**How to Cite:**
Jáquez, A.D.B.; Herrera, M.T.A.; Celestino, A.E.M.; Ramírez, E.N.; Cruz, D.A.M. Extension of LoRa Coverage and Integration of an Unsupervised Anomaly Detection Algorithm in an IoT Water Quality Monitoring System. Water 2023, 15, 1351. https://doi.org/10.3390/w15071351

water

Article

# Extension of LoRa Coverage and Integration of an Unsupervised Anomaly Detection Algorithm in an IoT Water Quality Monitoring System

**Armando Daniel Blanco Jáquez [1], María T. Alarcon Herrera [1], Ana Elizabeth Marín Celestino [2], Efraín Neri Ramírez [3] and Diego Armando Martínez Cruz [4,*]**

[1] Centro de Investigación en Materiales Avanzados, Departamento de Ingeniería Sustentable, Calle CIMAV 110, Ejido Arroyo Seco, Durango 34147, Mexico; armando.blanco@cimav.edu.mx (A.D.B.J.); teresa.alarcon@cimav.edu.mx (M.T.A.H.)
[2] CONACYT-Instituto Potosino de Investigación Científica y Tecnológica, A.C. División de Geociencias Aplicadas, Camino a la Presa San José 2055, Col. Lomas 4ta Sección, San Luis Potosí 78216, Mexico; ana.marin@ipicyt.edu.mx
[3] Facultad de Ingeniería y Ciencias, Universidad Autónoma de Tamaulipas (UAT), Centro Universitario Victoria, Adolfo López Mateos S/N, Ciudad Victoria 87120, Mexico; eneri@docentes.uat.edu.mx
[4] CONACYT-Centro de Investigación en Materiales Avanzados, S.C. Calle CIMAV 110, Ejido Arroyo Seco, Col. 15 de Mayo, Durango 34147, Mexico
* Correspondence: diego.martinez@cimav.edu.mx; Tel.: +52-614-439-4898

**Abstract:** High cost, long-range communication, and anomaly detection issues are associated with IoT systems in water quality monitoring. Therefore, this work proposes a prototype for a water quality monitoring system (IoT-WQMS) based on IoT technologies, which include in the system architecture a LoRa repeater and an anomaly detection algorithm. The system performs the data collection, data storage, anomaly detection, and alarm sending remotely and in real-time for the information to be captured by the multisensor node. The LoRa repeater allowed the spatial coverage of the LoRa communication to extend, making it possible to reach a place where originally there was no coverage with a single LoRa transmitter due to topography and line of sight. The prototype performed well in terms of packet loss rate, transmission time, and sensitivity, extending the long-range wireless communication distance. Indoor multinode testing validation for 29 days of the mean absolute error for average relative errors of water temperature, pH, turbidity, and total dissolved solids (TDS) were 0.65%, 0.30%, and 14.33%, respectively. The anomaly detector identified all erroneous data events due to node sensor recalibration and water recirculation pump failures. The IoT-WQMS increased the reliability of monitoring through the timely identification of any sensor malfunctions and extended the LoRa signal range, which are relevant features in the scope of in situ and real-time water quality monitoring.

**Keywords:** LoRa; anomaly detection; water quality; IoT; real-time monitoring

check for updates

**Citation:** Jáquez, A.D.B.; Herrera, M.T.A.; Celestino, A.E.M.; Ramírez, E.N.; Cruz, D.A.M. Extension of LoRa Coverage and Integration of an Unsupervised Anomaly Detection Algorithm in an IoT Water Quality Monitoring System. *Water* **2023**, *15*, 1351. https://doi.org/10.3390/w15071351

Academic Editor: Roko Andricevic

Received: 22 February 2023
Revised: 26 March 2023
Accepted: 27 March 2023
Published: 1 April 2023



## 1. Introduction

Traditional water quality monitoring mainly uses personnel to take manual water samples, which causes several limitations. These often include high costs, in terms of time and human resources. It can also be difficult to collect samples from remote or hard-to-reach locations [1–3]. Additionally, manual sampling can only provide a snapshot of the water quality at a specific time, rather than continuous monitoring. There is an inability to conduct trend analysis based on historical data and it is difficult to monitor and control sewage discharge in real-time [4,5]. In response to the above problems, researchers have proposed applications for advancements in water quality sensors, wireless communication networks (WSN), and internet of things technology in the field of water quality monitoring to achieve automatic and remote online monitoring of water quality [6–8]. However, IoT solutions for water quality monitoring are generally expensive [9,10], with high energy consumption [11,12], low computational efficiency [13,14], distance communication issues [15], and outlier detection methods issues for the generated time series [16,17].

ZigBee is often used in water quality monitoring as a local network communication and GPRS technology for remote network communication [18,19]. ZigBee transmission distances of 10 to 100 m cannot meet the requirements for monitoring large areas. Increasing the ZigBee communication distance by adding additional routing nodes will increase the power consumption of the device and reduce latency and reliability. GPRS communication is billed on a per-data throughput basis, meaning the costs of this transmission are high, and the power consumption of the device is significant.

LoRa is a low-power wide area network communication technology, which was developed based on the lower frequency band below sub-GHz, with long-distance communication capabilities, low power consumption, and low complexity. The maximum transmission distance in the open visual distance can reach 15 km [5]. LPWAN (low-power, wide-area network) technology allows for long-range communication with low-power devices, making it ideal for IoT devices in remote or hard-to-reach areas [20,21]. By applying LoRa technology, the transmission radius of the water quality nodes is extended, similar to other technologies, such as ZigBee and WIFI, where the radius is smaller. As a result, it can improve the high costs and energy requirements associated with IoT systems. These characteristics make LPWAN an essential element for the internet of things in environmental monitoring.

IoT water quality monitoring system is a field data observation technology that collects extensive, continuous, complex, and diverse data in real time. However, due to the technical limitations of the sensors and environmental disturbances in the field, abnormalities in sensor readings and null values are inevitable. In view of the above, IoT monitoring systems have incorporated real-time data processing services to recognize anomalous events that allow cleaning and transform them to achieve continuous and accurate observations of water quality conditions in the monitored area. The processing of massive amounts of data continuously obtained from IoT systems is an aspect that still needs to be addressed [22]. Automatic anomaly detection methods are mainly used for this task and are mainly divided into two methods: statistical and physical [17,23]. Among the statistical methods, the most common is determining whether the observed values exceed the range of three times the standard deviation [24]. The physical methods are mainly distance-based or density-based anomaly detections [25]. In addition, in recent years, with the development of data mining algorithms, artificial neural networks, support vector machines, Bayesian networks, and other intelligent recognition algorithms have been widely used [16,26–28].

In this work, we developed a water quality monitoring prototype (IoT-WQMS) based on IoT technologies for surface water bodies. We included a module for anomaly detection based on unsupervised machine-learning techniques for automatic and real-time monitoring of water quality parameters. The prototype allows in situ pH, total dissolved solids, and temperature measuring. This was achieved through the implementation of a low-power, wide-area network (LPWAN), which allows the transmission of information over a distance of approximately 6.5 km through radiofrequency, before sending this information through the Internet. Machine-learning-based algorithm services were incorporated into the prototype to analyze the information captured by the sensors in real time and detect anomalies in their behavior. To alert users of such anomalies, the program sends messages through the Telegram application, and they are also displayed on the webpage, depicting the value and date of the detected event.

The WQSM was experimentally tested so that it was carried out indoor multinode testing, anomaly detection module validation, and outdoor long-distance communication tests, while the constructed system was experimentally validated and demonstrated good system performance in the implemented systems. The coverage of the LPWAN network was extended with the addition of a repeater to the main gateway, thus, covering areas where the signal of a single gateway could not reach. As result, system performance was satisfactory at a lower cost compared to similar alternatives on the market.

## 2. Materials and Methods

### *2.1. The General Design of Water Quality Monitoring System (IoT-WQMS)*

In Durango, Mexico, continuous, onsite, and real-time monitoring of the water quality in the city's dams, one of the primary sources of water supply in this arid city, was being sought. Therefore, a LoRa-based IoT-WQMS was designed and developed for monitoring water bodies within a range of approximately 4 to 7.5 km from the gateway (Figure 1).

