The Automated Paper Strip Reader (APSR): A Reliable Low-Cost Solution for Colorimetric Detection

Tejal Shinde, Prachi Bhoir, Jagannath Nirmal, and Ninad Mehendale

***Abstract*—Manual interpretation of colorimetric test strips is prone to substantial variability due to human error and sensitivity to surrounding illumination conditions, resulting in limited diagnostic reliability at the point-of-care (POC) (only 60.7% accuracy in our evaluation). To overcome these issues, we designed an automated paper strip reader (APSR) built on a Raspberry Pi 4B platform. The device integrates a closed chamber with uniform LED illumination, an USB camera, and Python-based image processing algorithms to quantify 15 water quality parameters, including pH, nitrate, chloride, and selected heavy metals. Images are captured under controlled lighting and analyzed through parameter-specific calibration curves, with results displayed via a touchscreen interface. Validation using diverse water samples showed that APSR achieved an overall accuracy of 88.7%, outperforming human interpretation (60.7%). In particular, the APSR provided perfect detection accuracy (100%) for iron and copper, compared to only 63% and 40% by human readers, respectively. Each test requires approximately 5 s, with the device costing less than U.S. 150 to assemble, and offering one-click data export. This cost-effective automated system provides a robust solution for reliable POC testing in low-resource environments, aligning with the World Health Organization (WHO) ASSURED criteria for diagnostic devices.**

***Index Terms*—Digital diagnostics, image processing, microcontroller, resource-limited settings.**

I. Introduction

IN CONTEMPORARY healthcare, the examination of body fluids is central to both diagnosis and therapeutic decision-making. During initial patient assessment, tests of fluids such as blood and urine are frequently performed, as the accuracy and timeliness of these results directly influence subsequent medical interventions. Paper strip-based readers have gained prominence in this setting, offering a simple yet effective means of quantifying chemical constituents in various fluids. Such devices are capable of measuring indicators, including pH, glucose, protein, and nitrite levels in biological samples like blood, urine, sweat, and saliva. Beyond clinical applications, similar strip-based technologies are also employed in environmental monitoring, for instance, in the evaluation of water quality by measuring hardness, chlorine, bromine, and nitrite concentrations.

The COVID-19 pandemic underscored the urgent requirement for rapid, accessible, and reliable diagnostic solutions. With healthcare systems across the globe facing unprecedented patient loads, the need for technologies that reduce pressure on hospitals while maintaining diagnostic accuracy became evident. Even outside pandemic conditions, barriers such as cost, limited time, and logistical inconvenience often discourage individuals from undergoing routine check-ups, which may delay detection and treatment of underlying health conditions. This challenge is more severe in low- and middle-income countries (LMICs), where large populations lack access to primary healthcare services and insurance coverage. Consequently, there is a growing demand for diagnostic tools that are accurate, affordable, and suitable for point-of-care (POC) use. While conventional paper test strips remain inexpensive and widely available, they are subject to limitations, including dependence on ambient light and fading of reagents, which can reduce reliability and lead to errors during manual interpretation. Automating the diagnostic process mitigates these issues, ensuring more consistent and objective outcomes by reducing human error. To guide the development of such technologies, the World Health Organization (WHO) introduced the ASSURED framework, which highlights the importance of devices being affordable, sensitive, specific, user-friendly, rapid and robust, equipment-free, and deliverable.

Recent advances in fabrication technologies, particularly 3-D printing, provide promising opportunities for the development of next-generation POC diagnostic devices. Additive manufacturing enables rapid prototyping and the production of lightweight, low-cost, and durable systems. These features are particularly relevant in rural and resource-constrained settings, where access to healthcare infrastructure is limited. The present research seeks to address the limitations of current diagnostic methods by designing an automated paper strip reader (APSR) that adheres to the WHO ASSURED guidelines. The proposed device emphasizes affordability, accuracy, and ease of use, making it suitable for deployment in both household and clinical contexts. By integrating advanced image processing algorithms with modular 3-D-printed hardware, the system aims to deliver rapid, reliable, and user-friendly diagnostics without the need for specialized expertise.

