



Artículo original / Original article

Monitoreo de la calidad del agua para zonas rurales mediante un sistema de sensores de bajo costo basados en IoT

Water quality monitoring for rural areas using a low-cost IoT-based sensor system

Andrea Alexandra Garcia-Gonzales ¹; Cristian García-Estrella ^{1*}

¹ Universidad Nacional de San Martín, Tarapoto, Perú

Recibido: 13/11/2026

Aceptado: 21/01/2026

Publicado: 30/01/2026

*Autor de correspondencia: cgarcia@unsm.edu.pe

Resumen: El acceso limitado a tecnologías de monitoreo continuo de la calidad del agua en zonas rurales dificulta la gestión eficiente del recurso hídrico y expone a la población a riesgos sanitarios. Este estudio presenta el diseño e implementación de un sistema de monitoreo basado en sensores de bajo costo integrados a una arquitectura de Internet de las Cosas (IoT), con capacidad de transmisión de datos en tiempo real. La investigación se desarrolló en la comunidad rural de Chuina, ubicada en el distrito de Morales, región San Martín, Perú. Se implementó un sistema embebido basado en Arduino Mega, integrando sensores de pH, turbidez, temperatura, sólidos disueltos totales y conductividad eléctrica, junto con módulos de comunicación y almacenamiento. Se recopilieron 2880 registros automáticos durante 60 días de monitoreo. Los resultados evidenciaron que los parámetros fisicoquímicos se mantuvieron mayoritariamente dentro de los límites normativos y se identificaron patrones mediante estadística descriptiva y modelos de aprendizaje no supervisado mediante K-Means. Se concluye que el sistema presenta un funcionamiento robusto y confiable, siendo una alternativa replicable para zonas rurales con limitado infraestructura y útil para fortalecer la vigilancia ambiental.

Palabras clave: monitoreo ambiental; sistemas embebidos; aprendizaje no supervisado; sensores fisicoquímicos; zonas rurales

Abstract: Limited access to continuous water quality monitoring technologies in rural areas hinders efficient water resource management and exposes the population to health risks. This study presents the design and implementation of a monitoring system based on low-cost sensors integrated into an Internet of Things (IoT) architecture, capable of real-time data transmission. The research was conducted in the rural community of Chuina, located in the Morales district of the San Martín region, Peru. An embedded system based on Arduino Mega was implemented, integrating sensors for pH, turbidity, temperature, total dissolved solids, and electrical conductivity, alongside communication and storage modules. A total of 2,880 automatic readings were collected over a 60-day monitoring period. The results showed that physicochemical parameters generally remained within regulatory limits; patterns were identified using descriptive statistics and K-Means unsupervised learning models. The study concludes that the system operates robustly and reliably, offering a replicable alternative for rural areas with limited infrastructure and serving as a useful tool for strengthening environmental monitoring.

Keywords: environmental monitoring; embedded systems; unsupervised learning; physicochemical sensors; rural areas

1. Introducción

El agua potable constituye un recurso esencial para la salud pública, el bienestar de las poblaciones y el desarrollo socioeconómico, debido a su importancia para el consumo humano y para diversas actividades productivas. Sin embargo, garantizar su disponibilidad en condiciones adecuadas de calidad continúa siendo un desafío, particularmente en zonas rurales donde existen limitaciones de infraestructura, acceso y capacidad de monitoreo. En este contexto, el desarrollo de tecnologías orientadas a la evaluación continua del recurso hídrico adquiere especial relevancia para fortalecer su gestión sostenible (Gomes et al., 2024; Kumar et al., 2020). Asimismo, las desigualdades en el acceso al agua segura continúan siendo un problema relevante en el contexto peruano (Paredes & Woolley, 2022).

Entre las tecnologías emergentes aplicables a esta problemática destaca el Internet de las Cosas (Internet of Things [IoT]), cuya capacidad para integrar sensores, dispositivos de procesamiento y sistemas de comunicación permite recopilar y transmitir información de manera automatizada. Su creciente aplicación en diferentes sectores ha favorecido el desarrollo de soluciones orientadas al monitoreo ambiental y sanitario, proporcionando información continua para apoyar la gestión de los recursos naturales (Nieto Pacheco et al., 2023). En el ámbito hídrico, las arquitecturas IoT ofrecen la posibilidad de supervisar variables relacionadas con la calidad del agua mediante sistemas automatizados y de bajo costo (Karthikeyan et al., 2024).

En el Perú, a pesar de los avances alcanzados en materia de abastecimiento, persisten desigualdades espaciales y socioeconómicas en el acceso al agua potable segura, especialmente en poblaciones rurales y territorios con mayores dificultades de conectividad (Al-kassab-Córdova et al., 2023). Esta situación representa un problema de salud pública, debido a que el consumo de agua en condiciones inadecuadas se encuentra asociado con diferentes enfermedades de origen hídrico (Jaramillo Vera & Marines Batalla, 2023). La Organización Mundial de la Salud [OMS] (s. f.) advierte, asimismo, sobre los riesgos sanitarios derivados del acceso insuficiente a servicios seguros de agua. Frente a este escenario, las tecnologías IoT constituyen una alternativa para fortalecer el control de la calidad del agua y mejorar la disponibilidad de información para la toma de decisiones (Salgado et al., 2022).

Esta problemática adquiere particular importancia en territorios donde las condiciones ambientales y las limitaciones de infraestructura dificultan la vigilancia permanente de las fuentes de abastecimiento. En el contexto peruano, la gestión de los recursos hídricos requiere mecanismos que contribuyan al seguimiento de su calidad, mientras que las brechas existentes en el acceso al agua y saneamiento evidencian la necesidad de fortalecer las capacidades de monitoreo, principalmente en las poblaciones más vulnerables (Autoridad Nacional del Agua [ANA], s. f.; Instituto Nacional de Estadística e Informática [INEI], s. f.). Esta situación también tiene implicancias para la población infantil, particularmente expuesta a los efectos derivados de servicios deficientes de agua, saneamiento e higiene (Fondo de las Naciones Unidas para la Infancia [UNICEF], s. f.).

En respuesta a estas limitaciones, los sistemas de monitoreo basados en sensores de bajo costo representan una alternativa tecnológica con potencial para ampliar la vigilancia de la calidad del agua. La combinación de IoT y estrategias participativas de monitoreo permite generar información con mayor continuidad y facilitar su comunicación a los actores involucrados en la gestión del recurso hídrico (Amador-Castro et al., 2024). De manera similar, los sistemas IoT de bajo costo han demostrado posibilidades de aplicación para medir continuamente parámetros fisicoquímicos y proporcionar información sobre las variaciones experimentadas por las fuentes de agua (Lal et al., 2024).

En la comunidad rural de Chuina, ubicada en el distrito de Morales, región San Martín, persisten limitaciones para la medición, seguimiento y monitoreo continuo de parámetros críticos relacionados con la calidad del agua. La dependencia de evaluaciones puntuales dificulta identificar oportunamente variaciones asociadas con las condiciones ambientales y características

de la fuente hídrica. La evidencia disponible señala que las aguas superficiales y subterráneas pueden presentar variaciones fisicoquímicas relacionadas con las características propias del entorno (Carvalho et al., 2022; Ghachoui et al., 2024). Además, las tecnologías de monitoreo continuo ofrecen posibilidades para detectar cambios en la calidad del agua y generar alertas frente a variaciones relevantes (Gupta et al., 2023; Qian & Liu, 2024), aspecto especialmente importante cuando las deficiencias en los servicios de agua pueden afectar a poblaciones vulnerables (Montero Contreras, 2022).

