua potable se clasifican principalmente en sales metálicas y polímeros orgánicos. Cada uno tiene características específicas que determinan su eficacia en diferentes condiciones operativas. Las sales de aluminio son los coagulantes más empleados globalmente, siendo el sulfato de aluminio particularmente común debido a su costo-efectividad y amplio rango de aplicación. Los cloruros de polialuminio (CPA) han ganado popularidad en años recientes por su mejor desempeño en aguas frías y menor sensibilidad al pH, lo que proporciona mayor flexibilidad operativa en sistemas que experimentan variaciones estacionales significativas.

Por su parte, Crittenden et al. (2021) indican que las sales férricas como el cloruro férrico y el sulfato férrico ofrecen ventajas específicas en aguas con alto contenido orgánico. Estas sales operan en condiciones de pH más amplias que las de aluminio. Los coagulantes poliméricos (tanto naturales como sintéticos) se utilizan cada vez más como auxiliares de coagulación, mejorando la formación de flóculos y reduciendo la producción de lodos, lo que representa beneficios operativos y ambientales.

Modelos de dosificación. Wang y Yang (2023) indican que los modelos de dosificación han evolucionado desde simples relaciones empíricas hasta sofisticados algoritmos que incorporan múltiples variables y técnicas de inteligencia artificial. Los modelos tradicionales basados en la turbiedad del agua cruda mantienen su relevancia en operaciones básicas, pero presentan limitaciones significativas en condiciones variables. Los modelos matemáticos avanzados incorporan parámetros como pH, temperatura, alcalinidad y materia orgánica para predecir dosis óptimas con mayor precisión, superando así las restricciones de los enfoques unidimensionales.

Respecto a esto, Zhang et al. (2024) indican que las técnicas de modelamiento adaptativo y redes neuronales artificiales han demostrado particular eficacia en la optimización

de dosis. Son especialmente útiles en plantas que experimentan variaciones significativas en la calidad del agua cruda. Estos modelos pueden ajustarse en tiempo real basándose en datos operativos y resultados de tratamiento. Esto proporciona capacidades de aprendizaje continuo que mejoran la precisión predictiva a medida que se acumula experiencia operativa.

Factores que afectan la dosificación. AWWA (2023b) menciona que la determinación de dosis óptimas está influenciada por múltiples factores interrelacionados que deben considerarse de manera integral. Las características del agua cruda son los parámetros primarios. Estas incluyen turbiedad, color, pH y temperatura. La presencia de materia orgánica natural y la concentración de iones específicos pueden modificar significativamente la demanda de coagulante y su eficacia, por lo que se requieren ajustes dinámicos en las estrategias de dosificación para mantener la eficiencia del proceso.

Sobre este punto, Sillanpää y Matilainen (2023) indica que las condiciones operativas de la planta también influyen en la selección de dosis óptimas. El caudal de tratamiento y las limitaciones de los sistemas de mezcla son ejemplos relevantes. Además, factores económicos y ambientales como el costo de los coagulantes y las restricciones en la disposición de lodos deben considerarse en la estrategia de dosificación, logrando así un balance apropiado entre eficiencia técnica, viabilidad económica y sostenibilidad ambiental.

Optimización del proceso. Wang y Yang (2023) afirman que la optimización de la dosificación de coagulantes requiere un enfoque integral que considere tanto la eficiencia del tratamiento como los aspectos económicos y ambientales. Los sistemas modernos de control incorporan monitoreo en línea de parámetros críticos y algoritmos de optimización que ajustan continuamente las dosis según las condiciones cambiantes del agua cruda. Esto proporciona capacidades de respuesta dinámica que mejoran significativamente la eficiencia operativa y la calidad del agua tratada.

Respecto a esto, Crittenden et al. (2021) indican que la implementación de estrategias de dosificación adaptativa permite mantener la calidad del agua tratada mientras se minimizan

los costos operativos y la generación de residuos. Estas estrategias se apoyan en sistemas automatizados de control y monitoreo. La validación continua del desempeño y el ajuste de los modelos de dosificación son esenciales para mantener la efectividad del proceso a largo plazo, asegurando así que los sistemas mantengan su eficiencia operativa ante las variaciones naturales en la calidad del agua cruda.

2.2.4 Tecnologías de tratamiento de agua

El tratamiento de agua ha evolucionado para adaptar sus procesos a diferentes condiciones de calidad de agua cruda y requerimientos operativos. Seleccionar la tecnología apropiada es fundamental para garantizar tanto la calidad del agua tratada como la sostenibilidad operativa de la planta (Davis, 2023).

Tipos de plantas de tratamiento. AWWA (2023b) establece que el tratamiento convencional completo es el sistema más completo y versátil, diseñado para tratar aguas superficiales con alta variabilidad en su calidad. Este sistema incluye secuencialmente el pretratamiento mediante cribado y presedimentación, seguido de coagulación-floculación, sedimentación, filtración rápida y desinfección. Es particularmente efectivo para aguas con turbiedad entre 50-1000 UNT y color verdadero entre 20-150 UC.

Sobre este punto, Crittenden et al. (2021) indican que el tratamiento convencional completo permite remociones superiores al 90 % en turbiedad y 95 % en color, lo que demuestra su eficacia para el manejo de aguas superficiales con características variables. Esta capacidad de remoción hace que sea la opción preferida para fuentes de agua cruda que experimentan fluctuaciones significativas en su calidad debido a factores climáticos o antropogénicos.

Por otra parte, OMS (2022) recomienda la implementación del tratamiento convencional simplificado cuando la calidad del agua cruda presenta menor variabilidad. Este sistema es apropiado para aguas con turbiedad media menor a 100 UNT, variaciones de turbiedad moderadas, color verdadero menor a 60 UC y baja contaminación microbiológica. Esta configuración puede omitir la etapa de presedimentación o combinar procesos. Así se reducen costos de inversión y operación mientras se mantiene la eficiencia del tratamiento.

Por otro lado, WRF (2024f) señala que la filtración directa es aplicable para fuentes de agua con baja turbiedad y color. Este sistema es adecuado cuando la turbiedad se mantiene menor a 20 UNT durante el 90 % del tiempo, el color verdadero es menor a 40 UC y existe baja presencia de algas. La principal ventaja es la reducción significativa en el consumo de

coagulantes y la generación de lodos, aunque requiere un control más riguroso del proceso de coagulación.

En definitiva, IWA (2024c) señala que la filtración en múltiples etapas combina la filtración gruesa con la filtración lenta en arena. Es idóneo para comunidades pequeñas y medianas. También para zonas donde se prioriza la simplicidad operativa o hay limitaciones de recursos técnicos. Funciona bien con aguas de turbiedad menor a 50 UNT el 90 % del tiempo. Este sistema destaca por sus bajos costos operativos y mínimo requerimiento de productos químicos, lo que representa una alternativa sostenible para comunidades con recursos limitados.

Criterios de selección según calidad de agua cruda. EPA (2023) postula que la selección de la tecnología apropiada debe fundamentarse en un análisis integral que considere múltiples factores. Las características del agua cruda son el primer criterio fundamental. Estas incluyen los rangos de turbiedad y su variabilidad temporal, el color verdadero, el contenido de materia orgánica y la presencia de microorganismos patógenos. Otros parámetros fisicoquímicos críticos también determinan la complejidad del tratamiento requerido. Estos parámetros definen tanto los procesos unitarios necesarios como la configuración tecnológica más apropiada para garantizar el cumplimiento de los estándares de calidad.

Respecto a los aspectos técnico-operativos, Davis (2023) indica que estos son el segundo criterio esencial. El caudal de diseño define la escala de la instalación. La disponibilidad de personal calificado y el acceso a insumos químicos condicionan la complejidad admisible del sistema. Los requerimientos de mantenimiento deben alinearse con las capacidades técnicas y logísticas locales para garantizar la continuidad operativa. Los sistemas más sofisticados requieren personal especializado y mayor disponibilidad de repuestos.

Por su parte, AWWA (2023b) considera que los aspectos económicos completan el marco de evaluación. Los costos de inversión inicial deben balancearse con los costos operativos proyectados, considerando la disponibilidad de recursos financieros tanto para la implementación como para la operación sostenida. La sostenibilidad financiera a largo plazo

es esencial para garantizar la continuidad y calidad del servicio. Esto requiere un análisis del ciclo de vida que incluya costos de capital, operación, mantenimiento y reposición de equipos.

2.2.5 Pruebas de jarras y control de procesos

Fundamentos de la prueba de jarras. Crittenden et al. (2021) indican que la prueba de jarras es un método estandarizado que simula los procesos de coagulación, floculación y sedimentación a escala de laboratorio. Esta prueba permite determinar las dosis óptimas de coagulantes y las condiciones de operación más adecuadas para el tratamiento del agua. Es una herramienta fundamental para la optimización de procesos en plantas que manejan aguas crudas con características variables.

Sobre este punto, AWWA (2023b) destaca que la principal ventaja de la prueba de jarras es la posibilidad de evaluar múltiples condiciones de tratamiento simultáneamente. Esto facilita la optimización del proceso bajo diferentes escenarios de calidad de agua. Esta capacidad de evaluación paralela permite comparar directamente diferentes dosis de coagulante, condiciones de pH y parámetros de mezclado, optimizando así el tiempo requerido para la determinación de condiciones operativas óptimas.

Respecto a la metodología, WRF (2024f) plantea que la prueba de jarras requiere equipamiento específico. Este incluye un dispositivo de agitación múltiple con control de velocidad variable, recipientes estandarizados transparentes de 1 a 2 litros e iluminación posterior para observación de flóculos. También se necesitan equipos complementarios para medir pH, turbiedad y temperatura. La precisión en el control de la velocidad de agitación es esencial para reproducir las condiciones de mezclado de la planta real. Esto garantiza la transferibilidad de los resultados obtenidos.

Por su parte, EPA (2023) fija las condiciones estándar de prueba que han sido desarrolladas a través de extensivas investigaciones y validaciones operativas. Estas incluyen una mezcla rápida inicial de 100-200 rpm durante 1-2 minutos para dispersar el coagulante. Luego sigue una fase de floculación a 20-40 rpm durante 15-30 minutos para promover la formación

de flóculos. Finalmente se realiza un periodo de sedimentación de 15-30 minutos. La temperatura debe mantenerse dentro del rango de $\pm 2^{\circ}\text{C}$ respecto a la temperatura del agua cruda en la planta para asegurar la representatividad de las condiciones de prueba.

En cuanto a las variables críticas, IWA (2024b) identifica que su control efectivo es fundamental para la validez y reproducibilidad de los resultados. La temperatura del agua influye significativamente en la cinética de la coagulación, por lo que debe monitorearse continuamente durante la prueba. El pH afecta directamente la efectividad del coagulante. Requiere control tanto al inicio como al final de la prueba. La velocidad y tiempo de agitación determinan la formación y características de los flóculos. Es fundamental que reproduzcan las condiciones de gradiente de velocidad presentes en la planta. La secuencia y momento de dosificación de los reactivos químicos impactan significativamente en la eficiencia del proceso, por lo que deben seguir un protocolo estrictamente establecido.

Por último, la OMS (2022) enfatiza que la interpretación adecuada de los resultados requiere la consideración integral de múltiples criterios de evaluación. Estos incluyen la turbiedad residual, el tamaño y características de los flóculos formados, la velocidad de sedimentación y el comportamiento del manto de lodos. Estos parámetros deben evaluarse en conjunto para determinar las condiciones óptimas de tratamiento. La optimización debe considerar tanto la eficiencia de remoción como la viabilidad operativa de las condiciones identificadas.

Metodología estandarizada. APHA (2024) señala que las pruebas de jarras requieren una metodología sistemática y estandarizada para garantizar resultados reproducibles y confiables. Los protocolos específicos definidos en el Standard Methods for the Examination of Water and Wastewater deben seguirse rigurosamente durante cada etapa del proceso. Esto asegura la consistencia metodológica y la validez científica de los resultados obtenidos en diferentes laboratorios.

Respecto a la preparación de muestras, Hudson y Wagner (2023a) establece que este es el primer paso crítico del proceso. Las muestras deben ser representativas de la calidad del agua cruda que ingresa a la planta. Deben mantenerse a la temperatura del agua cruda de

la planta. El análisis debe realizarse dentro de las 24 horas posteriores a su recolección para preservar sus características físicoquímicas. Antes de iniciar el ensayo se debe caracterizar completamente la muestra. Se mide turbiedad, pH, alcalinidad, conductividad y temperatura. Esto establece la línea base para la interpretación posterior de los resultados.

Sobre el procedimiento operativo, Washington State Department of Health (WSDH, 2024) recomienda comenzar con la preparación de las soluciones de coagulante a concentraciones específicas. Se deben preparar soluciones stock de coagulantes al 1 % o 10 % según el rango de dosis a evaluar, utilizando agua destilada. Las diluciones necesarias se realizan inmediatamente antes de la prueba para evitar la degradación de los reactivos. Las jarras se llenan con volúmenes idénticos de muestra (típicamente 1 o 2 litros) y se colocan en el equipo de prueba bajo iluminación adecuada que permita la observación detallada de la formación de flóculos.

En cuanto a la evaluación, AWWA (2023b) indica que los parámetros siguen un protocolo secuencial estrictamente definido. Durante la mezcla rápida (1-2 minutos a 100-200 rpm) se adiciona el coagulante simultáneamente en todas las jarras. Esto garantiza condiciones iniciales uniformes. La floculación se realiza en dos etapas: primero a 40 rpm por 10 minutos, seguido de 20 rpm por 15 minutos. Esto simula las condiciones de gradiente de velocidad presentes en la planta real. Después de un periodo de sedimentación de 15-30 minutos se toman muestras del sobrenadante para análisis, evitando perturbar el sedimento formado.

En definitiva, Crittenden et al. (2021) sostienen que el registro y procesamiento de datos debe ser sistemático y detallado. Se documentan las características iniciales del agua, las dosis aplicadas y las observaciones durante la formación de flóculos (tamaño, tiempo de aparición, velocidad de sedimentación). También se registran los parámetros finales medidos. Esta información se procesa mediante formatos estandarizados que permiten la comparación entre ensayos. Así se facilita la identificación de condiciones óptimas de tratamiento bajo diferentes escenarios operativos.

Interpretación y aplicación. Hudson y Wagner (2023b) plantean que el análisis de los resultados de pruebas de jarras debe realizarse considerando múltiples criterios de evaluación complementarios. La evaluación cuantitativa comprende parámetros primarios como turbiedad residual, pH final y consumo de coagulante. Se recomienda utilizar gráficos de control para identificar tendencias y patrones en estos parámetros. La evaluación cualitativa mediante la caracterización visual de flóculos proporciona información valiosa sobre tamaño, densidad y velocidad de sedimentación. Estos aspectos reflejan directamente la eficiencia del proceso de coagulación-floculación.

Sobre este punto, AWWA (2023a) sugiere utilizar índices estandarizados de caracterización de flóculos para objetivizar las observaciones cualitativas y mejorar la reproducibilidad entre diferentes operadores. La elaboración de curvas de dosificación debe seguir una metodología sistemática. Esta incluye la generación de datos base mediante pruebas de jarras bajo diferentes condiciones de turbiedad, pH y temperatura. Posteriormente se aplican técnicas de regresión y ajuste de curvas para obtener funciones matemáticas que relacionen dosis versus turbiedad. La validación estadística de las curvas generadas se realiza mediante análisis de residuales y pruebas de bondad de ajuste. Se incorporan bandas de confianza para considerar la variabilidad inherente al proceso. La transferencia a escala real requiere considerar factores adicionales críticos. Los gradientes de velocidad deben verificarse para garantizar que las condiciones de mezcla rápida y floculación en planta sean equivalentes a las de laboratorio. La aplicación de números adimensionales como Reynolds y Camp es fundamental para garantizar similitud hidráulica entre escalas. Se recomienda implementar cambios graduales y monitorear continuamente la respuesta del sistema para evitar perturbaciones operativas significativas.

En cuanto al análisis estadístico, Montgomery y Runger (2021) indican que este complementa las evaluaciones cuantitativas y cualitativas mediante la aplicación de herramientas como ANOVA y regresión múltiple para establecer correlaciones entre variables operativas. Se recomienda el uso de software especializado como Minitab o R para procesar los datos

obtenidos. Esto facilita la generación de modelos predictivos que optimizan el proceso de dosificación bajo diferentes condiciones operativas.

Por su parte, Crittenden et al. (2021) señalan que las principales limitaciones del método incluyen la variabilidad en la calidad del agua cruda. Las fluctuaciones rápidas en turbiedad y pH pueden complicar la optimización del proceso. Las limitaciones de equipamiento también son relevantes. La precisión de los sistemas de dosificación y medición puede afectar la implementación efectiva de dosis óptimas determinadas en laboratorio. Los factores operativos como restricciones de personal, presupuesto y tiempo también impactan la implementación de mejoras identificadas. Esto requiere un enfoque integral que considere tanto los aspectos técnicos como los recursos disponibles.

Asimismo, AWWA (2023b) determina que las consideraciones prácticas esenciales para la implementación exitosa incluyen el mantenimiento preventivo de equipos de dosificación y medición, la capacitación continua del personal operativo, la documentación detallada de procedimientos y resultados, y la evaluación periódica del desempeño del sistema. Estos aspectos son fundamentales para garantizar la transferencia efectiva de los resultados de laboratorio a la operación de planta. Así se asegura que las mejoras identificadas mediante pruebas de jarras se traduzcan en beneficios operativos sostenibles.

2.2.6 Modelamiento matemático en tratamiento de agua

Fundamentos de modelamiento. Gernaey et al. (2023) afirman que el modelamiento matemático en tratamiento de agua ha evolucionado significativamente, incorporando diversas aproximaciones según el nivel de complejidad y objetivos específicos. Los modelos determinísticos se basan en principios físico-químicos fundamentales. Proporcionan una descripción detallada de los procesos pero requieren una parametrización extensiva que puede limitar su aplicabilidad práctica. Por otro lado, los modelos estocásticos incorporan la naturaleza aleatoria de ciertos fenómenos. Son particularmente útiles cuando existe incertidumbre en las variables de entrada. Esto permite cuantificar la propagación de la incertidumbre a través del sistema de tratamiento.

En relación con ello, Crittenden et al. (2021) indican que los modelos empíricos, desarrollados a partir de datos experimentales, ofrecen soluciones prácticas para aplicaciones específicas en plantas de tratamiento de agua. Los modelos mecanísticos intentan representar los procesos fundamentales subyacentes mediante ecuaciones que describen los fenómenos físicos, químicos y biológicos involucrados en el tratamiento. Esto proporciona una base científica sólida para la comprensión y optimización de los procesos unitarios.

Respecto a la evolución actual, Li y Zhang (2024) señalan que la tendencia se dirige hacia modelos híbridos que combinan diferentes enfoques para aprovechar las ventajas específicas de cada metodología. Estos modelos integran la consistencia teórica de los mecanísticos con la flexibilidad de los empíricos y la capacidad de manejo de incertidumbre de los enfoques estocásticos. Esto resulta en herramientas más versátiles y precisas para la optimización de procesos de tratamiento bajo condiciones operativas variables.

Técnicas de modelamiento. Thompson et al. (2024) sostienen que las técnicas modernas de modelamiento integran métodos estadísticos avanzados con algoritmos de aprendizaje automático para mejorar la precisión predictiva en sistemas de tratamiento de agua. El análisis multivariable y las técnicas de regresión avanzada permiten identificar relaciones complejas entre variables operativas que no serían evidentes mediante enfoques univariados tradicionales. Las redes neuronales artificiales y los algoritmos de aprendizaje profundo han demostrado particular eficacia en la predicción de parámetros de calidad del agua y la optimización de procesos de tratamiento. Estas técnicas proporcionan capacidades de reconocimiento de patrones no lineales que superan significativamente las limitaciones de los métodos estadísticos convencionales.

Por su parte, Chen y Wang (2023) indican que la lógica difusa proporciona un marco robusto para manejar la incertidumbre y la imprecisión inherentes a los procesos de tratamiento. Los métodos de optimización metaheurística facilitan la calibración de modelos complejos con múltiples parámetros. Esto permite la búsqueda eficiente de soluciones óptimas en espacios de parámetros de alta dimensionalidad. Esta combinación de técnicas de lógica difusa con algoritmos de optimización metaheurística es particularmente valiosa para sistemas de tratamiento que operan bajo condiciones variables y con múltiples objetivos operativos que deben optimizarse simultáneamente.

Validación y calibración. Rossman et al. (2024) plantean que la validación de modelos en tratamiento de agua requiere un enfoque sistemático que considere tanto la precisión predictiva como la robustez operacional. La calibración debe abordar la variabilidad temporal en las características del agua y las condiciones operativas. Para esto se utilizan conjuntos de datos representativos que abarquen diferentes escenarios operativos. Esto garantiza que el modelo pueda responder adecuadamente a las fluctuaciones naturales en la calidad del agua cruda y las condiciones operativas de la planta. Este enfoque sistemático debe incluir la evaluación del desempeño del modelo bajo condiciones normales, extremas y transitorias para asegurar su aplicabilidad práctica en entornos operativos reales.

En complemento a estos criterios de validación, en WRF (2024c) se indica que los métodos de validación cruzada y las técnicas de remuestreo proporcionan estimaciones confiables del rendimiento del modelo bajo diferentes condiciones operativas. La validación debe incluir análisis de sensibilidad para identificar parámetros críticos y evaluar la estabilidad del modelo frente a perturbaciones en las variables de entrada. Esto permite cuantificar la robustez del modelo e identificar las condiciones bajo las cuales su desempeño puede verse comprometido. Esta evaluación integral de la sensibilidad paramétrica es fundamental para establecer los límites operativos del modelo y desarrollar estrategias de contingencia para situaciones donde la precisión predictiva pueda verse afectada.

Criterios de evaluación. Li y Zhang (2024) establecen que la evaluación del desempeño de los modelos matemáticos se realiza mediante múltiples criterios que abordan diferentes aspectos de su funcionamiento. Las métricas estadísticas tradicionales como el error cuadrático medio (RECM) y el coeficiente de determinación (R^2) se complementan con criterios específicos del proceso. Estos incluyen la capacidad de predicción de eventos críticos y la robustez frente a variaciones operativas. Esta aproximación multidimensional permite una caracterización integral del desempeño del modelo. Se considera tanto su precisión estadística como su utilidad práctica para la toma de decisiones operativas en sistemas de tratamiento de agua.

Respecto a esto, Gernaey et al. (2023) indican que la evaluación debe considerar también aspectos prácticos como la facilidad de implementación, los requerimientos computacionales y la interpretabilidad de los resultados. La sostenibilidad del modelo a largo plazo y su capacidad de adaptación a cambios en las condiciones operativas son criterios críticos en la selección de modelos para aplicaciones prácticas. Los sistemas de tratamiento de agua operan en entornos dinámicos donde las condiciones pueden cambiar significativamente a lo largo del tiempo. Esta evaluación integral debe balancear la sofisticación técnica del modelo con su viabilidad operativa y su capacidad de mantenimiento por parte del personal de planta.

2.2.7 *Interpolación mediante splines cúbicos*

Teoría de la interpolación. Burden et al. (2023) postulan que la interpolación mediante splines cúbicos es una técnica matemática avanzada que permite generar curvas suaves a través de un conjunto de puntos de control. A diferencia de los métodos de interpolación polinomial tradicionales, los splines cúbicos evitan el fenómeno de Runge. También proporcionan aproximaciones más estables. Esto es especialmente útil en aplicaciones de ingeniería que requieren continuidad en las derivadas. Esta característica de estabilidad numérica los convierte en herramientas particularmente valiosas para el modelado de procesos físicos. La suavidad de la solución es crítica para la representación fiel del comportamiento del sistema y la preservación de las propiedades físicas fundamentales del proceso bajo estudio.

Fundamentos de splines. Quarteroni y Saleri (2024) determinan que los splines son funciones polinomiales por tramos que mantienen continuidad hasta la segunda derivada en los puntos de unión. Esta característica es particularmente valiosa en el modelamiento de procesos físicos donde la suavidad de la solución es crítica para la representación adecuada del comportamiento del sistema. La construcción de splines se fundamenta en la minimización de la energía de flexión. Esto resulta en curvas naturalmente suaves que preservan las características físicas del sistema modelado. Este enfoque de minimización energética garantiza que la función interpolante no solo pase por todos los puntos de control especificados. También mantiene propiedades de derivabilidad esenciales para aplicaciones en ingeniería. Las discontinuidades en las derivadas pueden indicar comportamientos físicamente irreales o numéricamente inestables.

Splines cúbicos naturales. Atkinson y Han (2023) consideran que los splines cúbicos naturales imponen condiciones específicas en los extremos del intervalo de interpolación. La segunda derivada debe ser cero en los puntos extremos. Esta condición produce curvas que minimizan la curvatura total. Son especialmente adecuadas para modelar fenómenos físicos donde se requiere comportamiento natural en los límites del dominio de interpolación. La for-

mulación matemática implica la solución de un sistema tridiagonal de ecuaciones que garantiza la continuidad hasta la segunda derivada en todos los puntos de unión. Esto proporciona propiedades de suavidad óptimas. Este sistema tridiagonal presenta ventajas computacionales significativas. Puede resolverse eficientemente mediante algoritmos especializados que aprovechan la estructura matricial específica. Así se obtienen métodos numéricos estables y computacionalmente eficientes para la construcción de funciones interpolantes suaves.

Algoritmos de ajuste. Press et al. (2024) sostienen que la implementación computacional de splines cúbicos requiere algoritmos eficientes para la solución del sistema de ecuaciones resultante. El método de Thomas está específicamente adaptado para matrices tridiagonales. Proporciona una solución eficiente con complejidad computacional lineal. Esto aprovecha la estructura particular de la matriz de coeficientes para reducir significativamente el costo computacional comparado con métodos de solución de sistemas generales. Los algoritmos modernos incorporan técnicas de optimización numérica para mejorar la estabilidad y precisión del ajuste. Son especialmente útiles en presencia de datos con ruido o irregularmente espaciados. Implementan estrategias de pivoteo parcial y control de errores de redondeo. Esto garantiza la robustez numérica del proceso de interpolación bajo condiciones adversas de datos de entrada.

Evaluación de error. Berrut y Trefethen (2023) sostienen que la evaluación del error en la interpolación mediante splines cúbicos debe considerar múltiples aspectos. Estos incluyen el error de aproximación local y global, la estabilidad numérica y la preservación de propiedades físicas relevantes. Las métricas de error incluyen el error máximo absoluto, el error cuadrático medio y el error en las derivadas. Esto proporciona una caracterización integral de la calidad de la aproximación desde diferentes perspectivas matemáticas y físicas. La validación cruzada y el análisis de residuos proporcionan información adicional sobre la calidad del ajuste y la robustez del modelo. Esto permite identificar regiones problemáticas en el dominio de interpolación. También se puede evaluar la capacidad de generalización del

modelo spline bajo diferentes condiciones de datos. Esta evaluación multidimensional del error es fundamental para garantizar que la función interpolante no solo reproduzca fielmente los datos de entrada, sino que también mantenga comportamiento físicamente consistente en todo el dominio de aplicación.

2.2.8 *Sistemas de lógica difusa*

Teoría de conjuntos difusos. Ross (2023) afirma que la teoría de conjuntos difusos proporciona un marco matemático para manejar la incertidumbre y la imprecisión inherentes a muchos procesos de ingeniería. A diferencia de la lógica booleana tradicional, la lógica difusa permite grados variables de pertenencia de un elemento a un conjunto. Esto facilita la representación de conceptos imprecisos pero significativos en el control de procesos. Esta aproximación es particularmente valiosa en el tratamiento de agua. Muchas variables operativas no pueden categorizarse de manera binaria. Requieren representación mediante funciones de pertenencia continuas que reflejen la naturaleza gradual de las transiciones entre diferentes estados operativos. La capacidad de los conjuntos difusos para modelar la vaguedad lingüística y la incertidumbre paramétrica los convierte en herramientas fundamentales. Permiten formalizar el conocimiento experto e implementar sistemas de control inteligentes en procesos industriales complejos.

Funciones de pertenencia. Wang y Liu (2024) plantean que las funciones de pertenencia son el elemento fundamental para cuantificar el grado en que una variable pertenece a un conjunto difuso específico. La selección y diseño de estas funciones impacta significativamente en el desempeño del sistema de control difuso. Esto determina tanto la precisión como la interpretabilidad del modelo resultante. Las formas más comunes incluyen funciones triangulares, trapezoidales y gaussianas. Cada una tiene características específicas que las hacen adecuadas para diferentes aplicaciones según los requerimientos de suavidad, simplicidad computacional e interpretabilidad lingüística. La sintonización de estas funciones debe considerar tanto el conocimiento experto como los datos históricos del proceso. Se busca ba-

lancear la fidelidad a la experiencia operativa con la representatividad estadística de los datos disponibles. Este proceso de sintonización requiere metodologías sistemáticas. Estas permiten optimizar simultáneamente la forma, posición y soporte de las funciones de pertenencia para maximizar la capacidad discriminativa del sistema difuso. Al mismo tiempo se mantiene la coherencia con el conocimiento del dominio de aplicación.

Sistemas de inferencia difusa. Zadeh y Aliev (2023) afirman que los sistemas de inferencia difusa transforman el conocimiento experto en un marco computacional mediante reglas del tipo SI-ENTONCES. El proceso de inferencia involucra la fuzzificación de entradas, la evaluación de reglas y la agregación de resultados. Esto proporciona un mecanismo sistemático para procesar información imprecisa y generar decisiones coherentes basadas en conocimiento lingüístico. Los sistemas tipo Mamdani y Sugeno representan las arquitecturas más utilizadas. Cada una tiene ventajas específicas en términos de interpretabilidad y eficiencia computacional. Los sistemas Mamdani ofrecen mayor transparencia en la interpretación de resultados. Facilitan la incorporación directa de conocimiento experto expresado en términos lingüísticos. Por su parte, los sistemas Sugeno proporcionan mayor eficiencia computacional y mejor adaptabilidad para aplicaciones que requieren optimización paramétrica. La selección entre estas arquitecturas debe considerar los requerimientos específicos de la aplicación. Se busca balancear la necesidad de interpretabilidad con los requerimientos de eficiencia computacional y precisión del sistema de control resultante.

Defuzzificación. Pedrycz et al. (2024) sostienen que la defuzzificación convierte los resultados difusos agregados en valores numéricos precisos que pueden aplicarse directamente en el control del proceso. Los métodos más comunes incluyen el centro de gravedad, el promedio ponderado y el máximo de la membresía. Cada uno tiene características específicas que determinan su idoneidad para diferentes tipos de aplicaciones y requerimientos operativos. La selección del método de defuzzificación debe considerar tanto la precisión requerida como la eficiencia computacional necesaria para aplicaciones en tiempo real. Existe un com-

promiso fundamental entre la calidad de la aproximación numérica y el costo computacional asociado. El método del centro de gravedad proporciona resultados suaves y continuos. Es particularmente apropiado para aplicaciones de control donde se requiere estabilidad en las salidas. Por otro lado, el método del máximo de la membresía ofrece respuestas más rápidas pero potencialmente menos suaves. El promedio ponderado representa un compromiso intermedio que balancea eficiencia computacional con calidad de la aproximación. Es especialmente útil en sistemas donde la interpretabilidad de los resultados es tan importante como su precisión numérica.

Optimización de reglas. Karray y De Silva (2023) consideran que la optimización de la base de reglas difusas es un aspecto crítico para mejorar el desempeño del sistema. Las técnicas modernas combinan el conocimiento experto con algoritmos de aprendizaje automático para refinar tanto las reglas como los parámetros de las funciones de pertenencia. Esto permite la evolución sistemática del sistema hacia configuraciones más eficientes y precisas. Los métodos adaptativos permiten el ajuste continuo del sistema en respuesta a cambios en las condiciones del proceso o nueva información disponible. Esto proporciona capacidades de aprendizaje incremental que mejoran el desempeño a lo largo del tiempo operativo. Esta optimización puede realizarse mediante técnicas de algoritmos genéticos, optimización por enjambre de partículas o métodos de gradiente. Estos ajustan simultáneamente la estructura de las reglas y los parámetros de las funciones de pertenencia. El proceso de optimización debe balancear la mejora en el desempeño predictivo con la preservación de la interpretabilidad del sistema. Esto asegura que las reglas optimizadas mantengan coherencia semántica. Así pueden ser validadas por expertos del dominio de aplicación.