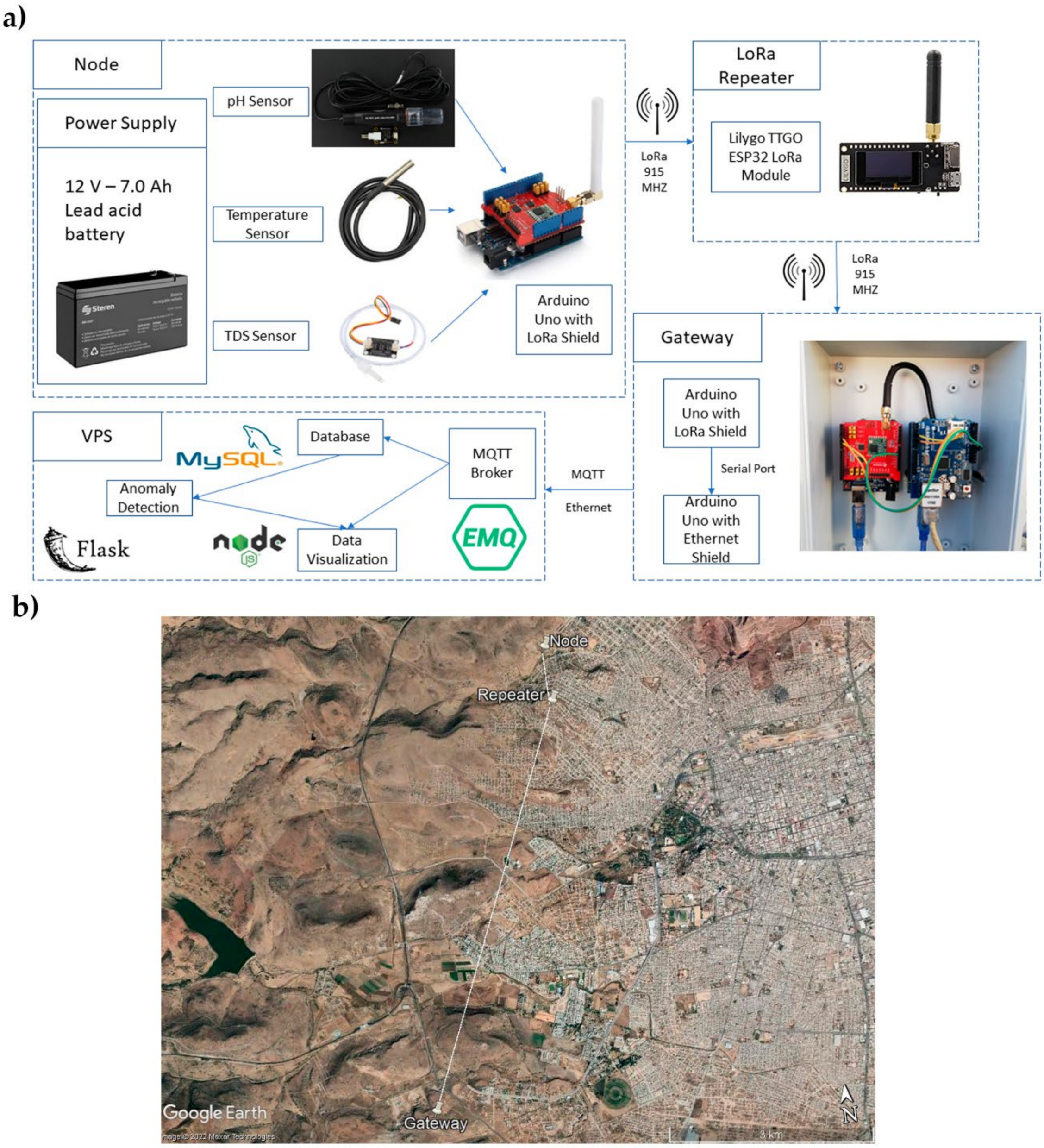


**Figure 1.** (**a**) The general system architecture diagram of IoT-WQSM; (**b**) onsite installation diagram of IoT-WQSM.

IoT-WQMS comprises four layers: monitoring node, LoRa transmitter, server, and user interface. The general structure of the system is shown in Figure 1. Water monitoring node includes a quality collection terminal integrated with water temperature, pH, TDS sensors, and an SX1276 LoRa module. Through the LoRa module, the data were collected by the sensors and sent to the LoRa repeater, and then, to the gateway at regular intervals. The gateway is the center of the star topology network architecture, and the data uploaded by multiple water quality collection terminals were packaged and sent to the server. The server was mainly responsible for managing the gateway, processing the data uploaded by the gateway, by the monitoring terminal ID into the corresponding database. Users login through the local monitoring computer on the webpage to enable real-time monitoring of the water quality data and assessment of the water quality status through the user interface. The system provided the possibility of generating early warnings in the presence of anomalous events or out-of-range values of monitored parameters.

### 2.2. Water Quality Monitoring Hardware Design

The hardware of the water quality collection terminal is shown in Figure 2. The collection terminal node consisted of a microcontroller unit (MCU), a sensor module, a LoRa RF module, and a power supply module. The LoRa repeater and gateway mainly consisted of a module for each MCU and LoRa RF. Among them, the microcontroller was an MCU Arduino Uno, which only has a power consumption of 10 mA current in power-down sleep, with 5 V-16 MHz Version, satisfying the low-power design scheme of this system [29]. The energy consumption of the whole acquisition terminal depended mainly on the working cycle of the system, during which the acquisition of sensor data, the processing of the microcontroller, and the sending of data by the RF module were completed. The hardware design process mainly adopted a polled wake-up mode of operation for the RF module to reduce the energy loss of the acquisition terminal.

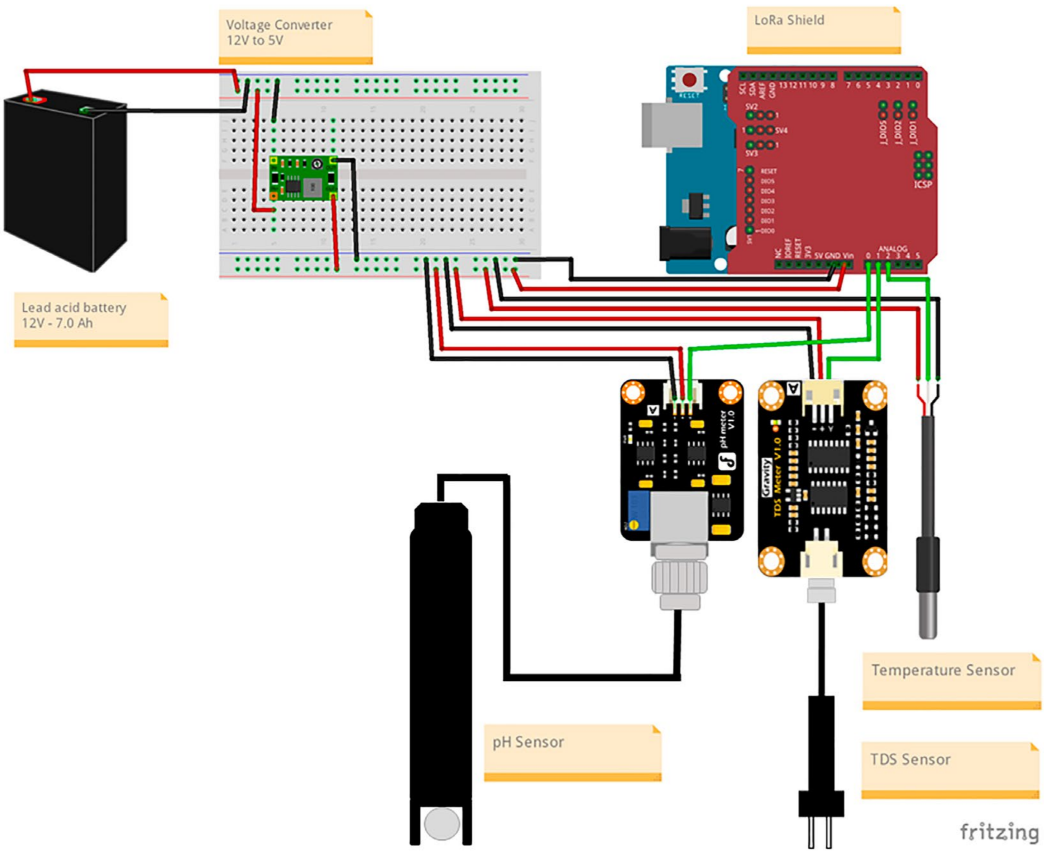


**Figure 2.** LoRA node connection diagram.

#### 2.2.1. Microcontroller Modules

The node had a central unit with an Arduino Uno board, which used the AT-mega328P microcontroller unit (MCU) (Arduino CC, Turin, Italy). A chip with an AVR CPU core and a maximum speed of 16 MHz. The product family features 32 kB of flash memory, multiple control peripherals, and a USB full-speed interface [30]. Using the chip's powerful storage and computing capabilities, the PA6, PA7, and PA8 analog PINS were set to correspond to the water quality sensor for analog to digital (A/D) acquisition, calculate the water quality PH value, temperature, and TDS value, the data were packaged and sent to the radio frequency module through the serial port. The LoRa repeater used a channel on the TTGO ESP32 development board; TTGO stands for "To Be The Best Go", which is a Chinese brand that produces various types of development boards based on the ESP32 chip [31]. The ESP32 is a low-cost, ultra-low power system on a chip, MCU with a Tensilica 32-bit Xtensa LX6 dual-core embedded microprocessor, which includes Wi-Fi and Bluetooth connectivity. The gateway consisted of two Arduino microcontrollers and one connected by a serial port. One was to establish communication over the LoRa layer and the other was for the Ethernet connection.