Received 21 November 2025; accepted 22 December 2025. Date of publication 24 March 2026; date of current version 11 May 2026. *(Corresponding author: Ninad Mehendale.)*

Tejal Shinde is with the Department of Microelectronics, TU Delft, 2628 CD Delft, The Netherlands (e-mail: tejal.shinde@hotmail.com).

Prachi Bhoir, Jagannath Nirmal, and Ninad Mehendale are with the Department of Electronics Engineering, K. J. Somaiya School of Engineering, Somaiya Vidyavihar University, Mumbai 400077, India (e-mail: prachi.bhoir@somaiya.edu; jhnirmal@somaiya.edu; ninad@somaiya.edu).

Digital Object Identifier 10.1109/JFLEX.2026.3677153

2768-167X © 2026 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

While several automated strip readers have been reported in recent literature [1], [2], significant gaps remain in their practical deployment for POC diagnostics, particularly in resource-limited settings. Most existing systems rely on smartphone cameras [3], [4], which introduce variability due to inconsistent ambient lighting and device-specific camera specifications, or require external computing infrastructure that limits portability [9], [12]. Furthermore, few studies have rigorously quantified the accuracy deficit of manual interpretation, a critical justification for automation. The present work addresses these limitations through the APSR, which provides three key innovations.

1) *Quantified Clinical Need:* Systematic validation against human readers (N = 10, 1800 measurements) demonstrating that manual interpretation achieves only 60.7% accuracy, with complete failure (0% accuracy) for critical parameters including nitrite, nitrate, and total chloride.
2) *Integrated Autonomous Design:* A self-contained system combining controlled illumination (5500 K, 800 $\pm$ 50 lux), Raspberry Pi-based processing, and battery power in a portable enclosure costing U.S. < $150, eliminating dependencies on smartphones or laboratory infrastructure.
3) *Multiplex Capability With Clinical Validation:* Simultaneous quantification of 15 water quality parameters validated against ISO-certified reference methods (ICP-MS, ion chromatography), achieving 88.7% overall accuracy with 97%–100% detection rates for heavy metals and nutrients. By addressing affordability, portability, and diagnostic reliability within a single integrated platform, APSR fulfills the WHO ASSURED criteria for deployment in low-resource environments where access to laboratory testing is limited.

II. Related Work

In recent years, POC colorimetric testing has attracted growing interest due to its potential to transform diagnostic practices. Multiple strategies have been proposed for analyzing assays and paper-based strips [5], with particular emphasis on developing cost-effective and user-friendly devices for health monitoring [1]. Image processing techniques combined with Raspberry Pi controllers have been widely adopted to achieve reliable measurements. For example, Paglinawan et al. [6] applied this approach for the evaluation of urine parameters, whereas similar methods have been reported for presumptive drug testing [7] and for estimating enzyme kinetic constants [3]. To improve measurement consistency, several systems employ 3-D-printed chambers to mitigate external lighting effects. In parallel, smartphone-assisted solutions have emerged as an alternative trend in colorimetric analysis. Jalal et al. [4] introduced a hybrid paper–plastic microfluidic lab-on-a-chip device that utilizes smartphone cameras for both image capture and processing. Carrio et al. [8] presented a saliva test reader integrating a light-controlled enclosure with a smartphone imaging unit, while Jang et al. [9] demonstrated a portable platform for progesterone detection in milk using ELISA methods.

Efforts to enhance detection sensitivity have also been reported. Heidari-Bafroui et al. [10] designed an infrared lightbox for phosphate quantification, while edge-detection algorithms have been employed to monitor blood coagulation and flow [11]. In addition, optical biosensors have been explored for glucose detection in saliva [12]. Several studies have investigated improved image analysis workflows. Wang et al. [13] proposed an SPSS-based image quantification approach, converting color data into hue values for pH prediction. More recently, deep learning models have been applied to classify assay types [2], illustrating the potential of artificial intelligence in colorimetric diagnostics. Research has also focused on disease-specific detection; for instance, Lin and Chang [14] developed a rapid and sensitive device for fecal occult blood testing, a critical tool for diagnosing gastrointestinal disorders and colorectal cancer. Despite this progress, it is evident that not all approaches fulfill the WHO's ASSURED criteria, which call for diagnostic devices that are affordable, sensitive, specific, user-friendly, rapid and robust, equipment-free, and deliverable.