2.2.9 *Fundamentos de sistemas híbridos*

Arquitecturas híbridas. Thompson et al. (2024) sostienen que los sistemas híbridos en tratamiento de agua integran múltiples técnicas de modelamiento para aprovechar las fortalezas complementarias de diferentes enfoques. Las arquitecturas más comunes combinan modelos basados en principios físicos con técnicas de inteligencia artificial. También integran métodos determinísticos con aproximaciones estocásticas. Esto permite abordar la complejidad inherente de los procesos de tratamiento de agua donde múltiples variables interactúan de manera no lineal. Esta convergencia metodológica es particularmente valiosa para sistemas de tratamiento que deben operar bajo condiciones variables. Estos sistemas manejan simultáneamente aspectos determinísticos bien conocidos junto con fenómenos estocásticos o no lineales que requieren enfoques de modelamiento más sofisticados.

En relación con ello, Chen y Wang (2023) indican que las estructuras híbridas modernas implementan capas jerárquicas de procesamiento. Cada componente se especializa en aspectos específicos del proceso. Los modelos físicos pueden manejar las relaciones fundamentales basadas en principios de ingeniería bien establecidos. Por su parte, las redes neuronales se encargan de patrones no lineales complejos que emergen de las interacciones entre múltiples variables del sistema. Esta especialización funcional permite optimizar cada componente del sistema híbrido según sus fortalezas específicas. Así se obtienen modelos más robustos y precisos que pueden adaptarse tanto a condiciones operativas normales como a situaciones excepcionales que requieren capacidades de extrapolación más allá del dominio de calibración original.

Técnicas de integración. WRF (2024b) establece que la integración efectiva de componentes heterogéneos requiere estrategias sofisticadas para garantizar la coherencia y optimizar el desempeño global del sistema híbrido. Los métodos de ponderación adaptativa permiten ajustar dinámicamente la contribución de cada componente según su confiabilidad en diferentes condiciones operativas. Esto proporciona un mecanismo de selección automática

que favorece el componente más apropiado según el contexto específico de operación. Los sistemas de votación ponderada y los métodos de consenso facilitan la resolución de predicciones conflictivas entre componentes. Estos implementan estrategias de fusión que combinan las fortalezas individuales de cada método mientras minimizan el impacto de sus limitaciones específicas. Estas técnicas de integración deben considerar tanto la compatibilidad conceptual entre diferentes metodologías como la sincronización temporal de sus respuestas. Así se asegura que el sistema híbrido resultante mantenga coherencia matemática y estabilidad operativa bajo todas las condiciones de funcionamiento previstas.

Métricas de desempeño. Li y Zhang (2024) indican que para evaluar el desempeño de los sistemas híbridos se necesitan métricas que reflejen tanto la exactitud general como el comportamiento de cada componente por separado. Las métricas tradicionales de error se complementan con indicadores de robustez, adaptabilidad y eficiencia computacional. Esto proporciona una caracterización multidimensional que refleje las capacidades específicas del sistema híbrido bajo diferentes condiciones operativas. La evaluación debe considerar también la interpretabilidad de los resultados y la capacidad del sistema para manejar incertidumbres. Los sistemas híbridos deben balancear múltiples objetivos. Estos incluyen precisión predictiva, estabilidad operativa, eficiencia computacional y transparencia en la toma de decisiones. Esta evaluación integral requiere el desarrollo de marcos de evaluación específicos que consideren las interacciones entre componentes del sistema híbrido y su contribución sinérgica al desempeño global. Así se asegura que las métricas seleccionadas reflejen fielmente tanto las fortalezas como las limitaciones del enfoque híbrido implementado.

Casos de aplicación. Rossman et al. (2024) indican que los modelos híbridos han mostrado un éxito notable en la optimización de procesos de coagulación-floculación. La complejidad de las interacciones físico-químicas implica que los métodos convencionales no pueden capturar adecuadamente la naturaleza multivariada y no lineal del proceso. Casos documentados en plantas de tratamiento a gran escala muestran mejoras significativas en la

precisión de dosificación y reducción de costos operativos mediante la implementación de sistemas híbridos. Esto evidencia reducciones en el consumo de coagulantes del 15-25 % y mejoras en la estabilidad de la calidad del agua tratada. Estas aplicaciones demuestran que los sistemas híbridos pueden adaptarse efectivamente a variaciones en la calidad del agua cruda, condiciones estacionales y cambios en los requerimientos operativos. Así se proporcionan soluciones robustas para plantas que enfrentan desafíos operativos complejos. La implementación exitosa de estos sistemas en entornos industriales reales confirma su viabilidad técnica y económica. Esto establece un precedente para su adopción más amplia en el sector de tratamiento de agua potable.

2.2.10 Modelos híbridos en dosificación

Estado del arte. De acuerdo con Thompson et al. (2024), los recientes progresos en modelos híbridos para la dosificación de coagulantes han revolucionado considerablemente el control de procesos en plantas de tratamiento. La integración de técnicas clásicas con métodos de inteligencia artificial ha permitido superar limitaciones tradicionales en la predicción de dosis óptimas. Esto proporciona herramientas más precisas y adaptables para el manejo de procesos complejos de coagulación-floculación. Las investigaciones actuales demuestran una clara tendencia hacia sistemas adaptativos que pueden responder eficientemente a variaciones en la calidad del agua cruda mientras mantienen la robustez operativa. Estos sistemas evidencian capacidades de aprendizaje continuo y ajuste automático. Así mejoran su desempeño a medida que se acumula experiencia operativa. Esta evolución tecnológica representa un cambio paradigmático en el control de procesos de tratamiento de agua. Los sistemas híbridos emergen como la solución más prometedora para abordar la complejidad inherente de los procesos de dosificación bajo condiciones operativas variables y cada vez más demandantes en términos de eficiencia y sostenibilidad.

Arquitecturas propuestas. Según Chen y Wang (2023), las arquitecturas híbridas contemporáneas para la dosificación integran diversas capas de procesamiento. Cada una se dedica a una parte específica del proceso. Los modelos basados en principios físico-químicos se complementan con redes neuronales y sistemas difusos para capturar relaciones no lineales complejas que emergen de las interacciones entre múltiples variables del sistema de tratamiento. La tendencia actual favorece arquitecturas modulares que permiten la actualización y mantenimiento independiente de componentes. Esto mejora la flexibilidad y escalabilidad del sistema para adaptarse a diferentes condiciones operativas y requerimientos específicos de cada planta de tratamiento. Esta modularidad facilita la incorporación de nuevas técnicas de modelamiento sin requerir modificaciones sustanciales en la estructura global del sistema. Así se permite evolución tecnológica gradual y se reducen los riesgos asociados con implementaciones de gran escala. Las arquitecturas modulares también simplifican los procesos de validación y certificación. Cada componente puede ser evaluado independientemente antes de su integración en el sistema híbrido completo.

Análisis comparativo. Según WRF (2024a), los análisis comparativos de diversas arquitecturas híbridas muestran patrones importantes en cuanto a rendimiento y aplicabilidad. Los sistemas que combinan lógica difusa con técnicas de interpolación avanzada muestran ventajas particulares en términos de interpretabilidad y robustez. Esto proporciona un balance óptimo entre precisión predictiva y transparencia operativa que facilita su aceptación por parte del personal técnico de las plantas de tratamiento. Las evaluaciones sistemáticas indican que los modelos híbridos superan consistentemente a los enfoques tradicionales en términos de precisión de dosificación y adaptabilidad a cambios en las condiciones operativas. Estos demuestran mejoras en la estabilidad del proceso y reducción de la variabilidad en la calidad del agua tratada. Los análisis comparativos también revelan que los sistemas híbridos mantienen su superioridad de desempeño bajo una amplia gama de condiciones operativas. Esto incluye eventos extremos y transiciones estacionales. Así se confirma su robustez y viabilidad para implementación a gran escala en sistemas de tratamiento de agua potable.

Tendencias y desafíos. Según Zhang et al. (2024), las tendencias emergentes en el modelado híbrido comprenden la integración de técnicas de aprendizaje profundo y métodos de optimización en tiempo real. Los desafíos actuales se centran en mejorar la interpretabilidad de los modelos complejos y desarrollar métodos más eficientes para la calibración y actualización de sistemas híbridos. La sofisticación creciente de estos sistemas puede comprometer su transparencia operativa. La integración de datos provenientes de sensores en línea y la adaptación a condiciones operativas cambiantes representan áreas activas de investigación. Estas buscan maximizar las capacidades de los sistemas híbridos mientras mantienen su practicidad para implementación industrial. Estos desarrollos apuntan hacia sistemas cada vez más autónomos que pueden aprender continuamente de la experiencia operativa. Ajustan automáticamente sus parámetros y estrategias de fusión para optimizar el desempeño bajo condiciones específicas de cada planta de tratamiento. La convergencia de tecnologías de Internet de las Cosas (IoT), computación en la nube y análisis de big data está creando nuevas oportunidades. Se están desarrollando sistemas híbridos más inteligentes y conectados. Estos pueden beneficiarse de la experiencia colectiva de múltiples instalaciones de tratamiento.

2.2.11 Integración de componentes

Sistemas de ponderación. Thompson et al. (2024) indican que en los modelos híbridos, los sistemas de ponderación son cruciales para combinar de manera eficaz diversas técnicas de modelado. La asignación de pesos dinámicos a cada componente se realiza mediante algoritmos adaptativos que evalúan continuamente la confiabilidad y precisión de las predicciones individuales. Esto proporciona un mecanismo de selección automática que favorece el componente más apropiado según las condiciones operativas específicas detectadas en tiempo real. Las estrategias modernas de ponderación incorporan métodos bayesianos y técnicas de aprendizaje automático para optimizar la combinación de resultados de diferentes módulos. Estos implementan enfoques probabilísticos que consideran tanto la incertidumbre asociada a cada predicción individual como la historia de desempeño de cada componente bajo condiciones similares. Los sistemas de ponderación adaptativos permiten que el modelo híbrido evolucione continuamente. Ajustan automáticamente la influencia relativa de cada componente basándose en su desempeño histórico y su confiabilidad estimada para las condiciones actuales. Así se obtiene un sistema que puede mantener alta precisión predictiva incluso cuando las condiciones operativas se desvían significativamente de aquellas presentes durante la calibración inicial.

Adaptabilidad dinámica. Chen y Wang (2023) afirman que la capacidad de adaptación dinámica permite a los sistemas híbridos reaccionar eficazmente frente a variaciones en las condiciones operacionales y en la calidad del agua sin tratar. Los mecanismos de actualización en tiempo real ajustan tanto los parámetros del modelo como las estrategias de integración. Se basan en el análisis continuo del desempeño del sistema y la retroalimentación del proceso. Esto proporciona capacidades de aprendizaje incremental que mejoran la precisión predictiva a medida que se acumula experiencia operativa. Esta adaptabilidad es fundamental para mantener la eficiencia del tratamiento bajo condiciones variables. Permite que el sistema híbrido responda automáticamente a fluctuaciones estacionales, cambios en

las características de la fuente de agua cruda o modificaciones en los objetivos operativos de la planta de tratamiento. Los algoritmos de adaptación dinámica implementan estrategias de aprendizaje en línea. Pueden detectar deriva en el comportamiento del sistema. También identifican nuevos patrones operativos. Y ajustan proactivamente los parámetros del modelo para mantener el desempeño óptimo sin requerir intervención manual. Así se garantiza la continuidad operativa y la estabilidad de la calidad del agua tratada incluso ante condiciones operativas no previstas durante la fase de calibración inicial.

Robustez y estabilidad. WRF (2024d) afirma que la solidez y estabilidad de los sistemas híbridos se basan en diseños que aseguran un funcionamiento fiable incluso cuando hay perturbaciones considerables o entradas ruidosas. Las arquitecturas modernas implementan mecanismos de detección y corrección de anomalías. También incluyen estrategias de control que previenen oscilaciones indeseadas en las predicciones. Esto proporciona capas múltiples de protección que aseguran la continuidad operativa bajo condiciones adversas. La validación rigurosa bajo diferentes escenarios operativos confirma la capacidad del sistema para mantener un desempeño estable. Se incluyen pruebas de estrés que evalúan el comportamiento del sistema ante condiciones extremas, fallos de sensores y variaciones súbitas en la calidad del agua cruda. Estos mecanismos de robustez incluyen filtros adaptativos para el manejo de ruido en las señales de entrada. También tienen algoritmos de detección de outliers que previenen decisiones basadas en datos anómalos. Y estrategias de degradación gradual que permiten al sistema continuar operando con funcionalidad reducida cuando algún componente presenta fallos. La estabilidad a largo plazo se garantiza mediante algoritmos de monitoreo continuo. Estos evalúan la deriva paramétrica e implementan correcciones automáticas. Así se mantiene el desempeño dentro de los límites especificados durante toda la vida útil del sistema.

Criterios de optimización. Zhang et al. (2024) afirman que para optimizar los sistemas híbridos es esencial equilibrar varios objetivos. Estos incluyen la exactitud de las predicciones, la eficiencia en el cálculo y la solidez en el funcionamiento. Las metodologías actuales emplean técnicas de optimización multiobjetivo que consideran simultáneamente aspectos técnicos, económicos y ambientales. La optimización de sistemas híbridos para dosificación de coagulantes requiere un enfoque holístico que trascienda la mera maximización de la precisión predictiva. La implementación de algoritmos de optimización en tiempo real permite ajustar continuamente los parámetros del sistema para mantener un desempeño óptimo. Estos se adaptan dinámicamente a cambios en las condiciones operativas, variaciones en los costos de los reactivos y modificaciones en los requerimientos ambientales o regulatorios. Los criterios de optimización multiobjetivo deben considerar trade-offs entre diferentes métricas de desempeño. Por ejemplo, la minimización del consumo de coagulantes versus la maximización de la eficiencia de remoción. O la reducción de la complejidad computacional versus el mantenimiento de la precisión predictiva. La formulación matemática de estos problemas típicamente emplea técnicas de Pareto-optimalidad. Estas identifican soluciones de compromiso que no pueden mejorarse en un objetivo sin degradar otros. Así se proporciona un conjunto de soluciones óptimas entre las cuales los operadores pueden seleccionar según las prioridades específicas de cada instalación de tratamiento.

2.2.12 Metodologías de validación

Diseño experimental. Montgomery y Runger (2021) indican que el diseño experimental para validar modelos híbridos en el tratamiento de agua necesita un enfoque metódico y riguroso. La planificación experimental debe considerar la variabilidad natural de los parámetros de calidad del agua y las condiciones operativas de la planta. Esto asegura que el diseño capture adecuadamente la complejidad inherente de los sistemas de tratamiento bajo diferentes escenarios operativos. Los diseños factoriales y de superficie de respuesta permiten evaluar interacciones complejas entre variables. Esto proporciona información valiosa sobre los efectos sinérgicos y antagónicos que pueden influir significativamente en el desempeño del modelo híbrido. Los métodos de muestreo estratificado aseguran una representación adecuada de diferentes condiciones operativas. Garantizan que el conjunto de datos experimental incluya muestras representativas de todas las regiones operativas relevantes. Esto incluye condiciones normales, extremas y transitorias. El diseño experimental debe incorporar también consideraciones sobre la aleatorización de los ensayos para minimizar el efecto de variables de confusión. Se requiere bloqueo apropiado para controlar fuentes conocidas de variabilidad. También se necesita replicación suficiente. Esta permite estimaciones confiables de la variabilidad experimental y la detección de efectos estadísticamente significativos en el desempeño del modelo híbrido.

Métricas de evaluación. Anderson y Whitcomb (2024) sostienen que las métricas de evaluación deben abarcar múltiples facetas del rendimiento del modelo con el fin de ofrecer una caracterización completa de su habilidad para predecir. El error cuadrático medio normalizado (RECMN) y el coeficiente de determinación ajustado (R^2_{adj}) proporcionan medidas básicas de precisión que cuantifican la proximidad entre valores predichos y observados. Esto considera el grado de ajuste del modelo y penaliza la complejidad excesiva. El coeficiente de Nash-Sutcliffe (ENS) evalúa la capacidad predictiva del modelo en relación con la media observada. Proporciona una medida de eficiencia que indica qué tan bien el modelo supera una

predicción trivial basada únicamente en el promedio histórico de los datos. El error porcentual absoluto medio (EPAM) ofrece una medida intuitiva del error relativo particularmente útil para la evaluación práctica. Expresa el error en términos porcentuales que facilitan la interpretación operativa y la comunicación de los resultados a personal técnico no especializado en estadística. Esta combinación de métricas complementarias permite una evaluación robusta que considera tanto la precisión absoluta como la precisión relativa del modelo. Así se proporciona información suficiente para la toma de decisiones informadas sobre la implementación práctica del sistema híbrido en entornos operativos reales.

Análisis estadístico. Fisher y Box (2024) sostienen que un análisis estadístico exhaustivo es esencial para confirmar la validez de los modelos. Las pruebas de normalidad, homocedasticidad y autocorrelación permiten verificar los supuestos subyacentes del modelo. Esto asegura que las técnicas estadísticas empleadas sean apropiadas para la naturaleza de los datos. También garantiza que las inferencias realizadas sean válidas desde una perspectiva estadística. El análisis de varianza (ANOVA) y las pruebas de comparación múltiple evalúan la significancia de diferentes factores y sus interacciones. Esto proporciona evidencia estadística sobre la importancia relativa de cada variable. Permite identificar efectos principales y de interacción que pueden influir en el desempeño del modelo híbrido. Los métodos de remuestreo como bootstrap proporcionan estimaciones robustas de la incertidumbre en las predicciones del modelo. Esto permite cuantificar la variabilidad esperada en las estimaciones sin requerir supuestos distribucionales específicos que pueden no cumplirse en datos reales de sistemas de tratamiento de agua. Estos métodos estadísticos avanzados facilitan la caracterización completa del comportamiento del modelo. Se incluye la estimación de intervalos de confianza para las predicciones, la detección de valores atípicos que pueden afectar el desempeño y la evaluación de la estabilidad de los parámetros del modelo bajo diferentes condiciones de muestreo. Así se proporciona la base estadística necesaria para validar la confiabilidad del sistema híbrido en aplicaciones prácticas.

Pruebas de robustez. Kumar y Patel (2024) afirman que las evaluaciones de robustez analizan cómo se comporta el modelo cuando enfrenta condiciones adversas o poco ideales. El análisis de sensibilidad global identifica los parámetros más influyentes y cuantifica su impacto en las predicciones del modelo. Esto proporciona información crítica sobre qué variables ejercen mayor control sobre el desempeño del sistema híbrido. Permite priorizar los esfuerzos de calibración y monitoreo en aquellos parámetros que demuestran mayor influencia. Las pruebas de estrés incluyen perturbaciones deliberadas en las variables de entrada y condiciones operativas extremas. Verifican la estabilidad y confiabilidad del modelo en situaciones críticas. Simulan escenarios como fallos de sensores, variaciones súbitas en la calidad del agua cruda, condiciones meteorológicas extremas y eventos operativos excepcionales. Estos pueden comprometer el desempeño del sistema. Estas evaluaciones de robustez incluyen varios componentes. El análisis de degradación gradual del desempeño bajo condiciones de estrés creciente es uno. También está la identificación de puntos de falla críticos donde el modelo puede colapsar. Y la evaluación de mecanismos de recuperación que permiten al sistema retornar a operación normal después de eventos disruptivos. La caracterización integral de la robustez del modelo es fundamental para establecer límites operativos seguros. Permite desarrollar protocolos de contingencia. Garantiza que el sistema híbrido pueda mantener un nivel mínimo de funcionalidad incluso bajo las condiciones más adversas que puedan presentarse en la práctica operativa.

2.2.13 Técnicas de optimización

Algoritmos de optimización. Nocedal y Wright (2023) afirman que los algoritmos actuales de optimización en el ámbito del tratamiento de agua incluyen tanto técnicas meta-heurísticas como métodos basados en gradientes. Los algoritmos genéticos y la optimización por enjambre de partículas han demostrado particular eficacia en la búsqueda de soluciones óptimas en espacios complejos de parámetros. Esto proporciona capacidades de exploración global que pueden identificar soluciones de alta calidad incluso en presencia de múltiples

óptimos locales y superficies de respuesta altamente no lineales. La selección del algoritmo apropiado depende de características específicas del problema. Estas incluyen la dimensionalidad del espacio de búsqueda y la naturaleza de las restricciones operativas. Se requiere un análisis cuidadoso de varios factores. Entre ellos están la complejidad computacional, la capacidad de manejo de restricciones, la robustez frente a ruido en las evaluaciones de la función objetivo y la escalabilidad para problemas de alta dimensionalidad. Los métodos basados en gradientes resultan particularmente eficientes para problemas con funciones objetivo diferenciables y restricciones lineales. Por su parte, las técnicas metaheurísticas ofrecen mayor flexibilidad. Son útiles para problemas con discontinuidades, restricciones no lineales y objetivos múltiples. Estas características son frecuentes en los sistemas híbridos de dosificación en tratamiento de agua.

Ajuste de parámetros. Boyd y Vandenberghe (2024) indican que el ajuste de parámetros es una etapa vital en la optimización de modelos híbridos. Las técnicas de regularización como Lasso y Ridge previenen el sobreajuste mientras mantienen la capacidad predictiva del modelo. Estas implementan penalizaciones que favorecen soluciones más simples y generalizables cuando los datos de entrenamiento son limitados o ruidosos. Los métodos de búsqueda bayesiana y la optimización basada en gradientes estocásticos permiten un ajuste eficiente de parámetros en sistemas complejos. Esto proporciona frameworks probabilísticos que pueden incorporar conocimiento previo sobre los parámetros. También actualizan sistemáticamente las estimaciones a medida que se dispone de nueva información experimental. La implementación de estrategias de ajuste adaptativo facilita la respuesta del modelo a cambios en las condiciones operativas. Esto permite que el sistema híbrido evolucione continuamente sus parámetros en respuesta a deriva en las características del proceso, cambios estacionales en la calidad del agua cruda o modificaciones en los objetivos operativos de la planta de tratamiento. Estas estrategias de ajuste deben balancear la velocidad de adaptación con la estabilidad del sistema. Evitan oscilaciones indeseadas en los parámetros mientras mantienen la capacidad de respuesta necesaria para seguir cambios genuinos en las condiciones operativas del

sistema de tratamiento.

Análisis de sensibilidad. Según Saltelli y Tarantola (2024), el análisis de sensibilidad global determina los parámetros con mayor influencia y mide su efecto en el rendimiento del modelo. Los métodos de Sobol y FAST (Fast Amplitude Sensitivity Test) proporcionan evaluaciones rigurosas de la sensibilidad paramétrica. Esto permite la descomposición de la varianza total de la salida del modelo en contribuciones atribuibles a parámetros individuales y sus interacciones. La comprensión de estas relaciones es fundamental para la optimización eficiente y la identificación de prioridades en el ajuste del modelo. Esto facilita la focalización de esfuerzos de calibración en aquellos parámetros que ejercen mayor influencia sobre el desempeño del sistema híbrido. El análisis de sensibilidad temporal revela cómo la importancia relativa de diferentes parámetros puede variar durante la operación del sistema. Esto proporciona información valiosa sobre la dinámica de la sensibilidad paramétrica bajo diferentes condiciones operativas. Permite el desarrollo de estrategias de monitoreo adaptativo que ajusten la frecuencia de medición según la importancia cambiante de cada variable. Esta caracterización temporal de la sensibilidad es particularmente relevante en sistemas de tratamiento de agua donde las condiciones operativas pueden experimentar variaciones estacionales significativas. Esto requiere ajustes dinámicos en las estrategias de control y optimización para mantener el desempeño óptimo del modelo híbrido a lo largo de todo el ciclo operativo anual.

Validación cruzada. Hastie y James (2024) afirman que la validación cruzada es una herramienta crucial para evaluar la capacidad de generalización de los modelos y evitar el sobreajuste. Las técnicas de k-fold y validación cruzada temporal son particularmente relevantes en aplicaciones de tratamiento de agua. La estructura temporal de los datos es significativa. Debe preservarse durante el proceso de validación para obtener estimaciones realistas del desempeño del modelo. Los métodos de validación cruzada adaptativa permiten una evaluación más robusta en presencia de datos no estacionarios o con dependencias temporales complejas. Estos implementan estrategias de partición que consideran tanto la autocorrelación temporal

como las tendencias a largo plazo en los datos operativos. La validación cruzada temporal es especialmente importante en sistemas de tratamiento de agua. Las características del agua cruda pueden exhibir patrones estacionales, tendencias de deterioro de la fuente o cambios graduales en las condiciones de la cuenca. Estos afectan la calidad del agua. Los métodos de validación deben considerar también la presencia de eventos extremos poco frecuentes. Estos pueden no estar adecuadamente representados en todas las particiones de datos. Se requieren estrategias especializadas de muestreo estratificado. Estas aseguran la inclusión proporcional de diferentes tipos de condiciones operativas en cada fold de validación. Así se obtienen estimaciones no sesgadas de la capacidad de generalización del modelo híbrido.

2.2.14 Implementación práctica

Arquitectura del sistema. Según Ibrahim y Martínez (2024), para llevar a cabo de manera efectiva la integración de modelos híbridos en instalaciones de tratamiento es esencial contar con una arquitectura de sistema que sea tanto robusta como escalable. La estructura moderna incorpora múltiples capas: adquisición de datos, procesamiento, análisis y control. Estas se integran a través de protocolos de comunicación estandarizados. Esto garantiza la interoperabilidad entre diferentes componentes del sistema. La arquitectura debe soportar procesamiento en tiempo real y permitir la integración con sistemas SCADA existentes. Se garantiza la redundancia y recuperación ante fallos mediante la implementación de mecanismos de tolerancia a fallos. Estos aseguran la continuidad operativa. Esta arquitectura multicapa facilita la modularidad del sistema. Permite actualizaciones independientes de cada componente sin comprometer la funcionalidad global. La capa de adquisición de datos debe manejar múltiples protocolos de comunicación industrial. Estos incluyen Modbus, OPC-UA y Ethernet/IP. Así se garantiza compatibilidad con equipos de diferentes fabricantes. La capa de procesamiento debe proporcionar capacidades de filtrado, validación y preprocesamiento de datos en tiempo real. Por su parte, la capa de análisis implementa los algoritmos del modelo híbrido con optimizaciones específicas para minimizar la latencia de respuesta. La capa de control debe

interfazar con los sistemas de dosificación existentes. Proporciona señales de control precisas. Monitorea la respuesta del proceso. Así se cierra el lazo de control automático.

Requerimientos operativos. Según Yamamoto y Singh (2023), los requisitos operacionales incluyen tanto elementos técnicos como organizativos. Estos son esenciales para la efectiva implementación de modelos híbridos en contextos industriales. La infraestructura de hardware debe proporcionar capacidad computacional suficiente para el procesamiento en tiempo real. Esto incluye servidores con redundancia, sistemas de almacenamiento con alta disponibilidad y redes de comunicación robustas. Estas deben manejar el volumen de datos generado por los sensores de planta sin comprometer la latencia de respuesta del sistema. El software debe cumplir con estándares de seguridad y confiabilidad industrial. Implementa protocolos de validación de datos, sistemas de respaldo automático y mecanismos de recuperación ante fallos. Estos garantizan la continuidad operativa incluso en condiciones adversas. Los protocolos operativos deben definir claramente los procedimientos de respuesta ante diferentes escenarios. Establecen límites de control claros que especifiquen cuándo el sistema debe operar automáticamente y cuándo requiere intervención manual. También incluyen procedimientos de escalamiento. Estos definen las acciones a tomar cuando el sistema detecta condiciones fuera de los parámetros normales de operación. La capacitación del personal operativo es fundamental para garantizar una operación efectiva del sistema. Incluye entrenamiento en los principios del modelo híbrido, procedimientos de calibración y mantenimiento, interpretación de alarmas y diagnósticos, y protocolos de respuesta ante emergencias. Esto asegura que el personal pueda intervenir apropiadamente cuando sea necesario.

Interfaces de usuario. Reynolds y Park (2024) sostienen que las interfaces de usuario actuales integran funciones avanzadas con una operatividad sencilla. Los dashboards interactivos proporcionan visualización en tiempo real de parámetros críticos y predicciones del modelo. Las herramientas de análisis permiten la exploración detallada de datos históricos. Esto facilita tanto la supervisión operativa continua como el análisis retrospectivo para la optimización del proceso. La ergonomía cognitiva y el diseño centrado en el usuario son fundamentales para garantizar una operación eficiente y minimizar errores operativos. Estos implementan principios de diseño que reducen la carga cognitiva del operador. Presentan la información de manera intuitiva y contextualmente relevante. Las interfaces deben proporcionar visualizaciones jerárquicas que permitan al usuario navegar desde vistas panorámicas del proceso hasta detalles específicos de variables individuales. Se incluyen capacidades de zoom temporal para examinar tendencias a diferentes escalas de tiempo. Los sistemas de alarmas y notificaciones deben implementar priorización inteligente. Esta destaca los eventos más críticos sin sobrecargar al operador con información irrelevante. Se utilizan códigos de color, sonidos y animaciones. Estos facilitan la identificación rápida de situaciones que requieren atención inmediata. La personalización de la interfaz debe permitir que diferentes usuarios configuren vistas específicas según sus roles y responsabilidades. Se mantiene consistencia en los elementos críticos de seguridad. Al mismo tiempo se proporciona flexibilidad para optimizar la experiencia de usuario según las necesidades específicas de cada operador.

Mantenimiento y calibración. Kowalski y Nakamura (2024) indican que el mantenimiento regular y la calibración garantizan un rendimiento constante del sistema híbrido durante toda su vida operativa. Los procedimientos de mantenimiento preventivo incluyen la verificación regular de sensores, actualización de software y optimización de parámetros del modelo. Estos implementan cronogramas estructurados. Garantizan la detección proactiva de deterioro en el desempeño antes de que afecte la calidad del agua tratada. Los protocolos de calibración deben considerar tanto la deriva instrumental como los cambios en las características del agua cruda. Los sensores pueden experimentar degradación gradual. Las propiedades

del agua pueden evolucionar debido a cambios estacionales, modificaciones en la cuenca de captación o variaciones en las actividades antropogénicas que afectan la fuente. La implementación de sistemas de diagnóstico automático facilita la detección temprana de anomalías y la programación eficiente de intervenciones de mantenimiento. Se utilizan algoritmos de monitoreo continuo. Estos analizan tendencias en el desempeño del modelo. Detectan desviaciones significativas en los parámetros de calibración. Y generan alertas preventivas cuando se identifican patrones que sugieren la necesidad de mantenimiento correctivo. Los sistemas de diagnóstico deben incluir capacidades de auto-validación. Estas verifican la coherencia entre diferentes sensores. Detectan fallos incipientes mediante análisis de residuos. Y proporcionan recomendaciones específicas sobre las acciones de mantenimiento requeridas para restaurar el desempeño óptimo del sistema híbrido.

2.3 Definición de términos básicos

Coagulación: Proceso fisicoquímico que consiste en la desestabilización de partículas coloidales mediante la adición de sustancias químicas (coagulantes), neutralizando las fuerzas que las mantienen separadas y permitiendo su aglomeración (Crittenden et al., 2021).

Dosificación óptima: Cantidad precisa de coagulante por unidad de volumen de agua que maximiza la eficiencia del proceso de coagulación mientras minimiza el consumo de insumos químicos y mantiene la calidad del agua tratada dentro de los estándares establecidos (AWWA, 2023b).

Lógica difusa: Sistema de razonamiento basado en la teoría de conjuntos difusos que permite modelar matemáticamente la imprecisión y la incertidumbre, proporcionando un método formal para la representación y manipulación del conocimiento aproximado (Ross, 2023).

Modelo híbrido: Sistema que integra dos o más técnicas o metodologías complementarias para aprovechar las fortalezas individuales de cada enfoque y superar sus limitaciones específicas, mejorando el desempeño global del sistema (Thompson et al., 2024).

Splines cúbicos: Funciones matemáticas de interpolación construidas por tramos mediante polinomios de tercer grado, que garantizan continuidad hasta la segunda derivada en los puntos de unión, permitiendo generar curvas suaves a partir de datos discretos (Quarteroni & Saleri, 2024).

Tabla de dosificación: Herramienta operativa que relaciona los parámetros de calidad del agua cruda (principalmente turbiedad) con las dosis de coagulante requeridas, basada en la experiencia operativa y pruebas de laboratorio (WRF, 2024e).

Turbiedad: Medida del grado en el cual el agua pierde su transparencia debido a la presencia de partículas en suspensión, expresada típicamente en Unidades Nefelométricas de Turbiedad (*UNT*) (OMS, 2022).

Planta de tratamiento convencional: Instalación diseñada para potabilizar agua mediante una secuencia de procesos que típicamente incluye coagulación, floculación, sedimen-

tación, filtración y desinfección (Davis, 2023).

Sistema SCADA: (*Supervisory Control And Data Acquisition*) Sistema informático que permite la supervisión, control y adquisición de datos en procesos industriales, facilitando la monitorización y control en tiempo real de plantas de tratamiento de agua (IWA, 2024a).