### 2.2.2. Sensor Module

The sensor module mainly included sensors and signal conditioning circuits, which were selected according to technical factors, such as the location of the sensors, cost, reliability, and power consumption in the actual application. In this system, three main types of sensors were connected: temperature, pH, and TDS.

The system uses the SKU DFR0198 temperature sensor from [32], which has the characteristics of small size, high accuracy, and low overhead hardware, while a waterproof stainless-steel package was selected for the system to measure the temperature of the water body, from −55 °C to 125 °C, with an accuracy of ±0.5 °C. The module had a configurable 9 to 12-bit resolution and the data were transmitted through the 1-Wire interface; further, it had a 64-bit unique identifier so that multiple sensors could share the same transmitting pin, a query time of less than 750 ms, and a current consumption of less than 3 µA. The pH sensor adopted the Gravity: Analog pH meter pro V2 composite electrode (Copyright Zhiwei Robotics Corp., Shanghai, China), with a pH detection range of 0 to 14 and a high measurement accuracy of 0.01, and stability. The composite electrode was a combination of a pH indicator electrode and a reference electrode, which output via a millivolt (mV) voltage signal by shifting the electrode's electric potential through a change in hydrogen ion activity. This signal was processed by an amplifier circuit based on the CA3140EZ operational amplifier, provided with the probe. This circuit amplified the output signal of the pH electrode from a millivolt voltage and converted it into the range of 0–5 V output voltage to reach the readable voltage required by the analog input of the microcontroller. The sensor can work in a voltage range of 3.3 to 5.5 V and handle an output signal of 0 to 3 V, with a current consumption of approximately 3 mA. The included software library allowed calibration using the two-point method, automatically detecting when the electrode was immersed in a pH solution of 4 or 7. The module had a response time of less than 1 min and a probe lifetime of 6 months when working for 24 h a day. The system used Tresd Print Tech online water quality TDS detection sensor (Copyright Tresd print tech Ltd., Guadalajara, Mexico) and the GravityTDS library (Copyright Zhiwei Robotics Corp., Shanghai, China), to convert the voltage to mg/L dissolved solids. This was used to detect the content of the total dissolved solids (TDS) in the water to determine the degree of cleanliness or pollution of the water. The operating voltage of the sensor was 3.3 to 5 V, with an output signal of 0 to 2.3 V, and a current consumption between 3 and 6 mA. The measurement range was 0 to 1000 mg/L, with an accuracy of ±10% at 25 °C. The GravityTDS library was used for voltage to mg/L conversion of dissolved solids. The estimated value was compensated based on the temperature measurement obtained by the system temperature sensor. The library allowed a one-point calibration of the sensor in a 47 mg/L solution.

### 2.2.3. LoRa Modules

In the node a Dragino Shield (Dragino Technology Co., Ltd., Shenzhen, China) was included, this included the SX1276 chip, which allowed us to send and receive messages by radio frequency at 915 MHz, using the LoRa protocol. The program included the

LoRa library for LoRa communication. Subsequently, a single-channel LoRa repeater was implemented in the TTGO development board based in the ESP32 MCU, which also had the SX176 chip. Finally, the gateway consisted of two Arduino Uno connected by a serial port. The first used the Draguino Shield for LoRa communication, similar to node configuration, and the other established an Ethernet connection.

#### 2.2.4. Power Modules

Considering the water quality collection terminal work outdoors, the power supply module used a 12 V lead acid battery 7.0 Ah power supply. The battery voltage was stepped down to 5 V through the Mini 560 DC-DC Converter Voltage. The 5 V powered the processor module and the wireless communication module.

The consumption of the entire system was 26 mA in active mode and 10 mA in sleep mode, which provided sufficient autonomy, since the selected battery provided approximately 30 days until it needed to be recharged.

### *2.3. System Software Architecture*

The system software architecture included water quality collection terminal software, LoRa transmission layer software, server application layer software, anomaly detection, and user interface software design.

#### 2.3.1. Water Quality Collection Terminal Software Design

After the acquisition terminal started and completed the system peripheral initialization pretest, the whole system entered into low-power sleep mode, the kernel clock stopped, the processor kernel itself, the memory system, the related controllers, and the internal buses also stopped working. The dynamic power consumption of these devices was reduced, and the CPU kernel woke from sleep mode whenever any enabled interrupt occurred. The PA0 pin rising edge signal was set to trigger an external interrupt. When the PA0 pin received a rising edge signal pulled high after the LoRa module AUX pin was pulled low for five minutes, it triggered the MCU's external interrupt, and the acquisition terminal restarted to complete data acquisition and transmission, then, entered sleep mode again and waited for the rising edge signal of the AUX pin to trigger the interrupt again.

#### 2.3.2. LoRa Transport Layer Software Design

The LoRa module sent the data, collected by the collection terminal, to the LoRa repeater. Afterward, the data were sent to the gateway, which included the data frames to the server fixed port through the MQTT protocol, and the server parsed the received data, which should include the information frames of the gateway alongside the collection terminal and the water quality data frames.

After the gateway started, it entered initialization and established a connection with the server through the MQTT protocol. After the connection was established, the working mode of the gateway was set to polling wake-up mode, and the gateway sent wake-up data to the LoRa module nodes periodically, and the second polling data was replaced when all nodes were polled with one data. End nodes were polled again in the next polling cycle. The wake-up period was 55 s. Gateway and LoRa repeater modules, after a successful connection, the gateway was set to 1 min as a cycle, sent wake-up data to the LoRa module, polling a wake-up at each collection terminal multichannel parallel collection and uploaded water quality data. After the gateway collected the data, the whole data frame was repackaged and uploaded to the MQTT broker in the server. After the end of polling, the water quality collection terminal went into the sleep state for 30 min, waiting for the gateway, and then, the second time to wake up the water quality collection terminal; meanwhile, the Gateway and LoRa repeater remained active in the case data reception from other nodes was required.

2.3.3. Server Application Layer Software Design

The server application consisted of all those services installed on the virtual private server (VPS), making it necessary to perform real-time monitoring of the values sent from the node (Figure 1a). The VPS was enabled on Amazon Web Services (Copyright Amazon.com, Inc., Seattle, WA, USA). The database software was MySQL (Copyright Oracle, Corp., Austin, TX, USA), and the control panel was Vesta CP (Copyright Vesta CP., San Fernando, CA, USA). The server programs were mainly used for background procedures, including socket network communication, database connection, port upload data processing, and storage, front-end interface loading, etc. EMQ X software, as the MQTT broker, coordinated the management of messages and devices. The server was the manager for the data collected by each terminal and stored in their corresponding database using Node.js algorithm, which allowed continuous connection to the broker and database. The user can monitor the terminal water quality parameters changes in real-time through Grafana (Copyright Raintank, Inc., New York, NY, USA). It can also allow the key parameters to be set to judge the conditions, when the water quality changes beyond the threshold conditions, triggering a mailbox or SMS alarm, and allowing the staff to respond promptly.

After the server established a connection with the gateway, it received the data in real time, and completed the parsing of the data packets. First, it judged whether the length of the data frame is larger than 18 bytes, and if it is smaller, it is judged to be a test data frame uploaded by the gateway and discards the current data frame. According to the data frame format, the validity of the data frame was verified, and the command word field can clarify whether the node is in a normal working state if it is in a non-normal working state, where the current data frame is discarded. When the data frame was complete and valid, the data frame was parsed, to determine the corresponding database according to the ID node of the node in the uploaded data frame, and then, the time and water quality parameters were added to the corresponding database.

*2.4. Multinode, Anomaly Detector, and LoRa Testing of the (IoT-WQMS)*

2.4.1. Indoor Multinode Testing

In order to verify that the monitoring node, server, and user interface of (IoT-WQMS) were all working correctly, indoor multinode testing was carried out. For this, IoT-WQMS was tested for 29 monitoring days on a continuous flow in a PVC pipe structure, with a pump for water recirculation (Figure 3). The PVC pipe structure had an individual sampling port suitable for each sensor and an extra one for taking control measurements. Control measurements were performed using the multiparameter meter (Thermo Scientific Orion Star A329) (Copyright Thermo Fisher Scientific Inc., Waltham, MA, USA).