The existing body of work reflects a wide variety of approaches to strip-reader design, differing in terms of controllers, imaging modules, illumination sources, computational methods, color models, and power options. Diagnostic strips can be classified as either singleplex—providing a binary readout for one analyte—or multiplex—yielding semiquantitative results across multiple analytes. Although visual comparison of test strips against reference color charts remains common, the adoption of automated methods is expanding. These include office scanner-based readers, IoT-enabled systems, and computer vision algorithms that reduce subjectivity and improve reproducibility. Python programming has frequently been employed for specialized tasks, such as blood coagulation assessment through edge detection [11], while android-based applications have been designed to interface with Raspberry Pi readers [10] and lab-on-a-chip devices [4]. Overall, while considerable progress has been achieved in POC colorimetric diagnostics, further development is required to create solutions that comprehensively satisfy the ASSURED framework while maintaining accuracy, speed, and ease of use across diverse clinical and environmental applications.

III. Methods

A. Concept Design

The APSR was conceptualized as a compact and integrated system capable of delivering precise colorimetric analysis while minimizing environmental variability. The overall structure can be described in terms of its hardware components and their interconnections.

As illustrated in Fig. 1, the Raspberry Pi 4B serves as the central controller and processing unit. A memory card preloaded with the Raspbian operating system functions as the onboard storage and database. A 5-in resistive TFT touchscreen is connected via a ribbon cable, acting as the primary user interface.

Image acquisition is performed by a USB webcam, which provides high-resolution images of the paper strips. Power is supplied by a rechargeable battery system, which simultaneously drives the Raspberry Pi and the LED illumination strip.

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

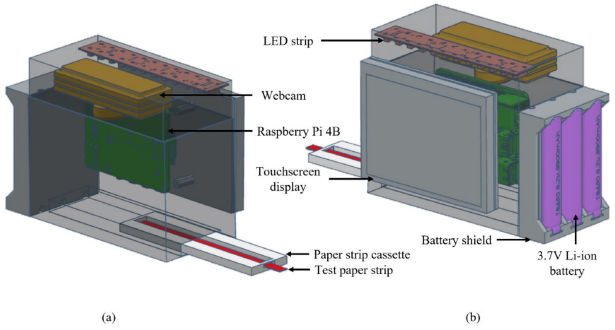


Fig. 1. Conceptual design diagram of the APSR. (a) Front view. (b) Back view.

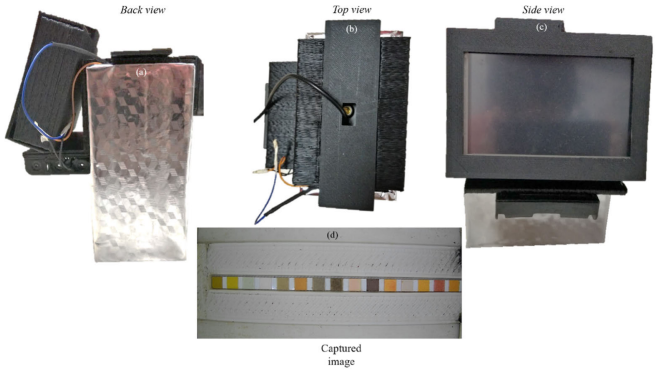


Fig. 2. Final APSR device: (a) back view, (b) top view, (c) side view, and (d) sample image captured by the webcam.

To ensure measurement consistency, the device is housed in a two-part box-like enclosure that isolates the testing environment from external light sources. The upper section holds the webcam and LED strip, while the lower section accommodates the Raspberry Pi board and a custom-designed cassette for holding the strips. This arrangement guarantees controlled illumination and reproducible imaging.

Externally, the APSR incorporates a battery shield supporting three 3.7-V cells, enabling portability and long operation. The touchscreen is mounted on the outer panel for straightforward interaction and result display. By combining hardware and software elements, this design enables a portable, user-friendly, and reliable diagnostic platform. Through digital image processing in a controlled setting, the APSR overcomes the limitations of manual strip interpretation and enhances POC diagnostic reliability.