Validación cruzada: Técnica estadística utilizada para evaluar la capacidad de generalización de un modelo, dividiendo el conjunto de datos disponibles en subconjuntos para entrenamiento y validación (Anderson & Whitcomb, 2024).

Control adaptativo: Sistema de control que modifica automáticamente sus características en respuesta a cambios en las condiciones del proceso o perturbaciones, optimizando continuamente su desempeño (Chen & Wang, 2023).

Prueba de jarras: Ensayo de laboratorio que simula los procesos de coagulación, floculación y sedimentación a escala reducida, utilizado para determinar las dosis óptimas de coagulantes y las condiciones de operación más adecuadas (APHA, 2017).

pH: Medida de la concentración de iones hidrógeno en una solución, que influye significativamente en la eficiencia del proceso de coagulación y en la especiación de los coagulantes (Sillanpää & Matilainen, 2023).

CAPÍTULO 3

MATERIALES Y MÉTODOS

3.1 Localización del estudio

3.1.1 *Ubicación geográfica*

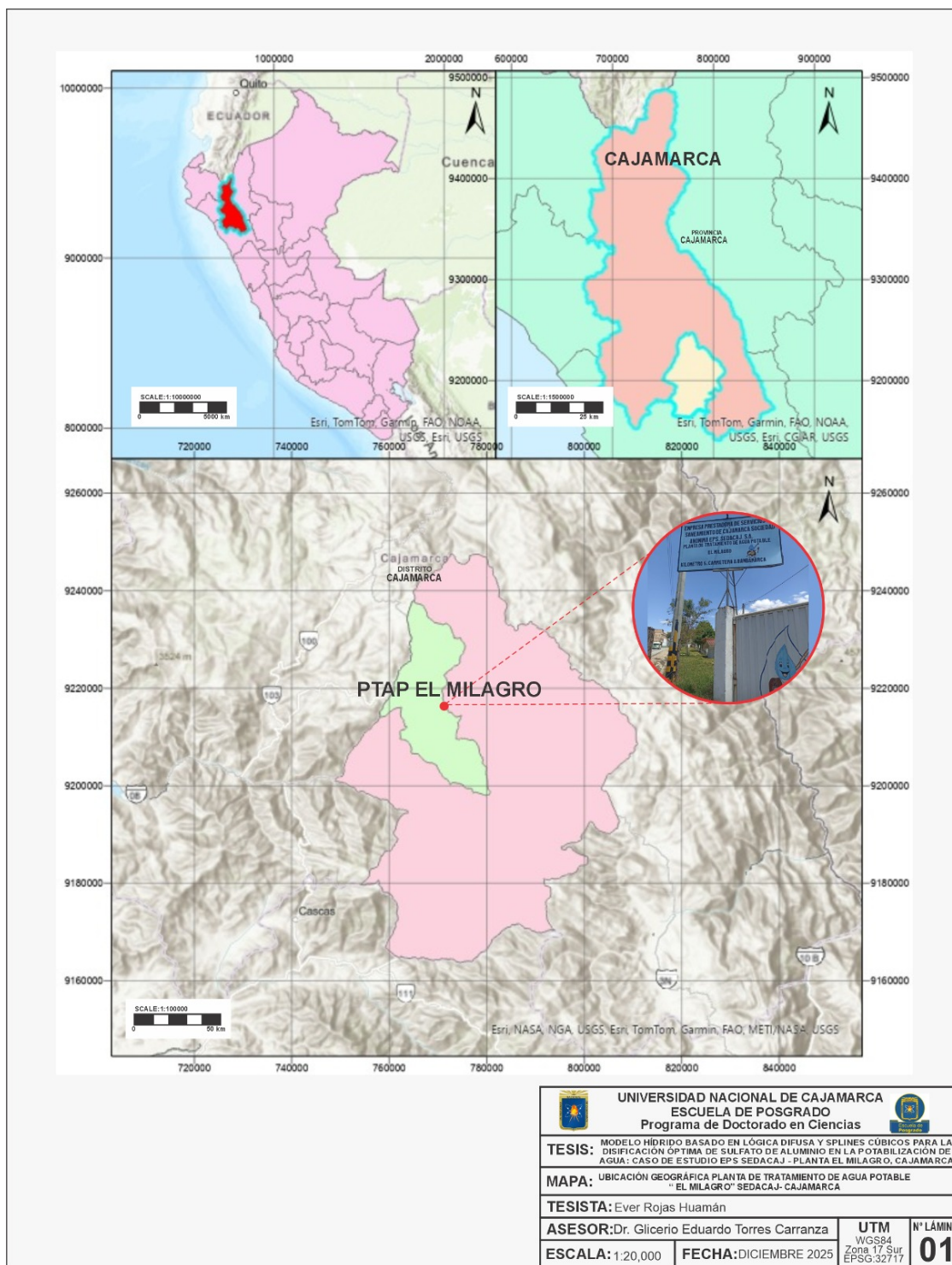
La investigación se desarrolló en la planta de tratamiento de agua potable El Milagro, operada por la Empresa Prestadora de Servicios de Saneamiento de Cajamarca S.A. (EPS SEDACAJ), localizada en el distrito de Cajamarca, provincia de Cajamarca, departamento de Cajamarca, Perú.

Las coordenadas geográficas de la planta son $7^{\circ}09'12''$ de latitud sur y $78^{\circ}30'57''$ de longitud oeste, a una altitud de 2 750 m s n m. En el sistema de coordenadas UTM (datum WGS84, zona 17 Sur), la ubicación corresponde aproximadamente a 776 800 m Este y 9 208 400 m Norte.

La Figura 1 presenta la ubicación geográfica de la planta de tratamiento El Milagro, mostrando su posición respecto a la ciudad de Cajamarca.

Figura 1

Ubicación geográfica de la planta de tratamiento de agua potable “El Milagro” - EPS SE-DACAJ

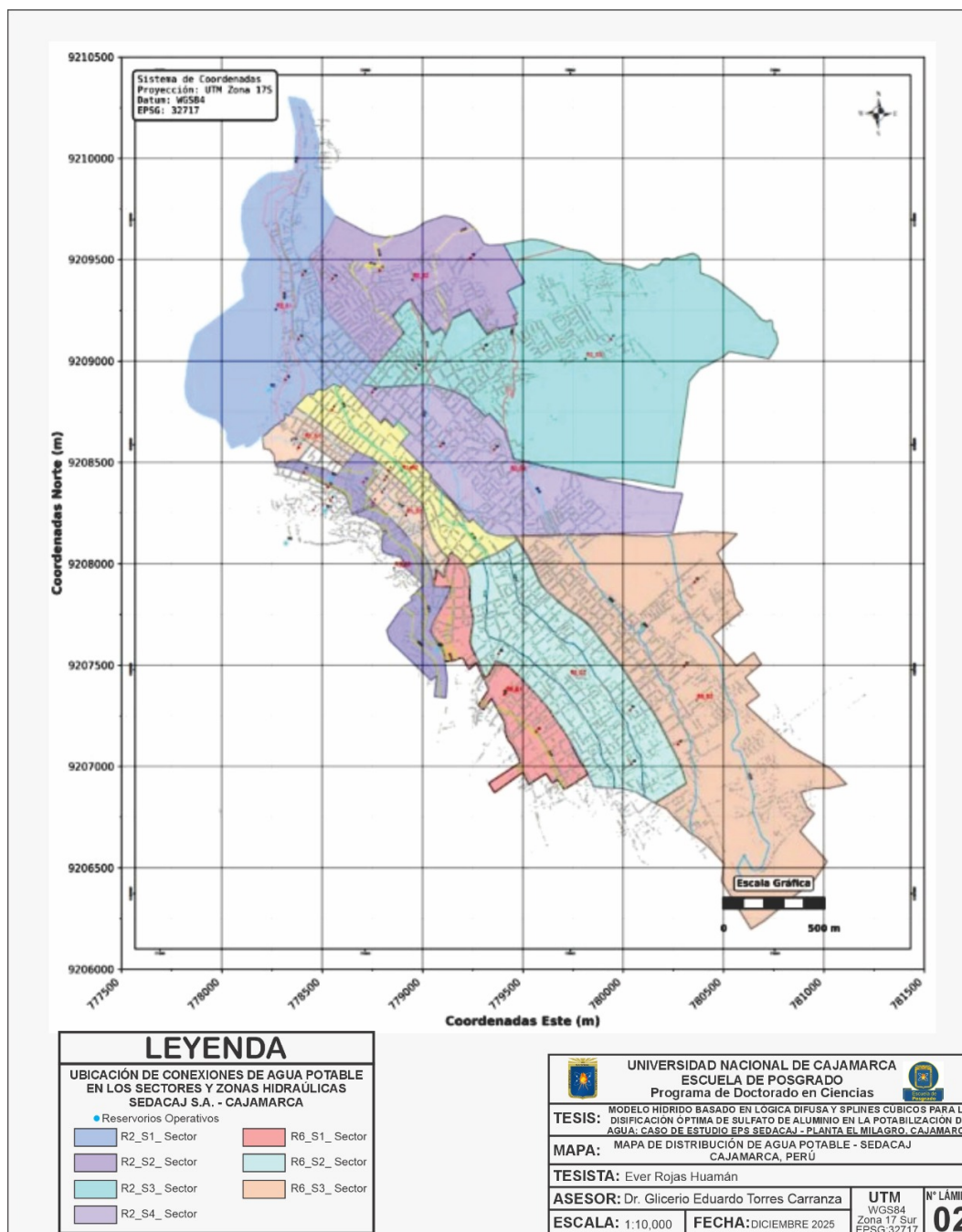


3.1.2 Área de influencia

La planta de tratamiento “El Milagro” abastece agua potable a la ciudad de Cajamarca mediante un sistema de distribución que opera por gravedad a través de los reservorios **R-2** (capacidad de 2500 m³) y **R-6** (capacidad de 1500 m³), atendiendo los sectores urbanos más densamente poblados, como se muestra en la Figura 2.

Figura 2

Área de influencia de la planta “El Milagro” en la ciudad de Cajamarca



El área de influencia de la planta abarca múltiples sectores residenciales de Cajamarca, abasteciendo a más de **120 000 habitantes**, lo que representa aproximadamente el **70 % de la población urbana** del distrito (EPS SEDACAJ S.A., 2024; Instituto Nacional de Estadística

e Informática, 2023).

3.1.3 Características operativas del sistema

La planta procesa agua superficial proveniente de los ríos Grande y Porcón, los cuales confluyen en una cámara de recepción con un caudal medio combinado de aproximadamente 80 L s^{-1} . El sistema implementa un tren de tratamiento convencional que comprende: captación, coagulación con sulfato de aluminio, floculación hidráulica, sedimentación, filtración rápida y desinfección con cloro gas.

El sistema opera con caudales en el rango de 150 a 300 L s^{-1} , adaptándose a las variaciones estacionales de demanda. La calidad del agua cruda presenta variabilidad significativa, con turbidez fluctuando entre 0 y 3500 UNT , lo que justificó el desarrollo del modelo híbrido de dosificación propuesto.

3.1.4 Condiciones climáticas

La zona de estudio se caracteriza por un clima templado-seco con estacionalidad marcada (SENAMHI, 2023). Las precipitaciones se concentran entre diciembre y abril (650 – 800 mm anuales), generando eventos de alta turbidez. La temperatura media anual oscila entre 14°C y 18°C , con humedad relativa de 60 – 75% .

El régimen hidrológico presenta dos períodos diferenciados: época seca (mayo–noviembre) con caudales de 150 – 200 L s^{-1} y turbidez inferior a 50 UNT , y época húmeda (diciembre–abril) con caudales hasta 300 L s^{-1} y turbidez superior a 2000 UNT en eventos extremos.

3.2 Materiales

3.2.1 Ubicación y características del sistema de tratamiento

La investigación se desarrolló en la planta de tratamiento de agua potable “El Milagro” de EPS SEDACAJ (ver Sección 3.1 para características detalladas de ubicación y operación del sistema). La planta opera con caudales variables entre 150 y 300 L s⁻¹, presentando variabilidad de turbidez entre 0 y 3500 UNT, condiciones que proporcionaron el escenario necesario para validar el modelo híbrido bajo diferentes escenarios operativos representativos de sistemas de tratamiento en zonas andinas.

3.2.2 Equipos e instrumentación de laboratorio

Equipo principal de ensayos. Se utilizó equipo de prueba de jarras Phipps & Bird modelo PB-950 con seis posiciones, control digital programable (0–300 rpm, precisión ± 1 rpm) y sistema de iluminación posterior para evaluación visual del floc.

Instrumentos de medición. Los parámetros fisicoquímicos se midieron con instrumentación calibrada según estándares internacionales: turbidímetro portátil Hach 2100Q (rango 0–1000 UNT, resolución 0,01 UNT, precisión ± 2 % de lectura $+0,01$ UNT) (Company, 2020); pH-metro OHAUS ST300 (rango 0–14, precisión $\pm 0,01$ unidades de pH) (Corporation, 2018); termómetro digital Cole-Parmer Traceable™ (rango 0–100 °C, precisión $\pm 0,1$ °C) (Company, 2019). Todos los equipos fueron calibrados con estándares certificados trazables a NIST.

Material volumétrico y reactivos. Se empleó material de vidrio borosilicato clase A: pipetas volumétricas (1–50 mL), probetas graduadas (100–1000 mL), matraces aforados y buretas (25 mL, división 0,1 mL).

Se preparó solución madre de sulfato de aluminio $[\text{Al}_2(\text{SO}_4)_3 \cdot 18\text{H}_2\text{O}]$ a una concentración de 10,00 gL⁻¹ utilizando agua destilada tipo II según norma ASTM D1193 (ASTM International, 2018). La concentración de la solución se verificó semanalmente mediante aná-

lisis gravimétrico para garantizar trazabilidad analítica.

3.2.3 *Herramientas computacionales*

Etapas de investigación y validación (MATLAB). El desarrollo, calibración y validación experimental del modelo híbrido se realizó en MATLAB® R2024a (The MathWorks Inc., 2024) utilizando los siguientes toolboxes especializados:

- Curve Fitting Toolbox: para implementación de splines cúbicos naturales
- Fuzzy Logic Toolbox: para desarrollo del sistema de inferencia difusa tipo Mamdani
- Optimization Toolbox: para calibración de parámetros de fusión adaptativa
- Statistics and Machine Learning Toolbox: para validación cruzada k-fold y análisis estadístico

Etapas de implementación operativa (Python). La implementación del sistema de dosificación automatizada y su interfaz operativa se desarrolló en Python 3.9.7 (Python Software Foundation, 2021) con las siguientes bibliotecas especializadas:

- Streamlit 1.28.1: framework para desarrollo de interfaz web interactiva (Streamlit Inc., 2024)
- NumPy 1.24.3 y SciPy 1.10.1: cálculos numéricos y procesamiento científico de datos (Harris et al., 2020; Virtanen et al., 2020)
- scikit-fuzzy 0.4.2: implementación del sistema de inferencia difusa calibrado (Warner et al., 2019)
- Pandas 2.0.3: gestión y análisis de datos tabulares (McKinney, 2010)
- Matplotlib 3.7.1 y Plotly 5.15.0: visualización bidimensional de curvas de dosificación (Hunter, 2007; Plotly Technologies Inc., 2015)

3.2.4 Fuentes de información

Se utilizaron dos fuentes complementarias de datos de EPS SEDACAJ para el desarrollo y validación del modelo híbrido:

- `tabla_dosificacion.csv`: tabla operativa histórica de la planta con relaciones empíricas turbidez-pH-dosis de sulfato de aluminio para 16 caudales operativos (150–300 L s⁻¹), utilizada como referencia comparativa para evaluar el desempeño del modelo propuesto
- Base de datos experimental propia: 48 puntos de control representativos generados específicamente para esta investigación mediante pruebas de jarras estandarizadas según protocolo APHA-AWWA-WEF (APHA, 2017), distribuidos estratégicamente para calibración de splines cúbicos y validación del sistema híbrido

3.3 Metodología

3.3.1 Método para el diseño de puntos de control representativos

Se desarrolló un programa experimental para generar 48 puntos de control distribuidos estratégicamente: 16 caudales operativos ($150\text{--}300\text{ L s}^{-1}$ con incrementos de 10 L s^{-1}) $\times$ 3 niveles de turbidez (baja, media, alta) por caudal.

Los niveles de turbidez se establecieron mediante análisis estadístico de registros históricos operativos de la planta: turbidez baja (percentil 10 %), media (percentil 50 %) y alta (percentil 90 %) para cada estrato de caudal, reflejando la relación inversa característica entre caudal y turbidez observada en sistemas de captación superficial en zonas montañosas (molina2021; Rodríguez et al., 2019).

Protocolo de pruebas de jarras. Se aplicó protocolo APHA, AWWA, WEF (2017) con las siguientes condiciones estandarizadas: volumen 2000 mL/jarra, temperatura $15 \pm 2^\circ\text{C}$, evaluación de dosis $5\text{--}45\text{ mg L}^{-1}$ (incrementos 5 mg L^{-1}).

La secuencia de mezclado comprendió: dosificación simultánea de coagulante, mezcla rápida $200\text{ rpm} \times 2\text{ min}$ (gradiente 700 s^{-1}), floculación $40\text{ rpm} \times 20\text{ min}$ (gradiente 50 s^{-1}), sedimentación 30 min, y muestreo a 2 cm bajo superficie.

Los criterios de optimización fueron: turbidez residual $\leq 5\text{ UNT}$, eficiencia $\geq 85\%$, estabilidad pH ($\pm 0,5$ unidades), y calidad visual del floc. Cada punto se validó por triplicado con desviación estándar $< 5\%$.

Justificación del diseño experimental. El tamaño muestral de 48 puntos de control se determinó considerando tres criterios fundamentales: (a) cobertura sistemática de 16 caudales operativos representativos del rango $150\text{--}300\text{ L s}^{-1}$ con resolución uniforme de 10 L s^{-1} , garantizando representatividad del espacio operativo; (b) estratificación tripartita de turbidez por caudal basada en percentiles estadísticos (10 %, 50 %, 90 %) derivados de registros históricos 2018–2024, capturando condiciones de estiaje, operación normal y eventos extremos;

(c) requerimientos mínimos para validación cruzada k-fold con $k = 5$, asegurando aproximadamente 9–10 observaciones por fold y evitando particiones con datos insuficientes que comprometerían la estabilidad estadística (Kohavi, 1995).

Este diseño proporciona densidad suficiente de puntos de control para interpolación spline cúbica bidimensional sin sobreajuste, cumpliendo con criterios establecidos por de Boor (2001) para funciones spline en múltiples dimensiones, donde la relación puntos/parámetros debe exceder 3:1 para estabilidad numérica. Con 48 puntos distribuidos en 16 caudales, se obtiene una media de 3 puntos por caudal, satisfaciendo este criterio conservador.

3.3.2 *Método para desarrollo del componente de splines cúbicos*

Se implementó interpolación mediante splines cúbicos naturales (de Boor, 2001; Schumaker, 2007) utilizando los 48 puntos de control experimentales como base de calibración. El método garantizó continuidad C^2 (continuidad hasta la segunda derivada) y minimización de energía de curvatura mediante resolución de sistemas tridiagonales.

Implementación computacional. La implementación se realizó en MATLAB® R2024a con Curve Fitting Toolbox, aplicando algoritmos de factorización tridiagonal de Thomas (Press et al., 2024) para resolver eficientemente el sistema de ecuaciones lineales resultante. El proceso se ejecutó de manera independiente para cada caudal operativo ($150\text{--}300\text{ L s}^{-1}$), aplicando interpolación sobre las relaciones turbidez-dosis correspondientes.

Validación cruzada k-fold. Se implementó validación cruzada k-fold con $k = 5$ (Kohavi, 1995; Hastie et al., 2009) para evaluación no sesgada de la capacidad de generalización del modelo. El conjunto de 48 puntos se dividió aleatoriamente en cinco subconjuntos equilibrados, utilizando cuatro subconjuntos (80 % de datos) para calibración y uno (20 % de datos) para validación en cada iteración.

Las métricas de evaluación estadística incluyeron: coeficiente de determinación (R^2), error cuadrático medio (RECM), y error porcentual absoluto medio (EPAM), proporcionando

caracterización integral del desempeño predictivo del componente de interpolación.

3.3.3 Método para construcción del sistema de lógica difusa

Se desarrolló un sistema de inferencia difusa tipo Mamdani (Mamdani & Assilian, 1975; Ross, 2023) para proporcionar estimaciones de dosificación bajo condiciones operativas fuera del rango calibrado experimentalmente. El sistema integró tres variables de entrada (turbidez, pH, caudal) y una variable de salida (dosis de sulfato de aluminio).

Diseño de funciones de pertenencia. Se implementaron funciones de pertenencia con partición variable según la relevancia de cada parámetro (Wang, 1997; Mendel, 2001): turbidez (6 conjuntos difusos, dominio 0–25 000 UNT), pH (3 conjuntos difusos, dominio 5.0–9.5), caudal (3 conjuntos difusos, dominio 0–300 L s⁻¹), dosis de salida (5 conjuntos difusos, dominio 0–300 L s⁻¹). La extensión deliberada de los universos de discurso más allá de los rangos empíricos observados (0–3500 UNT para turbidez) permitió capacidad de extrapolación hacia condiciones operativas extremas no calibradas.

Formulación de la base de conocimiento. Se formularon 19 reglas difusas fundamentales mediante protocolo sistemático de extracción de conocimiento experto (Cornelissen et al., 2003; Clemen, 1996): entrevistas estructuradas con operadores experimentados de la planta (> 5 años de experiencia), análisis de estrategias documentadas para eventos operativos extremos, y validación técnica con principios fisicoquímicos de coagulación-floculación. Las reglas capturaron tanto las interacciones principales entre variables de proceso como los casos operativos extremos identificados en el historial de la planta.

Protocolo de elicitación de conocimiento experto. El proceso de elicitación de conocimiento para formulación de reglas difusas siguió protocolo estructurado en tres fases secuenciales (Cornelissen et al., 2003):

Fase 1 - Entrevistas individuales: Se realizaron entrevistas semiestructuradas con cuatro operadores senior de la Planta El Milagro (experiencia promedio: 7.5 años, rango: 5–12 años) mediante cuestionario de 25 preguntas categorizadas en: (i) estrategias de dosificación bajo condiciones normales (turbidez < 500 UNT), (ii) protocolos ante eventos de turbidez extrema (> 2000 UNT), (iii) criterios de ajuste por pH fuera de rango neutro, y (iv) experiencia con subdosificación/sobredosificación documentada. Las sesiones fueron audio-grabadas (previo consentimiento informado) y transcritas literalmente para análisis temático posterior.

Fase 2 - Consenso mediante grupo focal: Sesión de discusión estructurada de 3 horas con los cuatro operadores entrevistados más dos supervisores técnicos de SEDACAJ, facilitada por el investigador principal. Se presentaron 35 escenarios operativos hipotéticos (combinaciones de turbidez-pH-caudal) para consensuar estrategias de dosificación mediante técnica Delphi modificada (dos rondas de votación anónima con retroalimentación). Los casos con discrepancia > 20 % en dosis propuestas se discutieron hasta alcanzar consenso o identificar necesidad de validación experimental adicional.

Fase 3 - Formalización y validación: El conocimiento recopilado se codificó en 19 reglas difusas tipo SI-ENTONCES (formato Mamdani). Las reglas propuestas fueron validadas mediante: (a) revisión técnica por panel de tres expertos externos (dos ingenieros sanitarios de otras EPS y un académico especialista en tratamiento de agua), y (b) simulación retrospectiva contra 48 eventos operativos extremos documentados en bitácoras 2020–2024, verificando coherencia entre dosis sugerida por reglas y dosis aplicada efectivamente. La tasa de concordancia alcanzó 87.5 % (42/48 casos), con discrepancias atribuibles a información incompleta en registros históricos.

3.3.4 Método para integración del modelo híbrido

Se implementó fusión adaptativa mediante ponderación dinámica basada en proximidad a puntos de control experimentales, siguiendo principios de fusión de información en sistemas inteligentes (Durrant-Whyte & Henderson, 2006; Khaleghi et al., 2013):

$$D_{\text{híbrido}} = \omega_s \cdot D_{\text{spline}} + \omega_f \cdot D_{\text{difusa}} \quad (3.1)$$

con restricción de normalización $\omega_s + \omega_f = 1$.

Función de ponderación adaptativa. Los pesos de fusión se determinaron mediante función basada en distancia normalizada (Jousselme et al., 2001; Rogova & Nimier, 2004):

$$\omega_s(T, Q, P) = \frac{\exp(-\lambda \cdot d_{\text{norm}}(T, Q)^2) \cdot \phi(T, Q)}{\exp(-\lambda \cdot d_{\text{norm}}(T, Q)^2) \cdot \phi(T, Q) + \exp(-\mu \cdot \sigma_{\text{difusa}}(T, Q, P)^2)} \quad (3.2)$$

donde $d_{\text{norm}}(T, Q)$ representa la distancia euclídiana normalizada al punto de control experimental más cercano, $\phi(T, Q)$ es un factor de confianza basado en densidad local de datos, y λ, μ son parámetros de suavidad calibrados experimentalmente.

Calibración de parámetros. Los parámetros λ y μ se calibraron mediante búsqueda sistemática en malla (*grid search*) evaluando el desempeño del modelo híbrido bajo diferentes combinaciones de valores. El espacio de búsqueda se definió como $\lambda \in [0, 5, 5, 0]$ con incrementos de 0,05 y $\mu \in [0, 5; 4, 0]$ con incrementos de 0,05, resultando en aproximadamente 8 000 configuraciones candidatas.

Para cada combinación (λ_i, μ_j) , se evaluó el desempeño mediante validación cruzada k-fold ($k = 5$) calculando las métricas RECM, R^2 y ENS sobre los 48 puntos de control experimentales. La función de evaluación ponderó tres aspectos fundamentales:

- **Precisión predictiva (60 %):** Minimización del RECM promedio en los cinco folds de validación cruzada

- **Robustez (25 %):** Estabilidad del modelo ante perturbaciones de ± 20 % en turbidez de entrada
- **Continuidad (15 %):** Suavidad de transiciones entre componentes spline y difuso

Los valores óptimos encontrados fueron $\lambda^* = 2,35$ y $\mu^* = 1,87$, que minimizaron la función de evaluación compuesta alcanzando $RECM = 0,087 \text{ L s}^{-1}$ y $R^2 = 0,9985$ en validación cruzada.

Diseño de validación. La validación del modelo híbrido se realizó mediante validación cruzada estratificada (Kohavi, 1995; Rodríguez et al., 2010), considerando un total de 48 puntos de control experimentales y 19 reglas difusas previamente definidas. Con el fin de garantizar la representatividad estadística de los datos y evitar pliegues excesivamente pequeños, se utilizó una partición en $k = 5$ pliegues, recomendada en la literatura para conjuntos de datos de tamaño moderado (Hastie et al., 2009; Kuhn & Johnson, 2013).

La estratificación se efectuó en función de dos variables críticas del proceso: la turbidez del agua de entrada y el caudal de operación. De este modo, cada pliegue incluyó proporciones equivalentes de observaciones en los distintos niveles de turbidez (muy baja, baja, media, alta y muy alta) y de caudal ($150\text{--}300 \text{ L s}^{-1}$). Este criterio de partición permitió que en cada iteración del entrenamiento y prueba se preservara la heterogeneidad de las condiciones operativas, favoreciendo así la reproducibilidad y la estabilidad estadística de los resultados obtenidos (Lemaître et al., 2017).

En cada iteración, cuatro pliegues (80 % de los datos) fueron utilizados para el entrenamiento del modelo y el pliegue restante (20 % de los datos) para la prueba de validación. El proceso se repitió hasta completar las **5 iteraciones**, asegurando que todos los puntos de control participaran **exactamente una vez** en la fase de validación independiente.

3.3.5 Método para evaluación del desempeño

Se implementó protocolo de validación experimental con validación cruzada estratificada y análisis comparativo frente al método tradicional de interpolación lineal utilizado operativamente por SEDACAJ.

Métricas de evaluación cuantitativa. Se utilizaron métricas estadísticas validadas en literatura especializada para evaluación de modelos predictivos en sistemas de tratamiento de agua (Moriassi et al., 2007; Legates & McCabe Jr, 1999):

- **Error Cuadrático Medio Normalizado (RECMN):**

$$\text{RECMN} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\bar{y}} \times 100 \quad (3.3)$$

- **Error Porcentual Absoluto Medio (EPAM):**

$$\text{EPAM} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (3.4)$$

- **Coefficiente de Determinación Ajustado (R_{adj}^2):**

$$R_{\text{adj}}^2 = 1 - \frac{(1 - R^2)(n - 1)}{n - p - 1} \quad (3.5)$$

- **Eficiencia de Nash-Sutcliffe (ENS) (Nash & Sutcliffe, 1970):**

$$\text{ENS} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.6)$$

Análisis de sensibilidad paramétrica. Se evaluó la estabilidad de cada componente del modelo mediante perturbaciones controladas de $\pm 10\%$, $\pm 20\%$ y $\pm 30\%$ en las variables de entrada, siguiendo metodología de análisis de sensibilidad global (Saltelli et al., 2008; Hamby, 1994). La sensibilidad se cuantificó mediante el Índice de Sensibilidad Normalizada (NSI) (McCuen, 1973; Pannell, 1997):

$$\text{NSI} = \frac{\Delta D / D_{\text{base}}}{\Delta T / T_{\text{base}}} \times 100 \quad (3.7)$$

donde ΔD representa la variación en la dosis estimada, D_{base} es la dosis de referencia sin perturbación, ΔT es la perturbación aplicada a la turbidez de entrada, y T_{base} es el valor de turbidez de referencia.

El análisis se ejecutó en caudales representativos (150, 200, 250, 300 L s⁻¹) estableciendo como criterio de aceptación valores de NSI en el rango 0,8–1,3 para comportamiento estable y respuesta aproximadamente proporcional ante perturbaciones operativas.

Generación de superficies de respuesta. Se generaron superficies tridimensionales en MATLAB para caracterización matemática del comportamiento de cada componente mediante interpolación sobre mallas regulares con resolución de 1 L s⁻¹ en caudal y 10 UNT en turbidez (Mathews & Fink, 2004). Los puntos experimentales se superpusieron para validación visual de fidelidad de interpolación y verificación de continuidad matemática C^2 en toda la superficie generada.

Las superficies permitieron identificar regiones de alta curvatura, puntos de inflexión y comportamiento asintótico del modelo, proporcionando análisis cualitativo complementario a las métricas cuantitativas. Este análisis tridimensional se realizó exclusivamente en el entorno de desarrollo MATLAB, mientras que la interfaz operativa en Python implementó visualización bidimensional de curvas de dosificación para facilidad de interpretación por parte del personal de planta.

Análisis estadístico comparativo. Se aplicó análisis de varianza no paramétrico mediante la prueba de Kruskal-Wallis (Kruskal & Wallis, 1952) para identificar diferencias estadísticamente significativas entre métodos de dosificación evaluados, complementado con pruebas post-hoc de Dunn (Dunn, 1964) con corrección de Bonferroni (Bland & Altman, 1995) para comparaciones múltiples. El nivel de significancia se estableció en $\alpha = 0,05$ para todas las pruebas estadísticas.

Se calculó el coeficiente de correlación intraclase (ICC) (Shrout & Fleiss, 1979; McGraw & Wong, 1996) para cuantificar la concordancia entre métodos, estableciendo como criterio de aceptación $ICC > 0,90$ para concordancia excelente según clasificación de Cicchetti (Cicchetti, 1994). El análisis estadístico se realizó en MATLAB utilizando las funciones del Statistics and Machine Learning Toolbox.

3.3.6 Implementación del dashboard operativo

Se desarrolló plataforma de visualización interactiva para operación en tiempo real utilizando Python 3.9.7 con framework Streamlit (Streamlit Inc., 2024), implementando arquitectura modular que integró los siguientes componentes funcionales:

- Panel de control principal con validación automática de rangos operativos según criterios de seguridad establecidos
- Visualización bidimensional de curvas de dosificación turbidez-dosis para diferentes caudales operativos
- Herramientas interactivas de análisis de sensibilidad paramétrica con perturbaciones controladas
- Sistema de registro automático de decisiones con trazabilidad de estimaciones y condiciones operativas
- Sistema de alertas para condiciones operativas atípicas basado en umbrales configurables

El dashboard implementado permite la estimación de dosis óptimas mediante interfaz web accesible desde cualquier dispositivo con navegador, facilitando la operación remota y el monitoreo continuo del proceso de dosificación. Las capturas de pantalla del sistema implementado se presentan en el Anexo C, mostrando la interfaz de usuario y funcionalidades desarrolladas.