En este contexto, persiste la necesidad de evaluar soluciones tecnológicas accesibles y confiables que permitan complementar los procedimientos convencionales de vigilancia de la calidad del agua en comunidades rurales. Por ello, el presente estudio tuvo como objetivo implementar y evaluar un sistema de bajo costo basado en IoT para el monitoreo de la calidad del agua en la comunidad rural de Chuina, integrando sensores para la medición de temperatura, pH, conductividad eléctrica, sólidos disueltos totales y turbidez. Asimismo, los datos obtenidos fueron analizados mediante estadística descriptiva y técnicas de aprendizaje no supervisado, con el propósito de identificar patrones en el comportamiento de los parámetros fisicoquímicos y generar información que contribuya a la vigilancia ambiental y a la gestión del recurso hídrico.

2. Materiales y métodos

La investigación se desarrolló en la comunidad rural de Chuina, ubicada en el distrito de Morales, región San Martín, Perú. La fuente hídrica evaluada corresponde a una vertiente utilizada como principal suministro de agua para consumo doméstico de la población local. Las condiciones climáticas de la región, marcadas por periodos de lluvias intensas, dificultan el acceso regular a la zona de captación y pueden influir en la variabilidad de los parámetros fisicoquímicos. La ejecución del estudio se realizó entre enero de 2025 y julio de 2025.

Muestreo y análisis de laboratorio para la línea base

Zona de captación: La captación se realizó en el punto natural de surgencia que constituye la principal fuente de abastecimiento de agua para la población local.

Recolección de muestras. Se extrajeron muestras de agua en volúmenes definido por el laboratorio, utilizando recipientes estériles, recolectadas bajo distintas condiciones climáticas, con el fin de evaluar la variabilidad natural de los parámetros fisicoquímicos (pH, turbidez, temperatura, sólidos disueltos totales y conductividad eléctrica).

Transporte de muestras. Las muestras se transportadas bajo condiciones controladas, minimizando el tiempo de traslado para preservar sus propiedades fisicoquímicas y evitar alteraciones por temperatura o manipulación.

Análisis en laboratorio. Los análisis se realizaron en laboratorio acreditado, siguiendo métodos analíticos convencionales y protocolos internos de control de calidad.

Marco normativo. La interpretación de los resultados se basó en estándares nacionales vigentes para agua de consumo humano y calidad ambiental.

Selección de Componentes

como criterios de selección la precisión, disponibilidad comercial, bajo consumo energético y facilidad de integración, y priorizando sensores de la línea Gravity por su confiabilidad en proyectos educativos y prototipos IoT (Lal et al., 2024; Lu et al., 2024). El sistema estuvo conformado por un sensor de conductividad eléctrica Gravity (K = 10), sensor de pH Gravity V2, sensor de sólidos disueltos totales (TDS) Gravity, sensor de turbidez Gravity y sensor de temperatura DS18B20. Para el procesamiento y control de los datos se utilizó una placa de desarrollo Arduino Mega 2560, mientras que la transmisión de la información se realizó mediante un módulo de comunicación GSM SIM800L y el almacenamiento local de los registros mediante un módulo MicroSD.

Arquitectura del hardware

La arquitectura del hardware se definió como un esquema de interconexión funcional, con el objetivo de establecer la lógica de comunicación, adquisición de datos y alimentación entre los sensores, los módulos y el microcontrolador Arduino Mega 2560 (Siren et al., 2024).

Este esquema lógico se empleó como base metodológica para el posterior ensamblaje físico del prototipo. De acuerdo con la Figura 1, El sensor de temperatura DS18B20 (1) se asignó al pin digital 2 mediante protocolo OneWire; el sensor de turbidez (2) se conectó al pin analógico A2; el sensor de TDS (3) al pin analógico A3; el sensor de pH (4) al pin analógico A4; y el sensor de conductividad eléctrica (5) al pin analógico A5.

El módulo de almacenamiento MicroSD (6) se conectó a los pines SPI del Arduino, conectando MISO al pin 50, MOSI al 51, SCK al 52 y CS al 53. El módulo de comunicación GSM SIM800L (7) se configuró a través de comunicación serial UART, enlazando su pin TX al pin 19 (RX1) y su RX al pin 18 (TX1) del microcontrolador.

Los sensores y el módulo de almacenamiento (1-6) compartieron una fuente común de alimentación de 5V (VCC) y conexión a tierra (GND). Por su parte, el componente (7) operó con una alimentación independiente de 3.7 V suministrada por una batería externa (8), seleccionada con el objetivo de asegurar estabilidad eléctrica durante los procesos de transmisión de datos.

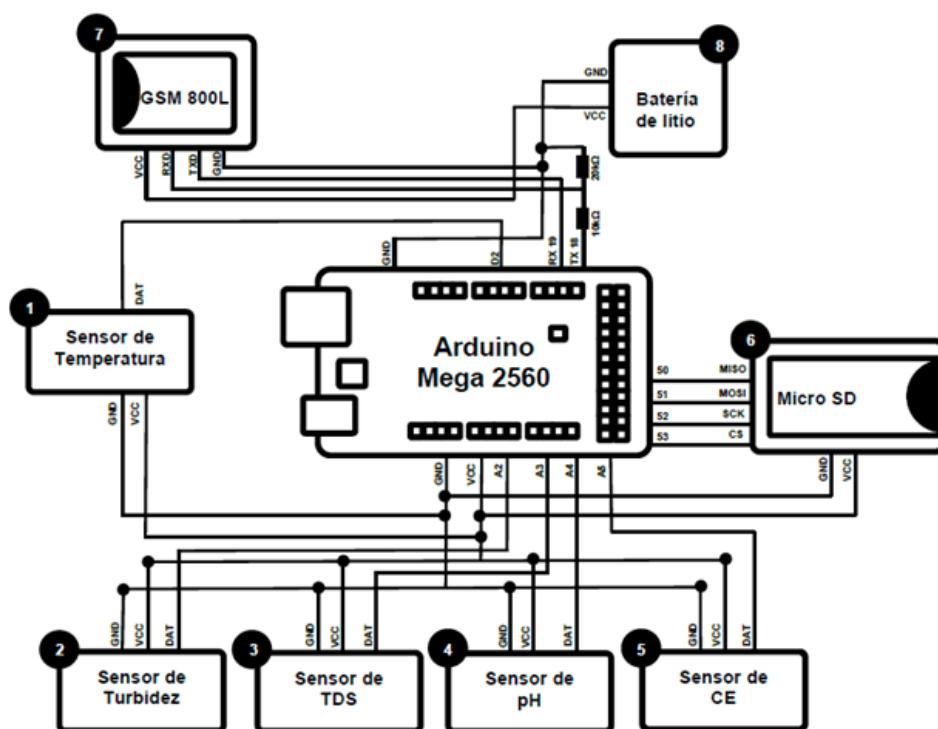


Figura 1. Esquema lógico del Hardware

2.4. Diseño del prototipo

El diseño del prototipo se desarrolló en dos etapas. En la primera, el sistema fue montado sobre dos placas protoboard, utilizadas como buses de alimentación común para facilitar la interconexión y realizar las pruebas iniciales de los sensores y módulos electrónicos. En la segunda etapa, el prototipo fue migrado a una placa de circuito impreso (PCB), para lo cual se diseñó el trazado de pistas y se implementaron conexiones soldadas, proporcionando una configuración más estable y adecuada para su funcionamiento en condiciones de campo (Venkatesh et al., 2024).