B. System Architecture

The APSR architecture adopts a three-layered design optimized for real-time colorimetric processing.

1) *Hardware Layer:* Provides physical interaction for image capture and environmental control. It includes the USB webcam for strip imaging, LED illumination with a 5500-K color temperature, and a 3-D-printed enclosure to block ambient light. The strip cassette maintains fixed positioning, while a battery shield ensures portability.
2) *Processing Layer:* Implements the computational workflow on the Raspberry Pi 4B. An eight-step algorithm is executed sequentially: 1) image capture with distortion correction; 2) grayscale conversion and Gaussian filtering; 3) strip localization via Hough transform; 4) pad segmentation using regions of interest (ROIs); 5) HSV-based color analysis; 6) calibration curve application; 7) quality assurance; and 8) output generation. Processing relies on Python libraries, with OpenCV for image operations and NumPy for numerical calculations, achieving analysis within the 5-s target.
3) *User Interface Layer:* Manages user interactions through a Tkinter-based touchscreen GUI. This layer supports real-time visualization, parameter selection, and control of the device. An integrated email export function enables immediate result sharing for clinical use.

Data flow proceeds from acquisition to processing to output display, with internal feedback loops for quality checks. If validation fails, the system automatically reacquires images up to three times before reporting an error. This modular design ensures robustness, simplifies troubleshooting, and allows for future upgrades such as cloud integration or machine learning without requiring major redesigns.

C. Hardware Design

Fig. 2 presents the assembled APSR prototype. The back view (a) displays the connectivity ports and structural layout; the top view (b) highlights the compact form factor; the side view (c) shows the 45° angled screen for ergonomic use; and Fig. 2(d) illustrates a captured image of a test strip, showing consistent illumination across all 15 pads.

The hardware is organized into two detachable modules. The upper part contains the webcam and LED strip, aligned for optimal illumination and focus. The lower module houses the Raspberry Pi board and a custom strip holder, ensuring proper positioning within the camera's field of view. Peripheral access for HDMI, USB, Ethernet, and power is maintained through dedicated openings. The enclosure is assembled using nuts, bolts, and washers, allowing easy maintenance and upgrades. The angled display enhances usability during operation.

This modular and minimalistic design guarantees controlled imaging conditions and user comfort. The clear strip visualization in Fig. 2(d) confirms that the enclosed setup effectively eliminates ambient light, thereby enabling reliable and repeatable measurements.

D. Sample Preparation and Characterization

Four synthetic water samples (W1–W4) were prepared in the laboratory to provide controlled validation across the full dynamic range of all 15 test parameters. Samples were prepared using analytical-grade reagents (≥99% purity, Sigma-Aldrich) dissolved in deionized water (resistivity >18.2 MΩ · cm, Millipore Milli-Q system).

Individual stock solutions (1000 mg/L) were prepared for each analyte and stored at 4 °C. For each sample, appropriate volumes were transferred to 1 L volumetric flasks using calibrated micropipettes (Eppendorf Research Plus) and diluted to volume. The samples were designed with progressively increasing concentrations from low (W1) to high (W4) to

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

TABLE I
COMPOSITION OF SYNTHETIC WATER SAMPLES USED FOR VALIDATION

Parameter	Chemical Used	W1 (Low)	W2 (Low-Med)	W3 (Med-High)	W4 (High)
Iron	$FeSO_4$	0.05 mg/L	0.2 mg/L	0.5 mg/L	1.0 mg/L
Copper	$CuSO_4$	0.1 mg/L	0.5 mg/L	1.0 mg/L	3.0 mg/L
Nitrite	$NaNO_2$	0.01 mg/L	0.05 mg/L	0.1 mg/L	1.0 mg/L
Nitrate	KNO_3	1 mg/L	10 mg/L	30 mg/L	50 mg/L
Total Chloride	NaCl	10 mg/L	25 mg/L	100 mg/L	250 mg/L
Chromium (VI)	$K_2Cr_2O_7$	0.05 mg/L	0.2 mg/L	0.5 mg/L	1.0 mg/L
Total Hardness	$CaCO_3$ + $MgSO_4$	20 mg/L	100 mg/L	200 mg/L	400 mg/L
Ammonium Chloride	NH_4Cl	0.05 mg/L	0.2 mg/L	2.0 mg/L	5.0 mg/L
Bromine	NaBr	0.2 mg/L	1.0 mg/L	2.0 mg/L	5.0 mg/L
Cyanuric Acid	Cyanuric acid	10 mg/L	50 mg/L	100 mg/L	150 mg/L
Fluoride	NaF	0.2 mg/L	1.0 mg/L	2.0 mg/L	3.0 mg/L
pH	Buffer/NaOH/HCl	5.0	6.5	7.5	9.0
Total Alkalinity	Na_2CO_3/$NaHCO_3$	40 mg/L	100 mg/L	150 mg/L	200 mg/L
Lead	$Pb(NO_3)_2$	0.05 mg/L	0.2 mg/L	0.5 mg/L	1.0 mg/L
Residual Chlorine	NaOCl	0.1 mg/L	0.5 mg/L	2.0 mg/L	4.0 mg/L