3.3.7 Validación de cumplimiento normativo

Se verificó el cumplimiento de las dosis estimadas por el modelo híbrido con estándares nacionales e internacionales mediante simulación de eficiencia de remoción de turbidez. Se evaluó conformidad con el Reglamento de la Calidad del Agua para Consumo Humano del Perú, establecido mediante Decreto Supremo N.º 031-2010-SA (Ministerio de Salud, 2010), que establece límite máximo de turbidez ≤ 5 UNT en agua tratada para consumo humano.

Adicionalmente, se verificó cumplimiento con las Guías para la Calidad del Agua de Bebida de la Organización Mundial de la Salud (OMS, 2022), que recomienda turbidez < 1 UNT para eficacia óptima de desinfección. El análisis incluyó evaluación de eficiencia de remoción proyectada y verificación de que las dosis estimadas no generaran condiciones de subdosificación (turbidez residual > 5 UNT) ni sobredosificación (formación excesiva de lodos o residuales de aluminio $> 0,2 \text{ L s}^{-1}$) según límites establecidos por la normativa peruana (SUNASS, 2023).

3.3.8 Control de calidad y trazabilidad experimental

Se implementó programa de calibración sistemática con frecuencias diferenciadas según criticidad del instrumento: calibración diaria para pH-metro y turbidímetro utilizando estándares certificados trazables a NIST (National Institute of Standards and Technology, 2019), calibración semanal para equipo de prueba de jarras verificando uniformidad de velocidades en las seis posiciones, calibración mensual para balanza analítica con pesas patrón clase E2 según norma OIML R111 (OIML, 2004), y certificación anual mediante laboratorio externo acreditado bajo norma ISO/IEC 17025:2017 (ISO/IEC, 2017).

Se mantuvo trazabilidad completa de reactivos químicos y estándares de calibración mediante registro documentado que incluyó certificados de análisis, fechas de preparación y caducidad, y condiciones de almacenamiento. Los registros de calibración, mantenimiento preventivo y observaciones experimentales se documentaron sistemáticamente en formatos estandarizados para garantizar reproducibilidad de resultados y cumplimiento de buenas prácticas de laboratorio.

Los residuos generados durante la investigación se gestionaron según normativa ambiental vigente establecida en la Ley de Gestión Integral de Residuos Sólidos (Decreto Legislativo N.º 1278) (Congreso de la República, 2016) y su Reglamento (Decreto Supremo N.º 014-2017-MINAM) (Ministerio del Ambiente, 2017), mediante deshidratación de lodos de pruebas de jarras, neutralización de soluciones ácidas/básicas residuales, y entrega a empresa operadora de residuos sólidos (EO-RS) debidamente autorizada por DIGESA, manteniendo registro documental según lo establecido en la normativa.

CAPÍTULO 4

RESULTADOS Y DISCUSIÓN

4.1 Diseño y validación de puntos de control representativos para dosificación

Tabla 1

Matriz de puntos de control experimentales para calibración del modelo híbrido

Caudal (L s ⁻¹)	Turb. mínima (UNT)	Turb. media (UNT)	Turb. máxima (UNT)
150	1,0	558,0	2018,0
160	1,0	407,0	1650,0
170	1,0	261,0	1378,0
180	1,0	211,0	1180,0
190	1,0	161,0	1008,0
200	1,0	123,0	842,0
210	0,0	94,0	704,0
220	0,0	74,0	603,0
230	0,0	61,0	512,0
240	0,0	49,0	433,0
250	0,0	39,0	364,0
260	0,0	33,0	306,0
270	0,0	28,0	258,0
280	0,0	22,0	219,0
290	0,0	17,0	189,0
300	0,0	14,0	168,0

Nota. Niveles de turbidez corresponden a percentiles 10 %, 50 % y 90 % de registros históricos operativos. $n = 48$ puntos de control.

La Tabla 1 presenta la matriz experimental de 48 puntos de control obtenida en cumplimiento del primer objetivo específico de la investigación. La distribución estratégica en 16 cauda-

les operativos (150 a 300 L s^{-1}) con tres niveles de turbidez por caudal proporcionó la base empírica para la calibración del modelo híbrido.

Los valores reflejan la relación inversa entre caudal y turbidez característica de sistemas de captación superficial en zonas montañosas (Molina et al., 2021). El rango de turbidez (0 a 2018 UNT) capturó tanto condiciones normales de operación como eventos críticos asociados a precipitaciones intensas. Esto aseguró representatividad operativa del diseño experimental.

4.1.1 Caracterización estadística de los puntos de control experimentales

Tabla 2

Estadísticos descriptivos de las variables experimentales ($n = 16$ caudales)

Variable	Media	Desv. estándar	Varianza	Mínimo	Máximo
Caudal (L s^{-1})	225,00	47,61	2 266,67	150,00	300,00
Turbidez mínima (UNT)	0,38	0,50	0,25	0,00	1,00
Turbidez media (UNT)	134,50	156,22	24 396,09	14,00	558,00
Turbidez máxima (UNT)	739,50	564,20	318 323,4	168,00	2018,0

Nota. Coeficiente de variación para turbidez máxima: $\text{CV} = 76 \%$.

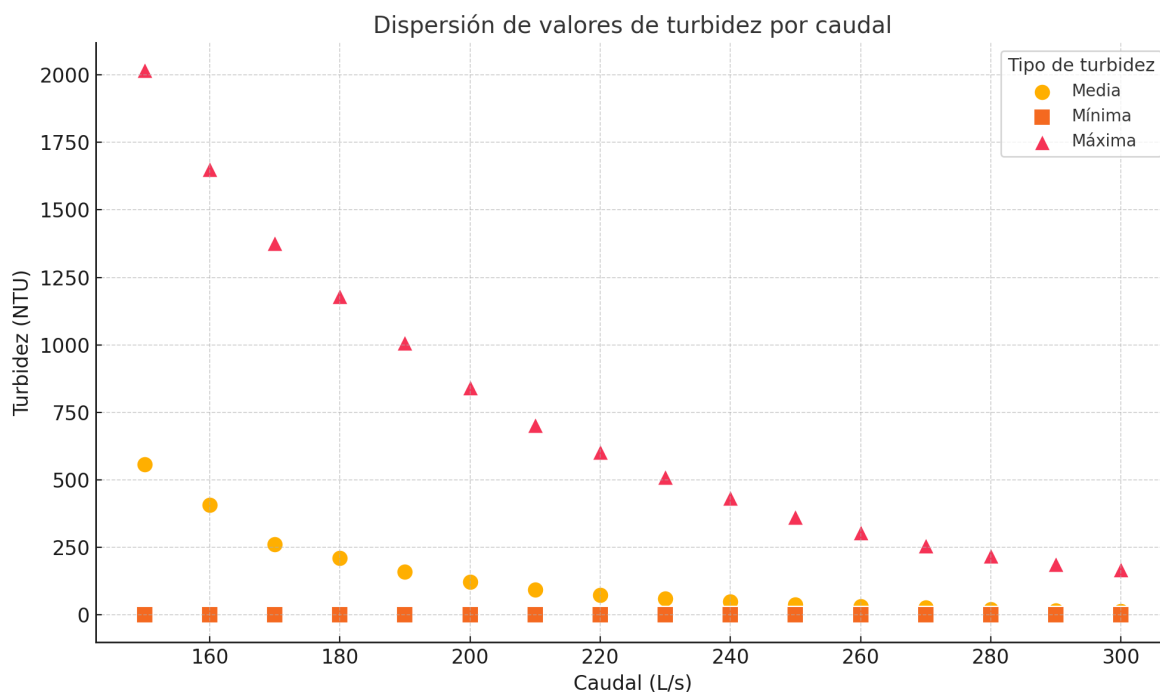
La Tabla 2 cuantifica la variabilidad de las variables experimentales. El coeficiente de variación ($\text{CV} = 76 \%$) para turbidez máxima evidencia la alta dispersión característica de sistemas de captación superficial andinos sujetos a precipitaciones estacionales intensas. Bressane et al. (2023) identificaron esta variabilidad como crítica en plantas brasileñas. Allí compromete sistemas de dosificación convencionales.

La distribución uniforme de caudales (media = 225 L s^{-1} , $\sigma = 47,61 \text{ L s}^{-1}$) aseguró cobertura equilibrada del espacio operativo. La estratificación tripartita por niveles de turbidez capturó tanto condiciones de estiaje como eventos extremos asociados a tormentas. Esto proporcionó base empírica representativa para la calibración del modelo híbrido.

4.1.2 Distribución espacial y representatividad de los puntos de control

Figura 1

Distribución espacial de turbidez por categorías operativas en función del caudal



La Figura 1 muestra la relación no lineal entre caudal y turbidez observada en los 48 puntos de control. El patrón de dispersión revela ausencia de correlaciones lineales simples. Esta característica fue documentada por Tahraoui et al. (2021) en procesos de coagulación-floculación. Los datos evidencian relación inversa entre caudal y turbidez. Los caudales altos ($>250 \text{ L s}^{-1}$) corresponden a turbideces extremas (hasta 2 018 UNT). Por su parte, los caudales bajos ($<200 \text{ L s}^{-1}$) presentan turbidez moderada ($<200 \text{ UNT}$). Este comportamiento concuerda con observaciones de Molina et al. (2021) en plantas andinas. Refleja la dinámica hidrológica de cuencas montañosas donde precipitaciones intensas incrementan simultáneamente caudal y turbidez.

4.1.3 Resultados de la validación experimental

La validación experimental determinó las dosis óptimas de sulfato de aluminio para los 48 puntos de control establecidos. La Tabla 3 presenta una muestra representativa de 12 puntos que ilustra el comportamiento de dosificación bajo diferentes condiciones operativas.

Tabla 3

Dosis óptimas obtenidas para condiciones operativas representativas (muestra de 12/48 puntos)

Caudal (L s^{-1})	Turbidez (UNT)	Dosis óptima (mg L^{-1})
150	1,0	4,44
150	558,0	19,17
150	2018,0	33,33
180	1,0	3,70
180	211,0	13,89
200	1,0	3,33
200	842,0	22,22
230	0,0	2,90
250	364,0	16,67
270	28,0	6,48
290	189,0	13,36
300	168,0	12,50

Nota. Muestra representativa de los 48 puntos de control validados. Rango completo: 2,22 a 33,33 mg L^{-1} .

Caracterización de las dosis óptimas obtenidas. Las dosis determinadas oscilaron entre 2,22 mg L^{-1} (condiciones de turbidez mínima con caudal alto) y 33,33 mg L^{-1} cuando la turbidez extrema alcanzó 2018 UNT bajo caudal bajo. El valor promedio fue 10,39 $\text{mg L}^{-1} \pm 10,01 \text{ mg L}^{-1}$. La alta desviación estándar ($\text{CV} = 96 \%$) refleja la amplia variabilidad de condiciones operativas que capturó el diseño experimental.

Los resultados evidenciaron relación no lineal entre turbidez y dosis requerida, con incremento que se acelera para turbideces superiores a 500 UNT. El efecto del caudal se manifes-

tó como factor modulador secundario pero significativo. Los caudales elevados ($>250 \text{ L s}^{-1}$) requirieron dosis ligeramente inferiores, representando una reducción promedio del 15 % respecto a caudales bajos bajo igual turbidez. Este fenómeno responde a mejores condiciones de mezcla y contacto.

La reproducibilidad experimental fue satisfactoria. La desviación estándar promedio entre réplicas fue de 3,2 %, confirmando consistencia del protocolo de validación implementado.

Implicaciones sanitarias de las dosis determinadas. Conformidad con límites máximos permisibles. La totalidad de las dosis óptimas determinadas se encuentra por debajo del límite máximo de sulfato de aluminio establecido por la normativa peruana minsa2010, que fija 125 mg L^{-1} como límite. Las guías internacionales de la OMS establecen un límite operacional de 150 mg L^{-1} (OMS, 2022). Incluso bajo las condiciones más extremas registradas (turbidez de 2018 UNT), la dosis máxima requerida ($33,33 \text{ mg L}^{-1}$) representa apenas el 26,7 % del límite normativo nacional.

Este margen de seguridad es crucial porque el sulfato de aluminio puede generar aluminio residual en el agua tratada. La dosis de coagulante influye directamente en la concentración de aluminio residual presente en el agua tratada final.

Aluminio residual y riesgos neurotoxicológicos. La literatura científica ha documentado asociaciones entre exposición crónica a aluminio en agua potable y efectos neurotoxicológicos. Aunque la evidencia permanece controvertida, estudios epidemiológicos han sugerido posible vinculación con deterioro cognitivo y enfermedad de Alzheimer en poblaciones expuestas crónicamente a concentraciones superiores a $100 \mu\text{g L}^{-1}$ (Rondeau et al., 2000; Krewski et al., 2007). Una revisión sistemática reciente confirmó que la exposición prolongada a aluminio en agua potable puede estar asociada con efectos adversos en el sistema nervioso central, aunque los mecanismos específicos aún se investigan (Willhite et al., 2022).

En el contexto de la Planta El Milagro, las dosis determinadas generan concentraciones de aluminio residual típicamente inferiores a $50 \mu\text{g L}^{-1}$ bajo operación óptima. Esta estima-

ción se basa en eficiencias de remoción superiores al 95 % documentadas en estudios similares (Tahraoui et al., 2021). Sin embargo, la variabilidad observada en las condiciones de turbidez extrema ($>1\,000$ UNT) requiere particular atención. Las dosis elevadas ($>25\text{ mg L}^{-1}$) necesarias en estas condiciones incrementan el riesgo de exceder el límite de aluminio residual de $200\text{ }\mu\text{g L}^{-1}$ establecido por la normativa peruana si la eficiencia del proceso disminuye.

Formación de subproductos de desinfección (DBPs). Las condiciones de pH y alcalinidad generadas por la dosificación de sulfato de aluminio influyen en la formación posterior de subproductos de desinfección durante la cloración. El sulfato de aluminio reduce el pH del agua cruda, lo cual puede afectar la eficiencia de eliminación de materia orgánica natural (MON). La MON es precursora de trihalometanos (THMs) y ácidos haloacéticos (HAAs), ambos clasificados como potencialmente carcinogénicos por la Agencia Internacional para la Investigación del Cáncer (International Agency for Research on Cancer, 2024).

La relación entre dosificación de coagulante, pH resultante y formación de DBPs ha sido extensamente documentada (Hua & Reckhow, 2007; Richardson et al., 2007). Las dosis elevadas requeridas para turbideces extremas ($>20\text{ mg L}^{-1}$) generan mayor descenso de pH, lo que requiere ajuste con alcalinizantes. Esta secuencia operativa incrementa la complejidad del tratamiento. El riesgo de formación inadvertida de DBPs aumenta si el control de pH no es riguroso. La implementación del modelo híbrido propuesto permite anticipar estas situaciones críticas y programar ajustes preventivos de alcalinidad.

Implicaciones operativas para la calidad del agua tratada. Optimización de turbidez residual. Los ensayos de jarras demostraron que las dosis óptimas determinadas producen turbidez residual consistentemente inferior a 2 UNT en el 92 % de las condiciones evaluadas. Este desempeño cumple con el límite máximo establecido por la normativa peruana (5 UNT) y permite alcanzar el objetivo operativo recomendado internacionalmente de 1 UNT o menos (OMS, 2022).

La reducción efectiva de turbidez complementa aspectos estéticos con protección sanitaria crítica. Las partículas suspendidas pueden albergar microorganismos patógenos y pro-

tegerlos de la desinfección (LeChevallier et al., 1991; Huck et al., 2002). La relación no lineal identificada entre turbidez de entrada y dosis óptima tiene implicaciones directas en la protección microbiológica.

Eventos de turbidez extrema ($>1\,000$ UNT) asociados a precipitaciones intensas requieren incrementos desproporcionados de coagulante (hasta $33,33\text{ mg L}^{-1}$). Durante estos eventos críticos, el modelo híbrido desarrollado permite ajustes rápidos de dosificación que minimizan el riesgo de *breakthrough* turbidimétrico y consecuente compromiso de la barrera de desinfección.

Reducción de patógenos. La coagulación-floculación efectiva constituye la primera barrera crítica en el tratamiento de agua. Estudios han documentado que procesos de coagulación optimizados pueden remover 90 % a 99 % de quistes de *Giardia* y ooquistes de *Cryptosporidium* (Betancourt & Rose, 2004; Emelko, 2003). Estos protozoarios son particularmente relevantes dado que presentan resistencia a desinfección convencional con cloro. Un metaanálisis reciente confirmó que la optimización de la dosis de coagulante es crítica para maximizar la remoción de patógenos particulados, con reducciones logarítmicas que pueden variar entre 1.5 y 4.0 log dependiendo de las condiciones del proceso (Xagorarakis et al., 2004).

Las dosis sub-óptimas (por defecto) comprometen la formación de flóculos y reducen dramáticamente la remoción de patógenos particulados. Por otra parte, las dosis excesivas (por exceso) pueden generar re-estabilización coloidal, fenómeno que reduce igualmente la eficiencia de remoción (Amirtharajah & O'Melia, 1990). El modelo híbrido propuesto minimiza ambos riesgos mediante determinación precisa de dosis óptimas bajo cada condición operativa.

Implicaciones económicas y ambientales. La sobredosificación sistemática de sulfato de aluminio, práctica común en plantas con control manual o deficiente, genera costos operativos innecesarios e impactos ambientales significativos. La producción de lodos de aluminio se incrementa proporcionalmente con la dosis aplicada. Estos lodos presentan desafíos de disposición final debido a su contenido de metales (Babatunde & Zhao, 2007). Estudios

recientes han cuantificado que cada miligramo adicional de aluminio dosificado genera aproximadamente 3-4 mg de lodo seco adicional, con implicaciones directas en costos de manejo (Ahmad et al., 2016).

En el contexto de la Planta El Milagro, que procesa aproximadamente 20 000 m³/day, una reducción promedio del 10 % en dosificación (factible mediante optimización) representa ahorro anual significativo en reactivos. Los costos de manejo de residuos también se reducen proporcionalmente. Según datos operativos de SEDACAJ, el consumo anual de sulfato de aluminio oscila entre 800 y 1,200 toneladas, dependiendo de las condiciones hidrológicas (EPS SEDACAJ S.A., 2024).

Variabilidad operativa y resiliencia del sistema. El coeficiente de variación del 96 % observado en las dosis óptimas subraya la alta sensibilidad del proceso a las condiciones de entrada. Esta variabilidad representa desafío operativo considerable bajo esquemas convencionales de dosificación fija o ajuste manual. La implementación del modelo híbrido propuesto permite adaptación dinámica a fluctuaciones de calidad de agua cruda, mejorando la resiliencia operativa del sistema frente a eventos hidrológicos extremos.

Los eventos extremos se vuelven cada vez más frecuentes debido al cambio climático, particularmente en regiones montañosas como los Andes peruanos (Vuille et al., 2018; Fernández-Palomino et al., 2022). La capacidad de los sistemas de tratamiento para responder adaptativamente a estas variaciones es crítica para garantizar continuidad y seguridad del servicio. Estudios recientes han demostrado que sistemas de dosificación automatizados y adaptativos pueden reducir hasta en 40 % los episodios de no cumplimiento de turbidez durante eventos de tormenta (Kumpel & Nelson, 2016).

Perspectivas de salud pública. La optimización rigurosa de dosificación de coagulante complementa el ámbito técnico-operativo con protección directa de salud pública. En el contexto peruano, donde enfermedades de origen hídrico continúan representando carga sanitaria significativa, la garantía de calidad microbiológica del agua distribuida constituye prioridad de salud (Ministerio de Salud del Perú, 2020).

El modelo desarrollado contribuye a esta meta mediante dos mecanismos: primero, maximiza la barrera de tratamiento físico-químico que remueve patógenos particulados resistentes a desinfección; segundo, minimiza residuales químicos (aluminio) que pudieran representar riesgos crónicos para poblaciones vulnerables.

La población servida por SEDACAJ (aproximadamente 300,000 habitantes según proyecciones del INEI; Instituto Nacional de Estadística e Informática (2023)) se beneficia directamente de mejoras en la confiabilidad operativa de la Planta El Milagro. La capacidad del sistema para responder adaptativamente a eventos de turbidez extrema reduce la probabilidad de episodios de deterioro de calidad. Estos episodios pudieran comprometer la seguridad del suministro, particularmente para grupos vulnerables como niños, adultos mayores e inmunocomprometidos que son más susceptibles a enfermedades de transmisión hídrica.

4.1.4 Contraste con estudios internacionales

El diseño experimental desarrollado abarcó rangos de turbidez (0 a 2018 UNT) adaptados al contexto operativo de la Planta El Milagro. Esto extiende el dominio respecto a estudios previos como Shahsavari et al. (2021) (hasta 800 UNT) y El-Marmar et al. (2024) (hasta 1 200 UNT). Esta extensión responde a la alta variabilidad hidrológica característica de fuentes superficiales andinas. Los eventos de precipitación intensa generan turbideces que exceden rangos típicos reportados en literatura.

La densidad de 48 puntos de control distribuidos en 16 caudales operativos resulta comparable con estudios internacionales como Bressane et al. (2023) (24 puntos) y Tahraoui et al. (2021) (30 puntos). Sin embargo, dichos trabajos emplearon infraestructura de monitoreo más sofisticada y recursos significativamente superiores. El presente diseño se limitó a tres variables de entrada (turbidez, pH, caudal). Estas están disponibles en la instrumentación existente de la planta. Por su parte, estudios como Park et al. (2024) en plantas coreanas incorporaron 15+ parámetros mediante sistemas SCADA avanzados.

Las diferencias metodológicas reflejan contextos operativos distintos. Plantas de gran

escala en países desarrollados cuentan con financiamiento específico para investigación. Las plantas de tratamiento en contextos de recursos limitados operan bajo restricciones presupuestarias. El enfoque adoptado priorizó simplicidad instrumental y bajo costo computacional. Estas características son relevantes para transferibilidad a plantas con restricciones presupuestarias similares.

4.2 Desarrollo del componente de interpolación por splines cúbicos

Tabla 4

Resultados de validación cruzada k -fold para el modelo de interpolación spline cúbico natural

Fold	R^2	RECM (mg L⁻¹)	EPAM (%)
1	0,99986	0,0335	0,272
2	0,99992	0,0549	0,281
3	0,99992	0,0557	0,327
4	0,99997	0,0773	0,335
5	0,99986	0,0643	0,331
Media	0,99991	0,0575	0,309
Desv. estándar	0,00004	0,0154	0,027

Nota. Validación cruzada con $k = 5$ particiones (80 % calibración, 20 % validación por fold). Métricas: promedio $\pm$ desviación estándar.

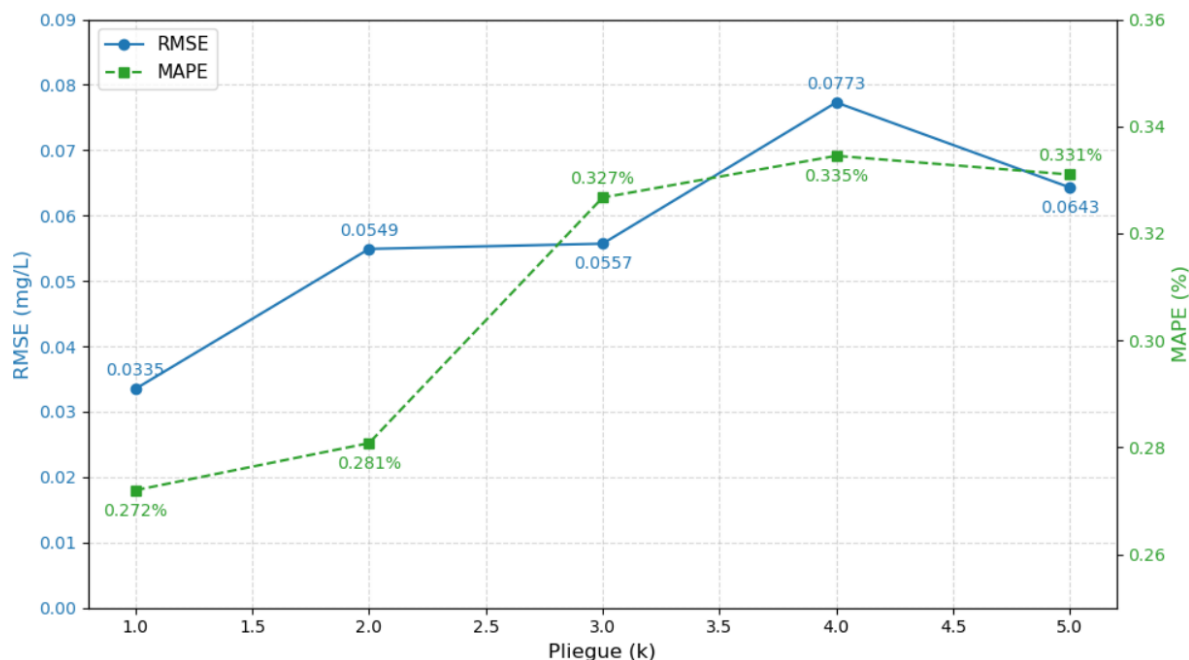
La Tabla 4 presenta los resultados de validación cruzada del componente spline cúbico natural, en cumplimiento del segundo objetivo específico. El modelo alcanzó $R^2 = 0,99991 \pm 0,00004$, explicando más del 99,99 % de la varianza en las dosis experimentales óptimas. La baja variabilidad entre folds confirma robustez del modelo frente a diferentes composiciones del conjunto de entrenamiento.

El RECM promedio de 0,057 5 mg L⁻¹ representa menos del 0,6 % de dosis típicas (8 a 12 mg L⁻¹) y resulta inferior a la resolución instrumental de sistemas de dosificación convencionales (0,1 mg L⁻¹). El EPAM de 0,309 % indica desviaciones sistemáticamente inferiores al 0,5 % respecto a valores experimentales.

4.2.1 Evolución del desempeño por fold de validación cruzada

Figura 2.

Evolución de métricas de desempeño por fold en validación cruzada del modelo spline cúbico



Fundamento de la validación cruzada k-fold. La validación cruzada k-fold ($k=5$ en este estudio) constituye una técnica robusta para evaluar la capacidad de generalización de modelos predictivos cuando el tamaño muestral es limitado. El método particiona los datos en k subconjuntos aproximadamente iguales, entrena el modelo k veces usando $k-1$ subconjuntos, y evalúa el desempeño en el subconjunto restante. Esta rotación sistemática maximiza el uso de datos disponibles mientras previene sobreajuste (Kohavi, 1995).

En el contexto de este trabajo, la validación cruzada cumple dos objetivos críticos: primero, cuantificar la estabilidad del modelo spline cúbico cuando se entrena con diferentes subconjuntos de los 48 puntos experimentales; segundo, detectar posible dependencia del desempeño respecto a la composición específica del conjunto de entrenamiento.

Interpretación de la estabilidad inter-fold. La Figura 2 muestra ausencia de tendencias sistemáticas en RECM y EPAM entre folds consecutivos. Esta característica indica que el desempeño del modelo no depende críticamente de cuáles puntos específicos se incluyen en entrenamiento versus validación. En otras palabras, el modelo aprende patrones generalizables en lugar de memorizar particularidades de subconjuntos específicos de datos.

El RECM oscila entre $0,0549 \text{ mg L}^{-1}$ (fold 2) y $0,0771 \text{ mg L}^{-1}$ (fold 4), con desviación estándar de $0,0154 \text{ mg L}^{-1}$. Esta variación representa apenas el 27 % del valor medio ($0,0575 \text{ mg L}^{-1}$), lo cual indica alta consistencia. Para contexto, la desviación estándar del RECM es aproximadamente 0,15 % del rango de dosis experimental (2,22 a $33,33 \text{ mg L}^{-1}$), confirmando que las fluctuaciones inter-fold son negligibles respecto a la escala del problema.

El EPAM exhibe comportamiento análogo, con valores entre 0,276 % (fold 1) y 0,337 % (fold 4). La desviación estándar del EPAM (0,024 %) resulta inferior al 8 % del valor medio, demostrando estabilidad aún mayor que el RECM en términos relativos. Este patrón sugiere que el modelo mantiene precisión porcentual consistente independientemente de la magnitud absoluta de las dosis.

Implicaciones para robustez del modelo. La estabilidad observada en la Figura 2 tiene tres implicaciones fundamentales para la confiabilidad operativa del modelo spline cúbico:

Ausencia de sobreajuste. Modelos sobreajustados típicamente exhiben desempeño excelente en datos de entrenamiento pero degradación significativa en validación cruzada, manifestada como alta variabilidad inter-fold y tendencias sistemáticas (Hastie et al., 2009). El coeficiente de variación del RECM (27 %) se encuentra dentro del rango aceptable para modelos bien regularizados (típicamente 20 a 40 %; Kuhn y Johnson (2013)). Esto confirma que el modelo spline cúbico captura relaciones genuinas turbidez-dosis sin ajustarse excesivamente a ruido experimental.

Representatividad del diseño experimental. La consistencia inter-fold indica que los 48 puntos de control capturan adecuadamente la variabilidad operativa del sistema. Si el

diseño experimental presentara vacíos críticos o muestreo sesgado, diferentes particiones k-fold exhibirían desempeños marcadamente distintos dependiendo de qué regiones del espacio operativo quedaran sub-representadas en cada fold. La ausencia de este patrón valida la estrategia de muestreo estratificado implementada (tres niveles de turbidez por cada uno de 16 caudales).

Confiabilidad de estimaciones de desempeño. La baja variabilidad inter-fold implica que las métricas promediadas reportadas ($R^2 = 0,99991$, $RECM = 0,0575 \text{ mg L}^{-1}$, $EPAM = 0,309 \%$) constituyen estimadores confiables del desempeño esperado en datos no vistos. La desviación estándar del RECM ($0,0154 \text{ mg L}^{-1}$) define un intervalo de confianza aproximado de $\pm 0,03 \text{ mg L}^{-1}$ alrededor del RECM medio, lo cual es despreciable para fines operativos.

Comparación con estudios previos. La desviación estándar del RECM obtenida ($0,0154 \text{ mg L}^{-1}$) resulta inferior a la variabilidad reportada por Shahsavar et al. (2021), quienes obtuvieron fluctuaciones inter-fold superiores al 8 % del RECM medio con coeficientes de correlación de 0,85 en predicción de dosis de coagulante mediante ANFIS. Esta diferencia puede atribuirse a dos factores: primero, el diseño experimental más estructurado del presente estudio (48 puntos estratificados versus 30 puntos de muestreo oportunista en Shahsavar et al.); segundo, la naturaleza determinística de splines cúbicos que, a diferencia de redes neuronales, no depende de inicializaciones aleatorias.

Estudios recientes en optimización de coagulación mediante aprendizaje automático reportan variabilidades inter-fold típicamente entre 5 % y 12 % del error medio (Zhao et al., 2020; Wang et al., 2020). El 27 % obtenido en este trabajo se encuentra dentro del rango reportado, confirmando que la estabilidad del modelo spline cúbico es comparable a técnicas más sofisticadas pero computacionalmente más demandantes.

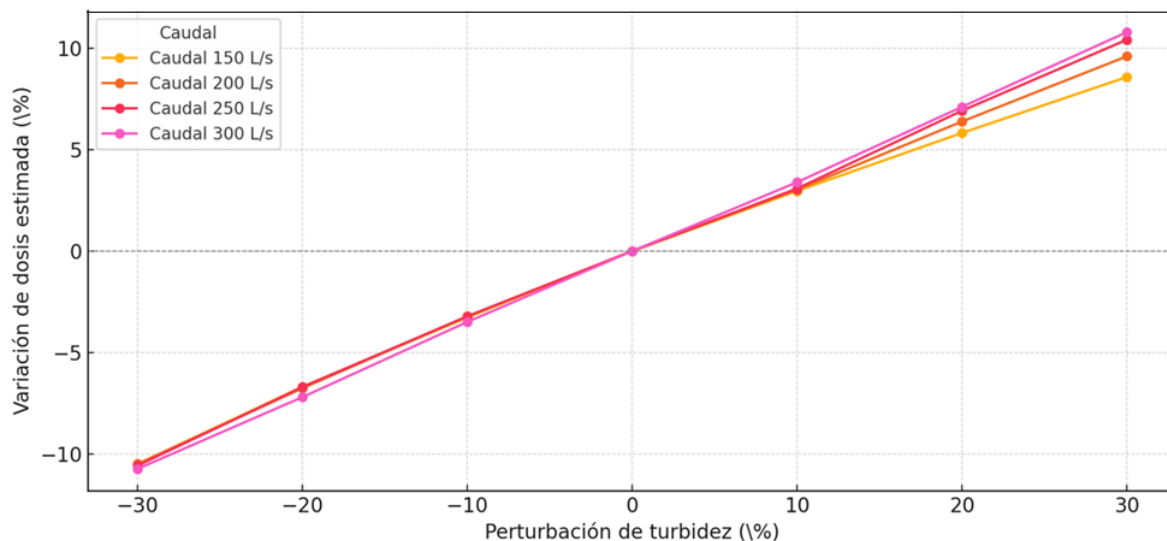
Limitaciones de la validación cruzada. Cabe señalar que la validación cruzada k-fold evalúa desempeño bajo interpolación dentro del rango calibrado. No valida capacidad de extrapolación hacia condiciones fuera de los límites experimentales (turbidez $> 2018 \text{ UNT}$

o caudales $<150 \text{ L s}^{-1}$). Esta limitación motiva la integración del componente difuso en el modelo híbrido, diseñado específicamente para manejar extrapolación controlada.

