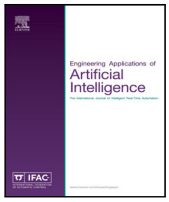
Engineering Applications of Artificial Intelligence 127 (2024) 107235

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



Water level identification with laser sensors, inertial units, and machine learning

Caetano M. Ranieri [a,*], Angelo V.K. Foletto [a], Rodrigo D. Garcia [a], Saulo N. Matos [a], Maria M.G. Medina [b], Leandro S. Marcolino [c], Jó Ueyama [a]

[a] Institute of Mathematical and Computer Sciences, University of São Paulo, Av. Trab. São Carlense, 400, São Carlos, 13566-590, SP, Brazil
[b] São Carlos School of Engineering, University of São Paulo, Av. Trab. São Carlense, 400, São Carlos, 13566-590, SP, Brazil
[c] School of Computing and Communication, Lancaster University, Bailrigg, Lancaster, LA1 4YW, United Kingdom

ARTICLE INFO

Keywords:
LiDAR
Ultrasonic sensor
Inertial measurement unit
Machine learning
Flood risk

ABSTRACT

Flood risk management usually hinges on accurate water level identification in urban streams such as rivers or creeks. Although research has emphasised the applicability of ultrasonic sensors as a contactless technology for sensor-based water level monitoring, Light Detection and Ranging (LiDAR) sensors are less sensitive to weather conditions that typically happen during flood events, such as dust, fog and rainfall. However, there has been little research on the applicability of LiDAR sensors in this field. No previous literature has analysed the impact of complicating variables on the quality of predictions or evaluated the possible benefits of using a combined approach with Inertial Measurement Units (IMU) and machine learning to produce superior predictions. In this work, we collected a dataset in a laboratory condition synchronising data from a LiDAR, an ultrasonic sensor and an IMU in an experimental device. We controlled the incidence angle, the distance, and the water turbidity to analyse their effect on the predictions. Traditional machine-learning techniques were evaluated as models to combine data from distance and inertial sensors, reducing the error rates compared to individual sensors' predictions. Results indicated a sharp drop in the mean absolute error, root mean squared error and coefficient of determination for all water turbidity and incidence angles considered, especially when tree-based ensembles were used. The ultrasonic sensor led to improved results for low water turbidity and increased incidence angle, but statistically significant differences were not found in the other cases.

1. Introduction

Proactive monitoring of water levels in urban waterways enables the early detection of potential flood-related hazards, allowing for the timely activation of damage control strategies (Brito et al., 2023; Ilhan, 2023; Kamoji and Kalla, 2023). Systems dedicated to this endeavour typically employ Internet of Things (IoT) devices, including interconnected sensors and cameras, which are integrated with a central server. This integration enables continuous monitoring of water levels in water streams, providing data for real-time analysis and decision-making (Furquim et al., 2016; Mamat et al., 2021).

Pressure transducers positioned at the base of a water body can be an effective tool for estimating water levels (Furquim et al., 2016; Yuliza et al., 2016; Schenato et al., 2021), relying on Stevin's law for hydrostatics pressure (Ramos et al., 2020). However, as pressure sensors need to be immersed in the water stream, the installation conditions could influence the measures by creating local pressure modifications. Moreover, they are vulnerable to damage or burial due to events such as floods, the presence of debris, or sedimentation. To account for these issues, a deployed system based on pressure sensors is subject to frequent operational interruptions and incurs additional maintenance costs.

Ultrasonic sensors present a potential solution to these challenges (Varun et al., 2018; Rocchi et al., 2019; Sahoo and Udgata, 2020; Yunita et al., 2020). They are the most used non-contact technology for measuring water levels, which does not require them to be submerged in water, thus causing fewer maintenance issues. Nonetheless, their performance can be impacted by factors such as rain, snow, fog, dust, and changes in temperature or humidity. This makes them less reliable for flood warning systems, where these conditions can change rapidly and dramatically (Bae and Ji, 2019; Panagopoulos et al., 2021).

In this paper, we are concerned with an alternative non-contact technology that is more robust to adverse weather conditions: Light

* Corresponding author.
E-mail address: cmranieri@usp.br (C.M. Ranieri).

https://doi.org/10.1016/j.engappai.2023.107235
Received 22 June 2023; Received in revised form 30 August 2023; Accepted 29 September 2023
Available online 11 October 2023

Detection and Ranging (LiDAR). Even though LiDAR sensors can be less accurate than ultrasonic sensors if the water turbidity is low or the incidence angle is high (Reymann and Lacroix, 2015; Song et al., 2002), they are more robust to environmental adversities. For this reason, they might be a promising technology for flood risk management.

While research has been performed to improve the performance of LiDAR sensors in water level identification (Paul et al., 2020; Tamari and Guerrero-Meza, 2016; Tamari et al., 2016), these works have been constrained to reducing predictive error based only on the own sensor's properties. To the best of our knowledge, existing literature has not yet assessed advanced methodologies to unveil intricate patterns among variables derived from LiDAR sensors and other complementary sensor technologies.

We addressed this research gap by presenting an innovative experimental device that integrates a LiDAR sensor with an Inertial Measurement Unit (IMU). We thoroughly assessed the impact on the predictive performance of three complicating factors: the incidence angle, the water turbidity, and the distance between the sensor and the water surface. To facilitate comprehensive evaluations, we also incorporated an ultrasonic sensor alongside the LiDAR.

We developed a laboratory data collection procedure to control the above-mentioned complicating factors. We analysed the performance of ultrasonic and LiDAR sensors from different perspectives, including results stratified per water turbidity. Then, we used these performances as baseline results for our proposed approach of using traditional machine-learning models to combine data from distance sensors to IMU to reduce predictive error.

Compared to the baselines, these models led to superior performance in all conditions considered. The overall mean average error (MAE) was reduced from 56.31 cm to 13.94 cm for the LiDAR and from 3.28 cm to 0.51 cm for the ultrasonic (distances in the dataset ranged in the $[50, 300]$ cm interval). The performance gap between the ultrasonic sensor and the LiDAR was significantly reduced through the proposed approach, with non-statistically significant differences for medium or high water turbidity, and tilt angles up to $7.5°$.

This work provided four distinct contributions: (i) detailing the design of an experimental device that integrates either a LiDAR or an ultrasonic sensor with an IMU to enhance water level detection; (ii) evaluating the independent performance of LiDAR and ultrasonic sensors for water level measurement within a controlled environment while regulating the tilt angle, the water turbidity, and the distance between sensor and water surface; (iii) investigating the use of machine learning models to fuse the IMU variables with the data from the LiDAR or ultrasonic sensors; and (iv) conducting a comprehensive assessment of each model, gauging their performance enhancement over individual sensors.

This paper unfolds as follows. Section 2 discusses the technical aspects of the sensors examined in this study. Section 3 reviews existing studies on water level detection using pressure, ultrasonic, and LiDAR sensors, emphasising the research gap we addressed. Section 4 outlines the data collection procedure, detailing the experimental apparatus and the tests performed in a controlled laboratory setting. Section 5 introduces the techniques for analysing sensor data, the machine learning methods used, and the additional analyses performed. Section 6 reports our results, presenting evaluation metrics, error distributions, and statistical significance tests. Section 7 offers an in-depth discussion of these findings, including implications and comparisons across different conditions. Finally, Section 8 offers closing thoughts and future research directions.

2. Background

This section elucidates the fundamental principles underlying the technologies employed in measuring water levels. These measurement techniques can be bifurcated into two distinct categories: contact and non-contact.

Contact measurement is typified by the presence of sensor interaction with the subject material, which could encompass both liquid and solid states. Pressure transducers constitute prime examples of contact measurement technologies (Hanni and Venkata, 2020). They use the Stevin principle to calculate water column height via hydrostatic pressure (Lipták and Venczel, 2016). To do this accurately, atmospheric pressure effects must be removed from the sensor readings. This is where vented cable technology can be used, as it can measure and adjust for atmospheric pressure (World Meteorological Organization (WMO), 2010; Panagopoulos et al., 2021). However, changes related to installation conditions like stilling or vortex emergence can change the pressure measurements independently of the water level. Besides, the sensor is in direct contact with water, making it vulnerable to being dislodged or moved by debris in floodwaters (Panagopoulos et al., 2021; Lo et al., 2015).

In this paper, we are concerned with non-contact level measurement devices, which encompass ultrasonic acoustic wave transmission, radio wave (i.e., radar) transmission, laser transmission technologies, and video cameras (World Meteorological Organization (WMO), 2010; Kashani et al., 2015; Rocchi et al., 2019; Fernandes Junior et al., 2021; Abu-Salih et al., 2023). The subsequent subsections delineate the two approaches we assessed in this work: ultrasonic and laser-based level measurement.

2.1. Ultrasonic-based measurement

Ultrasonic technology represents a non-contact approach to water level measurement, utilising a high-frequency acoustic transducer to propagate sound waves through the air towards the water surface (World Meteorological Organization (WMO), 2010). It may be applied in various fields, such as thickness measurement, mechanical properties, robotics, remote sensing and surface imaging (Burrascano et al., 2014). Additionally, it is capable of being utilised in the design of both Structural Health Monitoring (SHM) systems and self-aware smart structures (Shen and Giurgiutiu, 2014).

The ultrasonic ranging principle is based on converting electrical signals into sound energy using a piezoelectric crystal inside the transducer. The signal is emitted into the air, reflected off the target, and received back by the transducer (Milligan et al., 2013). In pulse-echo mode, distance (s) is calculated by measuring time t between transmitted and received signals using Eq. (1) (Qiu et al., 2022).

$$s = \frac{ct}{2} \tag{1}$$

where c is the sound velocity in the medium. The distance is divided by two because we want to calculate the one-way distance, not the round trip.

Hence, the water level is determined by calculating the wave's travelled distance. Ultrasonic water level meters showcase commendable performance in terms of measurement accuracy, and their installation proves comparatively straightforward in intricate environments. However, they are susceptible to environmental perturbations (Chen et al., 2022). As already stated, factors such as temperature, humidity, snow, rain, fog and dust can interfere with the transmitted and reflected signals (Toa and Whitehead, 2020; Panagopoulos et al., 2021; Bae and Ji, 2019). Moreover, it has been empirically substantiated that temperature and humidity can influence the measurements obtained from ultrasonic sensors (Panagopoulos et al., 2021).

2.2. Laser-based measurement

LiDAR constitutes an active remote sensing technology that utilises electromagnetic (EM) waves within the optical and infrared wavelength spectrum (McManamon, 2019). The system operates by transmitting a series of laser beams and subsequently gathering the received signals. Analogous to the ultrasonic method, the laser pulse's travel

distance (i.e., range) is computed via the wave's time-of-flight. This methodology exhibits significant potential in the domain of water level measurement (World Meteorological Organization (WMO), 2010).

The range measurement is influenced by the amount of received light, known as intensity. According to Kashani et al. (2015), several factors affect the measurement intensity, including the characteristics of the target surface, such as reflectance ρ. The reflectance of the target surface in water level measurement is directly influenced by the fluid's turbidity (Paul et al., 2020; Kashani et al., 2015).

Another factor is the data acquisition geometry, which includes the range r and incidence angle β. The incidence angle is the angle between the transmitted laser beam and the target surface (Yan et al., 2020; Kukko et al., 2008). Higher incidence angles generally lead to reduced backscattering of the incident laser energy towards the receiver, resulting in lower received optical power (Kashani et al., 2015).

These effects are modelled in Eq. (2), the LiDAR range Equation, adapted from technical and academic literature (Kashani et al., 2015; Jelalian, 1992; Paul et al., 2020), in which t is the timestamp, $P_a(t)$ is the power of the transmitted laser pulse, $P_b(t)$ is the laser power received from the water surface, T^2_{atm} is the transmission coefficient of the atmosphere, η_a and η_b are the optical transmission and reception efficiencies, and A_r is the area of the receptor.

$$P_b(t) = \frac{\rho P_a(t) T^2_{atm} \eta_a \eta_b A_r cos^2(\beta)}{\pi r^2}, \tag{2}$$

This equation relates the received optical power to the transmitted power and other parameters related to the system, acquisition geometry, environment, and target characteristics. It can be applied to water surfaces, allowing for the identification of water level (Allouis et al., 2007; Paul et al., 2020).

3. Related work

This section reviews water level measurement methodologies closely aligned with our study. Our focus was centred on design and systematic analysis employing ultrasonic and LiDAR sensors. Additionally, we examined research concentrated on minimising measurement inaccuracies, aligning with the primary objective of our proposed experimental device and machine learning models.

Water level measurement through ultrasonic sensors has been employed in different applications. Some works addressed level measurement and control in water tanks, such as Gondkar et al. (2022), Choubey et al. (2022), Jan et al. (2022) and Djalilov et al. (2023). A more critical environment was addressed by Sasikala et al. (2022), which proposed safeguard methodologies for dams and reservoirs based on water level measurement through ultrasonic sensors, with results showing a low error rate and short data transmission latency.