The absolute error is the result of the difference between the estimated value and the individual value resulting from a measurement [33] and is calculated by the following Equation (1):

$$\Delta a_n = \overline{a} - a_n \tag{1}$$

where $\Delta a_n$ is the absolute error, $\overline{a}$ is the reference value, and $a_n$ is the observed value. In this context, the reference value was the value obtained by the control instrument and the observed value was the value obtained by the sensors. The mean absolute error (MAE) was used as an estimator to calculate the error in the sensor measurements in reference to the control. Where $\overline{\Delta a}$ = mean absolute error and $\Delta a_i$ = absolute error (Equation (2)):

$$\overline{\Delta a} = \frac{1}{n}\sum_{i=1}^{n}|\Delta a_i| \tag{2}$$

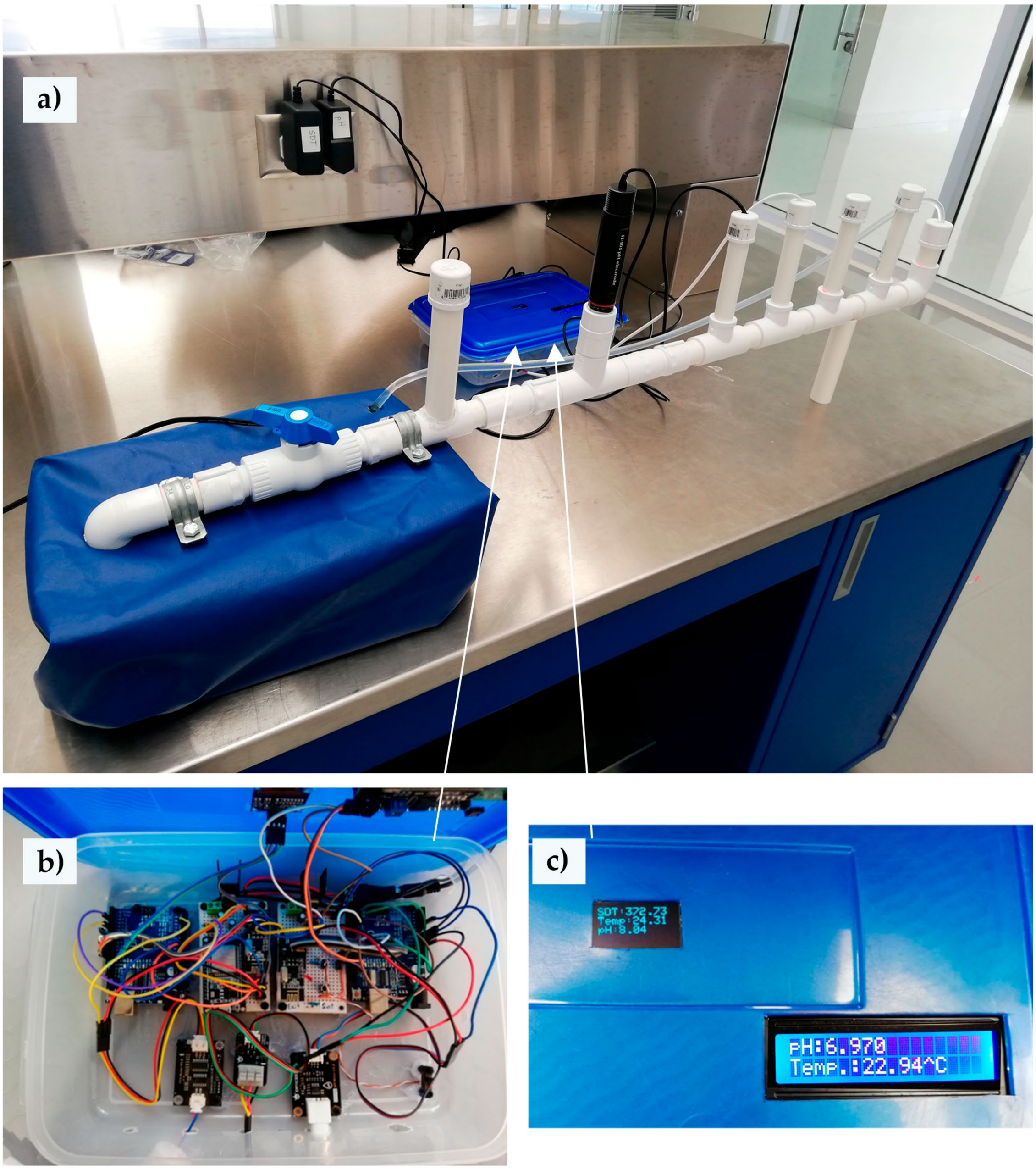


**Figure 3.** (**a**) PVC pipe structure with a pump for water recirculation for indoor multinode testing of IoT-WQMS, (**b**) IoT multinode designed for indoor testing, (**c**) OLED and LCD displays added for local verification of measurements.

### 2.4.2. Anomaly Detection Module Validation

The Anomaly Detection Toolkit package (ADTK) (Copyright, Arundo Analytics, Inc., Palo Alto, CA, USA), an anomaly detection module of unsupervised learning was integrated into WQSM. A detector block was used to analyze the time series and report anomalous data. This function is based on principal component analysis (PCA) and is recommended for simultaneous analysis of multiple time series, where they are known to be highly correlated, such as in the case of the water quality parameters.

The trained model was added to the general program, executed via Flask in the VPS. Figure 4 shows the program's stages for anomaly detection based on the unsupervised anomaly detection algorithm. This program allows a constant connection to the MQTT broker. Each timepoint was received from the node, where it connected to the database to obtain and analyze the values through the anomaly detector. If the values are detected as anomalous, an alert message was sent via Telegram and also displayed in the application control panel. On the other hand, if the data were correct, the system did not perform any action.

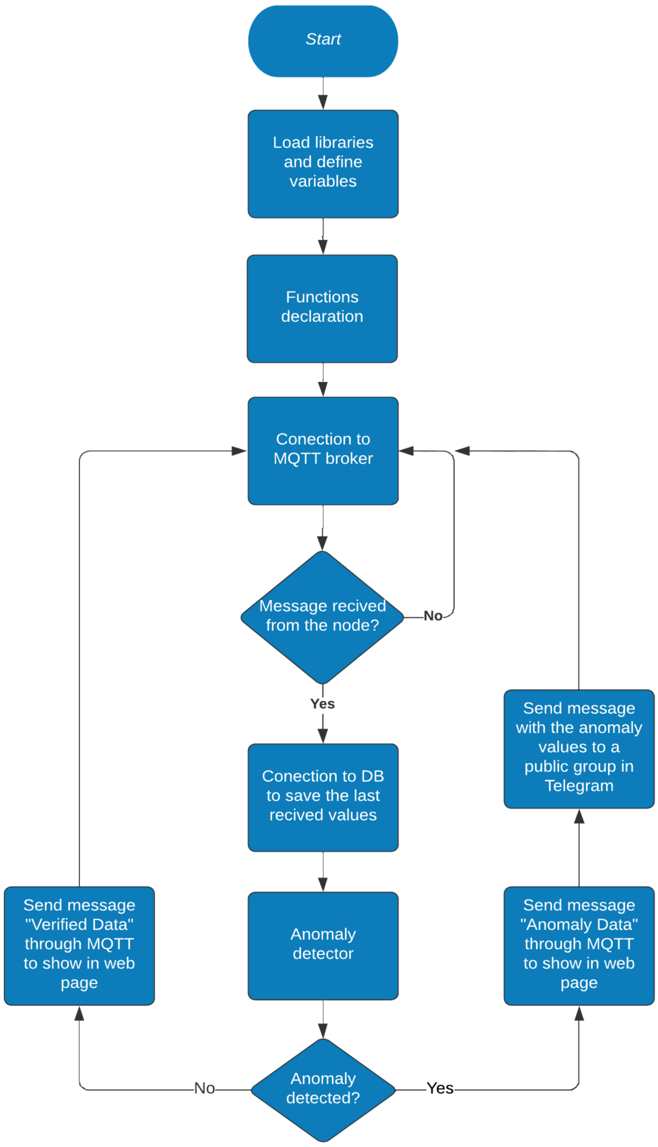


**Figure 4.** Flowchart of the process of anomaly detection module of unsupervised learning integrated with IoT-WQSM.

### 2.4.3. LoRa Validation

Using the LoRa communication protocol, the quality of the transmission data depends on the characteristics of the transmitter and receiver antennas, terrain characteristics, transmitter operating power, and receiver sensitivity. Therefore, it is helpful to model the transmission coverage to design the architecture needed to achieve connectivity for LoRa-based IoT systems. The radiofrequency transmission was modeled using the SPLAT software Vs. 1.4.2 (Copyright, SPLAT, John A. Magliacane, KD2BD) to visualize RSSI (received signal strength indication), according to the proposed transmitter and receiver locations and the parameters presented in Table 1. It allows for determining the feasibility of the communication of LoRa in the area to be implemented IoT monitoring system. A spatial model was obtained with the RSSI gradient approximating the power received in the area of interest. In order to test LoRa layer of IoT-WQMS, the radio frequency performance of the system was verified in the field by detecting the 0–6 km system RSSI (received signal strength indication) and PLR [27,28].