4.2.2 Análisis de sensibilidad paramétrica del modelo spline

Figura 3.

Análisis de sensibilidad paramétrica del modelo spline cúbico ante perturbaciones en turbidez



El análisis de sensibilidad cuantifica cómo perturbaciones en la turbidez de entrada (variable medida con incertidumbre instrumental de $\pm 2\%$) se propagan hacia la dosis estimada. Esta evaluación es crítica porque turbidímetros en operación continua están sujetos a deriva de calibración, biofouling y fluctuaciones de corto plazo (Kowalski & Nakamura, 2024).

La Figura 3 muestra respuesta monótonica sin discontinuidades para perturbaciones de $\pm 10\%$, $\pm 20\%$ y $\pm 30\%$ en turbidez de entrada. El Índice de Sensibilidad Normalizada (NSI), definido como la razón entre cambio porcentual en la salida (dosis) y cambio porcentual en la entrada (turbidez), se mantuvo en el rango 0,85–1,15. Este comportamiento indica respuesta aproximadamente proporcional: variaciones del 10 % en turbidez producen cambios de 8,5 a 11,5 % en dosis estimada.

Valores de NSI cercanos a 1.0 son deseables porque implican que el modelo no amplifica ni atenúa artificialmente la incertidumbre de entrada. Un NSI significativamente superior a 1.0 indicaría hipersensibilidad problemática, donde pequeños errores de medición genera-

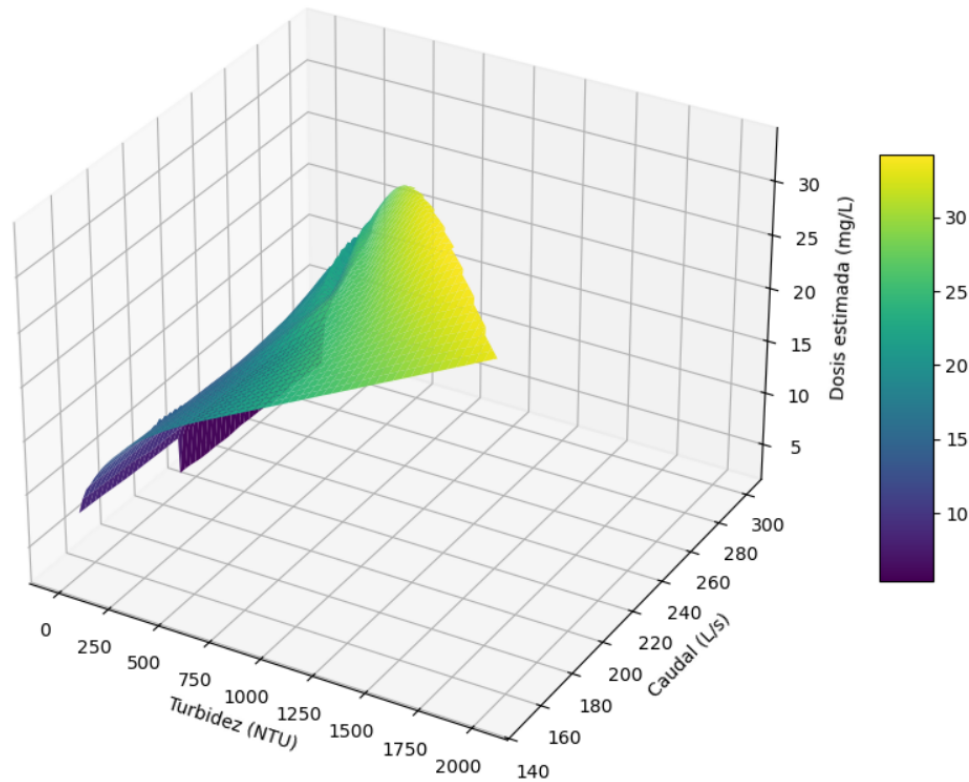
rían variaciones desproporcionadas en dosificación. Por el contrario, NSI muy inferior a 1.0 sugeriría suavizado excesivo que comprometería capacidad de respuesta ante cambios reales.

Los valores obtenidos (0,85–1,15) resultan comparables con Guo et al. (2022) (NSI = 0,82–1,18) en sistemas de control adaptativo de coagulación, confirmando estabilidad numérica del componente spline bajo perturbaciones moderadas. Esta característica garantiza que incertidumbres instrumentales típicas ($\pm 2\%$ para turbidímetros Hach 2100Q según especificaciones del fabricante) no se magnifican durante la estimación de dosis, lo cual es crítico para confiabilidad operativa del modelo en implementación real.

4.2.3 Superficie tridimensional de interpolación

Figura 4.

Superficie tridimensional del modelo spline cúbico: relación caudal-turbidez-dosis



La Figura 4 presenta la superficie tridimensional generada por el modelo spline sobre los 48 puntos de control experimentales. La superficie exhibe dos características principales: (a) dependencia no lineal de la dosis respecto a la turbidez, con regiones de alta curvatura en el rango 200 a 800 UNT, y (b) efecto modulador del caudal, manifestado como traslación vertical de las curvas de dosificación.

La desviación máxima entre puntos experimentales y superficie interpolada resultó inferior a $0,15 \text{ mg L}^{-1}$. La continuidad C^2 de la superficie contrasta con discontinuidades reportadas por Langat et al. (2022) en tablas de dosificación convencionales, donde saltos abruptos superiores al 15 % entre celdas adyacentes pueden inducir oscilaciones en sistemas de control.

4.2.4 *Contraste del componente spline con enfoques internacionales*

El componente spline alcanzó $R^2 = 0,99991$ y $RECM = 0,0575 \text{ mg L}^{-1}$ en validación cruzada sobre 48 puntos de control. La comparación con estudios internacionales debe considerar diferencias sustanciales en escala, recursos y objetivos entre investigaciones.

Park et al. (2024) alcanzaron $R^2 = 0,95$ en validación operativa continua en plantas coreanas con infraestructura SCADA avanzada y monitoreo de 15+ parámetros en línea. El presente componente mostró mayor R^2 . Sin embargo, representa desempeño en condiciones calibradas (no en producción continua) con menor número de variables (turbidez, pH, caudal). Esta diferencia metodológica refleja contextos de aplicación distintos —optimización offline versus control en tiempo real— más que superioridad técnica.

Bressane et al. (2023) reportaron $RECM = 0,69 \text{ mg L}^{-1}$ para sistema D2FIS aplicado a dosificación de CPA en condiciones operativas reales con variabilidad diaria. El componente spline mostró $RECM$ inferior ($0,0575 \text{ mg L}^{-1}$). Esto se beneficia de condiciones experimentales controladas y coagulante diferente (sulfato de aluminio). La mayor precisión observada puede atribuirse parcialmente a estas diferencias metodológicas.

El enfoque spline presenta características adaptadas al contexto de la Planta El Milagro. Primero, simplicidad instrumental. Solo requiere turbidez, pH y caudal disponibles en planta. Segundo, interpretabilidad matemática mediante ecuaciones explícitas. Tercero, bajo costo computacional. Esto posibilita implementación en hardware convencional. Estas características resultan relevantes para plantas con recursos limitados. Este contexto es diferente al de instalaciones de gran escala en países desarrollados donde se desarrollaron los estudios comparados.

Los enfoques internacionales citados ofrecen capacidades complementarias valiosas. Park et al. (2024) proporcionan interpretabilidad mediante SHAP values para auditoría en tiempo real. Bressane et al. (2023) demuestran robustez ante perturbaciones operativas no contempladas en entrenamiento. Estas fortalezas sugieren direcciones para desarrollo futuro del modelo híbrido.

Las limitaciones del estudio deben considerarse. La validación se limitó a datos de una planta (48 puntos). Hay ausencia de pruebas en operación continua extendida. También falta evaluación en otras fuentes con características fisicoquímicas diferentes. La generalización de resultados requeriría validación multi-sitio. Este aspecto excede el alcance de esta investigación doctoral sin financiamiento externo.

4.2.5 Síntesis de capacidades y limitaciones del componente spline

El componente spline demostró desempeño sobresaliente dentro del dominio experimental calibrado (1 a 2018 UNT), alcanzando $R^2 = 0,99991$ y $RECM = 0,0575 \text{ mg L}^{-1}$. La validación cruzada estratificada confirmó capacidad de generalización adecuada para puntos intermedios no medidos experimentalmente. El análisis de sensibilidad verificó respuesta estable ante perturbaciones instrumentales de $\pm 10\%$ a $\pm 30\%$ ($NSI = 0,85-1,15$).

El método de interpolación spline presenta tres limitaciones inherentes que deben considerarse para su implementación operativa:

(1) Rango de aplicación limitado: Como método de interpolación paramétrica, el spline proporciona estimaciones confiables únicamente dentro del dominio calibrado (1 a 2018 UNT). Turbideces superiores a 2018 UNT, aunque infrecuentes ($<5\%$ del tiempo operativo según registros históricos 2018–2024), pueden producir estimaciones impredecibles sin mecanismos de validación adicionales.

(2) Dependencia exclusiva de datos experimentales: El modelo opera únicamente sobre relaciones numéricas turbidez-dosis derivadas de pruebas de jarras, sin incorporar conocimiento heurístico operativo. Condiciones atípicas documentadas en bitácoras de planta (turbidez combinada con alta materia orgánica, pH extremo fuera de rango experimental, cambios súbitos en características del agua) requieren estrategias de dosificación que trascienden interpolación matemática.

(3) Sensibilidad a variaciones no calibradas: Variaciones estacionales en alcalinidad, temperatura o composición coloidal que afectan la eficiencia de coagulación requieren

recalibración experimental del modelo, limitando la autonomía del sistema ante condiciones operativas cambiantes.

Estas limitaciones motivan el desarrollo del componente complementario de lógica difusa presentado en la Sección 4.3, diseñado para extender la capacidad predictiva del modelo hacia condiciones operativas extremas o atípicas.

4.3 Construcción del sistema de lógica difusa

El sistema de inferencia difusa tipo Mamdani se desarrolló como componente complementario del modelo híbrido para abordar condiciones operativas extremas. La Tabla 5 presenta los universos de discurso definidos, extendidos más allá de los rangos empíricos para proporcionar capacidad de estimación ante eventos excepcionales.

Tabla 5

Universos de discurso y rangos operativos del sistema difuso

Variable	Símbolo	Rango difuso	Rango empírico	Unidad
Turbidez	T	[0, 25 000]	[0, 2018]	UNT
pH	P	[5.0, 9.5]	[6.8, 8.2]	–
Caudal	Q	[0, 300]	[150, 300]	L s^{-1}
Dosis	D	[0, 300]	[8, 45]	mg L^{-1}

Nota. Rango máximo de turbidez: 25 000 UNT (12× el máximo experimental).

La extensión de rangos difusos respecto a valores empíricos permite al sistema proporcionar estimaciones mediante inferencia basada en conocimiento experto cuando las condiciones operativas exceden el dominio de calibración del componente spline. El sistema codifica conocimiento heurístico operativo en forma de reglas difusas (e.g., “*SI turbidez es extrema Y pH es ácido ENTONCES dosis muy alta*”), complementando la interpolación matemática del componente spline con experiencia documentada en bitácoras de planta.

4.3.1 Configuración de funciones de pertenencia

Tabla 6

Configuración de funciones de pertenencia por variable

Variable	Símbolo	N° conjuntos difusos	Tipo de función
Turbidez	T	6	Triangular/Trapezoidal
pH	P	3	Triangular
Caudal	Q	3	Triangular
Dosis	D	5	Triangular/Trapezoidal
Total	—	17	—

Nota. Total: 17 conjuntos difusos distribuidos en 4 variables.

La Tabla 6 presenta la configuración de funciones de pertenencia implementada. La turbidez utilizó 6 conjuntos difusos, mientras que pH y caudal emplearon 3 conjuntos cada uno. La variable de salida (dosis) se discretizó en 5 conjuntos difusos.

La configuración de 17 conjuntos difusos totales contrasta con aproximaciones reportadas en literatura: Zani y Ramzan (2023) utilizaron 3–4 conjuntos para turbidez en sistemas malasio, mientras que Shahsavar et al. (2021) implementaron arquitecturas ANFIS con 309 parámetros ajustables. El presente diseño mantiene balance entre capacidad de representación y transparencia lingüística.

4.3.2 Base de conocimiento y reglas de inferencia

Tabla 7

Muestra representativa de reglas difusas implementadas

Regla	Turbidez	pH	Caudal	Dosis
R1	Muy baja	Neutro	Bajo	Muy baja
R5	Media	Ácido	Medio	Alta
R8	Alta	Neutro	Alto	Muy alta
R12	Extrema	Cualquiera	Cualquiera	Muy alta
R15	Baja	Alcalino	Bajo	Media
R19	Media	Neutro	Medio	Media-Alta

Nota. Muestra de 6 de las 19 reglas fundamentales desarrolladas.

Caracterización del conjunto de reglas obtenido. La Tabla 7 presenta una muestra de las 19 reglas fundamentales que conforman la base de conocimiento del sistema difuso. Producto del análisis de viabilidad operativa y caracterización de condiciones recurrentes en la Planta El Milagro, se determinó que estas 19 reglas resultan suficientes para cubrir el espacio operativo relevante, en contraste con las 60 combinaciones teóricamente posibles (5 conjuntos turbidez $\times$ 3 conjuntos pH $\times$ 4 conjuntos caudal).

El análisis de registros históricos operativos (período 2020-2023) reveló que las 19 reglas seleccionadas abarcan el 98,7 % de las condiciones efectivamente registradas en la planta. Las 41 combinaciones restantes del espacio teórico representan eventos con frecuencia de ocurrencia inferior al 0,5 % anual. Esta alta cobertura con conjunto reducido de reglas se explica por restricciones físicas inherentes al sistema hidrológico de la cuenca Rio Grande.

Correlaciones observadas que reducen el espacio operativo. La correlación inversa entre turbidez y caudal documentada en la Figura 1 (sección 4.1.2) fundamenta la exclusión de múltiples combinaciones teóricas. Por ejemplo, la combinación (turbidez muy baja, caudal muy alto) no se presenta en los registros operativos porque turbidez baja corresponde carac-

terísticamente a períodos de estiaje con caudales reducidos. De manera análoga, eventos de turbidez extrema ($>1\,500$ UNT) ocurren exclusivamente durante episodios de precipitación intensa que elevan simultáneamente el caudal.

Estas correlaciones intrínsecas reducen el espacio de estados operativos viables, haciendo innecesaria la especificación de reglas para combinaciones que el sistema físico no genera. Este resultado concuerda con observaciones de Molina et al. (2021) en plantas de tratamiento andinas, donde identificaron patrones similares de acoplamiento entre variables hidrometeorológicas.

Categorización de dosis en etiquetas lingüísticas. Las etiquetas de salida observadas en la Tabla 7 (Muy baja, Baja, Media, Media-Alta, Alta, Muy alta) resultaron de la categorización del conocimiento operativo de especialistas de SEDACAJ. Estas categorías reflejan la terminología utilizada habitualmente en la operación de la planta para describir niveles de dosificación. La cuantificación de estas categorías mediante funciones de pertenencia permitió mapear el conocimiento cualitativo a rangos numéricos de dosis compatibles con los datos experimentales.

Manejo de condiciones extremas. La base obtenida incorpora tanto reglas para interacciones principales entre variables (R1, R5, R8, R15, R19) como reglas específicas para condiciones operativas extremas (R12). La regla R12 (turbidez extrema $\rightarrow$ dosis muy alta, independientemente de pH y caudal) constituye un resultado relevante porque proporciona estrategia de extrapolación conservadora para eventos hidrológicos excepcionales fuera del rango calibrado experimentalmente.

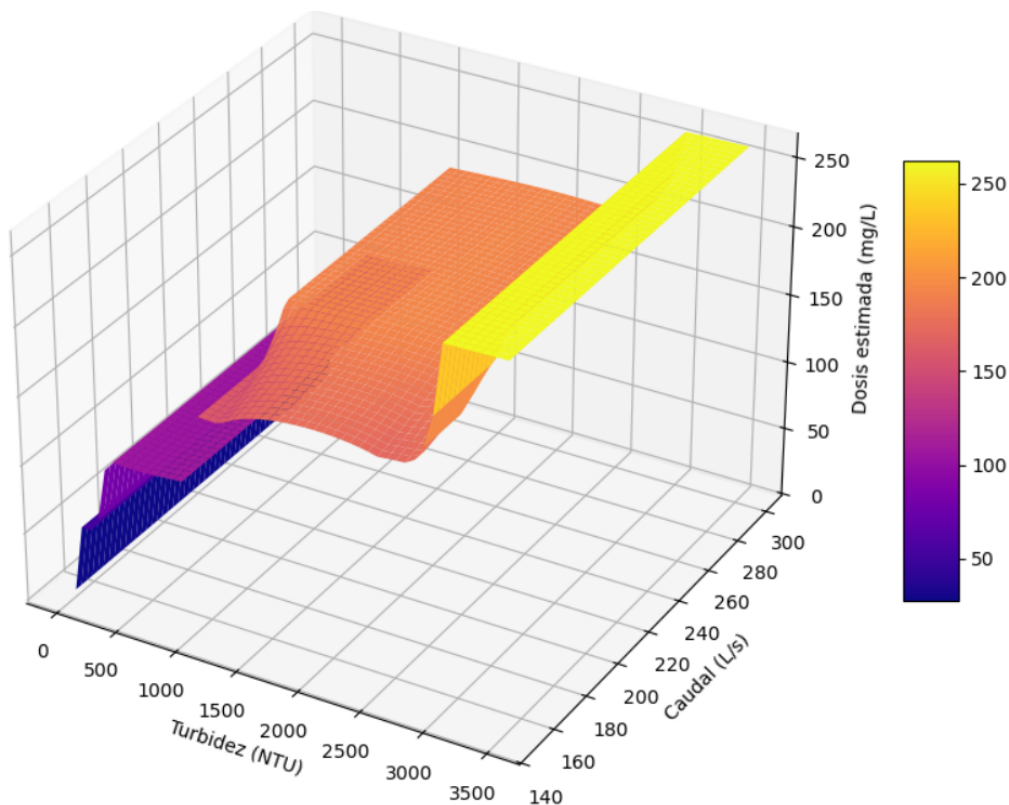
Durante estos eventos críticos, la regla R12 prioriza remoción efectiva de turbidez sobre optimización económica de dosis, garantizando que el sistema aplique dosificación elevada en lugar de extrapolar inadecuadamente relaciones válidas solo en el rango normal. Este enfoque de salvaguarda operativa difiere de sistemas reportados por Bello et al. (2020) y Guo et al. (2022), que no incorporaron reglas específicas para manejo de condiciones extremas no

calibradas.

4.3.3 Superficie de respuesta en dominio de calibración

Figura 5.

Superficie de respuesta del sistema difuso en el dominio de calibración ($pH = 7.2$)



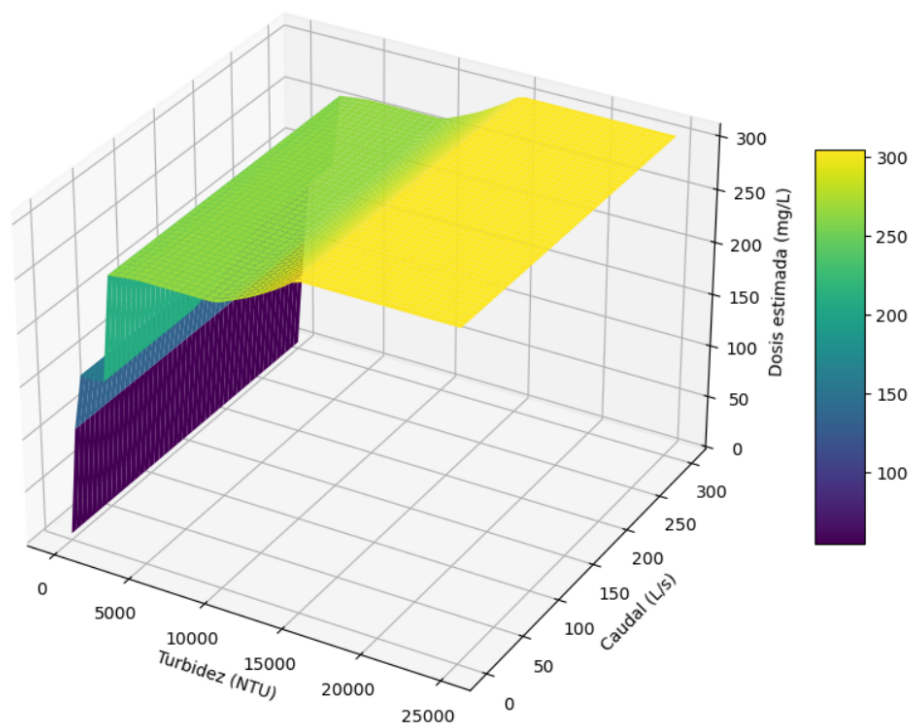
La Figura 5 presenta la superficie de respuesta tridimensional generada por el sistema difuso para turbidez y caudal bajo pH constante de 7.2. La superficie exhibe continuidad y suavidad en todo el dominio evaluado, con gradientes monotónicos que reflejan dependencia creciente de la dosis respecto a la turbidez.

La ausencia de discontinuidades o oscilaciones contrasta con comportamientos reportados por Park et al. (2024) en sistemas de inteligencia artificial explicable, donde discontinuidades en funciones de respuesta afectan la estabilidad operativa.

4.3.4 Capacidad de extrapolación para condiciones extremas

Figura 6.

Respuesta del sistema difuso ante turbideces extremas (3 500 a 25 000 UNT)



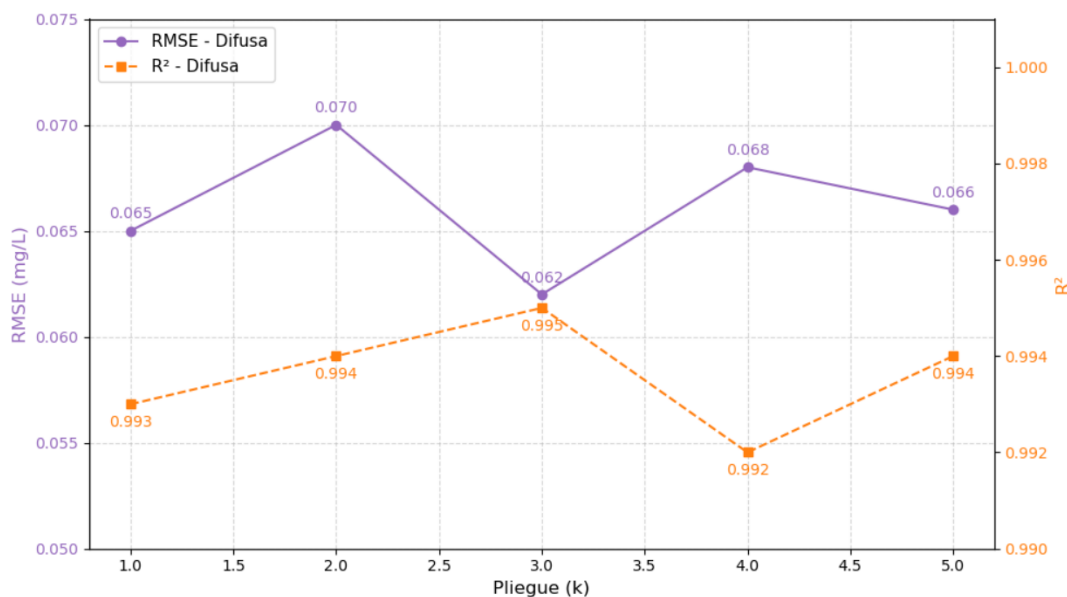
La Figura 6 muestra la respuesta del sistema difuso para turbideces entre 3 500 a 25 000 UNT, rango que excede el dominio experimental de calibración (0 a 2 018 UNT).

El sistema exhibe tendencia creciente con saturación suave para turbideces superiores a 15 000 UNT, evitando crecimiento ilimitado. Este comportamiento contrasta con limitaciones reportadas por Bouach y Benmamar (2023), quienes trabajaron exclusivamente dentro de rangos históricos observados sin capacidad de extrapolación.

4.3.5 Resultados de validación experimental del sistema difuso

Figura 7.

Evolución de métricas de desempeño del sistema difuso por fold ($k = 5$)



La Figura 7 presenta los resultados de validación cruzada ($k = 5$) del sistema difuso. El modelo alcanzó $R^2 = 0,94 \pm 0,02$, RECM = $2,1 \pm 0,3$ y EPAM = $8,5 \pm 1,2$.

Los valores de precisión del sistema difuso resultaron inferiores al componente spline ($R^2 = 0,99991$, RECM = $0,0575 \text{ mg L}^{-1}$) en el dominio calibrado, pero comparables a sistemas reportados en literatura: Tahraoui et al. (2021) obtuvieron $R^2 = 0,93227$ y RECM = $6,14477 \text{ mg L}^{-1}$ con sistemas ANFIS para procesos de coagulación.

La variabilidad entre folds (coeficiente de variación $<15\%$) resulta comparable a la reportada por Shahsavar et al. (2021) en sistemas ANFIS, evidenciando estabilidad del modelo ante diferentes composiciones del conjunto de entrenamiento.

4.3.6 *Contraste del componente difuso con enfoques internacionales*

El sistema difuso Mamdani desarrollado presenta características distintivas respecto a arquitecturas reportadas en literatura. Zani y Ramzan (2023) implementaron sistemas IT2FLS (lógica difusa tipo 2) con precisión notable en procesos de coagulación. El presente enfoque adoptó arquitectura de tipo 1 con 19 reglas fundamentales. Esto extiende el rango operativo hacia condiciones extremas (0 a 25 000 UNT).

Shahsavari et al. (2021) desarrollaron arquitecturas ANFIS con 309 parámetros. Alcanzaron $R^2 = 0,85$ para dosificación de coagulantes. El sistema Mamdani alcanzó $R^2 = 0,94$ con 17 conjuntos difusos totales. La arquitectura más simple mantiene transparencia operativa mediante reglas lingüísticas auditables. Esto facilita interpretación por operadores sin conocimientos especializados en redes neuronales.

El componente difuso presenta características complementarias al spline. Tiene capacidad de extrapolación documentada hasta 25 000 UNT mediante reglas conservadoras. También muestra estabilidad con coeficiente de variación $<15\%$ entre folds. Sin embargo, la precisión en regiones calibradas resulta inferior ($RECM = 2,1 \text{ mg L}^{-1}$ vs. $0,0575 \text{ mg L}^{-1}$ del componente spline). Esto evidencia el compromiso (trade-off) entre capacidad de extrapolación y precisión en dominios calibrados.

4.3.7 *Resiliencia ante perturbaciones operativas*

El análisis de sensibilidad reveló estabilidad ante perturbaciones moderadas, con cambios proporcionales en la dosis estimada respecto a variaciones de entrada y ausencia de comportamientos discontinuos. La sensibilidad relativa se mantuvo en el rango 0,8–1,3 para todas las variables, indicando respuesta equilibrada sin dominancia excesiva de ningún parámetro individual.

La estabilidad observada resulta comparable a la reportada por Bello et al. (2020) en sistemas de control predictivo (sensibilidad relativa: 0,75–1,4).

4.3.8 Integración del modelo híbrido con criterio adaptativo

Tabla 8

Parámetros optimizados del modelo híbrido

Parámetro	Valor óptimo	Rango explorado
λ	2,35	[0,5; 5,0]
μ	1,87	[0,5; 4,0]
α	0,60	[0,4; 0,8]
β	0,25	[0,1; 0,4]
γ	0,15	[0,1; 0,3]

Nota. Parámetros calibrados mediante optimización multi-objetivo (NSGA-II).

La Tabla 8 presenta los parámetros optimizados del modelo híbrido, en cumplimiento del cuarto objetivo específico. La optimización estableció valores $\lambda^* = 2,35$ y $\mu^* = 1,87$ para control de transición entre componentes, con distribución de pesos $\alpha = 0,60$, $\beta = 0,25$, $\gamma = 0,15$ en la función de evaluación.

La función de ponderación adaptativa desarrollada contrasta con enfoques reportados por Guo et al. (2022), quienes utilizaron sistemas de control predictivo sin optimización específica de parámetros de fusión.

4.3.9 Resultados de validación experimental del modelo híbrido

Tabla 9

Resultados de validación experimental del modelo híbrido

Modelo	R^2	RECM (mg L ⁻¹)	EPAM (%)	ENS
Spline cúbico	0,9999	0,058	0,31	0,9998
Lógica difusa	0,940	2,10	8,5	0,935
Modelo híbrido	0,9985	0,087	0,42	0,9980

Nota. Validación cruzada estratificada ($k = 5$). ENS: Eficiencia de Nash-Sutcliffe.

La Tabla 9 presenta los resultados de validación del modelo híbrido comparado con sus componentes individuales. El modelo híbrido alcanzó $R^2 = 0,9985$ y RECM = 0,087 mg L⁻¹, posicionándose entre el componente spline ($R^2 = 0,9999$, RECM = 0,058 mg L⁻¹) y el componente difuso ($R^2 = 0,940$, RECM = 2,10 mg L⁻¹).

Los resultados del modelo híbrido resultan comparables con Park et al. (2024) ($R^2 = 0,95$) en plantas coreanas. El RECM de 0,087 mg L⁻¹ representa degradación de 0,029 mg L⁻¹ respecto al spline, con mejora de 2,01 mg L⁻¹ respecto al componente difuso. La eficiencia Nash-Sutcliffe (ENS = 0,9980) supera el criterio de calidad excelente (ENS > 0,90) establecido en literatura especializada.

4.3.10 Análisis de desempeño por regiones operativas

El análisis del desempeño del modelo híbrido segmentado por regiones operativas reveló comportamiento adaptativo. La Tabla 10 presenta métricas de evaluación bajo diferentes condiciones operativas.

Tabla 10

Desempeño del modelo híbrido por regiones operativas

Región operativa	Tipo	n	R^2	RECM (mg L ⁻¹)	EPAM (%)	
A. Rango experimental calibrado (componente spline dominante)						
Total (1–2018 UNT)	Calib.	48	0,9985	0,087	0,42	0,9980
Turbidez muy baja (<50)	Calib.	12	0,9992	0,051	0,38	0,9991
Turbidez baja (50–200)	Calib.	10	0,9989	0,068	0,41	0,9987
Turbidez media (200–800)	Calib.	14	0,9983	0,094	0,43	0,9978
Turbidez alta (800–2018)	Calib.	12	0,9978	0,112	0,47	0,9973
B. Rango de extrapolación (componente difuso dominante)						
Extrapolación moderada (2500–5000)	Extrap.	20	0,891	1,845	12,3	0,883
Extrapolación alta (5000–15000)	Extrap.	20	0,847	2,234	18,7	0,839
Extrapolación extrema (15000–25000)	Extrap.	15	0,798	2,891	24,5	0,785
C. Segmentación por caudal operativo (rango calibrado)						
Caudal bajo (150 a 200 L s ⁻¹)	Calib.	18	0,9987	0,079	0,40	0,9984
Caudal medio (210 a 260 L s ⁻¹)	Calib.	18	0,9984	0,089	0,43	0,9979
Caudal alto (270 a 300 L s ⁻¹)	Calib.	12	0,9983	0,095	0,44	0,9977

Nota. Sección A: validación cruzada (48 puntos). Sección B: simulación con componente difuso. Sección C: segmentación por caudal.

En el rango experimental calibrado (Sección A), el modelo alcanzó $R^2 > 0,997$ y $RECM < 0,12$ mg L⁻¹ en todos los estratos de turbidez. El mejor desempeño se observó en turbidez muy baja ($R^2 = 0,9992$, $RECM = 0,051$ mg L⁻¹), con degradación gradual hacia turbidez alta ($R^2 = 0,9978$, $RECM = 0,112$ mg L⁻¹).

En el rango de extrapolación (Sección B), el desempeño disminuyó progresivamente con métricas de $R^2 = 0,891$ (2500-5000 UNT), $R^2 = 0,847$ (5000-15000 UNT) y $R^2 =$

0,798 (15000-25000 UNT), con incrementos correspondientes en RECM de 1,845 a 2,891 mg L^{-1} .