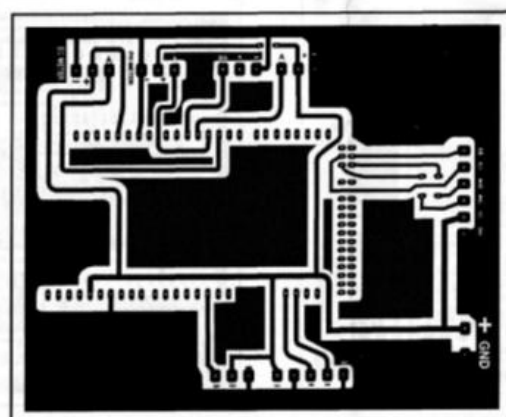


Figura 2. Prototipo Inicial y Diseño del PCB

Adicionalmente, se diseñó un encapsulado físico personalizado mediante tecnología de impresión 3D. El encapsulado tuvo como finalidad proteger los componentes electrónicos, facilitar el montaje ordenado de los sensores y módulos, y mejorar la portabilidad y reutilización del sistema [23]. Figura 3

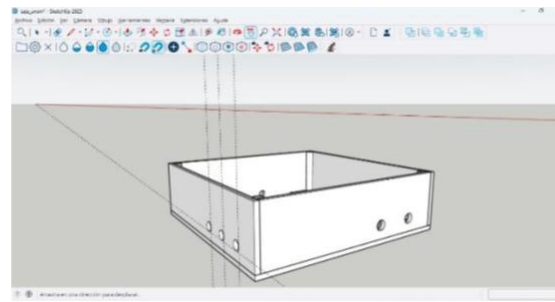
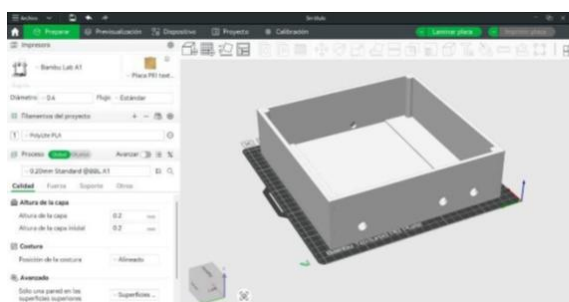


Figura 3. Diseño del encapsulado

2.5. Calibración

La calibración de los sensores se realizó siguiendo las recomendaciones del fabricante y los procedimientos establecidos en los manuales técnicos correspondientes. El sensor de pH fue calibrado utilizando soluciones buffer de pH 4.00 y pH 7.00, mientras que para el sensor de conductividad eléctrica se empleó una solución estándar de 12.88 mS/cm. En ambos casos, el procedimiento consistió en sumergir el electrodo en la solución patrón, esperar la estabilización de la lectura y ajustar los parámetros internos del módulo hasta obtener correspondencia con el valor de referencia. Este procedimiento permitió establecer condiciones adecuadas para la adquisición de los parámetros fisicoquímicos mediante el sistema IoT (Jabbar et al., 2024; Sugiharto et al., 2023).

Los sensores de turbidez, sólidos disueltos totales (TDS) y temperatura no fueron sometidos a un proceso de calibración manual, debido a que, de acuerdo con la documentación técnica considerada durante la implementación, disponían de calibración de fábrica y bibliotecas de conversión integradas. Por tanto, fueron utilizados directamente para la adquisición de datos durante el monitoreo ambiental. La configuración y el procedimiento de calibración implementados se presentan en la Figura 4.



Figura 4. Soluciones patrón para calibración

2.6. Sistema embebido

El sistema embebido fue desarrollado en el entorno Arduino IDE y programado para ejecutar un ciclo automático de adquisición, procesamiento, almacenamiento y transmisión de datos fisicoquímicos del agua.

El programa integró librerías específicas para la lectura de sensores de pH, conductividad eléctrica, temperatura, turbidez y TDS, así como para la gestión de comunicación GSM y almacenamiento en tarjeta microSD. Durante la inicialización, se configuraron los sensores, el módulo SIM800L y el sistema de almacenamiento local. La adquisición de datos se realizó mediante un esquema de muestreo periódico fijo. En cada ciclo, se leyeron las señales analógicas y digitales de los sensores y se convirtieron a valores fisicoquímicos mediante ecuaciones de conversión y bibliotecas del fabricante.

Los registros generados incluyeron temperatura, pH, conductividad eléctrica, TDS y turbidez, y fueron almacenados en formato CSV. Cuando existió conectividad, los datos se transmitieron mediante solicitudes HTTP a una hoja de cálculo en la nube (Google Sheets). En ausencia de señal GSM, los registros permanecieron almacenados localmente y se reenviaron automáticamente al restablecer la comunicación. El conjunto de datos obtenido fue utilizado posteriormente para el análisis estadístico y la aplicación de técnicas de aprendizaje no supervisado.

2.7. Machine Learning

Se aplicó el algoritmo de agrupamiento K-Means con el propósito de identificar patrones en el comportamiento conjunto de los parámetros fisicoquímicos del agua. El procedimiento analítico se estructuró en cinco etapas consecutivas: (i) obtención de datos, (ii) preprocesamiento, (iii) modelado, (iv) validación y (v) visualización, siguiendo un enfoque de análisis basado en técnicas de aprendizaje automático aplicadas a datos obtenidos mediante sistemas IoT (Arsene et al., 2022).

i. Obtención de datos. Se utilizó el conjunto de datos recolectado automáticamente en campo mediante el sistema IoT. Cada registro incluyó mediciones de temperatura (°C), pH, conductividad eléctrica (mS/cm), sólidos totales disueltos (mg/L) y turbidez (NTU). Los datos fueron almacenados en formato CSV y posteriormente importados a un entorno de análisis en Python para su procesamiento.

ii. Preprocesamiento de datos. Se evaluó la estructura y consistencia del conjunto de datos, así como la presencia de valores faltantes. Al no identificarse datos ausentes, no fue necesario aplicar procedimientos de imputación. Adicionalmente, se exploraron las relaciones entre las variables mediante una matriz de correlación y gráficos de dispersión, con el propósito de caracterizar el comportamiento de los parámetros antes de aplicar el algoritmo de agrupamiento.

iii. Modelado. Las variables fueron estandarizadas mediante StandardScaler, con la finalidad de evitar que las diferencias de escala entre los parámetros fisicoquímicos influyeran de manera desproporcionada en la formación de los grupos. Posteriormente, se aplicó el algoritmo K-Means sobre los datos estandarizados. El número óptimo de clústeres (k) se determinó mediante el método del codo, evaluando la variación de la inercia intraclúster para diferentes valores de k. La selección final se realizó identificando el punto de inflexión de la curva y considerando, complementariamente, la interpretabilidad de los patrones obtenidos.

iv. Validación. La calidad de la solución de agrupamiento se evaluó mediante el coeficiente de silueta (Silhouette Score), el cual permitió valorar simultáneamente la cohesión intraclúster y la separación entre los clústeres. Esta métrica se utilizó para determinar la consistencia de la estructura obtenida y respaldar la interpretación de los grupos como patrones diferenciados en el comportamiento de los parámetros de calidad del agua.

v. Visualización. Finalmente, se aplicó el Análisis de Componentes Principales (PCA) para reducir la dimensionalidad de los datos y proyectarlos en un espacio bidimensional. Esta representación permitió visualizar la distribución de los registros según los clústeres obtenidos y facilitar la interpretación de los patrones identificados durante el periodo de monitoreo.