**Table 1.** Parameters used for radiofrequency studies in SPLAT.

| Parameter | Value |
|---|---|
| Earth's dielectric constant (relative permittivity) | 15.000 |
| Earth conductivity (Siemens per meter) | 0.005 |
| Atmospheric refraction constant (N) | 301.000 |
| Frequency (MHz) | 915.000 |
| Radio climate | Continental temperature |
| Polarization | Vertical |
| Situation fraction | 0.50 |
| Time fraction | 0.90 |
| Effective radiated power (ERP) in Watts | 0.1 |
| Sensitivity (dBm) | −148 |
| Transmitter coordinates | 24.050934, −104.702537 |
| Receiver coordinates | 23.9917, −104.726421 |

## 3. Results

### 3.1. Indoor Multinode Testing Validation of the IoT-WQMS

Experimental validation of the IoT-WQMS, which included continuous flow, made it possible to approximate the natural field conditions, under which the system usually works. Figure 5 shows the pH, temperature, and TDS performance of the IoT-WQMS sensors compared to the control. The mean absolute error (MAE) for water temperature, pH, and turbidity and conductivity were 0.65, 0.30, and 14.3, respectively. Large deviations were found in the sensors used in the IoT-WQMS compared to the control. The TDS sensor showed two large overestimations on the 5th and 25th day. The pH sensor performed well during the test period, although it presented two significant underestimates around the 20th day. In general, the temperature sensor showed the best performance with slight mismatches compared to the control. It indicates the need to include sensor validation and calibration processes in IoT monitoring water quality systems, as reported in [34–36], Furthermore, the recommendations for selecting and maintaining the sensors should be considered [7,37,38]. At the same time, the inclusion of anomaly detection services in IoT systems improves the reliability of the operation of these systems by detecting when a sensor has a serious deviation [39,40].

### 3.2. Anomaly Detection

A database with 3681 observations was obtained during the 29 days of the IoT-WQSM test. This database included a column with the date and time (See Supplementary Materials, Table S1). The database generated during the first 10 days of the sensor validation experiment was used for training the unsupervised model. Finally, the trained model was included in the development of a general algorithm for the real-time operation of IoT-WQSM using the Python library Pickle (Python Vs. 3.8.12, n.d.). Pickle allows Python objects to be saved in a binary file, in this case, the previously trained model, and subsequently, load this file with all the object information. The anomalies detected by the program are of scalar type, which uses logical or Boolean values, i.e., a false-true (zero-one) classification, resulting in the data being divided into anomalous and normal.

The graphical output of the detection of the anomaly detection module integrated into IoT-WQSM is shown in Figure 6, where the anomalous events detected by the model can be seen in the red lines, while the blue, orange, and green lines are the generated time series for TDS, temperature, and pH, respectively, during the experiment.

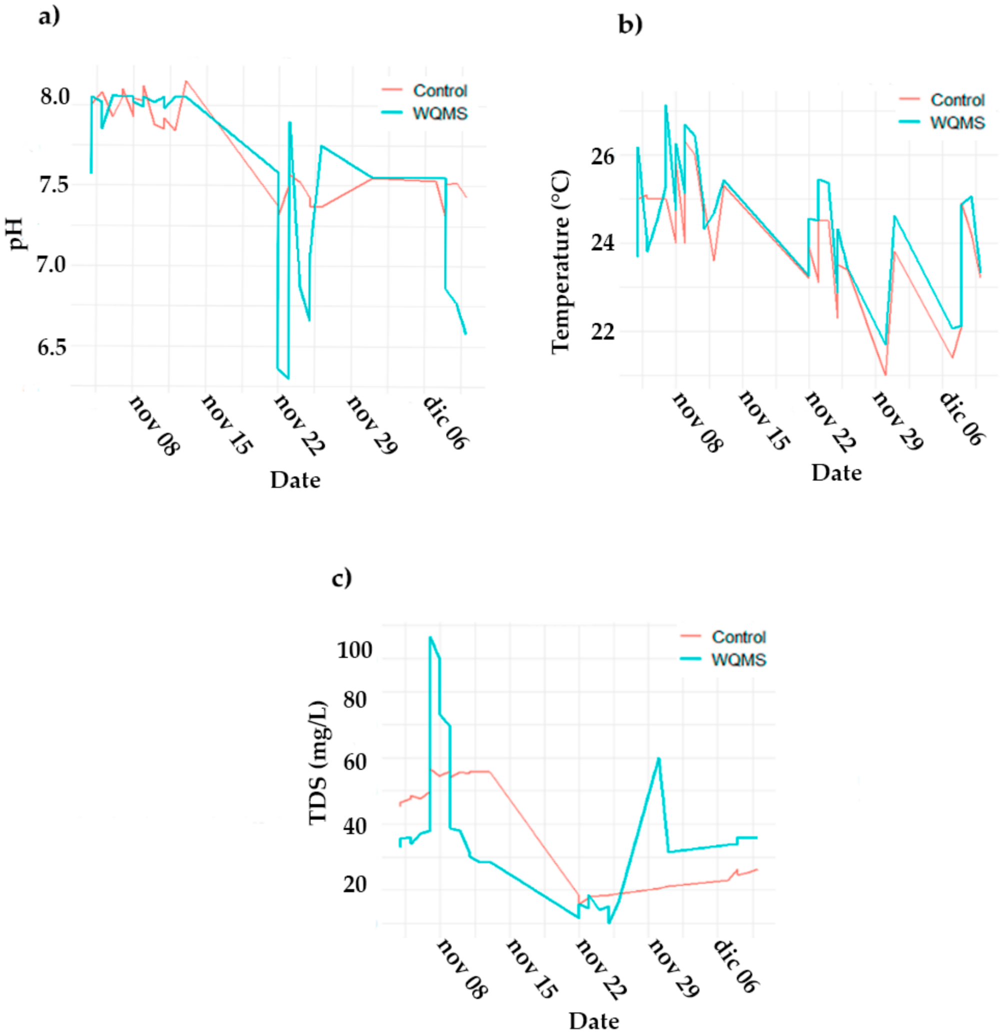


**Figure 5.** Validation of IoT-WQSM sensors, (**a**) pH, (**b**) temperature, and (**c**) total dissolved solids (TDS).

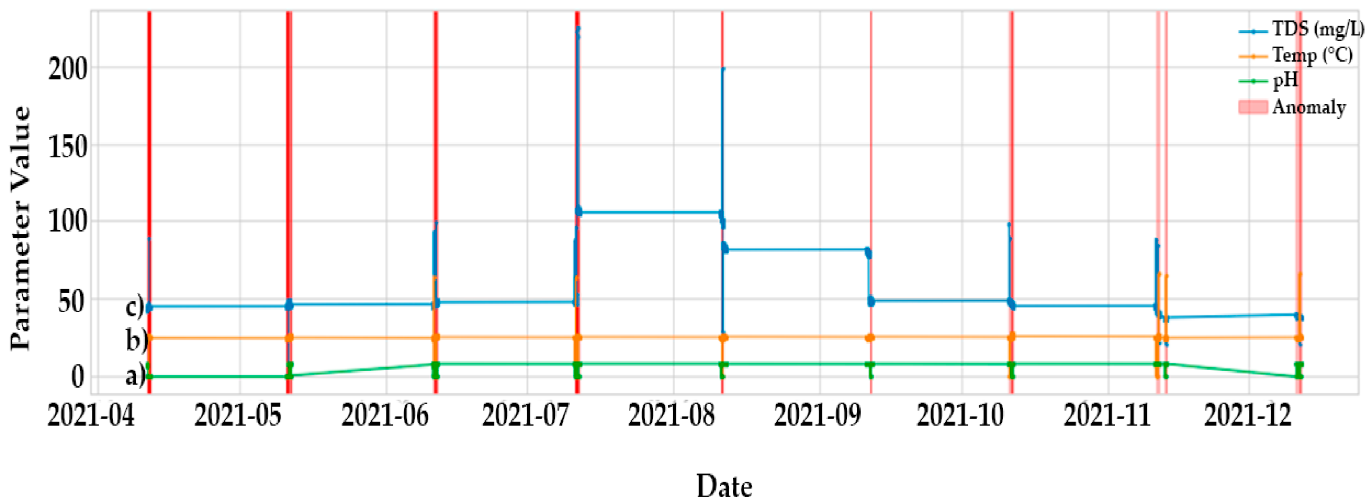


**Figure 6.** The graphical output of the anomaly detection module integrated with IoT-WQSM. Anomaly monitoring of (**a**) pH (lower green line), (**b**) temperature (middle orange line), (**c**) total dissolved solids (TDS, upper blue line), detected anomalies (red vertical lines).