Integrated systems for flood monitoring have been proposed using ultrasonic sensors, many of which focused on the IoT environment to ensure a reliable platform for water level monitoring. For example, Hanan and Sumadiyasa (2019) proposed a water surface-level detection system using the HC-SR04 ultrasonic sensors and the ESP8266-12E module. The system detects water height and sends the information to the module. Similarly, Prafanto and Budiman (2018) presented a wireless sensor network-based system for automatic, real-time detection of river water levels using ultrasonic sensors. The system transmits the water level data to a web server, allowing the public to monitor the river's height in real time.

A comparative analysis between ultrasonic sensors and pressure differential transducers for monitoring a river level was presented by Panagopoulos et al. (2021). The statistical and graph analysis proved that the evaluation metrics of the two sensors did not significantly differ. They observed that the stream river turbulence and the ambient temperature influenced the ultrasonic measurements, while the level measurement using pressure transducers did not fluctuate significantly. However, ultrasonic sensors were considered a complementary approach because they are not subject to damage risks inherent to the pressure sensors, which might be immersed in the water stream to generate predictions.

A related analysis was made by Wannoi and Wannoi (2022), who presented a prototype for early flood warning in a wide area. Two management interfaces were proposed to allow real-time monitoring and implement real-time alerts. A pressure transducer was compared to an ultrasonic sensor for assessing the water level of a two-metres depth river, with lower error values for the ultrasonic sensor in the conditions analysed.

Fewer studies have addressed laser-based methods for this task. Paul et al. (2020) investigated time-of-flight LiDAR for measuring water levels under various environmental conditions. The authors utilised a near-infrared (905 nm) LiDAR sensor, testing its performance under different conditions, such as measurement distance, surface roughness, air temperature, water turbidity, and measurement angle.

Tamari and Guerrero-Meza (2016) demonstrated the potential of inclined LiDAR to measure river levels, particularly in high turbidity conditions. Their methodology incorporated an industrial range-only LiDAR. The results provided evidence that an inclined LiDAR can be effectively used to monitor flash floods in the presence of turbid river water. Expanding this line of inquiry, Tamari et al. (2016) appraised the ability of LiDAR to measure river levels. These experiments, conducted under turbid water conditions, demonstrated that their method could measure the level with a reasonable uncertainty margin.

Martins et al. (2017) innovated a method to monitor the undular tidal bore of the Garonne River utilising a commercially available 2D LiDAR. This particular river is marked by non-hydrostatic wave behaviour and high water turbidity. A comprehensive wave-by-wave analysis was performed, with cross-comparisons between LiDAR, acoustic, and pressure-derived measurements to quantify the non-hydrostatic characteristics of the observed phenomenon. Although it is not a level measurement paper, it is a work that uses LiDAR to scan water surfaces.

Most of the aforementioned studies have implemented ultrasonic or LiDAR sensors for water level monitoring. Nonetheless, these works primarily relied on direct sensor readings without additional implementation of error minimisation methodologies. Their focus was predominantly centred on presenting application-specific Internet of Things architectures (Hanan and Sumadiyasa, 2019; Prafanto and Budiman, 2018) or directly evaluating sensor performance under varying conditions (Panagopoulos et al., 2021; Wannoi and Wannoi, 2022; Paul et al., 2020; Tamari and Guerrero-Meza, 2016; Tamari et al., 2016; Martins et al., 2017). Notwithstanding, other studies concentrated on refining the predictive capabilities of these sensors through diverse techniques.

Buhion et al. (2022) proposed an IoT system endowed with an SMS-based notification setup for flood monitoring. Two ultrasonic sensors were deployed (i.e., the URM07 and the HC-SR04 sensors), and the errors due to water turbulence were minimised by providing several iterations and discarding outliers.

Bae and Ji (2019) proposed a data processing algorithm focusing on outlier removal and smoothing for water level data collected by HC-SR04 ultrasonic sensors in stream-scale channels. The authors used modified Z-scores based on the median absolute deviation (MAD). The processed data are then smoothed using an exponentially weighted moving average (EWMA) method.

Rocchi et al. (2019) presented a sensing system for sea surface level measurement based on an ultrasonic sensor (SRF05) and the ATmega328p microprocessor. They proposed a low-cost ultrasonic sensor integrated into a floating device. The effects of climatic conditions were examined to determine the optimal operating range for the sensor and buoy architecture. The sensor exhibited signal anomalies at regular distance intervals. Hence, the authors combined multiple SRF05 sensors to optimise measurement accuracy. An analytical method based on ultrasonic signal reconstruction was presented to improve the measurement accuracy.

Table 1
Summary of features on related works on liquid level measurement systems. We assessed the use of LiDAR, ultrasonic sensors, error minimisation strategies, and auto-calibration methods.

Year	Work	LiDAR	Ultrasonic	Self adjustment	Multivariate approach
2022	Gondkar et al. (2022)	✗	✓	✗	✗
2022	Choubey et al. (2022)	✗	✓	✗	✗
2022	Jan et al. (2022)	✗	✓	✗	✗
2023	Djalilov et al. (2023)	✗	✓	✗	✗
2022	Sasikala et al. (2022)	✗	✓	✗	✗
2019	Hanan and Sumadiyasa (2019)	✗	✓	✗	✗
2018	Prafanto and Budiman (2018)	✗	✓	✗	✗
2021	Panagopoulos et al. (2021)	✗	✓	✗	✗
2022	Wannoi and Wannoi (2022)	✗	✓	✗	✗
2020	Paul et al. (2020)	✓	✗	✗	✗
2016	Tamari and Guerrero-Meza (2016)	✓	✗	✗	✗
2016	Tamari et al. (2016)	✓	✗	✗	✗
2017	Martins et al. (2017)	✓	✗	✗	✗
2022	Buhion et al. (2022)	✗	✓	✓	✗
2019	Bae and Ji (2019)	✗	✓	✓	✗
2019	Rocchi et al. (2019)	✗	✓	✓	✗
2020	Sahoo and Udgata (2020)	✗	✓	✓	✗
	Ours	✓	✓	✓	✓

Sahoo and Udgata (2020) proposed a system to increase the accuracy of water level measurements in storage tanks of varying depths using HC-SR04 ultrasonic sensors. To reduce measurement errors, the system applied a Levenberg–Marquardt Backpropagation Artificial Neural Network (LMBP-ANN).

Furthermore, flood warning systems can synthesise data from multiple river-level sensors to predict and manage flooding risks accurately. Alternative approaches can explore the Multiple Graph Learning Neural Networks (MGLNN) framework to integrate information from multiple graphs to optimise graph neural networks' learning and performance in semi-supervised classification tasks (Jiang et al., 2022). Another possibility is to apply object detection models to identify and calculate water level measurements based on images of fixed cameras, deployed alongside the experimental device described in this paper, and computer vision (Qiao et al., 2022; Roy et al., 2022).

Table 1 provides an overview of various relevant studies presented in this section. Most of the studies used ultrasonic sensors, some of which employed self-adjustment techniques. Others considered LiDAR sensors our focus area due to their reduced sensitivity to weather conditions. However, no previous work has utilised a multivariate approach with IMU sensors and a machine learning model to address the predictive inaccuracies inherent in LiDAR sensors. To our understanding, our research paper is pioneering in offering such an approach. For a more comprehensive evaluation, we carried out a systematic analysis, managing key factors that influence the performance of LiDAR sensors in water level measurement.

4. Dataset

We conducted thorough evaluations using collected data in a controlled environment to evaluate the methodology presented in this work, as described in Section 5. This approach allowed us to assess the methodology's performance under ideal conditions, minimising the impact of external factors such as weather conditions. By isolating variables such as sensor height relative to the water surface, tilt angle, and water turbidity, we could precisely determine their influence on the sensors and develop appropriate adjustments to account for them. The following subsections detail the data collection procedures and provide an exploratory inspection of the collected data.

4.1. Data collection

The data collection procedure was performed in the Hydraulics Laboratory of the São Carlos School of Engineering (EESC), University of São Paulo (USP). The experimental device, depicted in Fig. 1(a), comprised of a fixing plate equipped with two distance sensors (i.e., a LiDAR and an ultrasonic sensor) positioned symmetrically to enable direct comparisons. The LiDAR sensor utilised was the Tfmini LiDAR SRD8100, while the ultrasonic sensor was the HC-SR04. An Inertial Measurement Unit (IMU), in the form of an MPU6050 sensor, was incorporated to account for angle tilt. The plate was affixed to a ball head for precise control over the tilt angle.

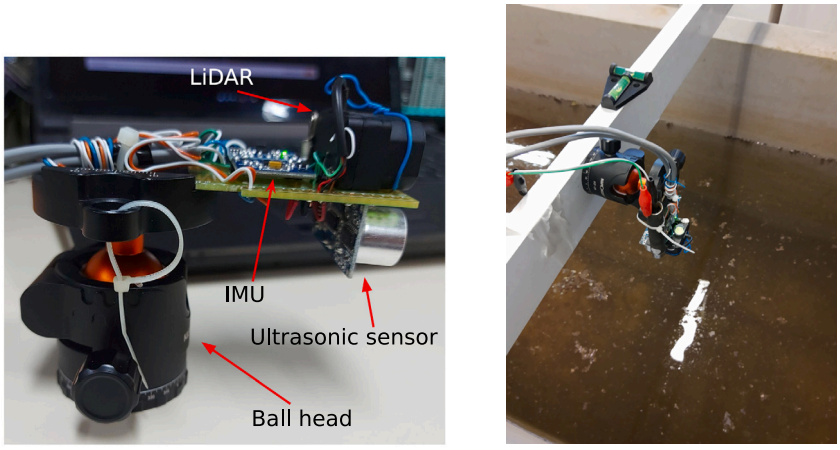
According to its specifications, the SRD8100 sensor can detect ranges up to 12 metres in ideal conditions. Such a short-range LiDAR sensor was employed to evaluate edge case scenarios since the laboratory was only four meters high. The distance between the sensors and the water bodies in the field might potentially reach significantly larger scales.

The device was affixed to a straight bar positioned orthogonal to the water surface within a laboratory tank, as illustrated in Fig. 1(b). To simulate varying water levels, the height of the bar was meticulously adjusted, maintaining its orthogonal orientation throughout the data collection process.

This setup provided us with control over three factors during data collection: the ground truth distance between the sensor and the water surface, the tilt angle, and the water turbidity. The distance between the sensor and the water surface, proportional to the water level, was represented by $h \in H$. This distance was measured along a line normal to the water surface plane and could be controlled by adjusting the height of a horizontal bar placed orthogonally to this plane. The tilt angle, denoted as $\theta \in \Theta$, refers to the rotation of the sensor in relation to the normal line. It was regulated using the ball head (see Fig. 1(a)). The turbidity, represented by $\tau \in T$, was regulated by introducing authentic soil samples to the tank and mixing them using a pump. The amount of soil added had a negligible volume compared to the volume of water in the tank. Hence, the distance h was not affected by this procedure. Fig. 2 illustrates this configuration.

We explored all possible combinations of these variables within the specified domain outlined in Eq. (3). The measurements in H were taken in centimetres, while the angles in Θ were measured in degrees. Although we measured the water turbidity in Nephelometric Turbidity Unit (NTU), we assigned categorical names to each of the three conditions considered, which composed the set T: low, medium, and high. These categories corresponded to specific turbidity measurements outlined in Table 2.

$$\begin{aligned} H &= \{50, 100, 150, 200, 250, 300\} \\ \Theta &= \{0, 2.5, 5, 7.5, 10\} \\ T &= \{low, medium, high\} \end{aligned} \tag{3}$$



(a) Experimental device. (b) Device placement.

Fig. 1. Setup for the data collection procedure. In (a), we show the experimental device consisting of one LiDAR sensor, one ultrasonic sensor, two IMUs, and a ball head, which controlled the sensors' tilt towards the water surface. In (b), we show the placement of the device in the water tank used for data collection.

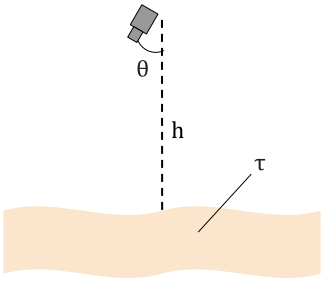


Fig. 2. Schematic representation of the controlled variables. The experimental device is positioned at a ground truth distance h from the water surface. The tilt angle θ was regulated using a ball head, while the water turbidity τ was manipulated by introducing soil into the water.

Table 2
Different categorical levels of water turbidity in T. Specifically, it provides the turbidity values in Nephelometric Turbidity Units for low, medium, and high categories.

Category	Value
Low	2.7 NTU
Medium	35.3 NTU
High	133.0 NTU

Let D be the set of all possible combinations of the controlled distances, tilt angles, and turbidity, as defined in Eq. (4).

$$D = \{(h, \theta, \tau) : h \in H, \theta \in \Theta, \tau \in T\} \tag{4}$$

A recording session k was performed for each $d_k \in D$. The total number of combinations in D is denoted by n so that $k \in \{1, \ldots, n\}$. Each recording session consisted of 300 data points, collected in a 1.5 Hz sampling rate and synchronised across all sensors. The set of data points corresponding to a tuple d_k is denoted by $X(d_k)$ in Eq. (5), where each point $\mathbf{x}_i^k \in X(d_k)$ is a multivariate structure with the different variables coming from each sensor.

$$\mathbf{X}(d_k) = \{\mathbf{x}_1^k, \ldots, \mathbf{x}_m^k\}, \quad d_k \in D \tag{5}$$

The data points comprised the variables outlined in Table 3. The LiDAR sensor provided IR_h and IR_s, representing the estimated distance and intensity of the returned light, respectively. The ultrasonic sensor yielded US_h, its estimate of distance. The IMU delivered 3D accelerometer readings as a_x, a_y, and a_z, along with 3D gyroscope readings as g_x, g_y, and g_z.