2.8. Evaluación de la efectividad del sistema IoT

Se aplicó un enfoque cuantitativo basado en estadística descriptiva. Los datos obtenidos del sistema IoT fueron exportados en formato .csv e importados al software IBM SPSS Statistics v28.0. Para cada parámetro fisicoquímico se calcularon los principales estadísticos descriptivos, incluyendo media, mediana, valores mínimo y máximo, desviación estándar, cuartiles (Q1 y Q3) y rango intercuartílico (IQR). Estos indicadores permitieron caracterizar la tendencia central y la dispersión de las mediciones, así como evaluar la estabilidad del sistema durante el periodo de monitoreo. De manera complementaria, se generaron histogramas de frecuencia para cada variable, con el objetivo de analizar la forma de las distribuciones y detectar posibles asimetrías o concentraciones de valores. Este análisis gráfico permitió reforzar la interpretación de los resultados estadísticos y evaluar visualmente la consistencia de los datos registrados por el sistema.

3. Desarrollo

3.1. Variabilidad observada de la línea base fisicoquímica del agua

Los resultados correspondientes a la variabilidad observada de la línea base fisicoquímica del agua, obtenidas a partir de muestras recolectadas en la fuente hídrica bajo diferentes condiciones climáticas durante los meses de marzo, abril y mayo de 2025, se visualiza en la Tabla 1.

Tabla 1. Parámetros fisicoquímicos de Chuina

Parámetros	Unidad	Muestra 1	Muestra 2	Muestra 3
Temperatura	°C	21,2	24,8	26,5
pH	-	6,72	7,05	7,25
Conductividad eléctrica	mS/cm	0,940	0,957	0,973
TDS	mg/L	616,44	632,29	638,91
Turbidez	NTU	1,10	0,65	0,45

Estos resultados permiten contrastar los valores obtenidos con los límites establecidos por la normativa nacional vigente. La Tabla 2 presenta los valores mínimos y máximos observados, así como su comparación normativa.

Tabla 2. Rango mínimo y máximo observado

Parámetros	Unidad	Valor Mínimo Observado	Valor Máximo Observado	Límite	Cumple
Temperatura	°C	21,2	26,5	15-30	SI
pH	-	6,72	7,25	6,5-8,5	SI

Conductividad eléctrica	mS/cm	0,940	0,973	< 1,5	SI
TDS	mg/L	616,44	638,91	< 1 000	SI
Turbidez	NTU	0,45	1,10	< 5	SI

Los resultados evidencian que los parámetros fisicoquímicos evaluados se mantuvieron dentro de los límites establecidos por la normativa nacional vigente para agua destinada al consumo humano. La temperatura del agua presentó valores entre 21.2 °C y 26.5 °C, asociados a las condiciones ambientales de cada fecha de muestreo, sin registrarse cambios abruptos durante el periodo de evaluación. El pH se mantuvo cercano a la neutralidad, con valores comprendidos entre 6.72 y 7.25, sin desviaciones significativas entre muestras.

Los sólidos totales disueltos (TDS) registraron concentraciones entre 616.44 mg/L y 638.91 mg/L, indicando una presencia moderada y estable de sales disueltas. La turbidez se mantuvo en niveles bajos, con valores entre 0.45 y 1.10 NTU, lo que sugiere una reducida presencia de partículas suspendidas. De manera similar, la conductividad eléctrica presentó valores entre 0.940 y 0.973 mS/cm, evidenciando una concentración iónica estable durante el periodo de muestreo.

3.2. Funcionamiento del sistema IoT y adquisición de datos

Como resultado de las etapas de selección de componentes, diseño de la arquitectura de hardware, ensamblaje del prototipo, calibración de sensores y desarrollo del sistema embebido, se obtuvo un sistema IoT funcional de bajo costo para el monitoreo continuo de la calidad del agua.

El sistema integró sensores de pH, conductividad eléctrica, sólidos totales disueltos, turbidez y temperatura, interconectados mediante una placa de circuito impreso (PCB) con conexiones soldadas y protegidos mediante un encapsulado fabricado por impresión 3D, diseñado para operar en condiciones de campo. Figura 5

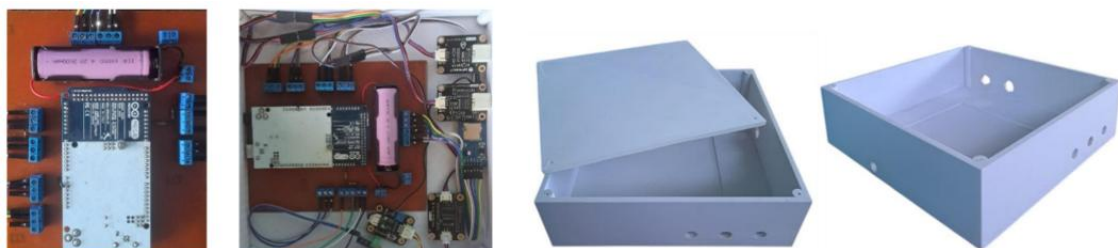


Figura 5. Prototipo final y Case 3D

Asimismo, Los sensores de pH y conductividad eléctrica fueron calibrados utilizando soluciones patrón proporcionadas por el fabricante. Tras la calibración, los errores promedio obtenidos se mantuvieron dentro del rango $\pm 5\%$, considerado aceptable para sensores de bajo costo empleados en aplicaciones de monitoreo ambiental. Estos resultados garantizan la confiabilidad de las mediciones utilizadas en el análisis posterior.

Durante la fase de operación, el sistema ejecutó correctamente el ciclo automático de adquisición, procesamiento, almacenamiento y transmisión de datos. Las mediciones recolectadas fueron almacenadas localmente en formato CSV y transmitidas de manera remota mediante el módulo GSM hacia una hoja de cálculo en la nube. La Tabla 3 Registro automático en Google Sheets presenta un extracto representativo de los registros transmitidos correctamente.

Tabla 3. Registro automático en Google Sheets

Timestamp	Temp (°C)	pH	EC (mS/cm)	TDS (mg/L)	Turbidez (NTU)
2025-05-03 8:00	21.46	6.68	0.97	638.91	0.45
2025-05-03 8:10	21.59	6.68	0.94	616.44	0.45
2025-05-03 8:20	21.62	6.78	0.96	622.93	0.42
2025-05-03 8:30	21.30	6.84	0.97	638.91	0.40

En ausencia de cobertura móvil, los datos fueron almacenados como respaldo en la tarjeta microSD del sistema, garantizando la continuidad y conservación de la información. Este mecanismo de doble almacenamiento permitió asegurar la disponibilidad de los registros para su posterior análisis estadístico y aplicación de técnicas de aprendizaje automático.

3.3. Análisis exploratorio y preprocesamiento de los datos

El conjunto de datos analizado estuvo conformado por 2880 registros, recolectados automáticamente por el sistema IoT durante un periodo de 60 días bajo diferentes condiciones ambientales.

La evaluación de la estructura del dataset permitió verificar la consistencia de las variables y la completitud de los registros. Tal como se muestra en la Tabla 4, ninguna de las variables presentó valores nulos o registros incompletos, por lo que no fue necesario aplicar técnicas de imputación de datos. Este resultado confirma la adecuada operación del sistema IoT y la confiabilidad del proceso automático de adquisición y almacenamiento de información.

Tabla 4. Estructura y completitud del conjunto de datos

Variables	Tipo	Registros validos	Valores nulos
Temp_C	float	2880	0
pH	float	2880	0
CE_mS	float	2880	0
TDS_mgL	float	2880	0
Turbidez_NTU	float	2880	0

Además, La Figura 6 muestra el mapa de correlación de Pearson entre los parámetros fisicoquímicos. Se identificó una fuerte correlación positiva entre la conductividad eléctrica y los sólidos totales disueltos ($r = 0.93$), relación consistente con el comportamiento fisicoquímico del agua. El resto de las variables presentó correlaciones débiles ($|r| < 0.30$), incluyendo el pH, que mostró un comportamiento práctica-mente independiente respecto a los demás parámetros. La temperatura exhibió una correlación negativa moderada con la turbidez ($r = -0.28$).