The anomalous events presented were erroneous data sent by the node and reached values of zero in some measurements. On day 7, there was a significant bias in the measurement of the TDS sensor. However, on day 8, when the sensor was recalibrated the TDS values returned to stable values, as at the beginning of the experiment. The anomaly detection module performed well in detecting large deviations, such as the one presented by the deactivation or failure of the pump, which allowed a continuous flow in the experiment.

The anomaly detector, implemented in the IoT-WQSM, connects the database via MQTT to the database. It allows queries to be made to the database in the service and alerts to be made through a chatbot in real time to the telegram groups or on the webpage. The use of supervised techniques for detecting anomalies in the IoT requires more human intervention, as they require the collection of representative training data and manual labeling [41]. The advantage of incorporating unsupervised classification algorithms is that they require minimal human intervention. Additionally, the anomaly detection system should be continually monitored and updated to ensure its effectiveness over time.

Only a few papers address anomaly detection in IoT systems for water quality monitoring. It is a complex task due to the many short- and long-term anomalies that may arise [42,43]. It is vital to address these issues because the quality of the prediction and warning applications in monitoring water resources depends on the data generated by IoT-based water quality monitoring systems.

To address these issues, researchers can explore various anomaly detection techniques or their combinations, including statistical analyses, Machine learning algorithm techniques on historical water quality data can identify anomalies in real-time [42–45]. It is important to note that no single anomaly detection technique is suitable for all IoT-based water quality monitoring scenarios. Therefore, researchers should evaluate the performance of multiple techniques and select the one that best suits the specific case. This IoT-WQSM offers the advantage of including both types of anomaly detection as a statistical process control (SPC) and an unsupervised anomaly detection algorithm. It is one of the advantages that allow the development of these ad hoc systems compared to similar commercial systems, where it is challenging and even illegal to incorporate various anomaly detection or modify the default configuration.

## 4. Outdoor Long-Distance Communication Test

The radiofrequency transmission of RSSI signals modeled from the node to the gateway of the IoT-WQSM system, in the proposed localizations, using the SPLAT software and RSSI average onsite provide similar results (Figure 7 and Table 2). The RSSI values modeled for these sites are within the range reported as suitable (up to 146 dbm) for establishing transmission between the node and the gateway [46]. However, onsite verification found that transmission was not possible at site 2, as indicated by the radiofrequency model. It was attributed to the topography, which caused it to be located below the line of sight to the gateway. The packet loss rate (PLR) was achieved in peripheral locations with low housing densities (sites 1 and 5) and urban site locations with housing densities (sites 3 and 4), except for the peripheral location, site 2, where it was not possible to receive an RSSI and PLR signal. Similar behavior for the LoRa radium signal has been reported in other studies [47–49].

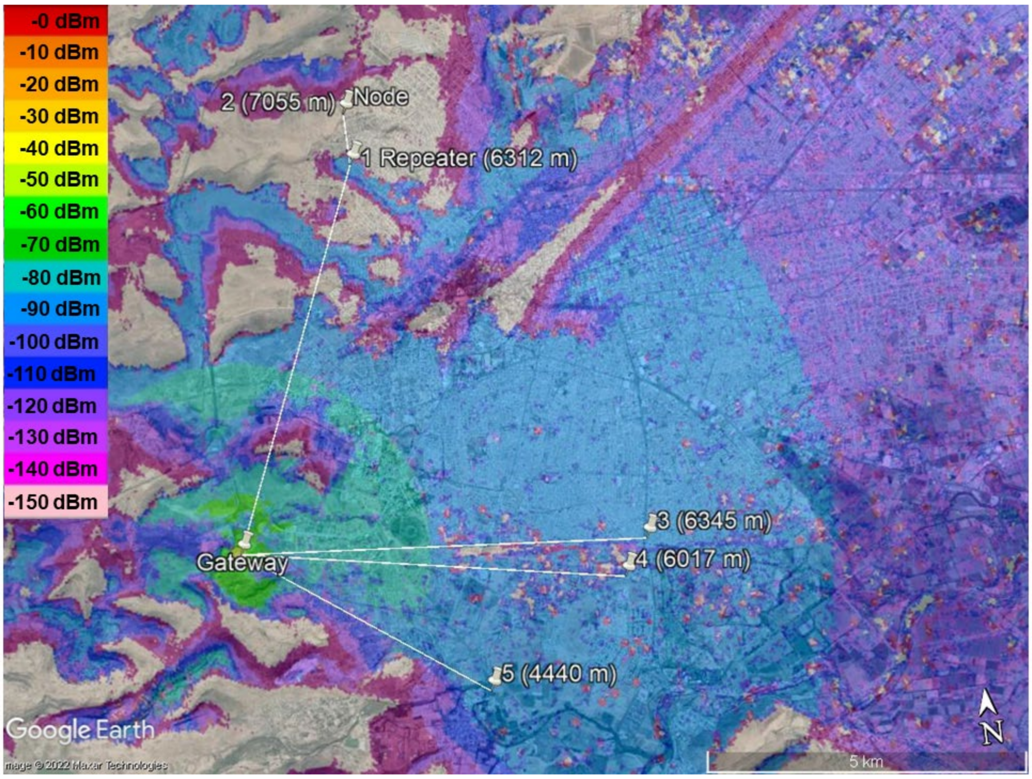


**Figure 7.** Modeling of RSSI (received signal strength indication) by SPLAT, according to the configuration of the proposed localization of the node and gateway for IoT-WQSM. (1,2,3,4,5) Locations of interest tested onsite with the WQSM system.

**Table 2.** Comparison of RSSI modeled with splat software and measurements obtained with two different LoRa configurations at the sites of interest.

| Site | Coordinates | | Distance ** (m) | PLR (%) | RSSI *** (dBi) | Modeled RSSI (dBi) |
|---|---|---|---|---|---|---|
| | Lat | Long | | | Single Lora | |
| 1 | 24.044 | −104.702 | 6312 | 4.2 | −109 | −90 |
| 2 | 24.051 | −104.702 | 7055 | N.S * | N.S | N.S |
| 3 | 23.986 | −104.702 | 6300 | 16.3 | −110 | −90 |
| 4 | 23.981 | −104.668 | 6000 | 4.8 | −113 | −90 |
| 5 | 23.968 | −104.690 | 4500 | 4.0 | −110 | −90 |

Notes: * No Signal; ** Distance from the node; *** RSSI average onsite.

Standards of power signal communication for IoT are divided into near-and far-range depending on distance, which determines the practical applications of IoT-based environmental monitoring. [12,50,51]. A close-range or WPAN allows transmission over 100 m maximum, while a far-range or LPWAN allows transmission over 100 km. Although LoRa is theoretically described as a technology that can transmit data from a few to ten kilometers, this is only sometimes available due to interference. Therefore, it is especially relevant when monitoring the water quality for freshwater resources in situ, where the availability of the mobile communication network is generally unavailable for remote areas. In order to overcome the transmission problems in LoRa coverage, on locations where the LoRa signal was not received, as with the Presa Del Hielo (site 2), we added a LoRa node repeater to the IoT-WQSM architecture. It allowed a place to be reached where originally there was no LoRa signal. A similar architecture network with a repeater has been used in LPWAN layers, as in Lora [52,53], Bluetooth [54], and NB-IoT [55], because including a repeater increases the expansion of the LPWAN network coverage.

The modeling results can help to assess the feasibility of different configuration options for LoRa-based IoT environmental monitoring networks in large areas or cities. They estimate if gateway localization can provide scalability, signal coverage, and functionality for the planned IoT network. This study indicates that the localization of CIMAV-DGO (https://cimav.edu.mx/ (accessed on 19 February 2023)), i.e., a location in a high and clear place, could maximize the coverage of the city. It has electrical power and Internet connectivity, and the gateway can be sheltered. The simulation, in this case, confirmed that the building of CIMAV-DGO is a highly feasible place to locate the gateway for a LoRa-based wireless sensor network in IoT for monitoring water and air quality in Durango.