Table 3
Detailed descriptions of the variables captured by the experimental device, including measurements from LiDAR, ultrasonic sensors, and 3D motion sensors.

Variable	Description
IR_h	Distance estimated by the LiDAR sensor
IR_s	Intensity returned from the LiDAR sensor
US_h	Distance estimated by the ultrasonic sensor
$\{a_x, a_y, a_z\}$	3D accelerometer measurements in x, y and z
$\{g_x, g_y, g_z\}$	3D gyroscope measurements in x, y and z

The 300 data points in $X(d_k)$ for each condition d_k were gathered in a dataset with $N = 31{,}500$ samples used in the analyses presented in the upcoming sections.

4.2. Exploratory inspection

To identify potential data loss resulting from noise, we initiated our analysis with an exploratory examination of the data. While the ultrasonic sensor remains unaffected by the optical characteristics of the water body, the same is not true for the LiDAR sensor, which returns $IR_h = 0$ when it fails to return an accurate reading. Hence, by considering the number of samples in which $IR_h = 0$, we could partially assess the presence of invalid readings from the LiDAR sensor. Fig. 3 presents the numbers of samples with zero-valued IR_h across all combinations of variables h, θ and τ. Since this phenomenon happened only with the LiDAR sensor (i.e., there were no zero-valued US_h for any of the circumstances analysed), we did not show a similar graph for the ultrasonic sensor.

As shown in Fig. 4(a), there is a complex correlation between the absolute error measurement produced by IR_h and the tilt angle, which differs according to the water turbidity. The graph illustrates the tilt angle θ plotted against the corresponding average absolute error between the predicted value IR_h and the actual distance h, measured in centimetres.

We provided two panels illustrating this relationship: one containing all samples and the other one only with samples in which $IR_h \neq 0$. These two views are complementary, as each of them has its own limitations. The top panel showed higher errors when the angle

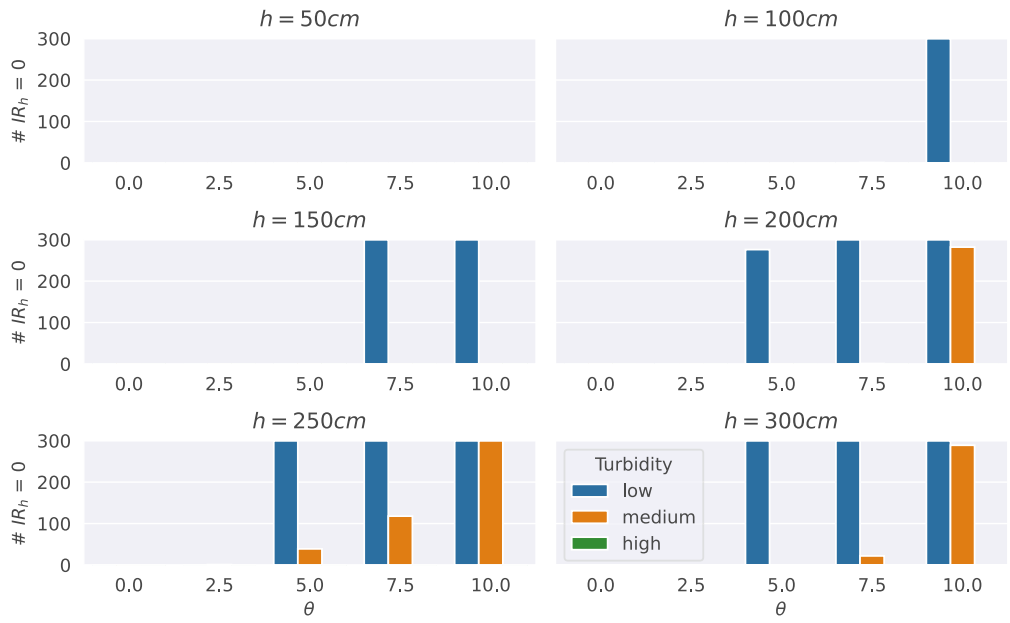


Fig. 3. Number of samples in which $IR_h = 0$ from the LiDAR sensor, for each distance h, turbidity τ, and angle θ. Lower values of τ and greater values of h make the LiDAR sensor more sensitive to θ.

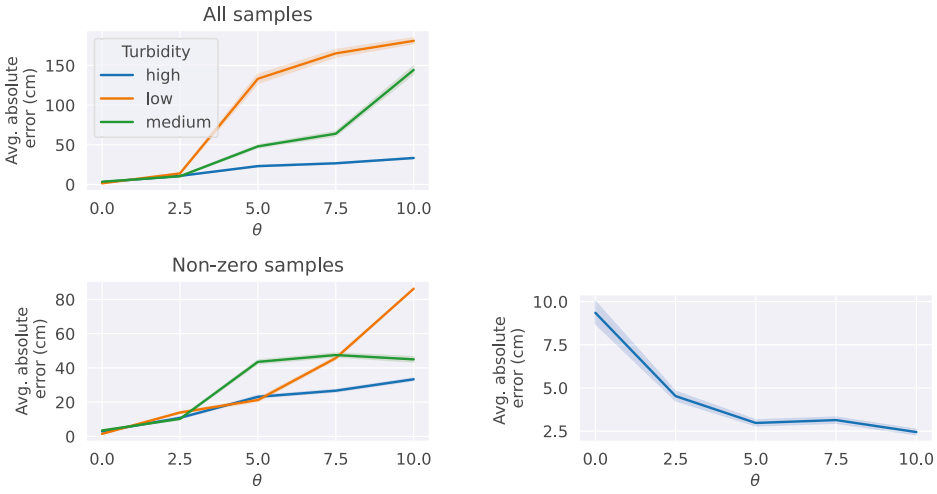


Fig. 4. Errors of the raw LiDAR and ultrasonar predictions per angle θ. The confidence interval is shown as shaded areas.

increased due to many samples where $IR_h = 0$. We added a bottom panel excluding these samples. This panel reveals that average error at higher θ values is skewed because of more samples with lower h values. This skewness results from discarding many higher h value samples due to sensor failures causing $IR_h = 0$ readings. Moreover, in both cases, the intricate relationship between the tilt angle and the absolute error is clearly visible.

Regarding the ultrasonic sensor, Fig. 4(b) shows the average absolute errors across all water turbidity conditions since this sensor is unaffected by this factor. As shown in the figure, the scale of the errors is one order of magnitude lower compared to the LiDAR. Interestingly, the error is lower when there is some inclination in the sensor, remaining as little as 2.5 cm for $\theta = 10°$.

The complex relationship between the variables and the need for a more sophisticated model to effectively perform error minimisation can also be verified by computing the Pearson correlation between each pair of variables in the dataset. We computed such a correlation matrix, illustrated as a heatmap in Fig. 5. Only correlation coefficients with absolute values exceeding a threshold of 0.25 are explicitly displayed.

Upon examination, it is evident that the variables US_h and h exhibit a robust positive correlation, approaching the coefficient of 1.0. Conversely, a relatively weaker correlation, with a coefficient around 0.5, is observed between IR_h and h, underscoring the importance of addressing the error introduced by the LiDAR sensor. As expected, a significant correlation emerges between the actual tilt angle θ and the accelerometer measurements a_x (close to -1) and a_y (approximately -0.7). Additionally, θ demonstrates a noticeable correlation with IR_s, with a coefficient around -0.6. This result suggests the potential of leveraging this variable, in conjunction with other factors like IMU readings, as a valuable input in a prediction model. The IMU variables

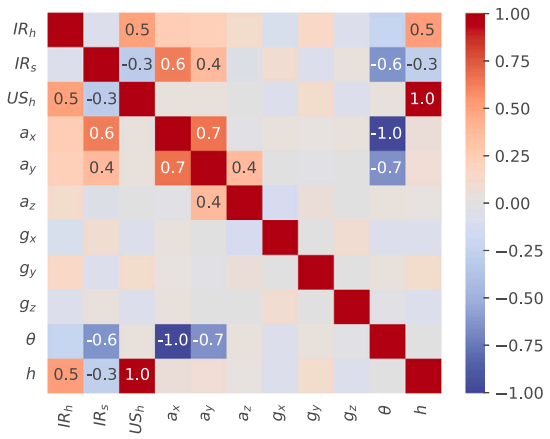


Fig. 5. Correlation matrix illustrating the pairwise Pearson correlation coefficients calculated among variables in the dataset.

a_x and a_y also have a non-trivial correlation with IR_s (0.6 and 0.4, respectively), while these two variables correlate among themselves with a 0.7 coefficient.

5. Water level identification

This section outlines the methods employed to evaluate and enhance the precision of the LiDAR and ultrasonic sensors in identifying water levels, which can be directly derived from the estimated distance between the sensor and the water surface produced by our approach. We introduce the analyses undertaken on the predictions from the individual sensors and the machine learning-based approach designed and employed to generate more accurate predictions using the LiDAR sensor and the IMU.

As already presented in Table 3, our dataset consisted of nine variables. Among them, two were associated with the LiDAR sensor (i.e., IR_h and IR_s), one with the ultrasonic sensor (i.e., US_h), and six with the IMU (i.e., a_x, a_y, a_z, g_x, g_y, and g_z). The inclusion of the ultrasonic sensor in our comparisons was motivated by its widespread usage as a non-immersive sensor for water level identification (Loizou and Koutroulis, 2016; Pasika and Gandla, 2020; Djalilov et al., 2023), with accurate results on controlled environments. For this reason, we employed it as a reference to evaluate the results of our models.

The problem was formulated as the regression task of estimating the ground truth distance h between the sensor and the water surface. Let $h'(\mathbf{X}(d_k))$ be a prediction of the ground truth value h, obtained from the set of data points $\mathbf{X}(d_k)$ corresponding to recording session k. To mitigate the influence of outliers present in the recording session, we opted to utilise the median value as the predicted value, as outlined in Eq. (6).

$$h'(\mathbf{X}(d_k)) = median(\mathbf{X}(d_k)[s_h]), \quad s_h \in \{IR_h, US_h\} \tag{6}$$

We calculated three regression metrics to assess the accuracy of the predictions (Glantz et al., 2001): the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the R^2 score (i.e., coefficient of determination). Additional analyses were performed in terms of the absolute error L computed for each prediction as in Eq. (7).

$$L(h'(\mathbf{X}(d_k))) = \left|\mathbf{X}(d_k)[h] - h'(\mathbf{X}(d_k))\right| \tag{7}$$

We began by performing a comparative analysis of measurements obtained from both the LiDAR and ultrasonic sensors. Our primary focus was on evaluating the performance of the individual sensors.

Table 4
List of input variables utilised by the models, categorised by the type of distance sensor—LiDAR and Ultrasonic.

Distance sensor	Input variables
LiDAR	$\{IR_h, IR_s, a_x, a_y, a_z, g_x, g_y, g_z\}$
Ultrasonar	$\{US_h, a_x, a_y, a_z, g_x, g_y, g_z\}$

Table 5
Hyper-parameters used for the machine-learning algorithms, cautiously tuned in exploratory tests.

Method	Parameter	Value
KNN	Number of neighbours	5
	Distance metric	Euclidean
Random Forest	Number of estimators	100
	Criterion	Squared error
AdaBoost	Number of estimators	150
	Learning rate	0.5
	Loss function	Linear
	Maximum depth	10

Hence, we assessed the predictions provided directly from variables IR_h and US_h to assess the performance of each sensor in its original configuration.

Afterwards, we constructed machine-learning models combining the variables from the LiDAR or ultrasonic sensors with those from the IMU. In the LiDAR sensor experiments, we excluded the US_h variable from the dataset and utilised the remaining variables. Conversely, in the ultrasonic sensor experiments, we excluded the IR_h and IR_s variables. We chose not to analyse models that combine LiDAR and ultrasonar because, in our controlled experimental scenario, models based only on the ultrasonic sensor and the IMU consistently yielded minimal errors (refer to Section 6), leaving little room for further improvement. The input variables utilised for the LiDAR and ultrasonic models are outlined in Table 4.

We conducted a thorough assessment of four machine-learning techniques: Linear Regression (LR), k-Nearest Neighbours (KNN), Random Forest (RF), and AdaBoost (Ada) (Hastie et al., 2009). The LR technique was used as our baseline regressor to establish a foundation for comparison. The KNN regressor was assessed as a simple yet effective instances-based approach. The RF and AdaBoost regressors, both based on decision tree ensembles, were evaluated as more sophisticated machine-learning models. Let A be the set of machine-learning techniques employed, given in Eq. (8).

$$A = \{\texttt{LR}, \texttt{KNN}, \texttt{RF}, \texttt{Ada}\} \tag{8}$$

Since the samples consisted of tabular data with independent samples, it would make little sense to employ deep learning algorithms that benefit from temporal or spatial dependencies (Goodfellow et al., 2016), such as Convolutional Neural Networks (CNN) or Long Short-Term Memory (LSTM). For this reason, we did not consider these approaches in our experiments.