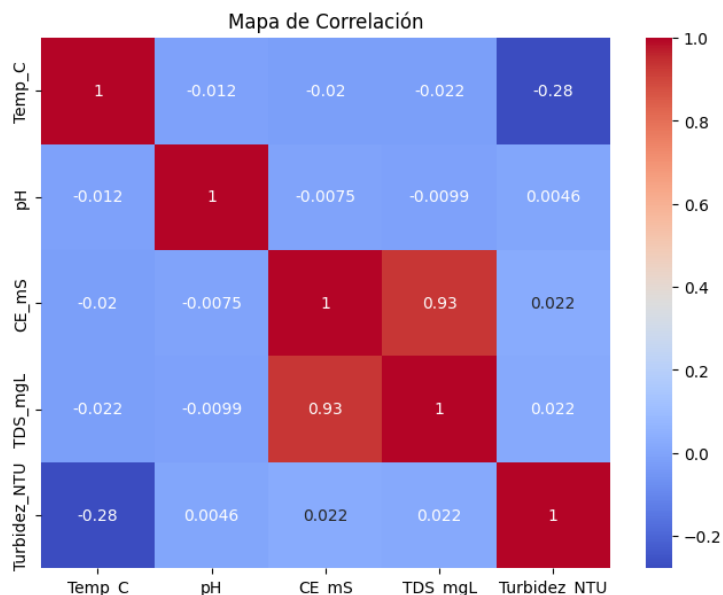


Figura 6. Mapa de correlación de Pearson

La Figura 7 presenta las relaciones bivariadas entre los parámetros mediante un diagrama de dispersión matricial. Los resultados confirman visualmente la relación lineal entre conductividad eléctrica y sólidos totales disueltos, mientras que para las demás combinaciones se observan nubes de puntos dispersas, sin tendencias lineales dominantes. Las distribuciones univariadas evidencian rangos de variación heterogéneos entre variables, lo que justifica la aplicación de procesos de estandarización previos al modelado de agrupamiento.

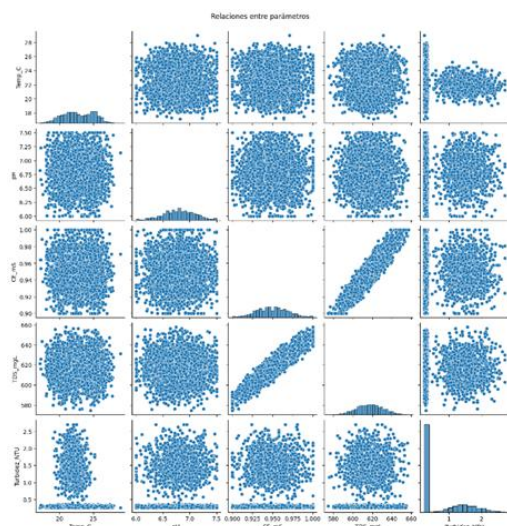


Figura 7. Diagrama de dispersión matricial

3.4. Agrupamiento de patrones mediante K-Means

Una vez preprocesados los datos, se aplicó el algoritmo de agrupamiento K-Means sobre las variables estandarizadas mediante el método StandardScaler, con el fin de asegurar una contribución equitativa de todos los parámetros fisicoquímicos.

La Figura 8 muestra la variación de la suma de errores cuadráticos intra-clúster (SSE) en función del número de clústeres (k), obtenida mediante el método del codo. Se observa una disminución marcada de la SSE para valores bajos de k , seguida de una reducción progresivamente menor a partir de $k = 3-4$, lo que indica una zona de inflexión amplia más que un punto único claramente definido.

Si bien el método del codo sugiere una mejora adicional al incrementar el número de clústeres, la diferencia marginal de la SSE entre $k = 3$ y $k = 4$ es relativamente reducida en comparación con las disminuciones observadas para valores menores de k . Se seleccionó $k = 3$ como número final de clústeres, priorizando la simplicidad del modelo y la interpretabilidad física de los grupos, sin una pérdida significativa en la compactación intra-clúster. Esta elección permitió identificar tres patrones claramente diferenciados de calidad del agua, evitando la fragmentación excesiva del conjunto de datos y facilitando el análisis e interpretación de los comportamientos dominantes durante el periodo de monitoreo.

```
kmeans = KMeans(n_clusters=k, random_state=42)
```

```
kmeans.fit(X_scaled)
```

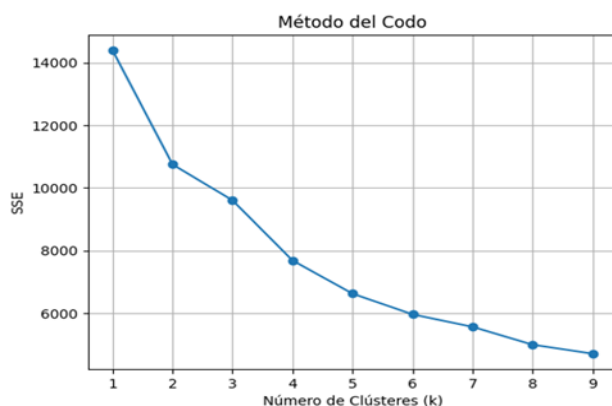


Figura 8. Gráfico del método del codo

3.5. Validación del agrupamiento

La calidad del agrupamiento obtenido con $k = 3$ fue evaluada mediante el coeficiente de silueta, el cual cuantifica simultáneamente la cohesión intra-clúster y la separación inter-clúster. El valor promedio del coeficiente de silueta obtenido fue ≈ 0.27 , indicando una estructura de agrupamiento moderada, característica de conjuntos de datos ambientales multivariados con transiciones graduales entre estados.

La Figura 9 presenta el gráfico de silueta para los tres clústeres identificados. En general, la mayoría de las observaciones presenta valores de silueta positivos, lo que evidencia una asignación adecuada de los registros a sus respectivos clústeres. Asimismo, los tres grupos muestran distribuciones de silueta relativamente equilibradas, sin clústeres dominados por valores negativos, lo que sugiere una separación consistente entre ellos

```
df['Cluster'] = kmeans.fit_predict(X_scaled)
```

```
score = silhouette_score(X_scaled, df['Cluster'])
```

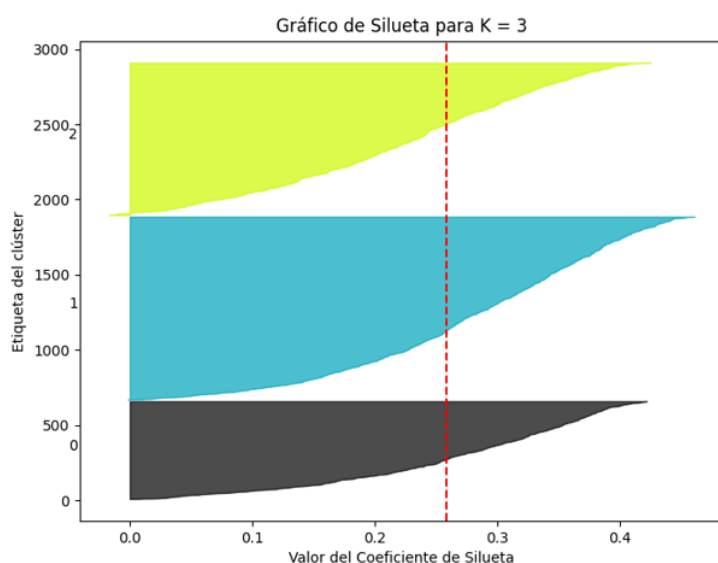


Figura 9. Gráfico de Silhouette

Si bien algunos registros presentan coeficientes de silueta cercanos a cero, estos corresponden a observaciones ubicadas en zonas de transición entre clústeres, un comportamiento esperado en sistemas naturales como la calidad del agua. Estos resultados confirman que la solución con $k = 3$ proporciona un compromiso adecuado entre cohesión, separación e interpretabilidad, respaldando su selección como número final de clústeres.