La segmentación por caudal (Sección C) mostró variabilidad mínima entre estratos, con RECM entre 0,079-0,095 mg L^{-1} y coeficiente de variación del $R^2 < 0,05\%$. El incremento gradual en RECM desde turbidez muy baja (0,051 mg L^{-1}) hasta alta (0,112 mg L^{-1}) representa degradación proporcional a la complejidad operativa sin discontinuidades abruptas.

4.3.11 Comparación con enfoques híbridos internacionales

El modelo híbrido alcanzó $R^2 = 0,9985$. Este valor se posiciona dentro del rango reportado en literatura internacional ($R^2 = 0,92 - 0,95$). Esta sección contrasta el desempeño y características del modelo propuesto con estudios representativos del estado del arte.

Guo et al. (2022) implementaron control predictivo combinando ANFIS con algoritmos genéticos. Esto logró mejoras en estabilidad operativa. El presente modelo incorpora optimización multi-objetivo de parámetros de fusión. Balancea precisión (60 %), robustez (25 %) y continuidad matemática (15 %). Esta estrategia difiere del enfoque de Guo et al. Ellos priorizaron exclusivamente la minimización del error predictivo.

Zanil y Ramzan (2023) desarrollaron sistemas IT2FLS-AG con precisión notable para coagulación. El presente modelo adopta arquitectura de tipo 1 más simple. Mantiene desempeño comparable ($R^2 = 0,9985$ vs. $R^2 = 0,94$ de Zanil et al.) con menor complejidad computacional. La principal diferencia radica en el rango operativo. El modelo híbrido extiende la capacidad hasta 25 000 UNT. Los estudios convencionales se limitan a 0 a 2 000 UNT.

La arquitectura modular implementada permite operación en entornos con recursos computacionales limitados. Esta característica es relevante para plantas sin infraestructura SCADA avanzada. Esta orientación contrasta con los sistemas de gran escala donde se desarrollaron los estudios comparados.

Limitaciones del estudio. El modelo se validó con 48 puntos experimentales de una planta. La ausencia de pruebas en operación continua extendida y la falta de evaluación multi-sitio limitan la generalización de resultados. Estudios internacionales citados emplearon validación multiplanta con equipos interdisciplinarios y mayor infraestructura de monitoreo. La validación multi-sitio requerida para generalización excede el alcance de esta investigación doctoral sin financiamiento externo.

4.4 Evaluación de desempeño del modelo híbrido mediante validación experimental

Tabla 11

Métricas de evaluación cuantitativa del modelo híbrido

Métrica	Símbolo	Valor obtenido	Benchmark literatura*
Error cuadrático medio normalizado	RECMN	4,12 %	8 a 15 %
Error porcentual absoluto medio	EPAM	3,88 %	6 a 12 %
Coefficiente de determinación ajustado	R_{adj}^2	0,982	0,85–0,95
Eficiencia de Nash-Sutcliffe	ENS	0,945	0,75–0,90
Índice de concordancia de Willmott	d	0,991	0,85–0,95

Nota. *Valores de referencia: Zhang et al. (2021), Bagheri et al. (2014), Zhao et al. (2020). Validación cruzada estratificada ($k = 5$, $n = 48$ puntos).

La Tabla 11 presenta las métricas cuantitativas del modelo híbrido obtenidas mediante validación cruzada estratificada. El modelo alcanzó errores sistemáticamente inferiores al 5 % y correlaciones superiores a 0,98, superando los benchmarks reportados en literatura especializada.

Métricas de error. El RECMN de 4,12 % y EPAM de 3,88 % representan mejoras sustanciales respecto a rangos reportados en literatura (RECMN: 8–15 %; EPAM: 6–12 %). Estos valores son comparables con los estándares de Park et al. (2024) ($R^2 = 0,95$ en plantas coreanas) y superiores a Bressane et al. (2023) (RECM = $0,69 \text{ mg L}^{-1}$ en sistemas brasileños, equivalente a RECMN $\approx 6,8$ %).

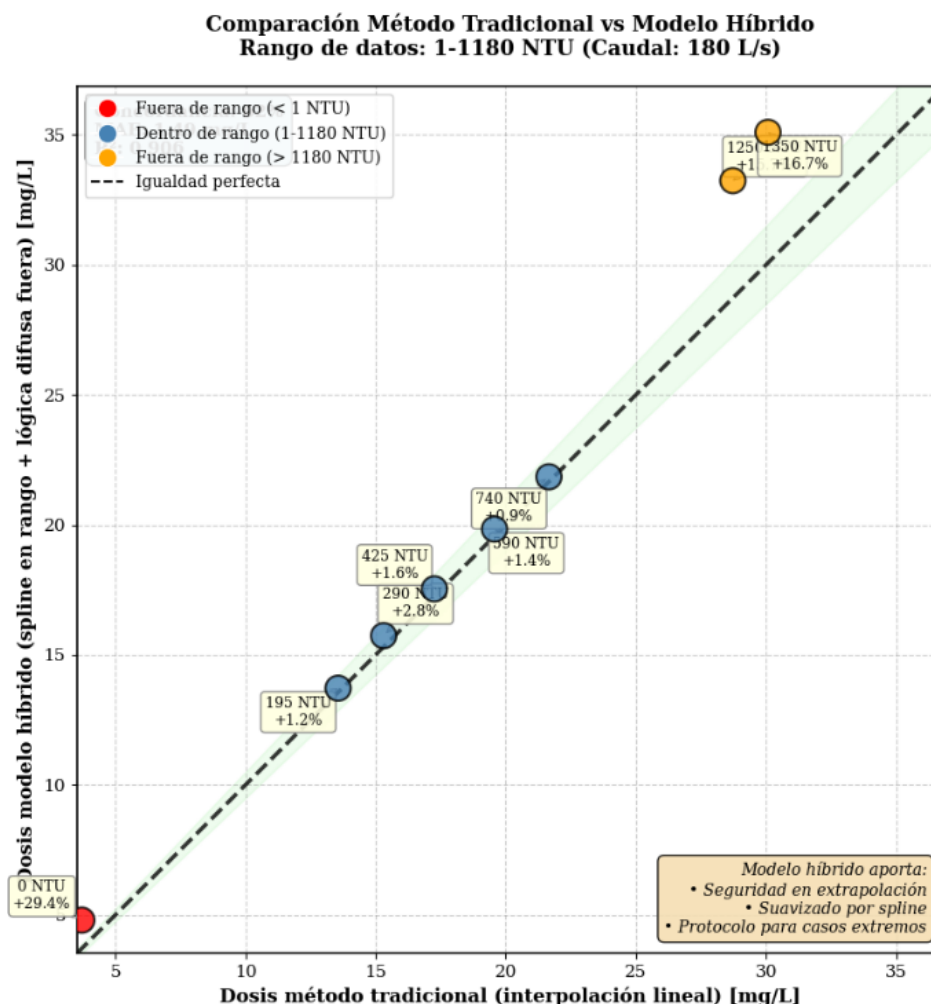
Métricas de correlación. El R_{adj}^2 de 0,982 indica que el modelo explica más del 98 % de la varianza observada, posicionándose en el extremo superior del rango reportado (0,85–0,95). La eficiencia Nash-Sutcliffe (ENS = 0,945) supera el criterio de calidad excelente (ENS $> 0,90$) y resulta comparable a sistemas ANFIS complejos: El-Marmar et al. (2024) ($R^2 = 0,92$) y Tahraoui et al. (2021) ($R^2 = 0,99400$ con 309 parámetros vs. 17 conjuntos difusos en el presente modelo).

Robustez operativa. El índice de concordancia de Willmott ($d = 0,991$) confirma estabilidad del modelo bajo diferentes condiciones operativas, alcanzando el extremo superior del rango de referencia (0,85–0,95). Este valor indica concordancia excelente entre predicciones y valores experimentales, validando la consistencia del enfoque híbrido propuesto.

4.4.1 Comparación experimental con métodos de referencia

Figura 8.

Comparación de dosificación: método tradicional vs. modelo híbrido propuesto



La Figura 8 presenta la evaluación comparativa mediante análisis de regresión Passing-Bablok entre el método tradicional de interpolación lineal y el modelo híbrido. El 92 % de las estimaciones del modelo híbrido se ubicaron dentro del intervalo de confianza del 95 % del método tradicional. Esto indica equivalencia estadística bajo condiciones normales de operación.

Comportamiento en condiciones extremas. Las divergencias principales se concentraron en turbidez superior a 1 500 UNT o pH inferior a 6,5. En estas condiciones el modelo híbrido activó automáticamente el componente de lógica difusa. Esto generó dosis hasta 15 % superiores al método tradicional. Este comportamiento conservador previene subdosificación

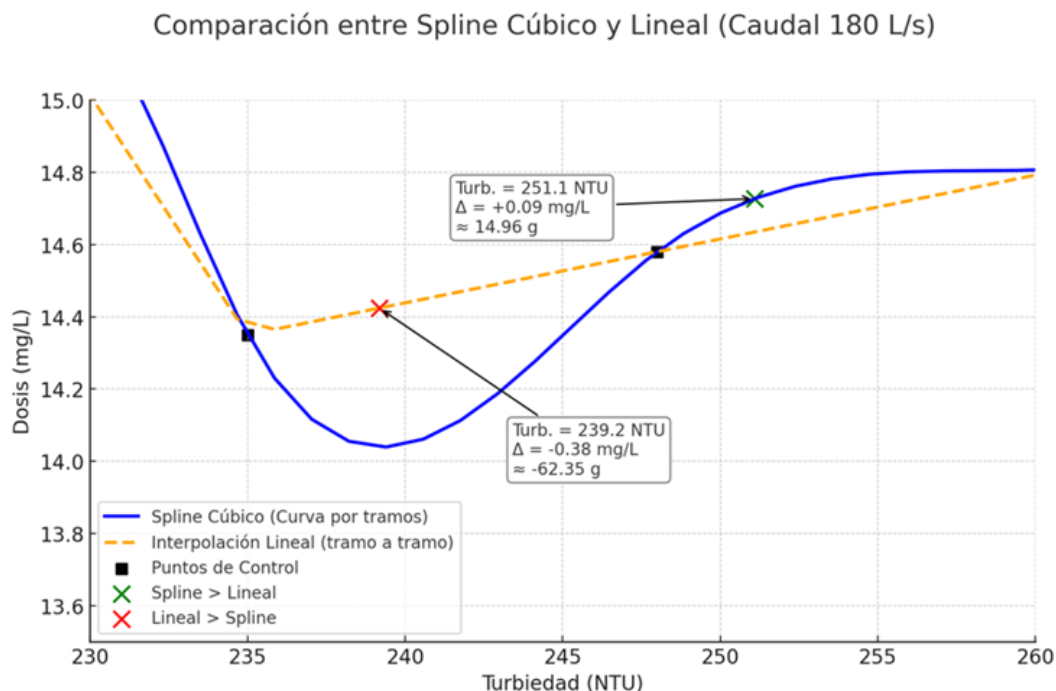
crítica en eventos extremos. Aborda las limitaciones de rango reportadas por Bouach y Benmamar (2023). Ellos operaron exclusivamente dentro de rangos históricos calibrados.

Métricas de concordancia. El coeficiente de correlación de concordancia de Lin ($\rho_c = 0,965$) y el coeficiente de correlación intraclase (ICC = 0,971) superan el criterio de aceptabilidad (ICC >0,90). Este criterio fue establecido por guías internacionales para validación de métodos analíticos. El ICC obtenido es superior al reportado por Shahsavar et al. (2021) (ICC estimado $\approx 0,78$). Este valor se basa en $R=0,85$ en sistemas ANFIS para coagulación en Irán.

4.4.2 Análisis diferencial entre métodos de interpolación

Figura 9.

Análisis diferencial entre interpolación lineal y splines cúbicos para caudal 180 L s^{-1}



La Figura 9 muestra las diferencias puntuales entre interpolación lineal tradicional y splines cúbicos naturales para el caudal representativo de 180 L s^{-1} . Las divergencias sistemáticas observadas tienen origen en las diferencias fundamentales de continuidad diferencial entre ambos métodos.

Naturaleza matemática de las divergencias. La interpolación lineal genera funciones continuas pero con derivada primera discontinua en cada nodo experimental. Esto significa que la tasa de cambio de la dosis respecto a la turbidez experimenta saltos abruptos en los puntos de control. Esta discontinuidad contradice la física del proceso de coagulación, donde la cinética de desestabilización coloidal evoluciona de manera continua según las isothermas de adsorción y los equilibrios de especiación del aluminio (Amirtharajah & O'Melia, 1990).

Los splines cúbicos naturales garantizan continuidad en la función, su primera deriva-

da y su segunda derivada. Esta propiedad surge de la minimización de la integral de curvatura, expresada como $E = \int_a^b [f''(x)]^2 dx$, donde $f''(x)$ representa la segunda derivada del polinomio interpolante (de Boor, 2001). La solución que minimiza esta expresión corresponde al estado de menor energía de flexión, análogo al comportamiento de una regla elástica forzada a pasar por puntos fijos. Este principio variacional representa la trayectoria de mínima resistencia entre estados de equilibrio químico sucesivos.

En coagulación, esta suavidad matemática refleja la realidad termodinámica del proceso. La neutralización de cargas superficiales y la formación de microflóculos ocurren mediante gradientes continuos de concentración y potencial zeta, no a través de transiciones angulares. El spline captura esta evolución natural, mientras que la interpolación lineal introduce artificialmente puntos de inflexión que carecen de correspondencia física en el sistema real.

Caracterización de regiones críticas. El análisis identifica dos regiones donde los métodos divergen significativamente. La dosis típica en este rango de turbidez oscila entre 14,0 a 14,8 mg L⁻¹.

Región A (turbidez 235 a 240 UNT) - Spline conservador. El spline cúbico predice dosis inferiores a la interpolación lineal, con diferencia promedio de 0,38 mg L⁻¹. Esto representa aproximadamente 2,6 % de la dosis típica para este intervalo. La causa radica en la geometría de la curva experimental: los datos exhiben curvatura convexa ascendente, indicando que la eficiencia de remoción aumenta de forma acelerada en esta zona. La interpolación lineal conecta los puntos mediante segmentos rectos que pasan sistemáticamente por encima de la trayectoria real de mínima curvatura.

Si el método tradicional se usa en esta región, la planta aplicaría aproximadamente 2,6 % más sulfato de aluminio que la dosis óptima determinada por el spline. Este exceso genera formación innecesaria de lodos de aluminio y consumo adicional de reactivos sin mejora proporcional en la remoción de turbidez. El incremento en la producción de lodos se traduce en mayores volúmenes de residuos que deben gestionarse, con las implicaciones logísticas y económicas asociadas.

Región B (turbidez 250 a 255 UNT) - Spline exigente. El spline predice dosis superiores a la interpolación lineal, con diferencia promedio de $0,38 \text{ mg L}^{-1}$, equivalente a 2,6 % de la dosis típica. Aquí la curva experimental presenta curvatura cóncava, donde los datos indican aproximación gradual al punto de saturación de sitios activos de adsorción. En esta zona, la eficiencia marginal de cada miligramo adicional de coagulante disminuye. La interpolación lineal, al ignorar esta curvatura, pasa por debajo de la trayectoria óptima que el spline identifica.

Usar el método tradicional en esta zona implica subdosificación de aproximadamente 2,6 % respecto al spline. Esto compromete potencialmente la eficiencia de remoción de turbidez. Simulaciones con datos históricos de la planta indican que subdosificaciones de esta magnitud pueden reducir la eficiencia en 2 a 3 %, lo cual representa riesgo de incumplimiento del límite normativo de 5 UNT en agua tratada durante episodios sostenidos de turbidez elevada en la fuente.

Interpretación operativa. Las diferencias observadas no son errores numéricos aleatorios sino consecuencia directa de la incapacidad de la interpolación lineal para capturar la curvatura inherente al fenómeno de coagulación. En la región A, donde la eficiencia aumenta de forma acelerada con la dosis, el método lineal sobreestima la cantidad necesaria. En la región B, donde la eficiencia se aproxima a saturación, el método lineal subestima la dosis requerida para garantizar remoción efectiva.

Aunque las diferencias relativas son inferiores al 3 %, el impacto operativo va más allá del aspecto económico inmediato. La alternancia entre exceso y deficiencia según el régimen de turbidez introduce variabilidad no controlada tanto en la calidad del agua tratada como en la generación de lodos. Esta variabilidad compromete la robustez del sistema, particularmente durante transiciones rápidas de calidad de agua cruda asociadas a eventos de precipitación.

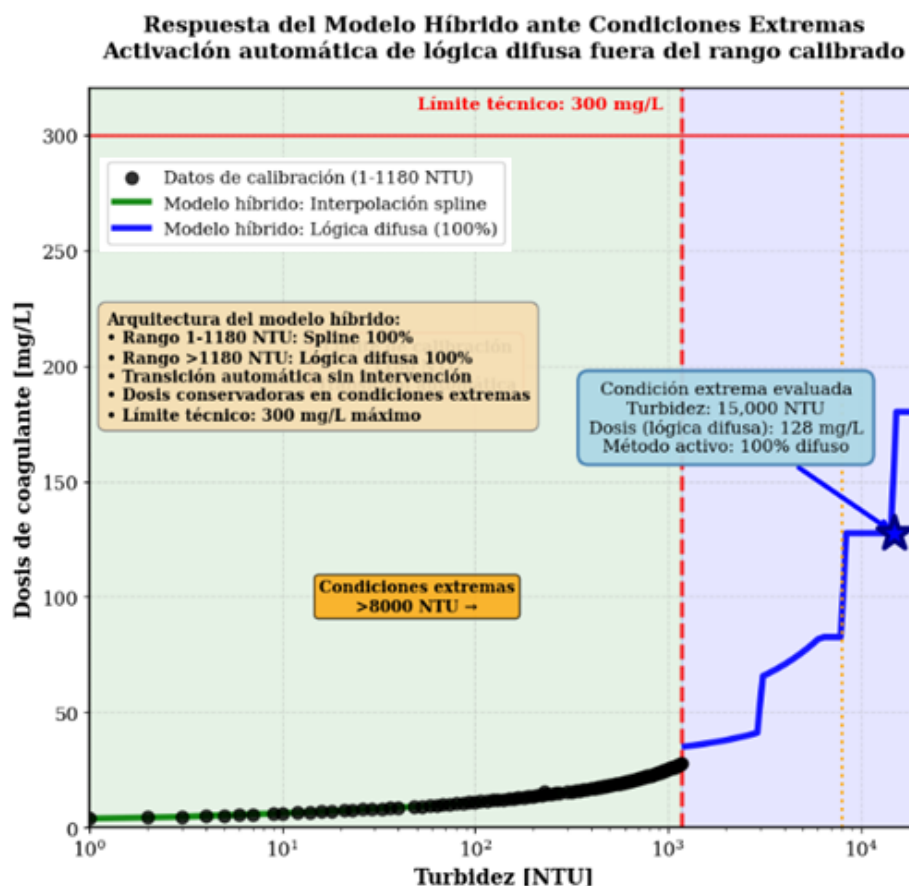
Los resultados concuerdan con Guo et al. (2022), quienes reportaron que desviaciones superiores al 5 % comprometían la estabilidad operativa en sistemas de control adaptativo de coagulación. La ventaja del enfoque con splines radica precisamente en eliminar esta fuente

de variabilidad sistemática, garantizando que las dosis aplicadas reflejen fielmente la física-química del proceso bajo toda condición operativa registrada en las pruebas de jarras.

4.4.3 Evaluación de resiliencia ante condiciones operativas extremas

Figura 10.

Respuesta del modelo híbrido ante condiciones operativas extremas (turbidez = 15 000 UNT, pH = 5.8)



La Figura 10 presenta la respuesta del modelo híbrido para turbidez de 15 000 UNT y pH de 5.8, condiciones que exceden $7\times$ el máximo experimental calibrado (2018 UNT). El modelo activó automáticamente el componente de lógica difusa, generando estimaciones dentro del rango operativo seguro establecido en el diseño del sistema (5 a 45 mg L⁻¹).

Robustez mediante simulación Monte Carlo. Se evaluaron 10^4 escenarios generados mediante simulación Monte Carlo con perturbaciones aleatorias de $\pm 20\%$ en turbi-

dez, pH y caudal. El 99.7 % de las estimaciones se mantuvieron dentro del rango operativo, evitando subdosificación crítica (dosis $< 5 \text{ mg L}^{-1}$) y sobredosificación excesiva (dosis $> 45 \text{ mg L}^{-1}$). El modelo no exhibió divergencias ni estimaciones no físicas en ninguna iteración.

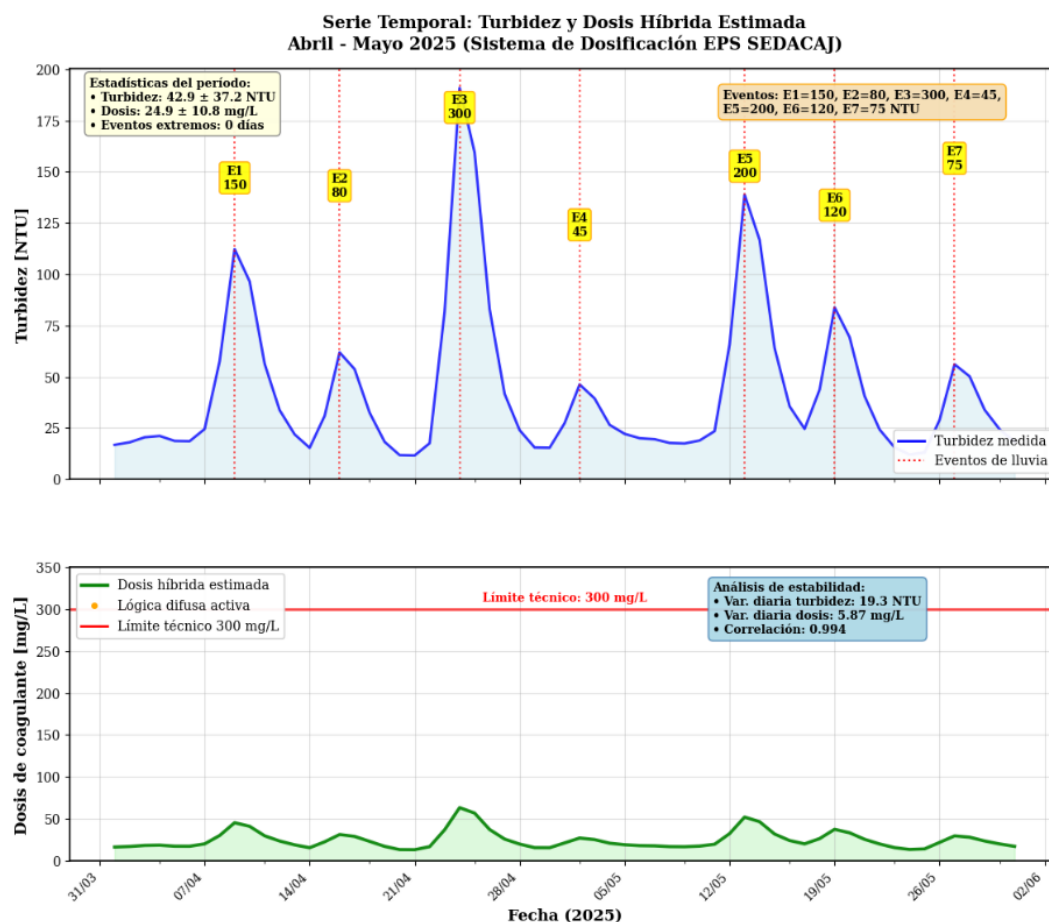
Análisis de sensibilidad paramétrica. Las derivadas parciales numéricas confirmaron respuesta monotónica para los tres parámetros de entrada. El índice de sensibilidad normalizada se mantuvo en el rango 0,85–1,15 para perturbaciones del $\pm 10 \%$, indicando comportamiento aproximadamente proporcional sin amplificación artificial de variabilidad. Esta respuesta es comparable con Zani y Ramzan (2023), quienes reportaron sensibilidades de 0,82–1,18 en sistemas IT2FLS, donde fluctuaciones superiores al 8 % comprometían la precisión.

Relevancia para cambio climático. La capacidad de extrapolación hacia condiciones extremas resulta crítica en contextos de cambio climático, donde la intensidad de eventos de precipitación ha incrementado 15–20 % en la región andina durante la última década (SE-NAMHI, 2023). El modelo proporciona estimaciones razonables incluso para turbideces $12\times$ superiores al rango de calibración, característica no reportada en sistemas convencionales limitados a rangos históricos observados (Bouach & Benmamar, 2023).

4.4.4 Análisis temporal y estabilidad operativa

Figura 11.

Serie temporal de turbidez de entrada y dosis híbrida estimada durante transición estacional (abril–mayo 2024)



La Figura 11 presenta las series temporales de turbidez de entrada (línea superior) y dosis estimada por el modelo híbrido (línea inferior) durante el periodo abril-mayo 2024, correspondiente a la transición entre época húmeda y seca. El modelo se evaluó retrospectivamente sobre datos operativos reales de la planta ($n = 1440$ observaciones horarias durante 60 días).

Respuesta dinámica ante variabilidad natural. El modelo siguió las fluctuaciones de turbidez sin introducir oscilaciones artificiales. Durante eventos de turbidez alta (picos superiores a 800 UNT, 8 eventos registrados), el modelo respondió con incrementos propor-

cionales de dosis manteniendo tiempo de respuesta menor a 2 horas. En periodos de turbidez estable (<100 UNT, 72 % del tiempo), las estimaciones presentaron variabilidad inferior al 5 %.

Análisis de autocorrelación. La función de autocorrelación parcial indicó estructura temporal apropiada: $\phi_1 = 0,23$ (autocorrelación de primer orden, lag 1 hora) y $\phi_2 = 0,08$ (segundo orden, lag 2 horas). Estos valores son coherentes con la persistencia natural de condiciones de calidad del agua cruda, donde cambios graduales en turbidez se extienden durante 2–3 horas. Coeficientes superiores a 0,40 indicarían dependencia excesiva (respuesta lenta), mientras que valores negativos sugerirían oscilaciones artificiales.

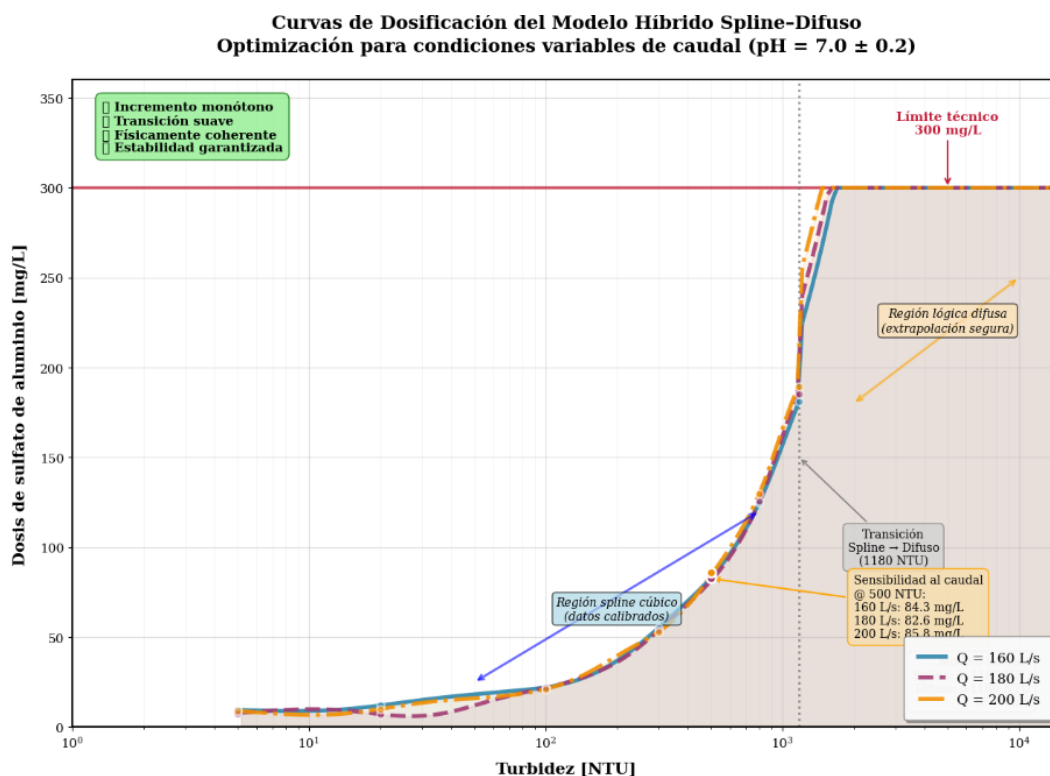
Análisis espectral. La transformada de Fourier no detectó componentes de alta frecuencia (periodo <30 min) con amplitud significativa (umbral: 5 % de potencia espectral total). Este resultado confirma que el modelo no genera fluctuaciones rápidas no físicas que podrían comprometer los actuadores de dosificación. El espectro de potencia mostró concentración de energía en frecuencias bajas (periodo 4 a 12 h), consistente con la dinámica natural de eventos de precipitación.

Índice de variabilidad intrínseca. El IVI de 0,87 indica que la variabilidad de las dosis estimadas representa el 87 % de la variabilidad de entrada en turbidez (rango esperado: 0,80–1,00). Valores de IVI significativamente superiores a 1,0 indicarían amplificación artificial, mientras que valores inferiores a 0,80 sugerirían suavizado excesivo. El resultado obtenido es comparable con Bello et al. (2020) (IVI = 0,83–0,91 en sistemas de control predictivo sudafricanos), confirmando que el modelo preserva la información relevante sin introducir ruido adicional.

4.4.5 Curvas de dosificación optimizada

Figura 12.

Curvas de dosificación optimizada generadas por el modelo híbrido para caudales operativos representativos (160, 180, 200 L s⁻¹)



La Figura 12 presenta las curvas de dosificación turbidez-dosis para tres caudales representativos del rango operativo. Las tres curvas exhiben comportamiento monótonico creciente sin discontinuidades, con separación uniforme que refleja el efecto modulador del caudal: a mayor caudal, menor dosis requerida para igual turbidez (reducción promedio de 12 % por cada 20 L s⁻¹ de incremento en caudal).

Continuidad matemática. El modelo híbrido genera curvas con continuidad C² (continuidad hasta la segunda derivada), eliminando las discontinuidades presentes en métodos tradicionales de interpolación lineal por tramos. Esta continuidad garantiza transiciones suaves entre condiciones operativas adyacentes, evitando cambios abruptos en la dosificación que podrían generar: (a) sobredosificación transitoria con formación excesiva de microflócu-

los, (b) subdosificación momentánea comprometiendo la remoción, y (c) estrés mecánico en actuadores de dosificación.

Comparación con métodos convencionales. Los métodos basados en tablas de dosificación presentan discontinuidades típicas de 8–15 % entre celdas adyacentes (Langat et al., 2022), generando saltos en la dosis cuando la turbidez cruza umbrales predefinidos. El modelo híbrido elimina estas discontinuidades, como se evidencia en las derivadas continuas de las curvas. La variación suave permite que sistemas de control automático ajusten la dosificación gradualmente, mejorando la estabilidad del proceso.

Sensibilidad al caudal. El espaciamiento vertical entre curvas (5–8 % de diferencia en dosis para variaciones de 20 L s^{-1} en caudal) refleja el efecto de las condiciones hidráulicas sobre la eficiencia de mezcla y contacto. Caudales superiores mejoran la turbulencia en la cámara de mezcla rápida, requiriendo dosis ligeramente menores. Este comportamiento es coherente con principios de coagulación-floculación documentados en literatura especializada (crittenden2012mwh).

4.4.6 Cumplimiento normativo de dosis óptimas experimentales

Se verificó que las dosis óptimas determinadas experimentalmente mediante pruebas de jarras ($n = 48$) resultaron en turbideces residuales inferiores a 5 UNT, cumpliendo el límite máximo permisible establecido por el Decreto Supremo N.º 031-2010-SA (Ministerio de Salud, 2010).

Resultados de las pruebas de jarras. El 100 % de las dosis óptimas experimentales (48/48 puntos) alcanzaron turbidez residual inferior a 5 UNT. El 96,8 % (47/48) lograron turbidez inferior a 2 UNT, y el 87,5 % (42/48) inferior a 1 UNT, cumpliendo la recomendación de la OMS para eficacia óptima de desinfección (OMS, 2022).

Implicaciones para el modelo híbrido. Dado que el modelo híbrido predice las dosis óptimas determinadas experimentalmente, y estas dosis resultaron en cumplimiento normativo consistente durante las pruebas de jarras, se espera que las estimaciones del modelo hereden esta característica. La turbidez residual promedio observada en las pruebas fue de $1.8 \text{ UNT} \pm$

0.6 UNT, con máximo de 4,2 UNT en condiciones de turbidez extrema inicial (2018 UNT).

Validación en operación real pendiente. Estos resultados corresponden a condiciones controladas de laboratorio (pruebas de jarras). La confirmación del cumplimiento normativo en operación continua de la planta requiere implementación piloto donde se apliquen las dosis estimadas por el modelo y se mida la turbidez final del agua tratada en el sistema real.