The LR model was optimised using the least squares method with no special parametrisation. Regarding the hyper-parameters of the other machine-learning methods, Table 5 shows the values used in our experiments. We configured the KNN regressor to make predictions based on 5 nearest neighbours, employing the Euclidean distance metric. The RF regressor was set up with 100 estimators and the squared error function to measure the quality of a split. The AdaBoost regressor was endowed with 150 decision tree estimators with a maximum depth equal to 10 and a learning rate of 0.5.

For splitting the data, we implemented measures to prevent any data leakage between the train and test sets. Since the controlled variables h, θ and τ remained consistent throughout each recording session, we assigned all data from each session exclusively to either the train or test set. This protocol was crucial to avoid inflated performance metrics when evaluating the test set.

We adopted a leave-one-out approach using each subset $\mathbf{X}(d_k)$, corresponding to an individual recording session k, as a test set, while the data points from the remaining recording sessions were used as the train set. Thus, the number of splits was equal to the total number of recording sessions, denoted by n. Since $length(H) = 6$, $length(\Theta) = 5$ and $length(T) = 3$, we ended up with $n = 6 \times 5 \times 3 = 90$ splits. The formal definition of the train and test sets, denoted as $train_k$ and $test_k$ for each k, is presented in Eq. (9).

$$\begin{cases} train_k = \mathbf{X} - \mathbf{X}(d_k) \\ test_k = \mathbf{X}(d_k) \end{cases} \quad k \in \{1, \dots, n\} \tag{9}$$

For each machine-learning architecture $a \in A$, we trained the models M_k^a using the train set of each split k. We generated the predictions $M_k^a(test_k)$ based on the corresponding test set. Afterwards, we aggregated the median predictions from all splits in a data frame H'_a, as defined in Eq. (10).

$$H'_a = median(M_1^a(test_1)) \cup \cdots \cup median(M_n^a(test_n)) \tag{10}$$

We calculated multiple evaluation metrics using H'_a. Additionally, we analysed the distributions under various conditions to evaluate the performance of different machine learning models and their behaviour across different values for water turbidity. The subsequent section presents these findings.

6. Results

In this section, we provide the outcomes of the methodologies detailed previously. The results include evaluation metrics, boxplots, significance tests, and relevant supplementary information required for their interpretation. Further analysis and discussions on these findings are also presented.

We conducted a two-stage assessment of the results. In the first stage, we evaluated the performance of individual sensors using only the variables IR_h and US_h, without employing any error minimisation techniques. In the second stage, we applied machine learning models and examined the enhancements observed under diverse conditions. To ensure consistent statistical significance across our analyses, we set $\alpha = 0.05$ for all conducted tests.

6.1. Individual sensors performance

The absolute errors h' from the LiDAR and ultrasonic sensors, measured in centimetres and categorised according to water turbidity, are represented by the boxplots in Fig. 6. It is worth mentioning that, in this particular figure, we are showing the distributions across all data points, not only the median values.

An independent t-test was applied for each water turbidity to assess the statistical significance of the results. The error was the dependent variable, while the type of sensor was the independent variable. The null hypothesis was that both the LiDAR and ultrasonic sensors led to similar absolute errors. The results are presented along with Fig. 6 using a star notation.[1][2]

The ultrasonar provided lower errors, with statistical significance, for all turbidity levels. This result was expected due to the robustness of the ultrasonic sensor to incidence angles, leading to better performance than the LiDAR regardless of the water turbidity.

Reminding the number of zero-valued readings of IR_h, shown in Fig. 3, we could verify that the LiDAR sensor encounters a higher failure rate under conditions of low water turbidity, especially when h and θ increase. This scenario influenced the individual sensors predictions shown in Fig. 6, in which a large number of invalid readings of the LiDAR sensor when $\tau = low$ resulted in many samples with $IR_h = 0$, increasing the overall error rates.

The evaluation metrics (i.e., MAE, RMSE, and R^2 score), computed based on the median values $h'(\mathbf{X}(d_k))$ and presented in Table 6, corroborate this analysis.

We analysed the influence of the turbidity level τ on the absolute error L for each sensor. To determine if there were any statistically significant differences among the three possible values in T, we performed a one-way ANOVA test. The results of this test are presented in Table 7. The independent variable was τ, the dependent variable was the absolute error L yielded by each sensor, and the null hypothesis was that the water turbidity did not affect the absolute error for a given sensor.

Fig. 7 presents the boxplots illustrating the absolute errors of both the LiDAR and ultrasonic sensors for each turbidity level τ. The figure also includes the outcomes of the post hoc independent t-tests conducted to determine pairwise statistically significant differences across the water turbidity values. We show these results only for the LiDAR sensor because the ANOVA test for the ultrasonic sensor yielded $p = 0.93 > 0.05$ (see Table 7), failing to reject the null hypothesis.

The null hypothesis of the ANOVA test was rejected for the LiDAR but not for the ultrasonic sensor. This outcome aligns with expectations since the ultrasonic sensor operates using sound waves, which remain unaffected by the water turbidity. The results from the post hoc t-tests applied to the LiDAR sensor, shown in Fig. 7, confirm the expected trend: the absolute error was significantly higher for $\tau = low$ than for the other conditions. No statistical significance was observed when comparing $\tau = medium$ to $\tau = high$. Moreover, by observing the scale of each plot, it becomes evident that, overall, the ultrasonic sensor led to better performance.

So far, we have presented an analysis of the predictions provided by the individual sensors without any error minimisation technique. As expected, the outcomes were unfavourable for the LiDAR sensor, as the ultrasonic sensor displayed reduced errors and greater stability across various turbidity levels and incidence angles.

In the subsequent subsection, we present the outcomes of our machine-learning approach, which integrated variables from both the LiDAR and the IMU. This approach aimed to bridge the disparity between the results obtained from these sensors, ultimately achieving comparable performance while leveraging the advantageous features of the LiDAR, such as its resilience to adverse weather conditions. As a reference for comparison, we also considered combining the ultrasonar with the IMU in the same conditions experimented for the LiDAR.

Table 6
Performance evaluation of LiDAR and Ultrasonic sensors across different turbidity levels, with metrics presented in centimetres (cm).

Modality	Turbidity	MAE	RMSE	R^2
LiDAR	Low	115.83	170.35	−1.90
	Medium	48.41	82.99	0.31
	High	21.13	27.23	0.93
Ultrasonar	Low	10.15	44.61	0.80
	Medium	12.57	45.50	0.79
	High	8.58	44.78	0.80

Table 7
One-way ANOVA test results comparing prediction errors of LiDAR and Ultrasonic sensors.

Sensor	F-value	p-value
LiDAR	11.86	$<10^{-4}$
Ultrasonar	0.07	0.93

[1] p-value annotation legend:
($ns \rightarrow p > 0.05$); ($* \rightarrow p \in [0.01, 0.05]$); ($** \rightarrow p \in [0.001, 0.01]$);
($*** \rightarrow p \in [0.0001, 0.001]$); ($**** \rightarrow p < 10^{-4}$).

[2] In this paper, we maintained a consistent statistical significance level by setting $\alpha = 0.05$ for all conducted tests.

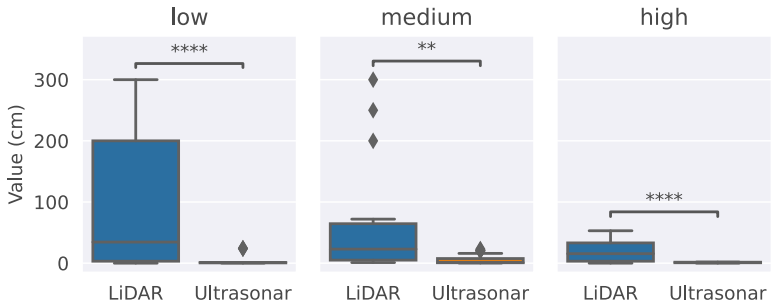


Fig. 6. Boxplots of the absolute errors from the LiDAR and ultrasonic sensors for each water turbidity considered. An independent t-test was applied for each water turbidity to evaluate the null hypothesis that both sensors lead to the same error. All data points were considered, not only the median values h'.

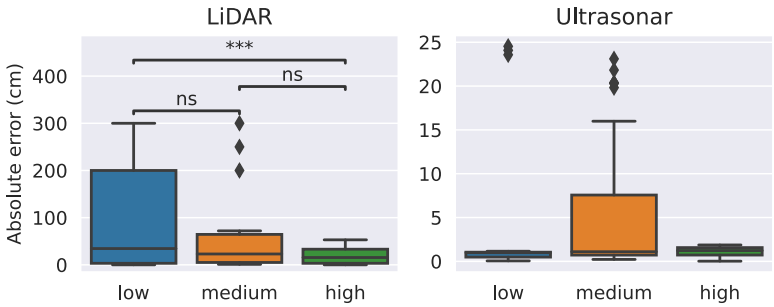


Fig. 7. Boxplots and results for the pairwise independent t-tests applied to the raw sensors' data. The independent variable was the absolute error, while the dependent variable was the water turbidity τ. Notice the different scales of each plot.

6.2. Machine-learning approach

In this subsection, we present the results of the machine learning-based approach proposed in this paper. We present a comparison between these results to those of the individual sensors, besides an assessment of the different machine-learning techniques considered.

6.2.1. Overall results

Since the RF models were the most accurate for both the LiDAR and ultrasonic sensors (see Section 6.2.2 for details), we compared their predictions with those of the individual sensors. The analyses were segmented for all turbidity levels and tilt angles. Henceforth, we refer to the RF model based on the LiDAR sensor as IR-RF and the RF model based on the ultrasonic sensor as US-RF. The individual predictions from the LiDAR and ultrasonic sensors (i.e., the median values of IR_h and US_h across each recording session) are denoted as IR and US, respectively.

Table 8 shows the overall evaluation metrics for the IR, US, IR-RF, and US-RF models. Besides, it depicts the metrics separated by water turbidity τ. Table 9 provides a similar comparative presentation for the tilt angle θ. In both cases, the measurements are shown in centimetres.

We computed one-way ANOVA tests to identify statistically significant differences between the absolute errors (dependent variable) yielded by each of the above-mentioned models (i.e., IR, US, IR-RF and US-RF). In Table 10, we show the results for the overall results and the results separated by water turbidity. In Table 11, we show the results separated by tilt angle.

Thereafter, we conducted post hoc tests using pairwise independent t-tests to further analyse the results. We did not perform post hoc t-tests for $\theta = 0$, for which the ANOVA test failed to reject the null hypothesis (see Table 11). The results, presented in centimetres, are displayed alongside the boxplots in Fig. 8.

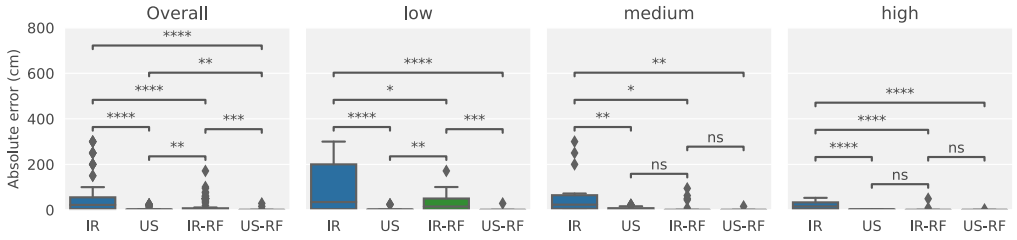
Through an examination of the performance metrics for the IR, US, IR-RF, and US-RF models categorised by water turbidity (Table 8) and tilt angle (Table 9), several observations emerge. The IR model yielded unfavourable outcomes for $\tau \in \{low, medium\}$ and $\theta \geq 5°$, as evidenced by R^2 scores falling below zero. However, applying the proposed machine-learning approach in the IR-RF model significantly mitigated these errors.

Table 8
Evaluation metrics for each model for measuring h across different turbidity levels, with results in centimetres.