## 5. Conclusions

A prototype was obtained for remote and real-time monitoring of water quality parameters (pH, temperature, and SDT) in surface water bodies, through a web application. The IoT-WQSM automatically performed the data collection, data storage, anomaly detection, and alarm sending, both remotely and in real time, for the information captured by the multisensor node. The LoRa repeater allowed the extension of the spatial coverage of the LoRa communication, making it possible to reach a place, where originally there was no coverage, with a single LoRa transmitter due to topography and line of sight. Coverage and line of sight studies using computer software allow the identification and selection of the best options to send the LoRa signal to its destination, according to the location of the spatial distribution of system architecture.

The deployment of LoRa networks typically involves setting up gateways that forward data from devices to the network server and requires proper planning, design, and management to ensure optimal performance and security. As LoRa is designed for specific cases, it is essential to evaluate if it is the right technology for the specific application and environment before deploying it. A network simulator can provide much flexibility in studying many different LoRa scenarios, as they allow modeling for different network topologies, device configurations, and environmental conditions. However, it is essential to keep in mind that the accuracy of the simulation results may be limited by the complexity of the models used in the simulation. For example, the simulator may not accurately model all of the factors that can affect the performance of a real-world LoRa network, such as interference from other wireless devices, atmospheric conditions, and the specific characteristics of the devices and gateways used in the network.

Integrating anomaly detection services to IoT-based water quality monitoring systems could expand capabilities in monitoring this valuable resource. Unlike most existing commercial systems, this system allows the configuration of algorithms for anomaly detection. The approach for anomaly detection in IoT-WQSM, where anomaly detection is based on the construction of numerical features of device profiles, can be more effective in detecting anomalies in IoT networks because it takes into account the unique patterns of the traffic generated by each device, rather than relying solely on aggregate statistical measures. It can also be more robust to changes in traffic patterns, as the numerical features of the device profiles can be updated dynamically to reflect changes in the behavior of individual devices, the determination of their normal values, and the search for deviations from the norm based on statistical methods.

The availability of sensors for in situ monitoring of various types of water quality parameters (e.g., E. coli biosensors, heavy metals, pesticides, organic load, anions, cations, etc.) is increasing in the market. Therefore, there is great variability in price, accuracy, and application ranges. To integrate more sensors into the system it is necessary to perform validations to evaluate their performance because it can vary from manufacturer to manufacturer.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/w15071351/s1, Table S1: Database of the 29-day WQSM test for pH, Temperature, and Total Dissolved Solids (TDS) determinations.

**Author Contributions:** Conceptualization, D.A.M.C. and A.D.B.J.; Formal analysis A.D.B.J., D.A.M.C. and M.T.A.H.; Methodology A.D.B.J., D.A.M.C., M.T.A.H., A.E.M.C. and E.N.R.; Visualization D.A.M.C., A.D.B.J., M.T.A.H., A.E.M.C. and E.N.R.; Writing—review & editing, D.A.M.C., A.D.B.J., M.T.A.H., A.E.M.C. and E.N.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** The present study was funded by the Consejo de Ciencia y Tecnologia del Estado de Durango (COCyTED, or the Council for Science and Technology of the State of Durango), under project number 63.

**Acknowledgments:** The authors would like to thank academic technicians Luis Arturo Torres Castañón and José Rafael Irigoyen Campuzano for their valuable collaboration in the analytical determinations.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Coraggio, E.; Han, D.; Gronow, C.; Tryfonas, T. Water Quality Sampling Frequency Analysis of Surface Freshwater: A Case Study on Bristol Floating Harbour. *Front. Sustain. Cities* **2022**, *3*, 791595. [CrossRef]
2. Shuo, J.; Yonghui, Z.; Wen, R.; Kebin, T. The unmanned autonomous cruise ship for water quality monitoring and sampling. In Proceedings of the 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC), Dalian, China, 25–27 December 2017; pp. 700–703.
3. Ayaz, M. Mobile unsupervised platform for real-time ocean water quality monitoring. *J. Control Eng. Appl. Inform.* **2019**, *21*, 79–88.
4. Park, J.; Kim, K.T.; Lee, W.H. Recent advances in information and communications technology (ICT) and sensor technology for monitoring water quality. *Water* **2020**, *12*, 510. [CrossRef]
5. Zhao, J.; Wei, S.; Wen, X.; Qiu, X. Analysis and prediction of big stream data in real-time water quality monitoring system. *J. Ambient. Intell. Smart Environ.* **2020**, *12*, 393–406. [CrossRef]
6. Ahmed, U.; Mumtaz, R.; Anwar, H.; Mumtaz, S.; Qamar, A.M. Water quality monitoring: From conventional to emerging technologies. *Water Supply* **2020**, *20*, 28–45. [CrossRef]
7. Jan, F.; Min-Allah, N.; Düştegör, D. Iot based smart water quality monitoring: Recent techniques, trends and challenges for domestic applications. *Water* **2021**, *13*, 1729. [CrossRef]
8. Martínez, R.; Vela, N.; El Aatik, A.; Murray, E.; Roche, P.; Navarro, J.M. On the use of an IoT integrated system for water quality monitoring and management in wastewater treatment plants. *Water* **2020**, *12*, 1096. [CrossRef]
9. Kumar, C.R.; Ibrahim, A. VLSI design of energy efficient computational centric smart objects for IoT. In Proceedings of the 2018 15th Learning and Technology Conference (L&T), Jeddah, Saudi Arabia, 25–26 February 2018; pp. 129–138.
10. Ishino, M.; Koizumi, Y.; Hasegawa, T. Leveraging proximity services for relay device discovery in user-provided IoT networks. In Proceedings of the 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT), Milan, Italy, 14–16 December 2015; pp. 553–558.
11. Gao, G.; Xiao, K.; Chen, M. An intelligent IoT-based control and traceability system to forecast and maintain water quality in freshwater fish farms. *Comput. Electron. Agric.* **2019**, *166*, 105013. [CrossRef]
12. Hossein Motlagh, N.; Mohammadrezaei, M.; Hunt, J.; Zakeri, B. Internet of Things [IoT] and the energy sector. *Energies* **2020**, *13*, 494. [CrossRef]
13. Zeba, O.; Hayashi, K.; Kizaki, K.; Saruwatari, S.; Watanabe, T. QuadScatter: Computational Efficiency in Simultaneous Transmissions for Large-Scale IoT Backscatter Networks. *IEEE Open J. Comput. Soc.* **2021**, *2*, 334–345. [CrossRef]
14. Verma, A.; Kaushal, S.; Sangaiah, A.K. Computational intelligence based heuristic approach for maximizing energy efficiency in internet of things. In *Intelligent Decision Support Systems for Sustainable Computing*; Springer: Berlin, Germany, 2017; pp. 53–76.
15. Xiao, F. Complex pignistic transformation-based evidential distance for multisource information fusion of medical diagnosis in the IoT. *Sensors* **2021**, *21*, 840. [CrossRef] [PubMed]
16. Sun, B.; Ma, L. An overview of outliers and detection methods in general for time series from IoT devices. In Proceedings of the International Conference on Computer Engineering and Networks, Xi'an, China, 16–18 October 2020; pp. 1180–1186.
17. Zhang, Y.; Meratnia, N.; Havinga, P. Outlier detection techniques for wireless sensor networks: A survey. *IEEE Commun. Surv. Tutor.* **2010**, *12*, 159–170. [CrossRef]
18. Adu-Manu, K.S.; Tapparello, C.; Heinzelman, W.; Katsriku, F.A.; Abdulai, J.-D. Water quality monitoring using wireless sensor networks: Current trends and future research directions. *ACM Trans. Sens. Netw. (TOSN)* **2017**, *13*, 1–41. [CrossRef]
19. Chen, Y.; Han, D. Water quality monitoring in smart city: A pilot project. *Autom. Constr.* **2018**, *89*, 307–316. [CrossRef]
20. Nair, K.K.; Abu-Mahfouz, A.M.; Lefophane, S. Analysis of the narrow band internet of things (NB-IoT) technology. In Proceedings of the 2019 Conference on Information Communications Technology and Society (ICTAS), Durban, South Africa, 6–8 March 2019; pp. 1–6.
21. Zaraket, C.; Papageorgas, P.; Aillerie, M.; Agavanakis, K.; Salame, C. Cyber security vulnerabilities of smart metering based on LPWAN wireless communication technologies. In *AIP Conference Proceedings*; AIP Publishing LLC: Melville, NY, USA, 2020; p. 020050.