3.6. Visualización e interpretación de clústeres mediante PCA

Al aplicar el PCA para reducir la dimensionalidad, la Figura 10 muestra la contribución de cada variable original a los componentes principales. El primer componente principal (PC1) está dominado por la conductividad eléctrica (CE) y los sólidos totales disueltos (TDS), con cargas elevadas y positivas (≈ 0.71), indicando que este componente representa principalmente la variabilidad asociada al contenido iónico del agua. En contraste, el segundo componente principal (PC2) está influenciado principalmente por la temperatura, con una carga positiva elevada (≈ 0.71), y por la turbidez, con una carga negativa significativa (≈ -0.70), lo que sugiere una inclinación asociada a condiciones físicas del sistema.

```
pca = PCA(n_components=2)
```

```
X_pca = pca.fit_transform(X_scaled)
```

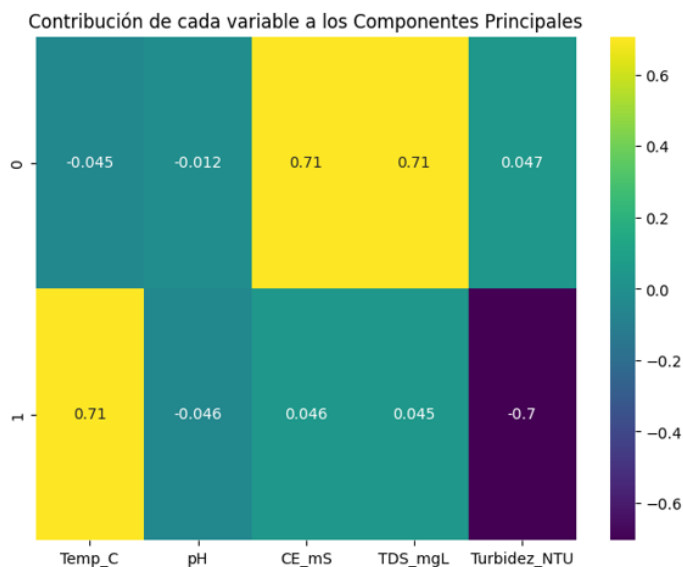


Figura 10. Mapa de cargas del Análisis de Componentes Principales

La Figura 11 muestra la proyección de los datos en el espacio bidimensional definido por los dos primeros componentes principales, junto con la asignación de clústeres obtenida mediante K-means con $k = 3$. Se observa una separación visual consistente entre los tres clústeres, con patrones diferenciados a lo largo de ambos componentes principales. Esta distribución confirma que la estructura del agrupamiento se mantiene tras la reducción de dimensionalidad y respalda la interpretación de los clústeres como comportamientos distintos de la calidad del agua durante el periodo de monitoreo.

```
kmeans_pca = KMeans(n_clusters=k, random_state=42)
```

```
clusters_pca = kmeans_pca.fit_predict(X_pca)
```

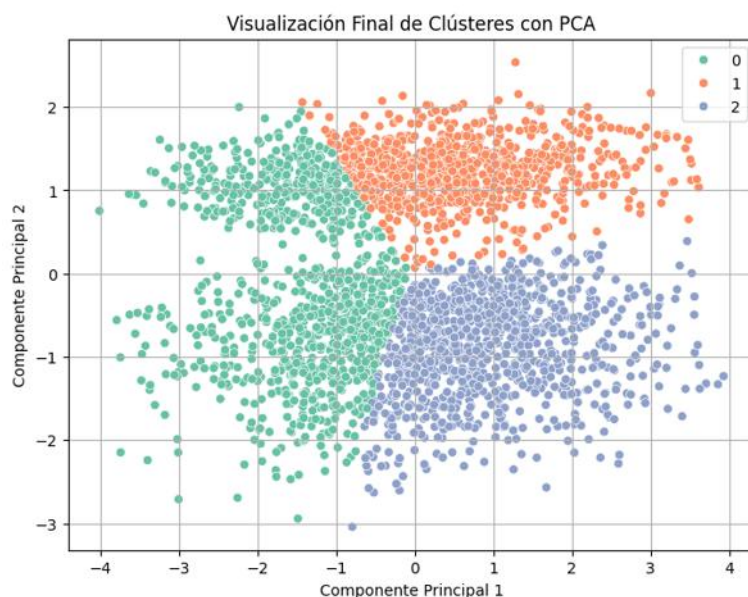


Figura 11. Proyección de clústeres en el espacio de Componentes Principales

3.7. Comportamiento del sistema IoT durante el monitoreo de la calidad del agua

Para evaluar la efectividad del sistema IoT desarrollado, se aplicó un enfoque cuantitativo basado en estadística descriptiva. La Tabla 5 Resumen de estadísticos descriptivos presenta el resumen

de los estadísticos descriptivos. Los valores medios obtenidos fueron 22.82 °C (temperatura), 6.80 (pH), 0.95 mS/cm (conductividad eléctrica), 617.29 mg/L (sólidos totales disueltos) y 0.75 NTU (turbidez). En general, los parámetros mostraron baja dispersión, lo que indica un comportamiento estable del sistema durante el periodo de monitoreo.

Tabla 5. Resumen de estadísticos descriptivos

Estadístico	Temp_°C	pH	CE_mS	TDS_mgL	Turbidez_NTU
N	2880	2880	2880	2880	2880
Mínimo	17,14	6	0,9	574,51	0,25
Máximo	28,97	7,5	1	657,64	2,71
Media	22,82	6,8	0,95	617,29	0,75
Desv. estándar	2,22	0,3	0,02	13,87	0,65

El análisis intercuartílico confirmó que la mayoría de las mediciones se concentró dentro de rangos acotados para todos los parámetros, evidenciando la consistencia de los datos registrados. La figura 12 muestra los histogramas de frecuencia, donde se observan distribuciones predominantemente unimodales para temperatura, pH, conductividad eléctrica y sólidos totales disueltos, mientras que la turbidez presenta una asimetría positiva asociada a eventos aislados.