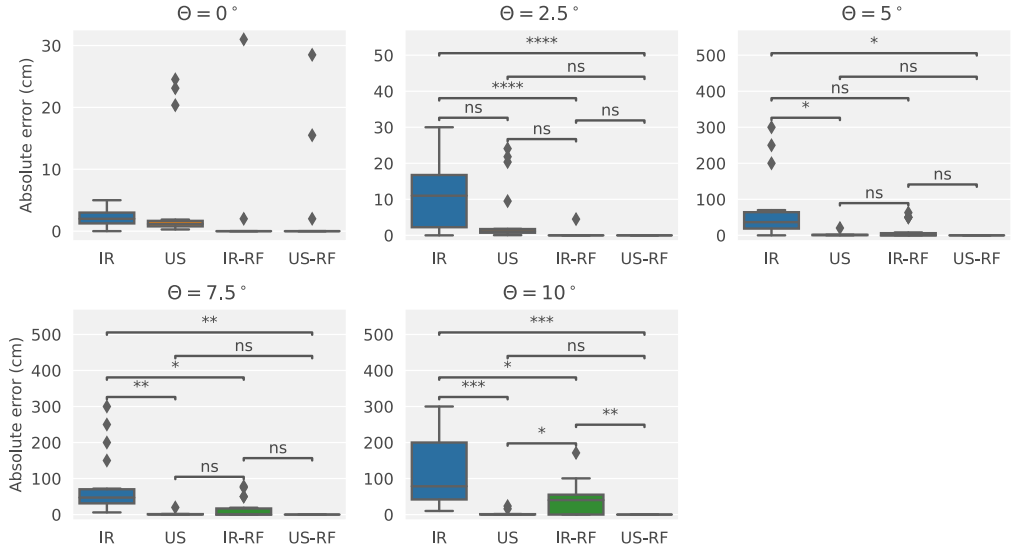
Turbidity	Model	MAE	RMSE	R^2
Overall	IR	56.31	99.96	−0.37
	US	3.28	7.31	0.99
	IR-RF	13.94	32.47	0.86
	US-RF	0.51	3.43	1.00
Low	IR	99.27	146.93	−1.96
	US	3.09	7.65	0.99
	IR-RF	31.26	50.01	0.66
	US-RF	0.95	5.20	1.00
Medium	IR	50.38	88.01	−0.06
	US	5.69	10.02	0.99
	IR-RF	8.65	24.08	0.92
	US-RF	0.52	2.83	1.00
High	IR	19.28	25.31	0.91
	US	1.07	1.20	1.00
	IR-RF	1.90	9.06	0.99
	US-RF	0.07	0.37	1.00

On the other hand, although the US-RF model exhibited some improvement towards the original US model, such an improvement was comparatively modest and, as shown in the analyses, not statistically significant. This result is related to the robustness of the ultrasonar to varying incidence angles. The ultrasonic sensor leads to low error rates even without any minimisation techniques (see Section 6.1), indicating that the potential for improvement is comparatively lower than that of the LiDAR sensor.

The boxplots presented in Fig. 8(a), and the corresponding pairwise t-tests provide further evidence supporting the patterns observed for the absolute errors. In the overall scenario, all models exhibited statistically



(a) Boxplots for overall results and for each τ.



(b) Boxplots for each θ.

Fig. 8. Boxplots and results for the pairwise independent t-tests applied to the different models for measuring the distance h, including the medians of the raw sensor data (i.e., IR and US) and the medians of the corresponding predictions of the Random Forest models (i.e., IR-RF and US-RF). The independent variable was the absolute error, and the null hypothesis was that each pair of models led to the same absolute error. Separate plots are shown for each water turbidity τ; and tilt angle θ.

Table 9
Evaluation metrics for each model measuring h, separated by angle, in centimetres.

Angle	Model	MAE	RMSE	R^2
0°	IR	2.33	2.65	1.00
	US	4.65	9.34	0.99
	IR-RF	1.83	7.32	0.99
	US-RF	2.56	7.66	0.99
2.5°	IR	11.39	14.41	0.97
	US	4.92	9.35	0.99
	IR-RF	0.25	1.06	1.00
	US-RF	0.00	0.00	1.00
5°	IR	67.22	108.89	−0.63
	US	1.98	4.90	1.00
	IR-RF	12.28	25.31	0.91
	US-RF	0.00	0.00	1.00
7.5°	IR	80.22	115.89	−0.84
	US	1.96	4.78	1.00
	IR-RF	15.74	30.97	0.87
	US-RF	0.00	0.00	1.00
10°	IR	120.39	156.38	−2.35
	US	2.91	6.76	0.99
	IR-RF	39.58	60.15	0.50
	US-RF	0.00	0.00	1.00

Table 10
Results for the one-way ANOVA test applied to the water-level estimation models, separated by water turbidity.

Turbidity	F-value	p-value
Overall	30.77	$<10^{-4}$
Low	18.32	$<10^{-4}$
Medium	10.57	$<10^{-4}$
High	28.02	$<10^{-4}$

Table 11
One-way ANOVA test results for evaluating differences in water-level estimation models at distinct angles.

Angle	F-value	p-value
0°	0.62	0.61
2.5°	13.60	$<10^{-4}$
5°	8.70	$<10^{-4}$
7.5°	12.59	$<10^{-4}$
10°	17.75	$<10^{-4}$

significant differences from each other. The US-RF model emerged as the top-performing model, closely followed by UR and IR-RF, which exhibited more outliers. The original IR model performed worse than all other models.

When considering the models' performance based on different turbidity levels, the overall trends resembled those observed for $\tau = low$. However, for $\tau \in medium, high$, the t-tests did not reject the null hypothesis when comparing IR-RF to US or US-RF. These findings highlight the significant error reduction achieved by the proposed approach for the LiDAR sensor, particularly in the case of $\tau = medium$. This improvement resulted in outcomes that closely aligned with those of the ultrasonic sensor, except for $\tau = low$.

By examining the outcomes categorised by tilt angle in Fig. 8(b), a discernible trend becomes apparent: the IR-RF model displayed no statistically significant differences when compared to the US and US-RF models for all conditions but the most extreme scenario with $\theta = 10°$. Under this particular condition, the US and US-RF models yielded lower absolute errors than the IR-RF model. It is worth recalling that the null hypothesis was not rejected during the ANOVA test conducted for $\theta = 0°$, indicating that the models did not exhibit statistically significant differences under this specific tilt angle.

The statistical analysis reveals that, for $\theta = 2.5°$, the IR model did not exhibit any significant differences compared to the US model. However, for $\theta \geq 5°$, the US model consistently yielded lower absolute errors. Additionally, the results indicated that while the US and US-RF models did not differ significantly in statistical terms, the evaluation metrics hinted at an improved performance for US-RF. Conversely, the IR-RF model consistently outperformed the IR model for $\theta \geq 2.5°$, except for $\theta = 5°$. This finding underscores the effectiveness of the proposed approach in reducing errors associated with the LiDAR sensor.

6.2.2. Models assessment

As already stated, we evaluated the four machine-learning models in A. We analysed the distribution of the prediction values besides computing the MAE, RMSE, and R^2 metrics. Since the data frame H'_a for each $a \in A$ contained predictions for each data point in the dataset, we could perform direct comparisons with the predictions from the individual sensors.

The evaluation metrics in centimetres for each model are shown in Tables 12 for the LiDAR sensor and 13 for the ultrasonic sensor. We provide both overall metrics and metrics categorised by water turbidity to analyse how each machine-learning model impacted the absolute error. We considered the different conditions of water turbidity for the LiDAR-based models. Since the ultrasonic sensor remains unaffected by water turbidity, we present only the overall results for the models based on it.

The evaluation metrics presented in Tables 12 and 13 reveal that the RF models consistently delivered the best overall performance for both sensors, closely followed by the AdaBoost models. These models led to R^2 scores higher than 90% and MAE values no higher than 10.58 cm for all conditions except for the LiDAR sensor with low water turbidity. For the ultrasonic sensor, the RF and AdaBoost models led to an overall error very close to zero.

While the KNN regressor exhibited inferior evaluation metrics compared to the ensemble-based models (RF and Ada) for all conditions except the LiDAR sensor with low water turbidity, it still effectively reduced errors compared to the LiDAR sensor's LR baseline. However, for the ultrasonic sensor, KNN did not outperform LR regarding error reduction.

In Table 14, we performed a repeated-measures ANOVA test to determine whether there were any statistically significant differences in the outcomes of each model. The independent variable was the machine learning model, the dependent variable was the absolute error, and the null hypothesis was that all trained models led to similar absolute errors.

Table 12
Evaluation metrics across different turbidity levels for each machine learning algorithm applied to data from the LiDAR sensor and IMU.

Turbidity	Model	MAE	RMSE	R^2
Overall	Linear Regression	59.68	72.90	0.27
	KNN	27.99	42.31	0.75
	Random Forest	15.04	34.92	0.83
	AdaBoost	15.70	35.81	0.82
Low	Linear Regression	74.86	88.87	−0.08
	KNN	34.67	49.80	0.66
	Random Forest	31.26	50.01	0.66
	AdaBoost	30.37	49.18	0.67
Medium	Linear Regression	56.06	68.51	0.36
	KNN	19.67	28.93	0.89
	Random Forest	8.65	24.08	0.92
	AdaBoost	10.58	26.76	0.90
High	Linear Regression	42.76	50.29	0.65
	KNN	16.17	22.97	0.93
	Random Forest	1.90	9.06	0.99
	AdaBoost	1.67	9.13	0.99

Table 13
Evaluation of machine learning algorithms based on data from the ultrasonic sensor and IMU.

Model	MAE	RMSE	R^2
Linear Regression	6.21	13.00	0.98
KNN	10.36	19.64	0.95
Random Forest	0.92	8.00	0.99
AdaBoost	1.17	12.85	0.98

Table 14
Repeated measures ANOVA test results for machine learning models by sensor type and turbidity level.

Sensor	Turbidity	F-value	p-value
LiDAR	Overall	63.63	$<10^{-4}$
	Low	16.07	$<10^{-4}$
	Medium	43.91	$<10^{-4}$
	High	39.43	$<10^{-4}$
Ultrasonar	Overall	26.92	$<10^{-4}$

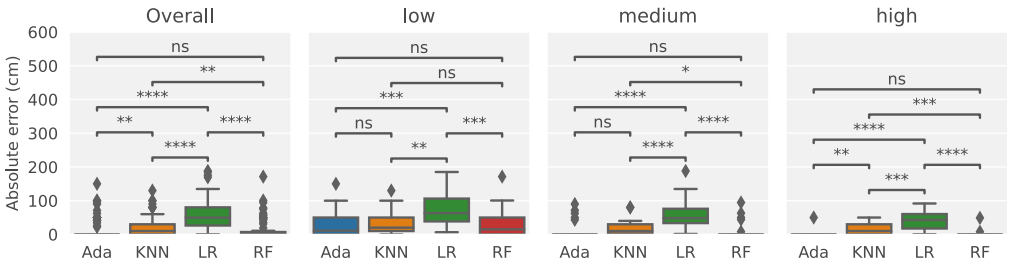
Since the null hypothesis was rejected for all cases, we conducted pairwise post hoc paired t-tests to identify which pairs of models differed from each other with statistical significance. The results from these tests are shown, along with the corresponding boxplots, in Fig. 9.

For both sensors, no statistically significant differences were observed in the absolute errors between the RF and AdaBoost models. However, the LR regressor exhibited larger errors than all other models when applied to the LiDAR sensor. Conversely, in the case of the ultrasonic sensor, the LR regressor only displayed a non-significant difference when compared to the KNN regressor. Thus, the apparent numerical advantage of LR over KNN could not be confirmed by significance tests.

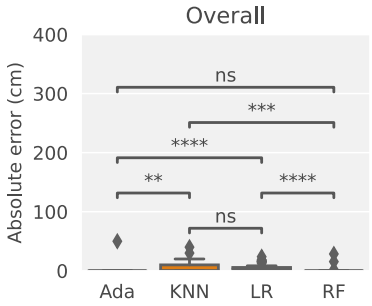
When considering the LiDAR sensor under conditions of low water turbidity, both the RF and AdaBoost models showed no statistically significant differences when compared to KNN. However, while AdaBoost still lacked statistical significance compared to KNN in the medium turbidity scenario, RF demonstrated a significant difference. In all other conditions, both the RF and AdaBoost models exhibited statistically significant improvements over KNN.

6.2.3. Feature importance

Besides being the most accurate, the tree-based models analysed (i.e., RF and Ada) allowed for computing importance values for each input variable based on the Mean Decrease in Impurity (MDI) (Scornet, 2023). The resulting values are shown in Fig. 10. To aid in visualising the contributions of the other variables, we presented the graphics on



(a) LiDAR sensor.



(b) Ultrasonic sensor.

Fig. 9. Boxplots and results for the pairwise paired t-tests applied to the machine learning models (dependent variable). The independent variable was the absolute error, and the null hypothesis was that each pair of models led to the same absolute error. For the LiDAR sensor, separate plots are shown for each sensor and water turbidity.

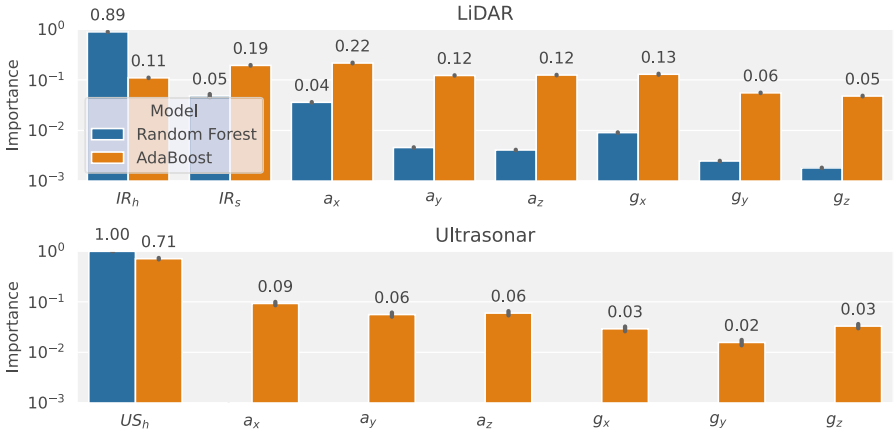


Fig. 10. Feature importance analysis for Random Forest and AdaBoost models using LiDAR and ultrasonic sensors. The figure showcases the logarithmic scale plots to highlight the significance of variables beyond IR_h, IR_s, and US_h. The values of contributions exceeding 10^{-2} are displayed, while contributions below 10^{-3} are omitted for clarity.

a logarithmic scale. Besides, for clarity purposes, we omitted negligible contributions with importance values below 10^{-3}.