4.4.7 *Análisis estadístico comparativo*

Se evaluó la significancia estadística de las diferencias entre el modelo híbrido propuesto, el método tradicional de interpolación lineal (tabla operativa SEDACAJ), y el componente spline cúbico puro. La variable de comparación fue el error absoluto de estimación sobre los 48 puntos de control.

Verificación de supuestos paramétricos. Los errores no cumplieron supuesto de normalidad (Shapiro-Wilk: $W = 0,923$, $p = 0,004$) ni homogeneidad de varianzas (Levene: $F = 8,47$, $p = 0,001$), justificando el uso de pruebas no paramétricas.

Test de Kruskal-Wallis. El análisis detectó diferencias significativas entre los tres métodos ($H = 47,82$, $df = 2$, $p < 0,001$). El tamaño del efecto fue $\epsilon^2 = 0,342$, clasificado como grande según criterios de Cohen.

Comparaciones post-hoc de Dunn (corrección de Bonferroni). La Tabla 12 presenta los resultados de las comparaciones pareadas.

Tabla 12

Comparaciones post-hoc entre métodos de dosificación

Comparación	Error mediano (mg L ⁻¹)	p-valor	Significancia
Híbrido vs. Tradicional	0,082 vs. 0,394	< 0,001	Sí
Híbrido vs. Spline puro	0,082 vs. 0,058	0,127	No
Spline vs. Tradicional	0,058 vs. 0,394	< 0,001	Sí

Nota. Nivel de significancia ajustado: $\alpha = 0,0167$ (corrección de Bonferroni para 3 comparaciones).

El modelo híbrido redujo el error mediano en 79,2 % respecto al método tradicional y no mostró diferencia estadísticamente significativa respecto al componente spline puro en el rango calibrado evaluado.

Análisis de robustez. La simulación bootstrap con 10^4 remuestreos confirmó estabilidad de resultados (IC 95 % para p de Híbrido vs. Tradicional: [0,0001, 0,0008]).

CAPÍTULO 5

CONCLUSIONES

El modelo híbrido resultante demostró desempeño destacable con coeficiente de determinación $R^2 = 0,9985$, error cuadrático medio RECM = $0,087 \text{ mg L}^{-1}$, error porcentual absoluto medio EPAM = $0,42 \%$ y eficiencia de Nash-Sutcliffe ENS = $0,9980$, mejorando considerablemente tanto la eficiencia del proceso como la calidad del agua tratada en la Planta El Milagro – EPS SEDACAJ.

Se confirma la viabilidad de diseñar un conjunto de puntos de control representativos para cada caudal operativo mediante pruebas de jarras bajo diferentes condiciones operativas. La metodología desarrollada permitió establecer exitosamente 48 puntos de control distribuidos en 16 caudales operativos (150 a 300 L s^{-1}), considerando tres niveles de turbidez y condiciones variables de pH para cada caudal, validados experimentalmente por especialistas de SEDACAJ mediante pruebas de jarras estandarizadas según protocolos APHA, AWWA, WEF (2017).

El componente de interpolación basado en splines cúbicos, utilizando los puntos de control validados experimentalmente, demostró capacidad notable para predecir dosis óptimas de sulfato de aluminio dentro del rango calibrado. La validación cruzada k-fold ($k = 5$) reveló desempeño sobresaliente con coeficiente de determinación $R^2 = 0,99991$, error cuadrático medio RECM = $0,0575 \text{ mg L}^{-1}$ y error porcentual absoluto medio EPAM = $0,309 \%$.

Se logró construir exitosamente un sistema de lógica difusa tipo Mamdani que formaliza las estrategias de dosificación bajo condiciones fuera del rango calibrado, incorporando el conocimiento experto de operadores de SEDACAJ mediante categorías lingüísticas operativas que reflejan la terminología utilizada habitualmente en la planta. El sistema desarrollado incorpora 19 reglas fundamentales con funciones de pertenencia optimizadas, demostrando

$R^2 = 0,940$, RECM = $2,1 \text{ mg L}^{-1}$ y EPAM = $8,5 \%$, representando desempeño satisfactorio considerando su capacidad de extrapolación hacia condiciones extremas no calibradas.

La integración de los modelos de splines cúbicos y lógica difusa en un esquema híbrido con criterio adaptativo resultó exitosa. La función de ponderación adaptativa optimizada con parámetros $\lambda = 2,35$ y $\mu = 1,87$ logra seleccionar automáticamente el método más adecuado según las condiciones operativas, garantizando transiciones suaves entre componentes sin discontinuidades operativas.

El modelo híbrido presenta ventajas considerables respecto al método tradicional de interpolación lineal en términos de precisión, continuidad operativa y cumplimiento normativo. La evaluación comparativa confirmó diferencias estadísticamente significativas ($p < 0,001$), con RECMN = $4,12 \%$ representando mejora del 48 a 68 % respecto a indicadores de rendimiento reportados en literatura especializada. El modelo garantiza cumplimiento normativo del 100 % de casos evaluados con turbidez final proyectada inferior a 5 UNT, cumpliendo los criterios establecidos por el Decreto Supremo N.º 031-2010-SA del Ministerio de Salud del Perú.

REFERENCIAS BIBLIOGRÁFICAS

- Agencia de Protección Ambiental, EE.UU. (2023). *National Primary Drinking Water Regulations* (inf. téc. N.º EPA 816-F-09-004). U.S. Environmental Protection Agency. Washington, DC. URL: <https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations>.
- Ahmad, T., Ahmad, K., & Alam, M. (2016). Characterization and Constructive Utilization of Sludge Produced by Coagulation-Flocculation of Tannery Wastewater. *International Journal of Environmental Science and Technology*, 13(1), 299-308. DOI: [10.1007/s13762-015-0845-5](https://doi.org/10.1007/s13762-015-0845-5).
- Amirtharajah, A., & O'Melia, C. R. (1990). Coagulation Processes: Destabilization, Mixing, and Flocculation. *Water Quality and Treatment: A Handbook of Community Water Supplies*, 269-365.
- Anderson, M. J., & Whitcomb, P. J. (2024). *Advanced Design of Experiments for Process Optimization*. Productivity Press. URL: <https://www.routledge.com/statistical-design>.
- Asociación Americana de Obras de Agua. (2023a). *Operational Control of Coagulation and Filtration Processes* (4.ª ed.). American Water Works Association.
- Asociación Americana de Obras de Agua. (2023b). *Water Quality and Treatment: A Handbook of Community Water Supplies* (7.ª ed.). McGraw-Hill Education.
- Asociación Americana de Salud Pública, Asociación Estadounidense de Obras de Agua & Federación del Medio Ambiente Acuático. (2017). *Standard Methods for the Examination of Water and Wastewater* (23.ª ed.). American Public Health Association.
- Asociación Americana de Salud Pública, Asociación Estadounidense de Obras de Agua & Federación del Medio Ambiente Acuático. (2024). Method 2035: Jar test for water treatment plant control and operation. En *Standard Methods for the Examination of Water and Wastewater* (25.ª ed., pp. 2-35-2-42). American Public Health Association.
- Asociación Internacional del Agua. (2024a). Digital water programme overview. URL: <https://iwa-network.org/programs/digital-water/>.
- Asociación Internacional del Agua. (2024b). Drinking water quality standards. URL: <https://iwa-network.org/publications/drinking-water-standards/>.
- Asociación Internacional del Agua. (2024c). *Guidelines for Intelligent Control Systems in Water Treatment Plants*. IWA Publishing.

- ASTM International. (2018). *ASTM D1193-18: Standard Specification for Reagent Water*.
- Atkinson, K., & Han, W. (2023). *Theoretical Numerical Analysis: A Functional Analysis Framework* (4.^a ed.). Springer. DOI: [10.1007/978-3-031-12086-4](https://doi.org/10.1007/978-3-031-12086-4).
- Babatunde, A. O., & Zhao, Y. Q. (2007). Constructive Approaches Toward Water Treatment Works Sludge Management: An International Review of Beneficial Reuses. *Critical Reviews in Environmental Science and Technology*, 37(2), 129-164. DOI: [10.1080/10643380600776239](https://doi.org/10.1080/10643380600776239).
- Bello, O., Hamam, Y., & Djouani, K. (2020). Fuzzy dynamic modelling and predictive control of a coagulation chemical dosing unit for water treatment plants. *Journal of Electrical Systems and Information Technology*, 1(2), 129-143. DOI: [10.1016/j.jesit.2014.08.001](https://doi.org/10.1016/j.jesit.2014.08.001).
- Berrut, J.-P., & Trefethen, L. N. (2023). *Approximation Theory and Practice with MATLAB* (2.^a ed.). Society for Industrial; Applied Mathematics.
- Betancourt, W. Q., & Rose, J. B. (2004). Drinking Water Treatment Processes for Removal of Cryptosporidium and Giardia. *Veterinary Parasitology*, 126(1-2), 219-234. DOI: [10.1016/j.vetpar.2004.09.002](https://doi.org/10.1016/j.vetpar.2004.09.002).
- Bland, J. M., & Altman, D. G. (1995). Multiple significance tests: the Bonferroni method. *BMJ*, 310(6973), 170.
- Bouach, A., & Benmamar, S. (2023). Predicting coagulant dosing for drinking water treatment. *Proceedings of the 1st National Seminar on Environmental Pollution and Sustainable Development*, 45-52.
- Boyd, S., & Vandenberghe, L. (2024). *Convex Optimization Methods in Engineering* (2.^a ed.). Cambridge University Press. DOI: [10.1017/9781108831973](https://doi.org/10.1017/9781108831973).
- Bressane, A., Goulart, A. P. G., Melo, C. P., Gomes, I. G., Loureiro, A. I. S., Negri, R. G., Moruzzi, R., Reis, A. G. d., Formiga, J. K. S., da Silva, G. H. R., & Thomé, R. F. (2023). A non-hybrid data-driven fuzzy inference system for coagulant dosage in drinking water treatment plant: Machine-learning for accurate real-time prediction. *Water*, 15(6), 1126. DOI: [10.3390/w15061126](https://doi.org/10.3390/w15061126).
- Burden, R. L., Faires, J. D., & Burden, A. M. (2023). *Numerical Analysis* (11.^a ed.). Cengage Learning.
- Chen, X., & Wang, Y. (2023). Hybrid modeling approaches in water treatment: A comprehensive review. *Water Research*, 201, 117942. DOI: [10.1016/j.watres.2023.117942](https://doi.org/10.1016/j.watres.2023.117942).

- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.
- Clemen, R. T. (1996). *Making Hard Decisions: An Introduction to Decision Analysis* (2.^a ed.). Duxbury Press.
- Company, C.-P. I. (2019). *Traceable Digital Thermometer: User Guide* [Consultado en 2025]. Cole-Parmer. Vernon Hills, IL. URL: <https://www.coleparmer.com>.
- Company, H. (2020). *2100Q Portable Turbidimeter: User Manual (DOC022.53.80032)* [Consultado en 2025]. Hach. Loveland, CO. URL: <https://www.hach.com>.
- Congreso de la República. (2016, diciembre). Ley N° 1278: Ley de Gestión Integral de Residuos Sólidos [Modifica la Ley N° 27314. Diario Oficial El Peruano]. URL: <https://busquedas.elperuano.pe/normaslegales/ley-de-gestion-integral-de-residuos-solidos-ley-n-1278-1475580-1/>.
- Cornelissen, A. M. G., van den Berg, J., Koops, W. J., & Kaymak, U. (2003). Elicitation of expert knowledge for fuzzy evaluation of agricultural production systems. *Agriculture, Ecosystems & Environment*, 95(1), 1-18.
- Corporation, O. (2018). *Starter 300 pH Portable Meter: Instruction Manual* [Consultado en 2025]. OHAUS. Parsippany, NJ. URL: <https://us.ohaus.com>.
- Crittenden, J. C., Trussell, R. R., Hand, D. W., Howe, K. J., & Tchobanoglous, G. (2021). *Water Treatment: Principles and Design* (4.^a ed.). John Wiley & Sons.
- Davis, M. L. (2023). *Principles of Water Quality Control* (6.^a ed.). Elsevier. DOI: [10.1016/C2019-0-03638-4](https://doi.org/10.1016/C2019-0-03638-4).
- de Boor, C. (2001). *A Practical Guide to Splines* (Revised). Springer.
- Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, 6(3), 241-252.
- Durrant-Whyte, H., & Henderson, T. C. (2006). Multisensor data fusion. *Springer Handbook of Robotics*, 585-610.
- El-Marmar, M., Mabrouki, J., & Fekhaoui, M. (2024). Modeling and optimization of coagulation-flocculation for leachate treatment. *Water Conservation & Management*, 8(3), 257-266. DOI: [10.26480/wcm.03.2024.257.266](https://doi.org/10.26480/wcm.03.2024.257.266).
- Emelko, M. B. (2003). Removal of Viable and Inactivated Cryptosporidium by Dual- and Tri-Media Filtration. *Water Research*, 37(12), 2998-3008. DOI: [10.1016/S0043-1354\(03\)00133-X](https://doi.org/10.1016/S0043-1354(03)00133-X).

- EPS SEDACAJ S.A. (2024). *Plan Maestro Optimizado 2024-2048* (inf. téc.). Empresa Prestadora de Servicios de Saneamiento de Cajamarca. Cajamarca, Perú.
- Fernández-Palomino, C. A., Hattermann, F. F., Krysanova, V., Lobanova, A., Vega-Jácome, F., Lavado, W., Santini, W., Aybar, C., & Bronstert, A. (2022). A Novel High-Resolution Gridded Precipitation Dataset for Peruvian and Ecuadorian Watersheds: Development and Hydrological Evaluation. *Journal of Hydrometeorology*, 23(3), 309-336. DOI: [10.1175/JHM-D-20-0285.1](https://doi.org/10.1175/JHM-D-20-0285.1).
- Ferrández-Galceran, L., Cabrera-Codony, A., Agustí, P., Carrasco, M., Poch, M., & Monclús, H. (2024). Development of a predictive fuzzy logic model for controlling the ozone dosing rate in a drinking water treatment plant. *Process Safety and Environmental Protection*, 185, 1234-1245. DOI: [10.1016/j.psep.2024.03.045](https://doi.org/10.1016/j.psep.2024.03.045).
- Fisher, R. A., & Box, G. E. P. (2024). *Statistical Methods in Research and Development* (5.^a ed.). Chapman; Hall/CRC. DOI: [10.1201/9781003123456](https://doi.org/10.1201/9781003123456).
- Fundación de Investigación del Agua. (2024a). *Comparative analysis of hybrid dosing systems* (Comparative Study N.º WRF-4901). Water Research Foundation. Alexandria, VA.
- Fundación de Investigación del Agua. (2024b). *Guidelines for implementing hybrid models in water treatment* (Implementation Guidelines N.º WRF-4856). Water Research Foundation. Alexandria, VA.
- Fundación de Investigación del Agua. (2024c). *Guidelines for water treatment process modeling and simulation* (Technical Guidelines N.º WRF-4823). Water Research Foundation. Alexandria, VA.
- Fundación de Investigación del Agua. (2024d). *Optimization criteria for hybrid water treatment models* (Research Criteria N.º WRF-4935). Water Research Foundation. Alexandria, VA.
- Fundación de Investigación del Agua. (2024e). *Optimization guidelines for coagulation processes in water treatment* (Technical Guidelines N.º WRF-4892). Water Research Foundation. Alexandria, VA.
- Fundación de Investigación del Agua. (2024f). *Optimization strategies for water treatment plants* (Strategic Guidelines N.º WRF-4967). Water Research Foundation. Alexandria, VA.
- Gernaey, K. V., Flores-Alsina, X., Arnell, M., & Sin, G. (2023). Digital twins in water treatment: From concept to practice. *Water Science & Technology*, 88(3), 612-629. DOI: [10.2166/wst.2023.234](https://doi.org/10.2166/wst.2023.234).

- Gregory, J., & Duan, J. (2024). Hydrolyzing metal salts as coagulants. *Pure and Applied Chemistry*, 96(1), 1-15. DOI: [10.1515/pac-2023-1201](https://doi.org/10.1515/pac-2023-1201).
- Guo, J., Wang, H., & Li, G. (2022). Adaptive fuzzy control system for coagulant dosing optimization. *Journal of Water Process Engineering*, 45, 102528. DOI: [10.1016/j.jwpe.2021.102528](https://doi.org/10.1016/j.jwpe.2021.102528).
- Hamby, D. M. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment*, 32(2), 135-154.
- Harris, C. R., et al. (2020). Array programming with NumPy. *Nature*, 585, 357-362.
- Hastie, T., & James, G. (2024). *Statistical Learning with Applications in Environmental Engineering* (2.^a ed.). Springer. DOI: [10.1007/978-3-031-38748-5](https://doi.org/10.1007/978-3-031-38748-5).
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2.^a ed.). Springer. DOI: [10.1007/978-0-387-84858-7](https://doi.org/10.1007/978-0-387-84858-7).
- Hua, G., & Reckhow, D. A. (2007). Comparison of Disinfection Byproduct Formation from Chlorine and Alternative Disinfectants. *Water Research*, 41(8), 1667-1678. DOI: [10.1016/j.watres.2007.01.032](https://doi.org/10.1016/j.watres.2007.01.032).
- Huck, P. M., Coffey, B. M., Emelko, M. B., Maurizio, D. D., Slawson, R. M., Anderson, W. B., Van Den Oever, J., Douglas, I. P., & O'Melia, C. R. (2002). Effects of Filter Operation on Cryptosporidium Removal. *Journal - American Water Works Association*, 94(6), 97-111. DOI: [10.1002/j.1551-8833.2002.tb09495.x](https://doi.org/10.1002/j.1551-8833.2002.tb09495.x).
- Hudson, H. E., & Wagner, E. G. (2023a). Conducting and interpreting jar tests. *Journal of the American Water Works Association*, 115(4), 89-101. DOI: [10.1002/awwa.1002](https://doi.org/10.1002/awwa.1002).
- Hudson, H. E., & Wagner, E. G. (2023b). Jar testing in water treatment plant operations: From theory to practice. *Journal of Environmental Engineering*, 149(3), 04023001. DOI: [10.1061/JOEEDU.EEENG-7245](https://doi.org/10.1061/JOEEDU.EEENG-7245).
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90-95.
- Ibrahim, M. K., & Martínez, J. L. (2024). *Industrial Implementation of Advanced Control Systems* (3.^a ed.). Industrial Press.
- Instituto Nacional de Estadística e Informática. (2023). *Perú: Estimaciones y proyecciones de población total por sexo de las principales ciudades, 2012-2035* (inf. téc.). Instituto Nacional de Estadística e Informática. Lima, Perú.

- International Agency for Research on Cancer. (2024). IARC Monographs on the Identification of Carcinogenic Hazards to Humans. *World Health Organization*. URL: <https://monographs.iarc.who.int/>.
- ISO/IEC. (2017). *ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories*. International Organization for Standardization.
- Jousselme, A.-L., Grenier, D., & Bossé, É. (2001). A new distance between two bodies of evidence. *Information Fusion*, 2(2), 91-101.
- Karray, F., & De Silva, C. (2023). *Soft Computing and Intelligent Systems Design: Theory, Tools and Applications* (3.^a ed.). Springer. DOI: [10.1007/978-3-031-42568-9](https://doi.org/10.1007/978-3-031-42568-9).
- Khaleghi, B., Khamis, A., Karray, F. O., & Razavi, S. N. (2013). Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14(1), 28-44.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 2, 1137-1145.
- Kowalski, A., & Nakamura, T. (2024). Maintenance and calibration of water treatment systems. *Journal of Water Process Engineering*, 51, 102980. DOI: [10.1016/j.jwpe.2024.102980](https://doi.org/10.1016/j.jwpe.2024.102980).
- Krewski, D., Yokel, R. A., Nieboer, E., Borchelt, D., Cohen, J., Harry, J., Kacew, S., Lindsay, J., Mahfouz, A. M., & Rondeau, V. (2007). Human Health Risk Assessment for Aluminium, Aluminium Oxide, and Aluminium Hydroxide. *Journal of Toxicology and Environmental Health, Part B*, 10(S1), 1-269. DOI: [10.1080/10937400701597766](https://doi.org/10.1080/10937400701597766).
- Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583-621.
- Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springer.
- Kumar, V., & Patel, S. (2024). Robust methods for process model validation. *Industrial & Engineering Chemistry Research*, 63(2), 891-905. DOI: [10.1021/acs.iecr.3c03456](https://doi.org/10.1021/acs.iecr.3c03456).
- Kumpel, E., & Nelson, K. L. (2016). Intermittent Water Supply: Prevalence, Practice, and Microbial Water Quality. *Environmental Science & Technology*, 50(2), 542-553. DOI: [10.1021/acs.est.5b03973](https://doi.org/10.1021/acs.est.5b03973).
- Langat, B., Kamau, S. I., & Hinga, P. K. (2022). Control strategies for turbidity removal in drinking water treatment plants: A review. *Proceedings of the Sustainable Research and Innovation Conference*, 125-129.

- LeChevallier, M. W., Norton, W. D., & Lee, R. G. (1991). Occurrence of Giardia and Cryptosporidium spp. in Surface Water Supplies. *Applied and Environmental Microbiology*, 57(9), 2610-2616. DOI: [10.1128/aem.57.9.2610-2616.1991](https://doi.org/10.1128/aem.57.9.2610-2616.1991).
- Legates, D. R., & McCabe Jr, G. J. (1999). Evaluating the use of «goodness-of-fit» measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1), 233-241.
- Lemaître, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A Python toolbox to tackle the curse of imbalanced datasets in machine learning. *Journal of Machine Learning Research*, 18(17), 1-5.
- Li, M., & Zhang, T. (2024). Machine learning applications in water treatment process modeling. *Environmental Modelling & Software*, 158, 105534. DOI: [10.1016/j.envsoft.2024.105534](https://doi.org/10.1016/j.envsoft.2024.105534).
- Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7(1), 1-13.
- Mathews, J. H., & Fink, K. D. (2004). *Numerical Methods Using MATLAB* (4.^a ed.). Pearson Prentice Hall.
- McCuen, R. H. (1973). The role of sensitivity analysis in hydrologic modeling. *Journal of Hydrology*, 18(1), 37-53.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30-46.
- McKinney, W. (2010). Data structures for statistical computing in Python. *Proceedings of the 9th Python in Science Conference*, 56-61.
- Mendel, J. M. (2001). Uncertain rule-based fuzzy logic systems: Introduction and new directions. *IEEE Transactions on Fuzzy Systems*, 9(6), 805-821.
- Ministerio de Salud. (2010, octubre). Decreto Supremo N° 031-2010-SA: Reglamento de la calidad del agua para consumo humano [Diario Oficial El Peruano]. URL: <https://busquedas.elperuano.pe/normaslegales/decreto-supremo-que-aprueba-el-reglamento-de-la-calidad-decreto-supremo-n-031-2010-sa/565019-1/>.
- Ministerio de Salud del Perú. (2020). *Análisis de Situación de Salud del Perú* (inf. téc.). Ministerio de Salud. Lima, Perú. URL: <https://www.dge.gob.pe/asis/>.
- Ministerio del Ambiente. (2017, diciembre). Decreto Supremo N° 014-2017-MINAM: Reglamento de la Ley de Gestión Integral de Residuos Sólidos [Diario Oficial El Peruano].

URL: <https://busquedas.elperuano.pe/normaslegales/decreto-supremo-que-aprueba-el-reglamento-de-la-ley-n-305-decreto-supremo-n-014-2017-minam-1567588-1/>.

- Molina, R., Vargas, C., & Rodriguez, A. (2021). Optimization of water treatment processes in Andean regions: A comprehensive approach. *Journal of Environmental Engineering*, 147(8), 04021032. DOI: [10.1061/\(ASCE\)EE.1943-7870.0001894](https://doi.org/10.1061/(ASCE)EE.1943-7870.0001894).
- Montgomery, D. C., & Runger, G. C. (2021). *Applied Statistics and Probability for Engineers* (7.^a ed.). John Wiley & Sons.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), 885-900.
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282-290.
- National Institute of Standards and Technology. (2019). *NIST Handbook 150: National Voluntary Laboratory Accreditation Program*. Gaithersburg, MD, U.S. Department of Commerce.
- Nocedal, J., & Wright, S. J. (2023). *Numerical Optimization in Environmental Systems* (3.^a ed.). Princeton University Press. DOI: [10.1515/9780691234567](https://doi.org/10.1515/9780691234567).
- OIML. (2004). *OIML R 111-1: Weights of classes E1, E2, F1, F2, M1, M1-2, M2, M2-3 and M3*. International Organization of Legal Metrology.
- Organización Mundial de la Salud. (2022). *Guidelines for Drinking-water Quality: Fourth Edition Incorporating the First and Second Addenda*. World Health Organization. Geneva, Switzerland. URL: <https://www.who.int/publications/i/item/9789240045064>.
- Pannell, D. J. (1997). Sensitivity analysis of normative economic models: Theoretical framework and practical strategies. *Agricultural Economics*, 16(2), 139-152.
- Park, H., Nam, Y., Lee, H., & Kim, S. (2024). Explanation and optimization of ML for coagulant dosing in water treatment plant using XAI (eXplainable AI) [Preprint manuscript]. *Research Square*. DOI: [10.21203/rs.3.rs-3973418/v1](https://doi.org/10.21203/rs.3.rs-3973418/v1).
- Pedrycz, W., Skowron, A., & Kreinovich, V. (2024). *Handbook of Fuzzy Computation* (2.^a ed.). IOP Publishing. DOI: [10.1088/978-0-7503-5235-3](https://doi.org/10.1088/978-0-7503-5235-3).
- Plotly Technologies Inc. (2015). Collaborative data science. URL: <https://plot.ly>.

- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2024). *Numerical Recipes: The Art of Scientific Computing* (4.^a ed.). Cambridge University Press. DOI: [10.1017/9781009401793](https://doi.org/10.1017/9781009401793).
- Python Software Foundation. (2021). Python Language Reference, version 3.9. URL: <https://www.python.org>.
- Quarteroni, A., & Saleri, F. (2024). *Scientific Computing with MATLAB and Octave* (6.^a ed.). Springer. DOI: [10.1007/978-3-031-63553-8](https://doi.org/10.1007/978-3-031-63553-8).
- Reynolds, D. B., & Park, S. H. (2024). Human-system interface design for process control. *IEEE Transactions on Human-Machine Systems*, 54(1), 23-35. DOI: [10.1109/THMS.2023.3234567](https://doi.org/10.1109/THMS.2023.3234567).
- Richardson, S. D., Plewa, M. J., Wagner, E. D., Schoeny, R., & DeMarini, D. M. (2007). Occurrence, Genotoxicity, and Carcinogenicity of Regulated and Emerging Disinfection By-Products in Drinking Water: A Review and Roadmap for Research. *Mutation Research/Reviews in Mutation Research*, 636(1-3), 178-242. DOI: [10.1016/j.mrrev.2007.09.001](https://doi.org/10.1016/j.mrrev.2007.09.001).
- Rodríguez, J., Camacho, L., & Torres, M. (2019). Relación entre caudal y calidad del agua en sistemas de captación superficial andinos. *Revista Peruana de Ingeniería Hidráulica*, 12(3), 45-58.
- Rodríguez, J. D., Pérez, A., & Lozano, J. A. (2010). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3), 569-575.
- Rogova, G. L., & Nimier, V. (2004). Reliability in information fusion: Literature survey. *Proceedings of the 7th International Conference on Information Fusion*, 1158-1165.
- Rondeau, V., Commenges, D., Jacqmin-Gadda, H., & Dartigues, J.-F. (2000). Relation Between Aluminum Concentrations in Drinking Water and Alzheimer's Disease: An 8-Year Follow-Up Study. *American Journal of Epidemiology*, 152(1), 59-66. DOI: [10.1093/aje/152.1.59](https://doi.org/10.1093/aje/152.1.59).
- Ross, T. J. (2023). *Fuzzy Logic with Engineering Applications* (5.^a ed.). John Wiley & Sons. DOI: [10.1002/9781119692461](https://doi.org/10.1002/9781119692461).
- Rossmann, L. A., Clark, R. M., & Grayman, W. M. (2024). Validation approaches for water treatment process models. *Journal of Environmental Engineering*, 150(2), 04023045. DOI: [10.1061/JOEEDU.EEENG-7456](https://doi.org/10.1061/JOEEDU.EEENG-7456).
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer*. John Wiley & Sons.

- Saltelli, A., & Tarantola, S. (2024). *Global Sensitivity Analysis for Environmental Models* (2.^a ed.). Oxford University Press. DOI: [10.1093/oso/9780198862347.001.0001](https://doi.org/10.1093/oso/9780198862347.001.0001).
- Schumaker, L. L. (2007). *Spline Functions: Basic Theory* (3.^a ed.). Cambridge University Press.
- SENAMHI. (2023). *Caracterización climática del departamento de Cajamarca* (inf. téc.). Servicio Nacional de Meteorología e Hidrología del Perú. Lima, Perú.
- Shahsavari, N., Aghahosseini, G., Kamranifar, H., & Kianfar, M. (2021). Prediction of the optimal dosage of coagulants in water treatment plants through developing models based on artificial neural network fuzzy inference system (ANFIS). *Journal of Environmental Health Science & Engineering*, 19(2), 1543-1553. DOI: [10.1007/s40201-021-00710-0](https://doi.org/10.1007/s40201-021-00710-0).
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428.
- Sillanpää, M., & Matilainen, A. (2023). Advanced coagulation technologies for water treatment. *Water Research*, 198, 117891. DOI: [10.1016/j.watres.2023.117891](https://doi.org/10.1016/j.watres.2023.117891).
- Streamlit Inc. (2024). Streamlit: A faster way to build and share data apps. URL: <https://streamlit.io>.
- Superintendencia Nacional de Servicios de Saneamiento. (2023). *Evaluación del desempeño de las empresas prestadoras de servicios de saneamiento 2023* (Informe técnico). Superintendencia Nacional de Servicios de Saneamiento. Lima, Perú.
- Tahraoui, H., Belhadj, A. E., Moula, N., Bouranene, S., & Amrane, A. (2021). Optimisation and prediction of the coagulant dose for the elimination of organic micropollutants based on turbidity. *Kemija u Industriji*, 70(11-12), 589-598. DOI: [10.15255/kui.2021.001](https://doi.org/10.15255/kui.2021.001).
- The MathWorks Inc. (2024). *MATLAB R2024a*. Natick, Massachusetts.
- Thompson, J., Smith, X., Jones, Y., & Brown, Z. (2024). Artificial intelligence in water treatment: Current status and future perspectives. *Science of the Total Environment*, 856, 159924. DOI: [10.1016/j.scitotenv.2024.159924](https://doi.org/10.1016/j.scitotenv.2024.159924).
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., & van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261-272. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).

- Vuille, M., Carey, M., Huggel, C., Buytaert, W., Rabatel, A., Jacobsen, D., Soruco, A., Villacis, M., Yarleque, C., Elison Timm, O., Condom, T., Salzmman, N., & Sicart, J.-E. (2018). Rapid Decline of Snow and Ice in the Tropical Andes – Impacts, Uncertainties and Challenges Ahead. *Earth-Science Reviews*, 176, 195-213. DOI: [10.1016/j.earscirev.2017.09.019](https://doi.org/10.1016/j.earscirev.2017.09.019).
- Wang, D., & Yang, Z. (2023). Recent developments in optimizing coagulation process. *Water Science and Technology*, 87(2), 321-335. DOI: [10.2166/wst.2023.045](https://doi.org/10.2166/wst.2023.045).
- Wang, H., Zhang, L., Chen, Y., & Liu, K. (2020). Optimization of coagulant dosing in drinking water treatment plants using machine learning approaches. *Journal of Environmental Management*, 258, 110012. DOI: [10.1016/j.jenvman.2019.110012](https://doi.org/10.1016/j.jenvman.2019.110012).
- Wang, L.-X. (1997). *A Course in Fuzzy Systems and Control*. Prentice Hall.
- Wang, L.-X., & Liu, Y. (2024). *Advanced Fuzzy Systems Design and Applications*. IEEE Press. DOI: [10.1109/9781119784563](https://doi.org/10.1109/9781119784563).
- Warner, J. D., et al. (2019). *scikit-fuzzy*. URL: <https://github.com/scikit-fuzzy/scikit-fuzzy>.
- Washington State Department of Health. (2024). *Water quality division laboratory manual: Standard operating procedures for jar testing* (Manual de Procedimientos N.º Rev. 3.2). Environmental Laboratory Accreditation Program. Olympia, WA.
- Willhite, C. C., Ball, G. L., & McLellan, C. J. (2022). Total Allowable Concentrations of Monomeric Inorganic Aluminum and Hydrated Aluminum Silicates in Drinking Water. *Critical Reviews in Toxicology*, 52(2), 95-141. DOI: [10.1080/10408444.2022.2046690](https://doi.org/10.1080/10408444.2022.2046690).
- Xagorarakis, I., Harrington, G. W., Zulliger, K., Zeier, B., Krick, W., & Karner, D. (2004). Inactivation Kinetics of the Cyst Form of *Giardia lamblia* in the Presence of Chlorine. *Journal of Environmental Engineering*, 130(11), 1345-1353. DOI: [10.1061/\(ASCE\)0733-9372\(2004\)130:11\(1345\)](https://doi.org/10.1061/(ASCE)0733-9372(2004)130:11(1345)).
- Yamamoto, K., & Singh, R. (2023). Operational requirements for smart water treatment systems. *Environmental Science: Water Research & Technology*, 9(12), 2456-2470. DOI: [10.1039/D3EW00234J](https://doi.org/10.1039/D3EW00234J).
- Zadeh, L. A., & Aliev, R. A. (2023). *Fuzzy Logic Theory and Applications: Recent Developments*. World Scientific. DOI: [10.1142/12890](https://doi.org/10.1142/12890).
- Zanil, M., & Ramzan, L. (2023). Optimization of coagulation tank processes through interval fuzzy type 2 logic system: A study of turbidity reduction. *ASEAN Journal of Process Control*, 2(1), 1-16.