Note: Concentrations represent nominal target values. Actual concentrations validated by an ISO-accredited laboratory were within ±5% for 13 of 15 parameters.

systematically evaluate performance across the detection range (see Table I).

After preparation, samples were equilibrated at room temperature (23 ± 2 °C) for 2 h with magnetic stirring. pH values were adjusted using 0.1-M HCl or NaOH and verified with a calibrated pH meter (Hanna HI-2020, ±0.01 accuracy). Each sample was prepared in triplicate batches: one for APSR testing, one for human reader evaluation, and one for reference laboratory validation.

Reference validation was performed by an ISO/IEC 17025:2017 accredited laboratory using certified methods: ICP-MS for metals (EPA Method 200.8), ion chromatography for anions (EPA Method 300.0), EDTA titration for hardness (EPA Method 130.2), and standard titration/colorimetric methods for alkalinity and chlorine (Standard Methods 2320B and 4500-Cl). All measurements were performed in triplicate.

Laboratory validation confirmed that actual concentrations were within ±5% of target values for 13 of 15 parameters. Lead in sample W1 measured below the ICP-MS detection limit (<0.05 mg/L versus target 0.05 mg/L), while residual chlorine showed 20%–30% degradation during transport to the external laboratory. These validated reference values (not nominal targets) were used as ground truth for all accuracy calculations.

The use of synthetic samples ensured controlled validation without matrix interference effects, systematic coverage of the full concentration range, and reproducibility. This approach represents an idealized scenario; future work will validate performance using environmental water samples containing natural organic matter and interfering ions.

E. Reference Methods and Experimental Validation

Multiparameter water-quality test strips used in this study were commercially procured from JNW Direct (USA), "JNW Direct 15-in-1 Drinking Water Test Strips" (100 count, includes companion mobile app and eBook). Each strip comprises 15 reagent pads capable of semiquantitative colorimetric estimation of pH, total hardness, total alkalinity (TA), free and total chlorine, nitrite, nitrate, fluoride, iron, copper, lead, chromium (VI), bromine, cyanuric acid, and residual chlorine. Strips were used exactly as supplied, with a 2-s immersion time and 60-s color-development interval, under controlled illumination. No chemical modification or pretreatment was performed.

1) Determination of True Values: Ground-truth values for all analytes were established using certified laboratory protocols. pH values were measured via EPA Method 150.1 using a calibrated Hanna HI-2020 pH meter (±0.01). Heavy metals (Cu, Pb, and Cr) were quantified by ICP-MS (Agilent 7900, EPA Method 200.8). Chloride and nitrate concentrations were obtained by ion chromatography (Dionex ICS-5000, EPA Method 300.0). Total hardness was determined by EDTA titration (EPA Method 130.2), while iron was quantified using atomic absorption spectroscopy (EPA Method 236.1). Bromine and chlorine were analyzed using the DPD colorimetric method (Standard Method 4500-Cl). All measurements were repeated in triplicate in an ISO/IEC 17025:2017 accredited laboratory.