Despite their close performance, the feature importance values revealed distinct behaviours between RF and AdaBoost. RF showed a significant contribution from the variables IR_h and US_h. Specifically, the LiDAR sensor placed 89% importance on IR_h, with some consideration given to IR_s (5%) and a_x (4%). The remaining variables contributed less than 1% each. On the other hand, the ultrasonic sensor was largely influenced by the variable US_h, which accounted for 100% of the importance.

AdaBoost provided a more balanced distribution of importance among the variables, with a_x (22%) being more influential than IR_h (11%) and IR_s (19%). All accelerometer variables contributed more than 10%, as did g_x. For the ultrasonic sensor, although the most important variable was US_h (71%), all the remaining variables contributed more than 1%, and a_x had an importance of 9%.

The divergent behaviours observed between these two ensemble algorithms can be attributed to their underlying characteristics (Lakshmanan et al., 2020). RF operates on the principle of bootstrap aggregation (i.e., bagging), in which multiple weak learners are trained on different bootstrapped datasets to minimise variance and enhance generalisation. Given that IR_h and US_h directly estimate the target variable h, it is reasonable to expect that these variables offer the highest reduction in entropy during the training process.

In contrast, AdaBoost utilises a boosting technique that aims to minimise bias and optimise predictions by training subsequent learners

to rectify the mistakes of their predecessors. As a result, this approach incorporates a wider range of features to capture and understand the relationships between them.

7. Discussion

In the preceding sections, we presented the evaluation metrics, boxplots, and significance tests corresponding to our analyses, shedding light on the individual performance of each model. In this section, we provide some insights into the implications of these results.

Our study consisted of a data-driven approach to assess the influence of controlled factors in a water measurement system composed of a set of sensors and a machine-learning model. We collected the first dataset synchronised data from LiDAR and IMU devices that controlled angle, distance and water turbidity in the domain of water level measurement. The proposed approach consisted of fusing both sensors using traditional machine-learning techniques to provide better predictions when compared to the individual sensors.

For the LiDAR and ultrasonic experiments, the baseline models were the predictions provided by the individual sensors. We also provided a straightforward logistic regression model and a KNN approach as machine-learning baselines to compare to the tree-based models, which we expected to generate accurate predictions due to their known ability to model tabular data (Shwartz-Ziv and Armon, 2022).

There were no other models deployed in similar datasets in the literature. Our dataset can be a starting point to evaluate different types of models that deal with water level measurement with combinations of laser and ultrasonic sensors with inertial units. Although more sophisticated patterns, such as long-term temporal dependencies, can only be achieved in real-world environments, our dataset has stationary data that can be used to improve models for deployed systems.

Using contactless sensors can prevent interference in measurements and physical damage caused by debris during flood events. LiDAR equipment is highly resistant to environmental factors and requires minimal maintenance. Additionally, LiDAR sensors are less sensitive than ultrasonic sensors to interference from fog and rainfall. However, the extent of this interference for the specific application has not been quantified, making it an interesting area for future research.

As mentioned earlier, the angle at which the LiDAR sensors are positioned is crucial for accurate readings. Therefore, we were concerned about wind and soil vibrations caused, for instance, by vehicles passing by. To address this, we propose integrating the IMU to manage angle variations and predict and overcome potential wind and vibration-related issues. It is important to note that solid and stable support for the sensor equipment is essential for the successful implementation of this plan.

As depicted in Fig. 3, three factors may cause the LiDAR sensor to fail, yielding only zero-valued readings: increased tilt angle, increased distance and low water turbidity. Apart from using higher-end sensors or implementing redundancy with other sensors, there are a few alternative ways to handle each of these complicating factors.

The angle of a deployed sensor can be remotely monitored using the IMU data. Suppose such an angle increases systematically and prominently enough to harm the quality of the readings. In that case, the operating staff needs to perform maintenance on the physical installation of the device. Punctual fluctuations in the angle might be controlled by using more rigid structures to hold the system. Since this condition can be directly monitored through the IMU, it would be trivial to diagnose an increased tilt angle in case of a sensor failure.

Sensor failures due to an increased distance between the sensor and the water are not a critical concern because they happen when the water level is low. However, with the sensing apparatus proposed, there is not a trivial way to discern a failure due to increased distance than a failure due to low water turbidity. Even though the water turbidity usually increases in flooding scenarios, a deployed system must be able to work even if this is not the case.

Thus, another way to account for sensor failures in a changing environment, such as during heavy rainfall, is to use historical data and a forecasting model to identify missing values. This approach is not applicable in our controlled experiment, but it can provide robustness in a deployed system that records the sensor readings over time.

We generated a dataset on a controlled scenario with a discrete set of distances, angles and water turbidity levels. This configuration differs from a real-world scenario in which these values are unconstrained. The splits for training and testing were designed to evaluate the generalisation capacity by testing the recording session of one combination (h, θ, τ) while training with all others. Since we assigned a prediction per recording session, this procedure corresponds to a leave-one-out protocol. This approach was especially suitable because we had a relatively small number of recording sessions $n = 90$.

Overall, the results showed that the machine learning-based approach could effectively reduce the error from the LiDAR sensor whenever it did not completely fail (see Fig. 3). The ensemble-based models applied to the LiDAR sensor achieved comparable absolute errors to those obtained by the ultrasonic sensor. The overall results for IR-RF were accurate even for $\theta \in \{5°, 7.5°\}$, since it could work under these conditions provided $\tau \in \{medium, high\}$.

These results are highly promising in terms of practical applications. Rainfall is recognised for its tendency to elevate water turbidity, particularly during flash flood events (Lu et al., 2023). Consequently, given a certain increment in turbidity has happened, the IR-RF model can provide accurate predictions even with increased tilt angles (see Fig. 8(b)). Therefore, with careful consideration for tilt angles, a LiDAR can yield accuracy comparable to that of an ultrasonic sensor if implemented along with an IMU and using our machine learning-based approach.

It is important to note that, for $\tau = medium$, the LiDAR sensors encountered difficulties at $\theta = 10°$ (refer to Fig. 3), but only when the distance values exceeded $h \geq 200$. Interestingly, even for low water levels, the sensor's sensitivity to increased tilt angles was significantly influenced by the distance h. These findings emphasise the critical role of the sensor's range in ensuring its robustness when confronted with variations in the tilt angle, making it a crucial consideration for future experiments or practical applications.

The LiDAR-based level measurement has proven to be effective in measuring turbid fluids, but its performance is significantly affected when the turbidity level is reduced. This limitation arises from the impact of reflectance on the accuracy of LiDAR measurements (Kaasalainen and Malkamäki, 2020). It is worth noting that, in urban environments, water turbidity tends to increase during storm runoff events due to the introduction of dirt into the stream. The distance between the sensor and the water surface, another factor that affects LiDAR performance, decreases as the water level rises. The tilt angle of the sensor is relatively stable since it is fixed to a solid structure such as a pole. Therefore, the factors that hinder the performance of the LiDAR sensor are less prevalent in actual flooding scenarios.

Also, as previously mentioned, LiDAR sensors demonstrate robustness in various challenging conditions that commonly occur during flash floods, where ultrasonic sensors tend to fail. These conditions include heavy rainfall, dust, and fog, besides varying temperature and humidity (Toa and Whitehead, 2020; Panagopoulos et al., 2021; Bae and Ji, 2019). Through our approach, we have managed to mitigate the drawbacks of LiDAR sensors associated with water turbidity and tilt angle, making this technology highly valuable in real flooding scenarios.

Since we have collected data in a controlled environment, we trained the machine-learning models using data from an ideal condition in which the only failures come from the sensors' limitations. Hence, apart from the zero-valued readings of the IR_h variable that happened in the most challenging scenarios for the LiDAR sensor, there were no unexpected behaviours in the variables measured. However, in a real-world implementation, environmental conditions may damage

the sensors, making it crucial to implement monitoring methods for detecting failed sensors.

Our machine-learning approach requires that both LiDAR and IMU are functional so that their data needs to be introduced to the model. If the LiDAR data is missing, such as what happens with the zero-valued readings discussed in Section 4.1, then no LiDAR-based predictions can be provided, with or without the machine-learning model. If the IMU fails, on the other hand, we can still use the individual readings from the LiDAR sensor to provide a less accurate prediction.

In any case, we recommend employing redundancy to the system, with at least one additional distance sensor. This additional sensor can be, for example, an ultrasonic sensor, a pressure transducer, or a computer vision-based technique. The most cost-effective approach for including such redundancies is implementing a collection of sensors in each measuring station since implementing more stations incurs additional financial, operational and environmental costs.

As presented in Section 5, our results are the median values of the predictions over 300 data points, corresponding to a recording session. Each recording session lasted 200 s (i.e., 3 min and 20 s) as data was collected at a rate of 1.5 Hz. Therefore, each prediction was generated based on data from the past 200 s.

In real-world conditions, the water level changes in a matter of minutes, even in the most rapidly changing environments. Therefore, it is feasible to use a set of data points taken from the past few minutes to account for measurement errors due to the dynamic nature of a water stream.

Nevertheless, these environments provide time-related patterns that can be used to generate predictions with more information. More sophisticated models could also be deployed to make predictions that take into account the variations that happened during a previous period of time. Provided sufficient historical data is available, the sensor data could be used even to anticipate the water level in the upcoming moments, as done in other works in the literature (Wu et al., 2020; Piadeh et al., 2022).

Regarding the computational cost, because of the low dimensionality of the input data, the models built in this work run inferences in less than a second, even for low-end devices such as embedded computers. Our approach's most relevant performance concern is the communication infrastructure, which can vary according to the technologies employed. Wireless communication systems can face challenges such as latency, bandwidth limitations, connectivity issues, packet loss, and throughput restrictions, impacting wireless networks' overall performance and reliability (Ahmed et al., 2023).

For real-world deployment, an alternative machine learning approach would be considering the temporal relationships between the data points, which could be treated as time series and processed through deep learning models. However, since we had only data collected from a controlled environment, the environment as a whole was stationary, with few temporal patterns to be uncovered. Besides, if we turned each recording session into a time series, we would end up with only 90 samples to train the regression model since our experimental design consisted of 90 angle, distance, and turbidity permutations. Such a dataset would be prohibitively small to train a deep architecture such as a convolutional or recurrent neural network, especially considering we had no related data for transfer learning or semi/self-supervised representation learning.

A possible solution to address this problem would be to segment the time series into smaller segments through a sliding window approach. However, to effectively increase the number of data samples, we would need to either select tiny segments, incurring in low-dimensional samples, or provide segments with high overlap, resulting in low variability between segments.

For these reasons, we employed traditional machine-learning techniques to point-wise samples and used the resulting predictions over the recording session to infer the water level based on a central tendency measure such as the median. Ensemble tree models have been extensively used in tabular datasets, with even superior results compared to deep learning (Shwartz-Ziv and Armon, 2022) and improved explainability of the results (Xu et al., 2019). Therefore, we emphasised two of these models in our experiments: AdaBoost and Random Forest. We included linear regression and a KNN regressor as baseline models to highlight the suitability of the proposed approach.

A natural direction for future work is replicating this experiment using data from an urban creek. In this condition, rapidly changing environmental factors generate rich temporal features; hence, deep learning models might lead to improved results. Nonetheless, the scope of this work was analysing the impact of a specific set of factors (i.e., angle, distance and turbidity) in the sensors' predictions and the suitability of merging IMU data to the variables returned from these sensors using machine learning techniques. Such an analysis could only be performed in a controlled setting.

To assess the feasibility of employing the proposed approach in a deployed system, we consider the time complexities of the machine-learning algorithms (Kearns, 1990; James et al., 2013). For providing inference over a single input sample, the LR model consists of a linear combination between the m input variables and a weight vector; hence, its computational complexity is $O(m)$. The KNN model requires computing the distance between the input sample and each of the N samples in the training set, running in $O(m \cdot n)$ complexity (in our case, $N = 31{,}200$). Regarding the tree-based models (Chikalov, 2011), the RF model provides a worst-case time complexity of $O(k \cdot m)$, where k is the number of trees (in our case, 100). The AdaBoost model also runs in $O(k \cdot m)$, although we used 150 weak learners in this case.

Since the feature space consists of m scalar variables (i.e., $m = 8$ for the LiDAR and $m = 7$ for the ultrasonar), the input space has low dimensionality. Therefore, considering their time complexities at inference time, all machine-learning models built in this study can be performed even in embedded computers.

The memory footprint of the trained models might also not be a concern. Linear regression is defined by its coefficients, a weight vector the same size as the input samples. KNN has no weights to be stored since it consists of an instances-based technique. The tree-based ensembles could potentially grow larger if the input space had higher dimensionality since more variables mean that the decision trees can achieve increased depth. However, this was not the case for our application, in which we had no more than eight input variables.