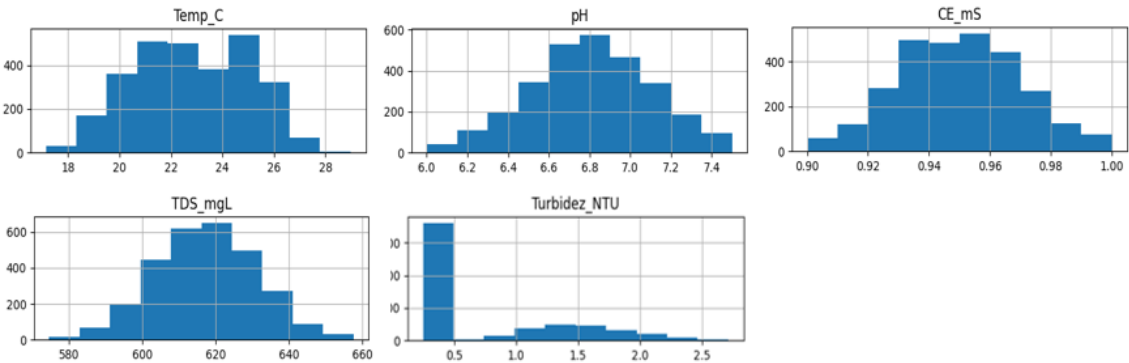


Figura 12. Histogramas de frecuencia

Discusión

Los resultados obtenidos evidencian que los parámetros fisicoquímicos del agua monitoreada en la comunidad rural Chuina presentan una variabilidad consistente con condiciones naturales propias de cuerpos de agua subterránea poco alterados. La temperatura registrada se mantuvo dentro de rangos comparables a los reportados por [22] en estudios de aguas subterráneas, lo que sugiere un comportamiento térmico acorde con contextos hidrogeológicos similares. Asimismo, [23] señalaron que variaciones en la temperatura pueden introducir sesgos en parámetros como el pH y la conductividad cuando las muestras no son analizadas oportunamente; en este sentido, la estabilidad observada respalda la efectividad de los procedimientos de muestreo y medición empleados en el presente estudio.

El pH mostró valores cercanos a la neutralidad, coherentes con los rangos documentados por [14,23] en cuerpos de agua naturales sin tratamiento. Estos autores destacan que procesos de almacenamiento prolongado pueden inducir acidificación del agua, lo cual no fue evidenciado en este estudio, reforzando la consistencia de las mediciones realizadas. De manera similar, los valores de sólidos disueltos totales se ubicaron dentro del rango inferior-medio reportado por [22], contrastando con los amplios intervalos observados [23] en aguas residuales, lo que sugiere una baja influencia antrópica en la zona evaluada.

La turbidez presentó valores predominantemente bajos, comparables a los reportados para agua potable en estudios previos [23], aunque con variaciones puntuales que concuerdan con lo descrito por [20], quienes atribuyen estos incrementos a eventos de arrastre superficial o remoción de sedimentos. En conjunto con la conductividad eléctrica, que mostró un comportamiento

estable y coherente con la composición iónica natural del agua [23], estos resultados confirman la consistencia interna del conjunto de datos y la capacidad del sistema para reflejar condiciones reales del entorno.

Desde el punto de vista tecnológico, el sistema IoT de bajo costo desarrollado incorpora elementos que lo diferencian de arquitecturas previamente reportadas. A diferencia de soluciones centradas únicamente en la adquisición de datos [22], el uso de una placa de circuito impreso y un encapsulado personalizado fabricado mediante impresión 3D contribuyó a una mayor robustez operativa en campo. Adicionalmente, las curvas de calibración obtenidas mostraron errores dentro de los márgenes aceptables para sensores de bajo costo y coeficientes de determinación cercanos a la unidad, superando la limitada caracterización metrológica reportada en estudios como el de [20].

La estrategia dual de transmisión y almacenamiento de datos implementada permitió garantizar la continuidad del registro incluso en escenarios de conectividad limitada, aspecto que no siempre es considerado en investigaciones similares [14]. Asimismo, la incorporación de técnicas de aprendizaje no supervisado permitió identificar patrones inherentes en los datos fisicoquímicos sin requerir etiquetado previo, complementando enfoques supervisados utilizados en trabajos recientes [22].

Finalmente, la evaluación normativa evidenció que la mayoría de los parámetros se mantuvieron dentro de los límites permisibles, mientras que la detección de valores fuera de rango en el pH demuestra la sensibilidad del sistema para identificar condiciones anómalas. Este comportamiento concuerda con observaciones realizadas por [24] quienes reportaron baja variabilidad y estabilidad en sistemas de monitoreo comparables. En conjunto, los resultados confirman que el sistema IoT propuesto constituye una herramienta viable para el seguimiento continuo de la calidad del agua en contextos rurales, aportando información confiable y consistente para la gestión del recurso hídrico.

4. Conclusiones

Los resultados obtenidos evidencian que los parámetros fisicoquímicos del agua monitoreada en la comunidad rural de Chuina presentaron una variabilidad consistente con las condiciones propias de los cuerpos de agua evaluados. La temperatura registrada se mantuvo dentro de rangos comparables con los reportados por Venkatesh et al. (2024) en sistemas de monitoreo de aguas subterráneas. Asimismo, Jabbar et al. (2024) destacan la importancia del monitoreo continuo de los parámetros de calidad del agua mediante sistemas IoT, particularmente en contextos rurales. En este sentido, la estabilidad observada durante el periodo de evaluación respalda la capacidad del sistema implementado para registrar las variaciones de las condiciones ambientales.

El pH presentó valores predominantemente cercanos a la neutralidad, comportamiento comparable con los rangos registrados mediante sistemas IoT para el monitoreo de cuerpos de agua (Lal et al., 2024; Jabbar et al., 2024). De manera similar, los sólidos disueltos totales mostraron un comportamiento relativamente estable durante el periodo de monitoreo. Estos resultados, considerados conjuntamente con los demás parámetros fisicoquímicos, sugieren que el sistema fue capaz de capturar las variaciones presentes en la fuente hídrica y generar registros continuos susceptibles de análisis.

La turbidez presentó valores predominantemente bajos, aunque se identificaron variaciones puntuales durante el periodo de observación. Este comportamiento resulta relevante debido a que los sistemas de monitoreo continuo permiten detectar cambios temporales que podrían pasar inadvertidos mediante mediciones aisladas (Lu et al., 2024). Por su parte, la conductividad eléctrica mostró un comportamiento relativamente estable, constituyendo un parámetro complementario para caracterizar las variaciones de la calidad del agua. La medición conjunta de

pH, TDS, temperatura y turbidez coincide con configuraciones empleadas en sistemas IoT orientados a la evaluación de la calidad del agua (Sugiharto et al., 2023).

Desde el punto de vista tecnológico, el sistema IoT de bajo costo desarrollado incorporó una arquitectura orientada a garantizar la adquisición continua de los parámetros fisicoquímicos. A diferencia de configuraciones basadas principalmente en redes de sensores para la adquisición y transmisión de información (Venkatesh et al., 2024), el prototipo desarrollado incorporó una placa de circuito impreso y un encapsulado personalizado fabricado mediante impresión 3D, lo que favoreció su integración física para las condiciones de operación en campo. Asimismo, las curvas de calibración obtenidas permitieron evaluar el comportamiento de los sensores utilizados y verificar su respuesta durante la etapa de implementación.

La estrategia de transmisión y almacenamiento de datos permitió mantener la continuidad de los registros ante las limitaciones de conectividad propias del entorno rural, aspecto especialmente relevante para la implementación de sistemas IoT en este tipo de escenarios. En este sentido, Lal et al. (2024) demostraron la viabilidad de sistemas IoT de bajo costo para el monitoreo continuo de la calidad del agua. Adicionalmente, en el presente estudio, la incorporación del algoritmo K-Means permitió identificar patrones en el comportamiento conjunto de los parámetros fisicoquímicos sin requerir categorías previamente etiquetadas, ampliando el análisis de los datos generados por el sistema.

Finalmente, la evaluación realizada permitió identificar tanto periodos de estabilidad como registros con variaciones en los parámetros monitoreados, particularmente en el pH. La capacidad de registrar estas fluctuaciones constituye una característica relevante de los sistemas de monitoreo en tiempo real, debido a que posibilita detectar cambios en las condiciones del agua durante el periodo de observación. En este sentido, Sugiharto et al. (2023) también emplearon sistemas IoT para el seguimiento en tiempo real de parámetros como pH, TDS, temperatura y turbidez. En conjunto, los hallazgos respaldan la viabilidad tecnológica del sistema IoT propuesto como herramienta de monitoreo continuo de la calidad del agua en contextos rurales, aunque su utilización para decisiones sanitarias requiere considerar los alcances de los sensores empleados y su validación frente a métodos de referencia.