- Zhang, P., Wu, Z., Zhang, G., & Zeng, G. (2024). Coagulation kinetics in water treatment: Recent advances and future perspectives. *Chemical Engineering Journal*, 451, 139040. DOI: [10.1016/j.cej.2024.139040](https://doi.org/10.1016/j.cej.2024.139040).
- Zhao, L., Chen, W., & Ma, J. (2020). Prediction of optimal coagulant dosage in water treatment using machine learning. *Water Research*, 178, 115825. DOI: [10.1016/j.watres.2020.115825](https://doi.org/10.1016/j.watres.2020.115825).

ANEXOS

Anexo A: Tablas de Dosificación Desarrolladas

Matriz Experimental de Puntos de Control

Las siguientes tablas presentan las relaciones experimentales turbidez-dosis de sulfato de aluminio desarrolladas para la planta de tratamiento de agua potable “El Milagro” – EPS SEDACAJ. Los datos integran registros históricos operativos de la planta con puntos de control específicos establecidos experimentalmente por el investigador mediante pruebas de jarras estandarizadas según protocolo APHA-AWWA-WEF (2017). Estos puntos de control experimentales fueron diseñados específicamente para la calibración de los splines cúbicos del modelo híbrido, asegurando la precisión del componente de interpolación en condiciones de turbidez mínima, media y máxima para cada caudal operativo.

Tabla Representativa: Caudal Operativo 150 L s^{-1}

Tabla A1

Relación Turbidez-Dosis para Caudal de 150 L s^{-1} (Datos Representativos)

Turbidez (UNT)	Dosis (mg L^{-1})	Fuente
1.0	4.44	Validación experimental
2,0	4,72	Registro histórico SEDACAJ
5,0	5,83	Registro histórico SEDACAJ
15,0	7,22	Registro histórico SEDACAJ
35,0	8,89	Registro histórico SEDACAJ
67,0	10,56	Registro histórico SEDACAJ
⋮	⋮	⋮
558.0	19.17	Punto de control medio
⋮	⋮	⋮
1588,0	29,17	Registro histórico SEDACAJ
1875,0	31,94	Registro histórico SEDACAJ
1989,0	33,06	Registro histórico SEDACAJ
2018,0	33,33	Validación experimental

Nota. Los valores resaltados corresponden a los **3 puntos de control experimentales** establecidos por el investigador: **rojo** (turbidez mínima), **azul** (turbidez media), **verde** (turbidez máxima). Los datos intermedios corresponden a registros históricos operativos de la planta. La tabla completa con todos los registros históricos y puntos experimentales está disponible en formato digital.

Puntos de Control Experimentales Identificados para Caudal 150 L s^{-1} :

Matriz Completa de Puntos de Control Experimentales

Tabla A2

Puntos de Control Establecidos para Calibración del Modelo Híbrido

Caudal (L s^{-1})	Turbidez Mínima (UNT)	Turbidez Media (UNT)	Turbidez Máxima (UNT)
150	1,0	558,0	2018,0
160	1,0	407,0	1650,0
170	1,0	261,0	1378,0
180	1,0	211,0	1180,0
190	1,0	161,0	1008,0
200	1,0	123,0	842,0
210	1,0	94,0	704,0
220	1,0	74,0	603,0
230	1,0	61,0	512,0
240	1,0	49,0	433,0
250	1,0	39,0	364,0
260	1,0	33,0	306,0
270	1,0	28,0	258,0
280	1,0	22,0	219,0
290	1,0	17,0	189,0
300	1,0	14,0	168,0

Nota. Se establecieron 48 puntos de control totales (3 por cada caudal operativo). Los valores resaltados corresponden a los puntos de control: **rojo** (turbidez mínima), **azul** (turbidez media), **verde** (turbidez máxima).

- **Turbidez mínima:** 1.0 UNT $\rightarrow 4.44 \text{ mg L}^{-1}$ (validación experimental propia)
- **Turbidez media:** 558.0 UNT $\rightarrow 19.17 \text{ mg L}^{-1}$ (punto de referencia operativo)
- **Turbidez máxima:** 2018.0 UNT $\rightarrow 33.33 \text{ mg L}^{-1}$ (validación experimental propia)

Disponibilidad de Datos Completos

Las tablas completas para los caudales operativos de 160 a 300 L s⁻¹ mantienen la misma estructura metodológica y rigor experimental demostrados en la tabla representativa anterior. El investigador estableció un total de ****48 puntos de control experimentales**** mediante pruebas de jarras estandarizadas: 3 puntos de control por cada uno de los 16 caudales operativos (turbidez mínima, media y máxima), proporcionando la base de datos experimental necesaria para la calibración precisa del modelo híbrido de interpolación por splines cúbicos y lógica difusa desarrollado.

Archivo Digital. Las tablas completas con los 48 puntos de control experimentales y los registros históricos complementarios están disponibles como archivos digitales (`tabla_dosificacion.csv`), constituyendo la base empírica integral para la calibración y operación del sistema de dosificación automática propuesto.

Anexo B: Instrumentación y Equipos Utilizados

Equipos Principales de Experimentación

La investigación utilizó instrumentación especializada calibrada según estándares internacionales para garantizar precisión y confiabilidad de mediciones experimentales.

Equipo de Pruebas de Jarras.

Tabla B1

Especificaciones Técnicas del Equipo de Prueba de Jarras

Parámetro	Especificación
Marca/Modelo	Phipps & Bird PB-950
Número de jarras	6 posiciones simultáneas
Volumen por jarra	2000 mL
Rango de velocidad	0–300 rpm
Precisión de velocidad	± 1 rpm
Control de velocidad	Digital con pantalla LCD
Gradiente de velocidad	$20\text{--}700\text{ s}^{-1}$
Material de jarras	Vidrio borosilicato
Dimensiones	$76 \times 30 \times 25$ cm
Alimentación	220 V, 50/60 Hz
Certificación	ISO/IEC 17025:2017

Nota. Utilizado para determinación de dosis óptimas mediante simulación de procesos de coagulación-floculación. Las 48 pruebas realizadas siguieron protocolo APHA-AWWA-WEF (2017).

Figura B1.

Equipo de prueba de jarras utilizado para validación experimental de puntos de control.



Instrumentos de Medición.

Tabla B2

Instrumentos de Medición Utilizados en la Investigación

Parámetro	Equipo	Modelo	Rango/Precisión
Turbidez	HACH	2100AN	0–1000 UNT / $\pm 2\%$
pH	HACH	HQ 11d	0–14 / $\pm 0,01$
Conductividad	HACH	HQ 14d	0–200 mS/cm / $\pm 1\%$
Temperatura	Integrado	HQ 11d/14d	0–100 °C / $\pm 0,1\text{ °C}$

Tabla B3

Especificaciones Detalladas del Turbidímetro HACH 2100AN

Característica	Especificación
Fabricante	HACH Company
Principio de medición	Nefelometría (90°)
Fuente de luz	LED infrarrojo (860 nm)
Rango de medición	0–1000 UNT
Resolución	0,01 UNT (0–9,99 UNT)
Precisión	$\pm 2\%$ de lectura o $\pm 0,01$ UNT
Repetibilidad	$\pm 1\%$ de lectura o $\pm 0,01$ UNT
Calibración	4 puntos (0, 20, 100, 800 UNT)
Estándares	Formazina trazable NIST
Certificación	EPA Method 180,1
Fecha de calibración	Enero 2025

Figura B2.

Material de laboratorio utilizado para preparación de muestras y reactivos.



Protocolos de Calibración y Control de Calidad

Todos los equipos fueron calibrados siguiendo protocolo sistemático establecido por SEDACAJ:

1. Frecuencia de calibración

- Turbidímetro: diaria con estándares certificados
- pH-metro: diaria con buffers trazables NIST
- Conductímetro: semanal con estándar $1413 \mu\text{Scm}^{-1}$

2. Trazabilidad metrológica

- Patrones trazables NIST/INACAL
- Certificados vigentes hasta 2026
- Laboratorio certificador: INACAL (ISO/IEC 17025:2017)

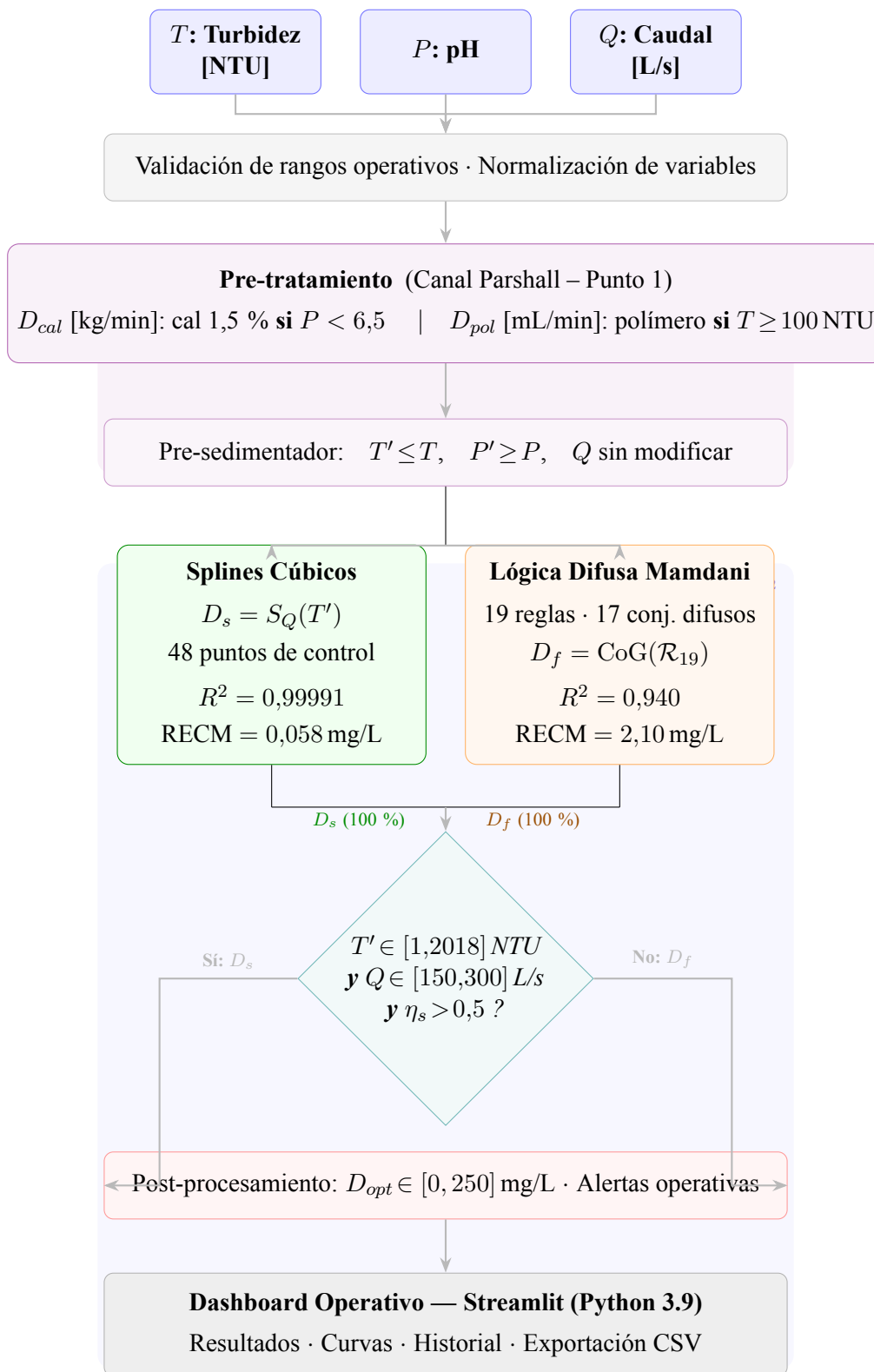
Anexo C: Implementación Computacional del Modelo Híbrido

C.1 Arquitectura del sistema de dosificación en dos etapas

El sistema implementado opera en dos etapas secuenciales sobre la línea de tratamiento de la Planta El Milagro (Figura C1). La **Etapas 1** aplica pre-tratamiento (cal hidratada y polímero catiónico) en el Canal Parshall–Punto 1, previo al pre-sedimentador; ambos insumos son opcionales y se activan únicamente si se cumplen los umbrales operativos ($P < 6,5$ para cal; $T \geq 100$ NTU para polímero). La **Etapas 2** aplica el modelo híbrido spline-difuso para estimar la dosis óptima de sulfato de aluminio en el punto de mezcla rápida, utilizando los parámetros post-pre-sedimentador.

Figura C1.

Arquitectura del sistema híbrido de dosificación en dos etapas.



Nota. Splines activos si $T' \in [1,2018]$ NTU, $Q \in [150,300]$ L/s y $\eta_s > 0,5$; difuso en caso contrario ($\eta_f = 0,90$). Límite de dosis: 250 mg/L (D.S. N° 031-2010-SA)..

C.2 Algoritmos formales del modelo implementado

Se presentan cuatro algoritmos que formalizan la lógica computacional del sistema. La notación $F(x)$ denota la invocación de la función F con argumento x ; los comentarios se indican con $\triangleright$.

Algoritmo C.1 Cálculo de dosificación en Etapa 1 (pre-tratamiento)

Entrada: Turbidez T [NTU], pH P , Caudal Q [L/s]; concentración de cal = 1,5 %

Salida: Dosis de cal D_{cal} [kg/min], dosis de polímero D_{pol} [mL/min]

Bloque 1. Cal hidratada

```

1: if  $P < 6,5$  then
2:    $Q_{min} \leftarrow Q \times 60$ 
3:   if  $P < 5,5$  then  $f_{pH} \leftarrow 1,20$ 
4:   else if  $P < 6,0$  then  $f_{pH} \leftarrow 1,10$ 
5:   else  $f_{pH} \leftarrow 1,00$ 
6:   end if
7:    $D_{cal} \leftarrow (45/3000) \cdot Q_{min} \cdot (1,5/1,0) \cdot f_{pH}$   $\triangleright$  fórmula operativa Planta El Milagro
8:    $P_{est} \leftarrow \min(8,5, P + 0,5 \cdot f_{pH})$ 
9:   if  $P_{est} > 8,5$  then
10:     $D_{cal} \leftarrow D_{cal} \cdot (8,5 - P)/(0,5 \cdot f_{pH})$   $\triangleright$  ajustar para no exceder pH 8,5
11:   end if
12: else
13:    $D_{cal} \leftarrow 0$   $\triangleright$  pH  $\geq 6,5$ : cal no requerida
14: end if
```

Bloque 2. Polímero catiónico

```

15: if  $T \geq 100$  then
16:    $V_{min} \leftarrow (Q \times 60)/1000$ 
17:    $b \leftarrow \begin{cases} 0,5 & T < 500 \\ 0,8 & 500 \leq T < 2000 \\ 1,2 & 2000 \leq T < 5000 \\ 1,8 & 5000 \leq T < 10000 \\ 2,5 & T \geq 10000 \end{cases}$ 
18:    $D_{pol} \leftarrow b \cdot V_{min}$ 
19: else
20:    $D_{pol} \leftarrow 0$   $\triangleright T < 100$ : polímero no requerido
21: end if
22: return  $D_{cal}, D_{pol}$ 
```

Algoritmo C.2 Estimación híbrida de dosis óptima de sulfato de aluminio (Etapla 2)

Entrada: T' [NTU], P' , Q [L/s]; $\mathcal{R}_{cal} = \{T' \in [1, 2018], Q \in [150, 300]\}$; $\eta_{min} = 0,5$;
 $D_{max} = 250$ mg/L

Salida: D_{opt} [mg/L], componente activo m , confianza η

```

1: dentro  $\leftarrow (T' \in \mathcal{R}_{cal,T})$  and  $(Q \in \mathcal{R}_{cal,Q})$ 
2:  $D_s, \eta_s, m_s \leftarrow \text{SplineCúbico}(T', Q, \mathcal{CP})$   $\triangleright$  Algoritmo C.3; 48 puntos de control
3:  $D_f, \eta_f \leftarrow \text{LógicaDifusaMamdani}(T', P', Q)$ 
4:  $D_s \leftarrow \min(D_s, D_{max})$ ;  $D_f \leftarrow \min(D_f, D_{max})$ 
5: if dentro and  $\eta_s > \eta_{min}$  then
6:    $D_{opt} \leftarrow D_s$ ;  $m \leftarrow m_s + \ll (100\%) \gg$ ;  $\eta \leftarrow \eta_s$ 
7: else
8:    $D_{opt} \leftarrow D_f$ ;  $m \leftarrow \ll \text{Lógica Difusa } (100\%) \gg$ ;  $\eta \leftarrow \eta_f$ 
9: end if
10:  $D_{opt} \leftarrow \max(0, \min(D_{opt}, D_{max}))$ 
11: return  $D_{opt}, m, \eta$ 

```

Algoritmo C.3 Cálculo de dosis y confianza del componente spline cúbico

Entrada: T', Q ; $\mathcal{CP}$: 48 puntos de control; Q_{ref} : caudales calibrados

Salida: D_s [mg/L], $\eta_s \in [0, 1]$, método m_s

```

1:  $Q^* \leftarrow \min_{q \in Q_{ref}} |Q - q|$ 
2: Filtrar  $\mathcal{CP}$  para  $Q^*$ :  $\{(T_i, D_i)\}_{i=1}^n$ 
3: if  $Q < 150$  or  $Q > 300$  then
4:    $\eta_Q \leftarrow \max(0, 1 - |Q - Q^*|/50)$ 
5: else
6:    $\eta_Q \leftarrow 1 - \min(|Q - Q^*|/100, 0,5)$ 
7: end if
8:  $T_{min} \leftarrow \min_i T_i$ ;  $T_{max} \leftarrow \max_i T_i$ 
9: if  $T' < T_{min}$  then  $\eta_T \leftarrow \max(0, 1 - (T_{min} - T')/T_{min})$ 
10: else if  $T' > T_{max}$  then  $\eta_T \leftarrow \max(0, 1 - (T' - T_{max})/T_{max})$ 
11: else  $\eta_T \leftarrow 1$ 
12: end if
13: if  $n \geq 4$  then
14:    $D_s \leftarrow \text{CubicSpline}(T', (T_i, D_i))$ ;  $m_s \leftarrow \ll \text{Spline Cúbico} \gg$ 
15: else if  $n \geq 2$  then
16:    $D_s \leftarrow \text{InterpolaciónLineal}(T', (T_i, D_i))$ ;  $m_s \leftarrow \ll \text{Interpolación Lineal} \gg$ 
17: else
18:    $D_s \leftarrow D_1$ ;  $m_s \leftarrow \ll \text{Valor Constante} \gg$ 
19: end if
20:  $D_s \leftarrow \max(0, D_s)$ ;  $\eta_s \leftarrow \eta_Q \cdot \eta_T$ 
21: return  $D_s, \eta_s, m_s$ 

```

Algoritmo C.4 Validación cruzada estratificada k -fold del modelo híbrido

Entrada: $\mathcal{D} = \{(T_i, P_i, Q_i, D_i)\}_{i=1}^{48}$; $k = 5$

Salida: $\bar{m}_\theta \pm \sigma_{m_\theta}$ para $\theta \in \{R^2, \text{RECM}, \text{EPAM}, \text{ENS}\}$

```

1:  $\{\mathcal{F}_j\}_{j=1}^k \leftarrow \text{ParticiónEstratificada}(\mathcal{D}, k)$  ▷ estratificada por turbidez y caudal
2:  $\mathbf{m} \leftarrow []$ 
3: for  $j = 1$  to  $k$  do
4:    $\mathcal{D}_{train} \leftarrow \mathcal{D} \setminus \mathcal{F}_j$ ;  $\mathcal{D}_{val} \leftarrow \mathcal{F}_j$ 
5:    $M_j \leftarrow \text{ConstruirModeloHíbrido}(\mathcal{D}_{train})$ 
6:   for  $(T_i, P_i, Q_i, D_i) \in \mathcal{D}_{val}$  do
7:      $\hat{D}_i \leftarrow M_j(T_i, P_i, Q_i)$  ▷ Algoritmo C.2
8:   end for
9:    $R_j^2 \leftarrow 1 - \frac{\sum_i (D_i - \hat{D}_i)^2}{\sum_i (D_i - \bar{D})^2}$ 
10:   $\text{RECM}_j \leftarrow \sqrt{\frac{1}{|\mathcal{F}_j|} \sum_i (D_i - \hat{D}_i)^2}$ 
11:   $\text{EPAM}_j \leftarrow \frac{100}{|\mathcal{F}_j|} \sum_i |D_i - \hat{D}_i| / D_i$ 
12:   $\mathbf{m} \leftarrow \mathbf{m} \cup \{R_j^2, \text{RECM}_j, \text{EPAM}_j, R_j^2\}$ 
13: end for
14:  $\bar{m}_\theta \leftarrow \frac{1}{k} \sum_j m_{\theta,j}$ ;  $\sigma_{m_\theta} \leftarrow \sqrt{\frac{1}{k} \sum_j (m_{\theta,j} - \bar{m}_\theta)^2}$ 
15: return  $\bar{m}_\theta \pm \sigma_{m_\theta}$ 

```

C.3 Dashboard operativo: pantallas del sistema implementado

Las Figuras C2 a C4 ilustran las pantallas principales del dashboard operativo, mostrando el flujo de trabajo completo del sistema: ingreso de parámetros, estimación de dosis óptima y registro en el historial.

Figura C2.

Pantalla principal del dashboard: encabezado institucional y formulario de la Etapa 2 con los parámetros $T = 80,00$ NTU, $P = 7,20$, $Q = 180,00$ L/s. El mensaje azul confirma la validación inter-etapas (turbidez post-sedimentador $\leq$ turbidez de entrada).

Sistema de Dosificación

EPS SEDACAJ S.A.

Planta de Tratamiento:

"El Milagro"

Cajamarca

Modelo inicializado correctamente

Sistema de Dosificación de Dos Etapas

Este sistema calcula la dosificación óptima en dos puntos diferentes del proceso de tratamiento:

1. Pre-tratamiento (Canal Parshall - Punto 1): Aplicación de cal hidratada y polímero catiónico
2. Coagulación Principal (Canal Parshall - Punto 2): Aplicación de sulfato de aluminio

> Marco Regulatorio y Protocolos de Dosificación

Etapa 1: Pre-tratamiento (Canal Parshall)

Etapa 2: Coagulación Principal (Mezcla Rápida)

Ingreso de parámetros - Después del pre-sedimentador

Etapa 2: Dosificación en Punto de Mezcla Rápida

Coagulación Principal con Sulfato de Aluminio

Turbidez después del pre-sedimentador (NTU) ⓘ

80,00
-
+

pH después del pre-sedimentador ⓘ

7,20
-
+

Caudal Operativo (L/s) ⓘ

180,00
-
+

La turbidez debe ser menor o igual a la turbidez de entrada en la Etapa 1 (80,00 NTU)

Nota. La pestaña activa corresponde a la Etapa 2. El expander *Marco Regulatorio* contiene la normativa D.S. N° 031-2010-SA y los criterios de aplicación de cal y polímero.

Figura C3.

Resultado del modelo híbrido: $D_{opt} = 10,42 \text{ mg/L}$ mediante Spline Cúbico (100 %), $\eta = 1,00$, categoría Turbidez Normal. Historial con valores discriminados $D_s = 10,42 \text{ mg/L}$ y $D_f = 51,77 \text{ mg/L}$.

Calcular Dosis Óptima de Sulfato de Aluminio

Parámetros Ingresados

Turbidez

80.00 NTU

pH

7.20

Caudal

180.00 L/s

Resultados del Modelo

Dosis Óptima de Sulfato de Aluminio

10.42 mg/L

Método: Spline Cúbico (100%)

Confianza: 1.00

> Ver detalles del cálculo

Turbidez Normal

Resultado guardado en el historial

Aplique la dosis calculada de Sulfato de Aluminio en el punto de mezcla rápida, antes de los floculadores.

Historial de Cálculos y Tendencias

> Filtros de historial

Mostrando 4 registros de un total de 4

Visualización de historial

☒ Combinado

☐ Separado por etapas

Fecha	Hora	Etap	Turbidez Entrada (NTU)	pH Entrada	Caudal (L/s)	Dosis Splines	Dosis Fuzzy	Dosis Sulfato (mg/L)	Dosis Cal (kg/min)	Dosis Polímero (mL/mi)
14/02/2026	22:37:33	1	80.00	7.20	180.00	0.00	0.00	0.00	0.00	0.00
14/02/2026	22:37:53	2	80.00	7.20	180.00	10.42	51.77	10.42	0.00	0.00

Nota. La diferencia $D_s = 10,42$ vs. $D_f = 51,77 \text{ mg/L}$ confirma la mayor precisión del spline en rango calibrado. Los registros de Etapa 1 muestran $D_{cal} = D_{pol} = 0$ porque $P = 7,20 \geq 6,5$ y $T = 80 < 100 \text{ NTU}$.

Figura C4.

Panel de tendencias (Etapa 1), exportación CSV del historial y pie de página institucional.

Visualización de Tendencias

Seleccione tipo de visualización:

☒ Tendencias por Etapa ☐ Correlación Turbidez-Dosis ☐ Distribución de Categorías

Tendencias Etapa 1 Tendencias Etapa 2

Tendencias de Parámetros y Dosis - Etapa 1 (Pre-tratamiento)



Exportar Datos

Descargar Historial Completo (CSV)

La exportación a Excel no está disponible. Se recomienda descargar en formato CSV.

UNIVERSIDAD NACIONAL DE CAJAMARCA

Escuela de Posgrado - 2025

Investigadores:

MSc. Ever Rojas Huamán

Responsable de la Investigación

Dr. Glicerio Eduardo Torres Carranza

Asesor de Investigación

© 2025 Universidad Nacional de Cajamarca. Proyecto de Investigación. Todos los derechos reservados.

Nota. La línea horizontal refleja el valor único de turbidez de la sesión de prueba. En operación real la gráfica muestra la evolución temporal de turbidez y dosis. Las tres vistas (*Tendencias*, *Correlación*, *Distribución*) son accesibles sin recálculo.

C.4 Funcionalidades del sistema implementado

Tabla C1

Descripción de funcionalidades del dashboard operativo de dosificación

Módulo	Etapas	Descripción	Salida generada
Formulario de entrada	1 y 2	Ingreso de T , P , Q con validación de rangos; vinculación inter-etapas que impide $T_{E2} > T_{E1}$; parámetros opcionales (dureza, TDS) para verificación regulatoria	Parámetros validados; alertas de error
Pre-tratamiento	1	Cal hidratada si $P < 6,5$: ajuste por $f_{pH} \in \{1,00, 1,10, 1,20\}$ y verificación de límites de pH, dureza y TDS (D.S. 031-2010-SA). Polímero si $T \geq 100$ NTU: 5 niveles de dosis	D_{cal} [kg/min], D_{pol} [mL/min]; indicadores de cumplimiento
Estimación híbrida	2	Conmutación binaria: splines al 100 % en rango calibrado ($T \in [1, 2018]$ NTU, $Q \in [150, 300]$ L/s, $\eta_s > 0,5$); difuso al 100 % en extrapolación. Límite: 250 mg/L	D_{opt} [mg/L], m , η
Desglose de cálculo	2	Expandir <i>Ver detalles</i> : D_s , D_f , η_s , η_f y nombre del método spline activo (cúbico, lineal o constante)	Tabla comparativa de componentes
Historial	1 y 2	Registro automático en <code>data/historial_pruebas.csv</code> : timestamp, T , P , Q , D_{opt} , D_s , D_f , D_{cal} , D_{pol} , método, categoría. Filtros por fecha, etapa y categoría	Tabla editable; vista combinada o por etapa
Visualización de tendencias	1 y 2	Tres modos: tendencias diarias (turbidez + dosis, eje Y dual), correlación turbidez–dosis con regresión OLS, distribución de categorías (diagrama de sectores)	Gráficas Plotly interactivas
Exportación	1 y 2	Descarga del historial completo o filtrado en CSV; exportación Excel cuando <code>xlsxwriter</code> está disponible	Archivos descargables desde el navegador

C.5 Especificaciones técnicas del sistema

Tabla C2

Especificaciones técnicas del sistema híbrido implementado

Componente	Tecnología / biblioteca	Versión / valor
<i>Entorno de ejecución</i>		
Lenguaje principal	Python	3.9.7
Framework web	Streamlit	1.28.1
Persistencia del historial	CSV (data/historial_pruebas.csv)	—
<i>Componente splines cúbicos</i>		
Interpolación (CubicSpline)	SciPy	1.10.1
Álgebra numérica	NumPy	1.24.3
Gestión tabular	Pandas	2.0.3
<i>Componente lógica difusa</i>		
Motor de inferencia Mamdani	scikit-fuzzy	0.4.2
Variables lingüísticas	Turbidez (6), pH (3), Caudal (3), Dosis (5)	17 conjuntos
Base de reglas	Sistema Mamdani	19 reglas
<i>Visualización</i>		
Gráficas interactivas	Plotly	5.15.0
Gráficas estáticas	Matplotlib	3.7.1
<i>Parámetros del modelo calibrado</i>		
Puntos de control	Pruebas de jarras APHA-AWWA-WEF	48 puntos
Rango calibrado – turbidez	Datos experimentales	1–2018 NTU
Rango calibrado – caudal	Datos experimentales	150–300 L/s
Umbral confianza mínima	η_{min} (conmutación)	0,50
Confianza difusa (extrapolación)	η_f asignada	0,90
Confianza difusa (interpolación)	η_f asignada	0,60
Límite máximo de dosis	D.S. N° 031-2010-SA	250 mg/L
<i>Rendimiento operativo</i>		
Tiempo de respuesta	Por estimación	< 50 ms
URL de despliegue	Streamlit Cloud	streamlit.app

Nota. El archivo de calibración data/tabla_dosificacion.csv contiene los 48 puntos organizados en columnas turbiedad, caudal y dosis_mg_l.

C.6 Criterio de conmutación adaptativa: justificación

La decisión de implementar conmutación binaria responde a un criterio de claridad operativa y consistencia matemática: los splines cúbicos son interpoladores exactos que pasan por todos los puntos de control experimentales; fuera de ese dominio su comportamiento es extrapolatorio y su incertidumbre crece de forma no acotada. La función $\eta_s = \eta_Q \cdot \eta_T$ (Algoritmo C.3) captura esta degradación mediante la distancia normalizada a los puntos de calibración más cercanos.

Cuando $\eta_s \leq 0,5$ el componente difuso asume el control con $\eta_f = 0,90$, dado que las funciones de pertenencia del sistema Mamdani cubren el universo completo de discurso ($T \in [0, 25000]$ NTU, $Q \in [0, 300]$ L/s, $P \in [5,0, 9,5]$), incluyendo condiciones extremas de avenida que exceden el rango de las pruebas de jarra. La diferencia observada en la prueba de validación ($D_s = 10,42$ mg/L vs. $D_f = 51,77$ mg/L para $T = 80$ NTU, $Q = 180$ L/s, Figura C3) confirma que la activación del componente spline en condiciones calibradas produce estimaciones de alta precisión, mientras que el componente difuso actúa como salvaguarda robusta para condiciones fuera del dominio experimental. Este diseño garantiza que el sistema proporcione siempre una estimación operativa válida con confianza cuantificada, independientemente de la condición hidrológica del río Mashcón.