Related work that relies on LiDAR readings for water level measurement either addresses the individual sensor performance under certain environmental conditions (Paul et al., 2020) or evaluates inclined sensor readings with previously known angles and adjusted with standard trigonometric relationships (Tamari et al., 2016).

Our experiments contributed to the state-of-the-art by providing a systematic analysis of the influence of three factors on the performance of a water level measurement system based on a LiDAR sensor. The machine-learning techniques combined the sensor's readings and the IMU data to assess whether combining these modalities could lead to results comparable to an ultrasound sensor, a technology more commonly employed for measuring the height of water bodies (Panagopoulos et al., 2021).

Although distance sensors and inertial units have been combined in other works in the literature (Liu et al., 2023), these works were from different domains, and their results are not comparable to ours. Since this was the first attempt to combine LiDAR sensors and IMU data through machine learning for water level measurement and evaluate the corresponding results in a controlled dataset, no results in the literature compare directly to ours.

Moreover, our contribution resides in the systematic evaluations and data analyses performed. Although the machine-learning models effectively improved the errors in the measurements, they were based on traditional machine-learning algorithms. They were employed as a means to evaluate our hypotheses, not as an innovation by themselves.

A system that aims at monitoring a large area requires deploying a set of sensors that adequately covers the area of interest. Thus, physical or data-driven flood models (Piadeh et al., 2022) can be developed to predict flooding events before they occur. In any case, each sensor's location is paramount and might be cautiously designed.

A communication protocol is essential for transmitting sensor data for applications in large areas that employ multiple sensors across systems. The experimental device presented in this work is a fixed multi-sensor system without network communication. However, it can be integrated with protocols such as Long Range (LoRa) or Message Queuing Telemetry Transport (MQTT) for lightweight messaging, enabling data relay through a centralised and shared server. Mobile network services such as 4G/5G can also be used. These technologies enable the monitoring of flood detection across diverse locations (Shinji and Shiomoto, 2022; Khuen and Zourmand, 2020).

It is worth noticing that, as a requirement, the proposed solution might be deployed at fixed locations at poles next to the water stream. Therefore, their locations are usually known during deployment and may change only through extensive maintenance. Nonetheless, the positions can be reassessed depending on the communication technology employed. Mobile networks such as 4G/5G can provide a reasonable estimate of the device's position, while LoRa-based localisation has also been well documented (Furquim et al., 2016).

Considering the possibilities of sensor failure and the limitations of each approach in a deployed system, using a variety of sensors to analyse different level measurement methods may be beneficial. Contact methods, such as using a pressure transducer and thermal differential, and non-contact methods, such as a level measure based on machine vision and radar, may be integrated into a Wireless Sensor Network (WSN) to assess advantages and disadvantages. Additionally, the level measurements obtained by these sensors can be used as input parameters of a machine learning algorithm and improve its performance.

Water level identification, as part of flood risk management systems, is intended to work continuously, simultaneously on several points, and integrated with analytical tools to improve the decision process. We created and tested a system designed to accurately and reliably detect water levels in urban streams. We assessed LiDAR, IMU, and machine learning technologies in a controlled environment. This system shows potential to be implemented in real-world applications, but further development is required to address any challenges that may arise.

To integrate the proposed system into an existing flood monitoring system, it needs to include a communication protocol to send information to a remote control system. This could involve using existing Internet of Things (IoT) communication protocols such as MQTT and LoRaWAN via a client–server architecture (Kassab and Darabkh, 2020). From a scalability perspective, multiple water level devices can be integrated into the system, improving reliability by adding redundant sensors and expanding monitoring coverage with new measuring stations. Updated and accurate water level information allows the implementation of decision support tools such as risk prediction models and alarm systems (Wu et al., 2020; Piadeh et al., 2022).

Some of the advantages of LiDAR sensors and their association with IMU are their low financial and operational costs. When setting up a station, the primary expense is the necessary supportive structures to place the devices above the stream. The other costs are relative to powering the system, which can be done using urban infrastructure or solar power panels and ensuring stable communication between the system and the monitoring devices. All these structures can be shared with other purposes, such as meteorology or environmental monitoring, to dilute costs. To integrate the system, low-cost platforms such as ESP32 (Babiuch et al., 2019) can be used.

Environmental impacts of the stations in urban areas are expected to be very low. The sensors are not immersed in the water, and the energy consumption is derisory. On the other hand, the proposed system supporting or improving flood management strategies aimed to reduce the substantial financial, environmental and social cost of floods: there are estimations of 600 billion USD lost, 2.8 billion people suffering health impacts and 300,000 people suffering physical injuries worldwide from 2001 to 2018 from water-related disasters, which floods are the more common events (Lee et al., 2020).

8. Conclusion

Considering the necessity to accurately estimate the level of water streams to tackle flood risk in urban environments, this research paper assessed the feasibility of identifying the height of a water body using a LiDAR sensor combined with an IMU through traditional machine-learning techniques. We thoroughly evaluated the advantages and limitations of using a LiDAR sensor for water level measurements in a controlled environment, juxtaposing it with an ultrasonic sensor. We engineered an experimental device, incorporating these sensors with an IMU, and gathered a dataset where we regulated factors such as the device's tilt angle, water turbidity, and the distance between the device and the water surface. Each sensor's performance was analysed individually, followed by an assessment of machine-learning models designed to integrate the data from these sensors with the IMU variables.

The variables registered by the LiDAR or ultrasonic sensors were concatenated to the variables of the IMU, preprocessed and fed to the machine-learning model. We evaluated four types of regressors: linear regression, K Nearest Neighbours, Random Forest and AdaBoost. The tree ensembles (i.e., Random Forest and AdaBoost) led to the lowest error rates, effectively mitigating errors originating from individual sensors. The performance gap between the ultrasonic sensor and LiDAR has narrowed, with statistically significant differences observed only for low water turbidity and increased tilt angle.

The results in these conditions were also the most affected by the only situation in which insufficient information was available for our approach to be helpful: when the LiDAR sensor failed outright, yielding only zero-valued readings. The most straightforward procedure to circumvent this issue is applying higher-end LiDAR sensors with increased range. Other solutions address the factors that cause the zero-valued readings: increased tilt angle, increased distance and low water turbidity. Utilising more rigid structures can control angle fluctuations. Regarding failures due to an increased distance between the sensor and water and low water turbidity, historical data and forecasting models can aid in identifying missing values during changing environments such as heavy rainfall. These factors are not present in controlled experiments such as ours but are relevant in real-world scenarios.

Moreover, these unfavourable circumstances are not likely in real-world flood scenarios, in which water turbidity typically rises, and while the tilt angle may vary over time, it can be effectively regulated through ongoing maintenance. Although the ultrasonic sensor is naturally robust to these factors, it is still vulnerable to weather conditions that are likely to happen in actual flooding scenarios (i.e., varying temperature and humidity, fog, rain or dust). For this reason, provided the LiDAR sensor has sufficient capacity and is self-calibrated using our approach, it may be an important additional modality for building a resilient system.

The experiments performed in this study were restricted to the data collected in a laboratory environment. Although such an assessment helps evaluate the influence of controlled variables in the measurements and the effectiveness of the proposed solution in mitigating the error rates, deploying a similar approach in a real-world environment would not necessarily yield the same results. Environmental variables can only be assessed in field experiments, which cannot be performed in a laboratory. Besides, the dynamics of actual water streams can provide additional information that can be used to improve the prediction models, especially considering historical data.

In other words, the models presented in this work are optimised to controlled conditions that do not always resemble those found in nature. They are unable to identify the dynamics of actual water

streams, which could be helpful inclusively to improve the error rates and make the system less vulnerable to factors such as the zero-valued reading that occurs in extreme situations for the LiDAR range. Also, the combined approach with distance sensor and IMU through machine learning can only work if both the distance sensor and the IMU are working properly, which suggests that redundancies with different types of sensors are desirable for critical situations such as flood-prone regions being monitored.

Future endeavours may involve deploying a refined version of the experimental device to actual urban water streams supplemented with cameras and gauges to evaluate the predictions' precision. Gathering such data from multiple locations during a rainy season could yield a valuable dataset that enables the replication of the experiments conducted in this study but in a more complex setting. This approach will permit us to scrutinise the impact of adverse weather conditions on the performance of each sensor and explore how machine learning models can adeptly tackle these challenges.

CRediT authorship contribution statement

Caetano M. Ranieri: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Angelo V.K. Foletto:** Conceptualization, Software, Investigation, Data curation. **Rodrigo D. Garcia:** Validation, Writing – original draft, Visualization. **Saulo N. Matos:** Validation, Writing – original draft, Visualization. **Maria M.G. Medina:** Conceptualization, Investigation, Resources, Writing – review & editing, Supervision. **Leandro S. Marcolino:** Conceptualization, Validation, Writing – review & editing. **Jó Ueyama:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Caetano Mazzoni Ranieri reports financial support was provided by State of Sao Paulo Research Foundation.

Angelo Victor Kraemer Foletto reports financial support was provided by Coordination of Higher Education Personnel Improvement. Jo Ueyama reports financial support was provided by National Council for Scientific and Technological Development.

Data availability

Data will be made available on request.

Declaration of Generative AI and AI-assisted technologies in the writing process

Statement: During the preparation of this work, the authors used Grammarly and ChatGPT in order to improve readability and language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgements

This research was funded by the São Paulo Research Foundation (FAPESP), grants 2021/10921-2 and 2022/09644-7. It was carried out using the computational resources of the Centre for Mathematical Sciences Applied to Industry (CeMEAI) funded by FAPESP, grant 2013/07375-0. We also acknowledge the Coordination for the Improvement of Higher Education Personnel (CAPES) and the National Council for Scientific and Technological Development (CNPq).

References

Abu-Salih, B., Wongthongtham, P., Coutinho, K., Qaddoura, R., Alshaweesh, O., Wedyan, M., 2023. The development of a road network flood risk detection model using optimised ensemble learning. Eng. Appl. Artif. Intell. 122, 106081. http://dx.doi.org/10.1016/j.engappai.2023.106081.

Ahmed, R., Mahmood, M., Matin, M.A., 2023. Challenges in meeting qos requirements toward 6 g wireless networks: A state of the art survey. Trans. Emerg. Telecommun. Technol. 34 (2), e4693.

Allouis, T., Bailly, J.-S., Feurer, D., 2007. Assessing water surface effects on LiDAR bathymetry measurements in very shallow rivers: a theoretical study. In: Second ESA Space for Hydrology Workshop. CHE, Geneva, pp. 12–14.

Babiuch, M., Foltỳnek, P., Smutnỳ, P., 2019. Using the esp32 microcontroller for data processing. In: 2019 20th International Carpathian Control Conference (ICCC). IEEE, pp. 1–6.

Bae, I., Ji, U., 2019. Outlier detection and smoothing process for water level data measured by ultrasonic sensor in stream flows. Water 11 (5), 951. http://dx.doi.org/10.3390/w11050951.

Brito, L.A., Meneguette, R.I., De Grande, R.E., Ranieri, C.M., Ueyama, J., 2023. Floras: urban flash-flood prediction using a multivariate model. Appl. Intell. 53 (12), 16107–16125. http://dx.doi.org/10.1007/s10489-022-04319-0.

Buhion, J., Buhion, J., Ocon, J., 2022. The development of a real-time, interactive water level monitoring system through SMS with AC load control. Tech. Rom. J. Appl. Sci. Technol. http://dx.doi.org/10.47577/technium.v4i5.6698.

Burrascano, P., Callegari, S., Montisci, A., Ricci, M., Versaci, M., 2014. Ultrasonic Nondestructive Evaluation Systems: Industrial Application Issues. Springer.

Chen, C., Fu, R., Ai, X., Huang, C., Cong, L., Li, X., Jiang, J., Pei, Q., 2022. An integrated method for river water level recognition from surveillance images using convolution neural networks. Remote Sens. 14 (23), 6023. http://dx.doi.org/10.3390/rs14236023.

Chikalov, I., 2011. Average Time Complexity of Decision Trees. Springer.

Choubey, S., William, P., Pawar, A., Jawale, M., Gupta, K., Parganiha, V., 2022. Intelligent water level monitoring and water pump controlling system using IoT. In: 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC). IEEE, pp. 423–427. http://dx.doi.org/10.1109/ICESC54411.2022.9885358.

Djalilov, A., Sobirov, E., Nazarov, O., Urolov, S., Gayipov, I., 2023. Study on automatic water level detection process using ultrasonic sensor. In: IOP Conference Series: Earth and Environmental Science, Vol. 1142. IOP Publishing, 012020. http://dx.doi.org/10.1088/1755-1315/1142/1/012020.

Fernandes Junior, F.E., Nonato, L.G., Ranieri, C.M., Ueyama, J., 2021. Memory-based pruning of deep neural networks for IoT devices applied to flood detection. Sensors 21 (22), 7506. http://dx.doi.org/10.3390/s21227506.

Furquim, G., Pessin, G., Faiçal, B.S., Mendiondo, E.M., Ueyama, J., 2016. Improving the accuracy of a flood forecasting model by means of machine learning and chaos theory: a case study involving a real wireless sensor network deployment in Brazil. Neural Comput. Appl. 27, 1129–1141. http://dx.doi.org/10.1007/s00521-015-1930-z.