22. Al-amri, R.; Murugesan, R.K.; Man, M.; Abdulateef, A.F.; Al-Sharafi, M.A.; Alkahtani, A.A. A review of machine learning and deep learning techniques for anomaly detection in IoT data. *Appl. Sci.* **2021**, *11*, 5320. [CrossRef]
23. Vigoya, L.; Fernandez, D.; Carneiro, V.; Cacheda, F. Annotated dataset for anomaly detection in a data center with IoT sensors. *Sensors* **2020**, *20*, 3745. [CrossRef]
24. Pittino, F.; Puggl, M.; Moldaschl, T.; Hirschl, C. Automatic anomaly detection on in-production manufacturing machines using statistical learning methods. *Sensors* **2020**, *20*, 2344. [CrossRef]
25. Basora, L.; Olive, X.; Dubot, T. Recent advances in anomaly detection methods applied to aviation. *Aerospace* **2019**, *6*, 117. [CrossRef]
26. Chen, F.; Deng, P.; Wan, J.; Zhang, D.; Vasilakos, A.V.; Rong, X. Data mining for the internet of things: Literature review and challenges. *Int. J. Distrib. Sens. Netw.* **2015**, *11*, 431047. [CrossRef]
27. Souza, A.M.; Amazonas, J.R. An outlier detect algorithm using big data processing and internet of things architecture. *Procedia Comput. Sci.* **2015**, *52*, 1010–1015. [CrossRef]
28. Zhong, Y.; Fong, S.; Hu, S.; Wong, R.; Lin, W. A novel sensor data pre-processing methodology for the Internet of Things using anomaly detection and transfer-by-subspace-similarity transformation. *Sensors* **2019**, *19*, 4536. [CrossRef] [PubMed]
29. Pawar, N.; Bourgeau, T.; Chaouchi, H. Low-cost, Low-power Testbed for Establishing Network of LoRaWAN Nodes. In Proceedings of the EWSN, Lyon, France, 17–19 February 2020; pp. 192–194.
30. Restuccia, F.; Pagani, M.; Mascitti, A.; Barrow, M.; Marinoni, M.; Biondi, A.; Buttazzo, G.; Kastner, R. ARTe: Providing real-time multitasking to Arduino. *J. Syst. Softw.* **2022**, *186*, 111185. [CrossRef]
31. Podder, R.; Barai, R.K. Hybrid Encryption Algorithm for the Data Security of ESP32 based IoT-enabled Robots. In Proceedings of the 2021 Innovations in Energy Management and Renewable Resources, Kolkata, India, 5–7 February 2021; pp. 1–5.
32. DFRobot Forum. Waterproof DS18B20 Digital Temperature Sensor (SKU: DFR0198). 2019. Available online: https://wiki.dfrobot.com/Waterproof_DS18B20_Digital_Temperature_Sensor__SKU_DFR0198 (accessed on 19 February 2023).
33. Lin, J.-Y.; Tsai, H.-L.; Lyu, W.-H. An Integrated Wireless Multi-Sensor System for Monitoring the Water Quality of Aquaculture. *Sensors* **2021**, *21*, 8179. [CrossRef] [PubMed]
34. Goparaju, S.U.N.; Vaddhiparthy, S.S.S.; Pradeep, C.; Vattem, A.; Gangadharan, D. Design of an IoT System for Machine Learning Calibrated TDS Measurement in Smart Campus. In Proceedings of the 2021 IEEE 7th World Forum on Internet of Things (WF-IoT), New Orleans, LA, USA, 14–31 June 2021; pp. 877–882.
35. Singh, R.; Baz, M.; Gehlot, A.; Rashid, M.; Khurana, M.; Akram, S.V.; Alshamrani, S.S.; AlGhamdi, A. Water Quality Monitoring and Management of Building Water Tank Using Industrial Internet of Things. *Sustainability* **2021**, *13*, 8452. [CrossRef]
36. Wong, Y.J.; Nakayama, R.; Shimizu, Y.; Kamiya, A.; Shen, S.; Rashid, I.Z.M.; Sulaiman, N.M.N. Toward industrial revolution 4.0: Development, validation, and application of 3D-printed IoT-based water quality monitoring system. *J. Clean. Prod.* **2021**, *324*, 129230. [CrossRef]
37. Banna, M.H.; Imran, S.; Francisque, A.; Najjaran, H.; Sadiq, R.; Rodriguez, M.; Hoorfar, M. Online drinking water quality monitoring: Review on available and emerging technologies. *Crit. Rev. Environ. Sci. Technol.* **2014**, *44*, 1370–1421. [CrossRef]
38. Lambrou, T.P.; Anastasiou, C.C.; Panayiotou, C.G.; Polycarpou, M.M. A low-cost sensor network for real-time monitoring and contamination detection in drinking water distribution systems. *IEEE Sensors J.* **2014**, *14*, 2765–2772. [CrossRef]
39. González-Vidal, A.; Cuenca-Jara, J.; Skarmeta, A.F. Iot for water management: Towards intelligent anomaly detection. In Proceedings of the 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland, 21–24 April 2019; pp. 858–863.
40. Wang, M.; Fu, C.; Du, X. Decision-Tree Based Root Cause Localization for Anomalies in Smart IoT Systems. In Proceedings of the ICC 2021-IEEE International Conference on Communications, Xiamen, China, 28–30 July 2021; pp. 1–5.
41. Spanos, G.; Giannoutakis, K.M.; Votis, K.; Tzovaras, D. Combining statistical and machine learning techniques in IoT anomaly detection for smart homes. In Proceedings of the 2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Limassol, Cyprus, 11–13 September 2019; pp. 1–6.
42. Iyer, S.; Thakur, S.; Dixit, M.; Katkam, R.; Agrawal, A.; Kazi, F. Blockchain and anomaly detection based monitoring system for enforcing wastewater reuse. In Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 6–8 July 2019; pp. 1–7.
43. Bourelly, C.; Bria, A.; Ferrigno, L.; Gerevini, L.; Marrocco, C.; Molinara, M.; Cerro, G.; Cicalini, M.; Ria, A. A preliminary solution for anomaly detection in water quality monitoring. In Proceedings of the 2020 IEEE International Conference on Smart Computing (SMARTCOMP), Bologna, Italy, 14–17 September 2020; pp. 410–415.
44. Garmaroodi, M.S.S.; Farivar, F.; Haghighi, M.S.; Shoorehdeli, M.A.; Jolfaei, A. Detection of anomalies in industrial iot systems by data mining: Study of christ osmotron water purification system. *IEEE Internet Things J.* **2020**, *8*, 10280–10287. [CrossRef]
45. Ajaaj, A.A.; Mishra, A.K.; Khan, A.A. Evaluation of satellite and gauge-based precipitation products through hydrologic simulation in Tigris River basin under data-scarce environment. *J. Hydrol. Eng.* **2019**, *24*, 05018033. [CrossRef]
46. Tamang, D.; Pozzebon, A.; Parri, L.; Fort, A.; Abrardo, A. Designing a reliable and low-latency LoRaWAN solution for environmental monitoring in factories at major accident risk. *Sensors* **2022**, *22*, 2372. [CrossRef]
47. Amadou, I.; Foubert, B.; Mitton, N. LoRa in a haystack: A study of the LORA signal behavior. In Proceedings of the 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Barcelona, Spain, 21–23 October 2019; pp. 1–4.

48. Magrin, D.; Capuzzo, M.; Zanella, A. A thorough study of LoRaWAN performance under different parameter settings. *IEEE Internet Things J.* **2019**, *7*, 116–127. [CrossRef]
49. Qadir, Q.M. Analysis of the reliability of LoRa. *IEEE Commun. Lett.* **2020**, *25*, 1037–1040. [CrossRef]
50. Ray, P.P. A survey on Internet of Things architectures. *J. King Saud Univ. Comput. Inf. Sci.* **2018**, *30*, 291–319.
51. Samizadeh Nikoui, T.; Rahmani, A.M.; Balador, A.; Haj Seyyed Javadi, H. Internet of Things architecture challenges: A systematic review. *Int. J. Commun. Syst.* **2021**, *34*, e4678. [CrossRef]
52. Gia, T.N.; Qingqing, L.; Queralta, J.P.; Zou, Z.; Tenhunen, H.; Westerlund, T. Edge AI in smart farming IoT: CNNs at the edge and fog computing with LoRa. In Proceedings of the 2019 IEEE AFRICON, Accra, Ghana, 25–27 September 2019; pp. 1–6.
53. Awadallah, S.; Moure, D.; Torres-González, P. An internet of things [IoT] application on volcano monitoring. *Sensors* **2019**, *19*, 4651. [CrossRef] [PubMed]
54. Sreeram, M.; Sreeja, M. A novel architecture for IoT and smart community. In Proceedings of the 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 10–11 February 2017; pp. 486–491.
55. Ha, S.; Seo, H.; Moon, Y.; Lee, D.; Jeong, J. A novel solution for NB-IoT cell coverage expansion. In Proceedings of the 2018 Global Internet of Things Summit (GIoTS), Bilbao, Spain, 4–7 June 2018; pp. 1–5.