Financiamiento

Ninguno.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Contribución de autores

Los autores declaran haber realizado todas las etapas de la investigación y redacción del Manuscrito.

Referencias bibliográficas

Al-Kassab-Córdova, A., Silva-Perez, C., Robles-Valcarcel, P., Bendezu-Quispe, G., Ortiz, A. I., & Benites-Zapata, V. A. (2023). Spatial and socioeconomic inequalities in the access to safe drinking water in Peruvian households. *Journal of Water and Health*, 21, 525–535. <https://doi.org/10.2166/wh.2023.316>

Amador-Castro, F., González-López, M. E., Lopez-Gonzalez, G., Garcia-Gonzalez, A., Díaz-Torres, O., Carbajal-Espinosa, O., & Gradilla-Hernández, M. S. (2024). Internet of Things and

citizen science as alternative water quality monitoring approaches and the importance of effective water quality communication. *Journal of Environmental Management*, 352, 119959. <https://doi.org/10.1016/j.jenvman.2023.119959>

Arsene, D., Predescu, A., Pahonțu, B., Chiru, C. G., Apostol, E.-S., & Truică, C.-O. (2022). Advanced strategies for monitoring water consumption patterns in households based on IoT and machine learning. *Water*, 14, 2187. <https://doi.org/10.3390/w14142187>

Autoridad Nacional del Agua. (s. f.). ANA optimiza la gestión ambiental de los vertimientos por ríos más saludables en el Perú. Gobierno del Perú.

Carvalho, A. K. N. D., Souza, R. F. D., & De Oliveira, S. D. S. (2022). Qualidade de águas superficiais e subterrâneas para irrigação em um município do semiárido do estado do Rio Grande do Norte. *Pesquisas em Geociências*, 49, e119720. <https://doi.org/10.22456/1807-9806.119720>

Fondo de las Naciones Unidas para la Infancia. (s. f.). La triple amenaza: La combinación de las enfermedades, los riesgos climáticos y la precariedad de los servicios de agua, saneamiento e higiene puede tener repercusiones mortales para la infancia. UNICEF.

Ghachoui, E., Taouil, H., Naser, R. S. M., & Said, I. A. (2024). Spatial assessment of groundwater quality for drinking in Souk El Arbaa, Morocco. *Water Supply*, 24, 1267–1281. <https://doi.org/10.2166/ws.2024.059>

Gomes, R., Pereira, C., Galvão, J., & Ribeiro, V. (2024). Drinking water: Strategies for sustainable management and water for everyone. En J. R. D. C. S. Galvão, P. Brito, F. D. S. Neves, H. D. A. Almeida, S. D. J. M. Mourato, & C. Nobre (Eds.), *Proceedings of the 3rd International Conference on Water Energy Food and Sustainability (ICoWEFS 2023)* (pp. 709–715). Springer Nature Switzerland.

Gupta, M., Tanwar, S., & Jain, V. (2023). Potential of IoT for water monitoring to upgrade food quality. En A. E. Hassanien & M. Soliman (Eds.), *Artificial intelligence: A real opportunity in the food industry* (Vol. 1000, pp. 119–139). Springer International Publishing.

Instituto Nacional de Estadística e Informática. (s. f.). Perú: Formas de acceso al agua y saneamiento básico.

Jabbar, W. A., Mei Ting, T., I. Hamidun, M. F., Che Kamarudin, A. H., Wu, W., Sultan, J., Alsewari, A. A., & Ali, M. A. H. (2024). Development of LoRaWAN-based IoT system for water quality monitoring in rural areas. *Expert Systems with Applications*, 242, 122862. <https://doi.org/10.1016/j.eswa.2023.122862>

Jaramillo Vera, E. N., & Marines Batalla, P. (2023). Enfermedades hídricas que afectan a la población como efecto del uso y consumo de agua disponible: Una mirada desde las acciones de enfermería. *Más Vita. Revista de Ciencias de la Salud*, 5, 29–43. <https://doi.org/10.47606/ACVEN/MV0210>

Karthikeyan, V., Visu, Y. P., & Kalaimannan, E. (2024). Smart system for wastewater treatment (SSWWT): Internet of Things (IoT)-based intelligent framework for society. En *Sustainability principles and applications in engineering practice* (pp. 57–77).

Kumar, D. S., Askarunisa, A., & Kumar, R. M. (2020). Embedded processor based automated assessment of quality of the water in an IoT background. *Microprocessors and Microsystems*, 77, 103167. <https://doi.org/10.1016/j.micpro.2020.103167>

- Lal, K., Menon, S., Noble, F., & Arif, K. M. (2024). Low-cost IoT based system for lake water quality monitoring. *PLOS ONE*, 19, e0299089. <https://doi.org/10.1371/journal.pone.0299089>
- Lu, Q., Zou, J., Ye, Y., & Wang, Z. (2024). Design and implementation of a Li River water quality monitoring and analysis system based on outlier data analysis. *PLOS ONE*, 19, e0299435. <https://doi.org/10.1371/journal.pone.0299435>
- Montero Contreras, C. V. (2022). Impacto de los servicios de agua y saneamiento sobre las enfermedades diarreicas en los niños de la sierra del Perú. *Revista de Análisis Económico y Financiero*, 5, 1–8. <https://doi.org/10.24265/raef.2022.v5n1.45>
- Nieto Pacheco, A. J., Villegas Pilay, Y. E., & Galarza Alay, J. I. (2023). Avances actuales de las IoT y sus diversas aplicaciones. *ConcienciaDigital*, 6, 58–74. <https://doi.org/10.33262/concienciadigital.v6i1.2440>
- Organización Mundial de la Salud. (s. f.). Agua para consumo humano.
- Paredes, C. L., & Woolley, K. (2022). ¿Agua para todos? Differences in access to clean water and a bathroom at home by ethno-racial characteristics in contemporary Peru. *Sociology of Development*, 8, 450–484. <https://doi.org/10.1525/sod.2021.0044>
- Qian, P., & Liu, G. (2024). Application of high-frequency intelligent sensing network in monitoring and early warning of water quality dynamic change. *International Journal of Computational Intelligence Systems*, 17, 195. <https://doi.org/10.1007/s44196-024-00605-6>
- Salgado, J., Pizarro, C., Wong, L., & Castillo, J. (2022). IoT Watercare: Water quality control system in unofficial settlements of Peru based in an IoT architecture. En 2022 31st Conference of Open Innovations Association (FRUCT) (pp. 277–288). IEEE.
- Siren, S., Pec, R., Srun, C., Meas, S., Ny, V., Eang, S., & Sari, P. (2024). An IoT prototype for water quality monitoring system with case study in NPIC campus. *International Journal of Electrical and Electronics Research*, 12, 1407–1417. <https://doi.org/10.37391/ijeer.120435>
- Sugiharto, W. H., Susanto, H., & Prasetyo, A. B. (2023). Real-time water quality assessment via IoT: Monitoring pH, TDS, temperature, and turbidity. *International Journal of Safety and Security Engineering*, 28, 823–831. <https://doi.org/10.18280/isi.280403>
- Venkatesh, J., Partheeban, P., Baskaran, A., Krishnan, D., & Sridhar, M. (2024). Wireless sensor network technology and geospatial technology for groundwater quality monitoring. *Journal of Industrial Information Integration*, 38, 100569. <https://doi.org/10.1016/j.jii.2024.100569>