Glantz, S.A., Slinker, B.K., Neilands, T.B., 2001. Primer of Applied Regression & Analysis of Variance, Vol. 654. McGraw-Hill, Inc., New York.

Gondkar, S.S., Pardeshi, D., William, P., 2022. Innovative system for water level management using IoT to prevent water wastage. In: 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC). IEEE, pp. 1555–1558. http://dx.doi.org/10.1109/ICAAIC53929.2022.9792746.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

Hanan, A., Sumadiyasa, M., 2019. Water level detection system based on ultrasonic sensors HC-SR04 and ESP8266-12 modules with telegram and buzzer communication media. Instrum. Mes. Métrol. 18 (3), 305–309. http://dx.doi.org/10.18280/i2m.180311.

Hanni, J.R., Venkata, S.K., 2020. Does the existing liquid level measurement system cater the requirement of future generation? Measurement 156, 107594. http://dx.doi.org/10.1016/j.measurement.2020.107594.

Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Vol. 2. Springer.

Ilhan, A., 2023. Forecasting of volumetric flow rate of ergene river using machine learning. Eng. Appl. Artif. Intell. 121, 105983. http://dx.doi.org/10.1016/j.engappai.2023.105983.

James, G., Witten, D., Hastie, T., Tibshirani, R., et al., 2013. An Introduction to Statistical Learning, Vol. 112. Springer.

Jan, F., Min-Allah, N., Saeed, S., Iqbal, S.Z., Ahmed, R., 2022. Iot-based solutions to monitor water level, leakage, and motor control for smart water tanks. Water 14 (3), 309. http://dx.doi.org/10.3390/w14030309.

Jelalian, A.V., 1992. Laser Radar Systems. Artech House Publishers.

Jiang, B., Chen, S., Wang, B., Luo, B., 2022. Mglnn: Semi-supervised learning via multiple graph cooperative learning neural networks. Neural Netw. 153, 204–214. http://dx.doi.org/10.1016/j.neunet.2022.05.024.

Kaasalainen, S., Malkamäki, T., 2020. Potential of active multispectral lidar for detecting low reflectance targets. Opt. Express 28 (2), 1408–1416. http://dx.doi.org/10.1364/OE.379491.

Kamoji, S., Kalla, M., 2023. Effective flood prediction model based on Twitter text and image analysis using BMLP and SDAE-HHNN. Eng. Appl. Artif. Intell. 123, 106365. http://dx.doi.org/10.1016/j.engappai.2023.106365.

Kashani, A.G., Olsen, M.J., Parrish, C.E., Wilson, N., 2015. A review of LiDAR radiometric processing: from ad hoc intensity correction to rigorous radiometric calibration. Sensors 15 (11), 28099–28128. http://dx.doi.org/10.3390/s151128099.

Kassab, W., Darabkh, K.A., 2020. A–z survey of internet of things: Architectures, protocols, applications, recent advances, future directions and recommendations. J. Netw. Comput. Appl. 163, 102663. http://dx.doi.org/10.1016/j.jnca.2020.102663.

Kearns, M.J., 1990. The Computational Complexity of Machine Learning. MIT Press.

Khuen, C.K., Zourmand, A., 2020. Fuzzy logic-based flood detection system using lora technology. In: 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA). pp. 40–45. http://dx.doi.org/10.1109/CSPA48992.2020.9068698.

Kukko, A., Kaasalainen, S., Litkey, P., 2008. Effect of incidence angle on laser scanner intensity and surface data. Appl. Opt. 47 (7), 986–992. http://dx.doi.org/10.1364/AO.47.000986.

Lakshmanan, V., Robinson, S., Munn, M., 2020. Machine Learning Design Patterns. O'Reilly Media.

Lee, J., Perera, D., Glickman, T., Taing, L., 2020. Water-related disasters and their health impacts: A global review. Prog. Disaster Sci. 8, 100123.

Lipták, B.G., Venczel, K., 2016. Measurement and Safety: Volume I. CRC Press.

Liu, S., Chen, J., Wang, C., Lin, L., 2023. Ultrasonic positioning and imu data fusion for pen-based 3d hand gesture recognition. Multimedia Tools Appl. 1–19.

Lo, S.-W., Wu, J.-H., Lin, F.-P., Hsu, C.-H., 2015. Cyber surveillance for flood disasters. Sensors 15 (2), http://dx.doi.org/10.3390/s150202369.

Loizou, K., Koutroulis, E., 2016. Water level sensing: State of the art review and performance evaluation of a low-cost measurement system. Measurement 89, 204–214. http://dx.doi.org/10.1016/j.measurement.2016.04.019.

Lu, Y., Chen, J., Xu, Q., Han, Z., Peart, M., Ng, C.-N., Lee, F.Y., Hau, B.C., Law, W.W., 2023. Spatiotemporal variations of river water turbidity in responding to rainstorm-streamflow processes and farming activities in a mountainous catchment, Lai Chi Wo, Hong Kong, China. Sci. Total Environ. 863, 160759. http://dx.doi.org/10.1016/j.scitotenv.2022.160759.

Mamat, N.H., Othman, M.H., Othman, W.Z., Noor, M.F.M., 2021. Internet of things in flood warning system: An overview on the hardware implementation. In: Proceedings of the 1st International Conference on Electronics, Biomedical Engineering, and Health Informatics. Springer Singapore, pp. 269–279. http://dx.doi.org/10.1007/978-981-33-6926-9_23.

Martins, K., Bonneton, P., Frappart, F., Detandt, G., Bonneton, N., Blenkinsopp, C.E., 2017. High frequency field measurements of an undular bore using a 2D LiDAR Scanner. Remote Sens. 9 (5), 462. http://dx.doi.org/10.3390/rs9050462.

McManamon, P.F., 2019. Lidar technologies and systems.

Milligan, S., Vandelinde, H., Cavanagh, M., 2013. Understanding Ultrasonic Level Measurement. Momentum Press.

Panagopoulos, Y., Papadopoulos, A., Poulis, G., Nikiforakis, E., Dimitriou, E., 2021. Assessment of an ultrasonic water stage monitoring sensor operating in an urban stream. Sensors 21 (14), 4689. http://dx.doi.org/10.3390/s21144689.

Pasika, S., Gandla, S.T., 2020. Smart water quality monitoring system with cost-effective using IoT. Heliyon 6 (7), e04096. http://dx.doi.org/10.1016/j.heliyon.2020.e04096.

Paul, J.D., Buytaert, W., Sah, N., 2020. A technical evaluation of lidar-based measurement of river water levels. Water Resour. Res. 56 (4), http://dx.doi.org/10.1029/2019wr026810.

Piadeh, F., Behzadian, K., Alani, A.M., 2022. A critical review of real-time modelling of flood forecasting in urban drainage systems. J. Hydrol. 607, 127476.

Prafanto, A., Budiman, E., 2018. A water level detection: Iot platform based on wireless sensor network. In: 2018 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT). IEEE, pp. 46–49. http://dx.doi.org/10.1109/eiconcit.2018.8878559.

Qiao, G., Yang, M., Wang, H., 2022. A water level measurement approach based on yolov5s. Sensors 22 (10), http://dx.doi.org/10.3390/s22103714, URL https://www.mdpi.com/1424-8220/22/10/3714.

Qiu, Z., Lu, Y., Qiu, Z., 2022. Review of ultrasonic ranging methods and their current challenges. Micromachines 13 (4), 520.

Ramos, L., Rodrigues, F., Reis, C., Bozano, D., Reis, D., Goncalves, A., 2020. An experiment to observe Stevin's law with an arduino. Phys. Educ. 55 (3), 033004. http://dx.doi.org/10.1088/1361-6552/ab7085.

Reymann, C., Lacroix, S., 2015. Improving LiDAR point cloud classification using intensities and multiple echoes. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 5122–5128. http://dx.doi.org/10.1109/IROS.2015.7354098.

Rocchi, A., Santecchia, E., Ciciulla, F., Mengucci, P., Barucca, G., 2019. Characterization and optimization of level measurement by an ultrasonic sensor system. IEEE Sens. J. 19 (8), 3077–3084. http://dx.doi.org/10.1109/JSEN.2018.2890568.

Roy, A.M., Bose, R., Bhaduri, J., 2022. A fast accurate fine-grain object detection model based on YOLOv4 deep neural network. Neural Comput. Appl. 34 (5), 3895–3921. http://dx.doi.org/10.1007/s00521-021-06651-x.

Sahoo, A.K., Udgata, S.K., 2020. A novel ANN-based adaptive ultrasonic measurement system for accurate water level monitoring. IEEE Trans. Instrum. Meas. 69 (6), 3359–3369. http://dx.doi.org/10.1109/tim.2019.2939932.

Sasikala, G., Srinivasan, S., Navarajan, J., Theresa, M.M., 2022. Iot based water level monitoring and management in reservoir. In: 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC). IEEE, pp. 1763–1767. http://dx.doi.org/10.1109/ICESC54411.2022.9885714.

Schenato, L., Aguilar-Lopez, J.P., Galtarossa, A., Pasuto, A., Bogaard, T., Palmieri, L., 2021. A rugged FBG-based pressure sensor for water level monitoring in dikes. IEEE Sens. J. 21 (12), 13263–13271. http://dx.doi.org/10.1109/JSEN.2021.3067516.

Scornet, E., 2023. Trees, forests, and impurity-based variable importance in regression. In: Annales de l'Institut Henri Poincare (B) Probabilites et statistiques, Vol. 59. Institut Henri Poincaré, pp. 21–52. http://dx.doi.org/10.1214/21-AIHP1240.

Shen, Y., Giurgiutiu, V., 2014. Wfr-2d: an analytical model for pwas-generated 2d ultrasonic guided wave propagation. In: Health Monitoring of Structural and Biological Systems 2014, Vol. 9064. SPIE, pp. 307–321.

Shinji, Y., Shiomoto, K., 2022. Qos in iot-based river flood monitoring system using mqtt brokers in tandem connection. In: 2022 IEEE 8th World Forum on Internet of Things (WF-IoT). pp. 1–2. http://dx.doi.org/10.1109/WF-IoT54382.2022.10152261.

Shwartz-Ziv, R., Armon, A., 2022. Tabular data: Deep learning is not all you need. Inf. Fusion 81, 84–90.

Song, J.-H., Han, S.-H., Yu, K., Kim, Y.-I., 2002. Assessing the possibility of land-cover classification using LiDAR intensity data. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. 34 (3/B), 259–262.

Tamari, S., Guerrero-Meza, V., 2016. Flash flood monitoring with an inclined LiDAR installed at a river bank: proof of concept. Remote Sens. 8 (10), 834. http://dx.doi.org/10.3390/rs8100834.

Tamari, S., Guerrero-Meza, V., Rifad, Y., Bravo-Inclán, L., Sánchez-Chávez, J.J., 2016. Stage monitoring in turbid reservoirs with an inclined terrestrial near-infrared lidar. Remote Sens. 8 (12), 999. http://dx.doi.org/10.3390/rs8120999.

Toa, M., Whitehead, A., 2020. Ultrasonic Sensing Basics. Texas Instruments, Dallas.

Varun, K.S., Kumar, K.A., Chowdary, V.R., Raju, C., 2018. Water level management using ultrasonic sensor (automation). Int. J. Comput. Sci. Eng. 6 (6), 799–804.

Wannoi, N., Wannoi, C., 2022. A real-time prototype of a water level monitor and wide area early flood warning. Creative Sci. 14 (1), 19–24. http://dx.doi.org/10.55674/snrujst.v14i1.244485.

World Meteorological Organization (WMO), 2010. Manual on stream gauging volume I—Fieldwork WMO-No. 1044.

Wu, W., Emerton, R., Duan, Q., Wood, A.W., Wetterhall, F., Robertson, D.E., 2020. Ensemble flood forecasting: Current status and future opportunities. Wiley Interdiscip. Rev.: Water 7 (3), e1432.

Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., Zhu, J., 2019. Explainable ai: A brief survey on history, research areas, approaches and challenges. In: Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II 8. Springer, pp. 563–574.

Yan, S., Yang, G., Li, Q., Wang, C., 2020. Waveform centroid discrimination of pulsed LiDAR by combining EMD and intensity weighted method under low SNR conditions. Infrared Phys. Technol. 109, 103385. http://dx.doi.org/10.1016/j.infrared.2020.103385.

Yuliza, E., Salam, R., Amri, I., Atmajati, E., Hapidin, D., Meilano, I., Munir, M., Abdullah, M., et al., 2016. Characterization of a water level measurement system developed using a commercial submersible pressure transducer. In: 2016 International Conference on Instrumentation, Control and Automation (ICA). IEEE, pp. 99–102. http://dx.doi.org/10.1109/ICA.2016.7811483.

Yunita, A.M., Wardah, N.N., Sugiarto, A., Susanti, E., Sujai, L., Rizky, R., 2020. Water level measurements at the cikupa pandeglang bantendam using fuzzy sugenowith microcontroler-based ultrasonik sensor. In: Journal of Physics: Conference Series, Vol. 1477. IOP Publishing, 052048. http://dx.doi.org/10.1088/1742-6596/1477/5/052048.