2) Testing Conditions and Camera Settings: Experiments were carried out under controlled conditions: temperature 23 ± 2 °C, relative humidity 45 ± 5%, and ambient illumination <5 lux inside the chamber. LED illumination provided 5500-K color temperature, CRI > 90, and intensity 800 ± 50 lux at the sample plane. Images were acquired with a Logitech C920 webcam at 1920 × 1080 resolution, exposure 1/60 s, ISO 100, and white balance fixed at 5500 K. Test strips were placed at 15-cm distance with ±5° angular tolerance.

3) Algorithm Parameters and Quality Thresholds: The algorithm employed Gaussian blurring with $\sigma = 1.5$, Canny edge detection with thresholds (50, 150), and ROI extraction of 50 × 50 pixels per pad. Images were converted to HSV using OpenCV's `cv2.COLOR_RGB2HSV` and filtered with a 3 × 3 median kernel. Calibration curves were modeled using second-order polynomial regression with $R^2 > 0.95$ as a fit criterion. Quality checks included coefficient of variation (CV) <10% for pad color uniformity and Laplacian variance >100 for focus.

4) Study Design: Ten participants (five trained technicians and five untrained volunteers) manually interpreted four water samples across 15 parameters in triplicate, generating 1800 readings (10 × 4 × 15 × 3). In parallel, the APSR conducted three automated measurements per sample for comparison.

F. Statistical Analysis

Statistical methods were used to assess accuracy, reliability, and agreement between APSR and human interpretation.

1) *Accuracy:*

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{Total Observations}} \times 100$$

with 95% confidence intervals estimated by Wilson's score method.

2) *Interrater Reliability:* Assessed via Cohen's κ for categorical data and intraclass correlation coefficient (ICC) for continuous measurements.

3) *Method Comparison:* APSR and human accuracy were compared using paired t-tests with Bonferroni correction ($\alpha = 0.003$).

4) *Repeatability:* Calculated using CV (CV%):

$$\text{CV\%} = \left(\frac{\text{SD}}{\text{Mean}}\right) \times 100.$$

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

TABLE II
COMPARISON OF APSR AND HUMAN READER PERFORMANCE

Parameter	APSR Accuracy (95% CI)	Human Accuracy (95% CI)	p-value	ICC (Human)	CV% (APSR)
Total Hardness	95% (91.2–97.3)	76% (71.4–80.2)	<0.001	0.72	3.2
Bromine	97% (93.8–98.7)	94% (90.1–96.5)	0.082	0.88	2.1
Iron	100% (97.5–100)	63% (58.2–67.5)	<0.001	0.65	1.8
Copper	100% (97.5–100)	40% (35.4–44.8)	<0.001	0.51	2.5
Lead	79% (74.6–82.9)	100% (97.5–100)	<0.001	0.95	5.3
Nitrite	97% (93.8–98.7)	0% (0–3.7)	<0.001	N/A	2.8
Nitrate	98% (94.9–99.4)	0% (0–3.7)	<0.001	N/A	2.4

Note: ICC <0.5 = poor reliability; 0.5–0.75 = moderate; 0.75–0.9 = good; >0.9 = excellent.

5) *Bland–Altman Analysis:* Agreement was examined with Bland–Altman plots, defining limits of agreement (LoA) as

$$\text{LoA} = \text{mean difference} \pm 1.96 \times \text{SD of differences}.$$

IV. RESULTS

A. Performance Evaluation

Table II summarizes the comparative analysis between the APSR and manual human interpretation across several water quality parameters. Overall, the APSR achieved markedly higher diagnostic accuracy than human readers, with an average accuracy of 88.7% (95% CI: 86.2%–91.0%) compared to 60.7% (95% CI: 57.8%–63.5%; $p < 0.001$). The most notable improvements were observed for iron, copper, nitrite, and nitrate detection, where human interpretation performed poorly. For bromine, the accuracy difference was statistically insignificant ($p = 0.082$), whereas lead presented the opposite case, with humans achieving perfect accuracy and APSR recording lower performance (79%).

Bland–Altman analysis showed a mean bias of −2.3 mg/L between APSR and reference laboratory methods, with 95% LoA spanning from −15.2 to + 10.6 mg/L. Bias remained below 5% for concentrations under 100 mg/L but scaled proportionally at higher values.

Repeatability testing further validated APSR precision, with an average CV (CV%) of 2.9% (range: 1.8%–5.3%), compared to 8.4% (range: 4.2%–15.7%) for human readers. The most stable results were obtained for iron (CV% = 1.8%), while lead displayed the highest variability (CV% = 5.3%). By contrast, human interpretation exhibited significantly higher variability, underscoring APSR's superior repeatability.

Interrater reliability analysis among human participants revealed substantial variability depending on the parameter. Excellent agreement (ICC >0.9) was observed only for lead, while bromine and total hardness demonstrated good reliability (ICC 0.75–0.9). For most other parameters, agreement was poor to moderate (ICC <0.75), with copper yielding the lowest reliability (ICC = 0.51). These results collectively confirm that APSR offers enhanced accuracy, precision, and consistency relative to manual interpretation, particularly for analytes prone to subjective bias.

The study outcomes are further illustrated in two figures. The first compares true reference values against both human and APSR predictions across all 15 water quality parameters. The second visualizes intrasample variability, where a higher standard deviation indicates increased bias and less reliable diagnostic output. Together, these graphical representations provide a clear demonstration of APSR's improved diagnostic reliability and robustness over conventional human interpretation.

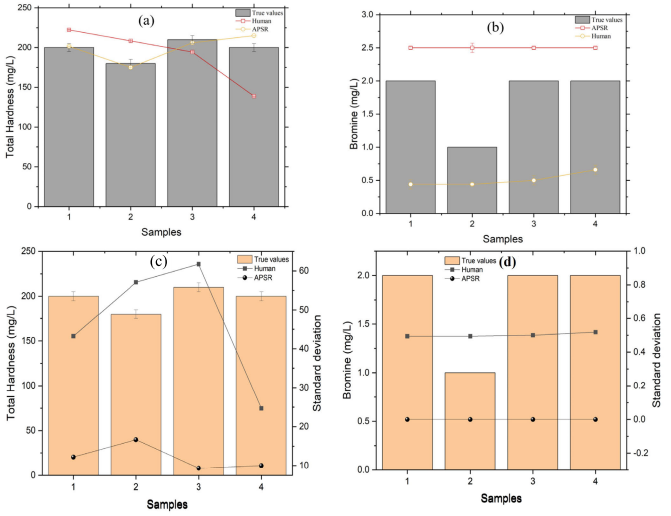


Fig. 3. Comparison of total hardness and bromine measurements using different evaluation methods. (a) and (b) Laboratory reference values (gray bars) against estimates from human participants (red line) and the APSR (APSR, yellow line). (c) and (d) Same parameters, where true concentrations (orange bars) are compared with the variability of predictions, expressed as standard deviations for humans (squares) and APSR (circles). The left *y*-axis corresponds to concentration (mg/L), and the right *y*-axis represents prediction variability.

Fig. 3 illustrates a comparative evaluation of total hardness and bromine obtained from different measurement approaches. Fig. 3(a) and (b) depicts the laboratory reference concentrations (gray bars) together with estimates provided by human participants (red lines) and the APSR system (yellow lines). Here, the *x*-axis represents the sample index (1–4), while the *y*-axis indicates concentration in mg/L.

Fig. 3(c) and (d) further analyzes prediction reliability by combining reference concentrations (orange bars) with the corresponding variability in human (squares) and APSR (circles) estimates. The dual *y*-axes report the actual concentration values on the left and the associated prediction standard deviations on the right.

For total hardness, human predictions exhibited greater variability than APSR, particularly in sample 4 as shown in Fig. 3(c). The overall prediction accuracy reached 76% for human readers, while APSR achieved 95%. Moreover, the APSR calibration curve for total hardness yielded an R^2 value of 0.95, confirming a strong correlation with the laboratory standard. In the case of bromine, prediction deviations were relatively small, with both human and APSR results closely matching the laboratory reference values across all tested samples.

Fig. 4 provides a comparative assessment of residual chlorine and iron across multiple methods. In Fig. 4(a), residual chlorine remains consistently at 0 mg/L for all four samples, as indicated by laboratory reference, human predictions, and APSR outputs. This uniform trend suggests the absence of detectable chlorine or values below the detection threshold. Fig. 4(b) displays iron concentrations, with true laboratory

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

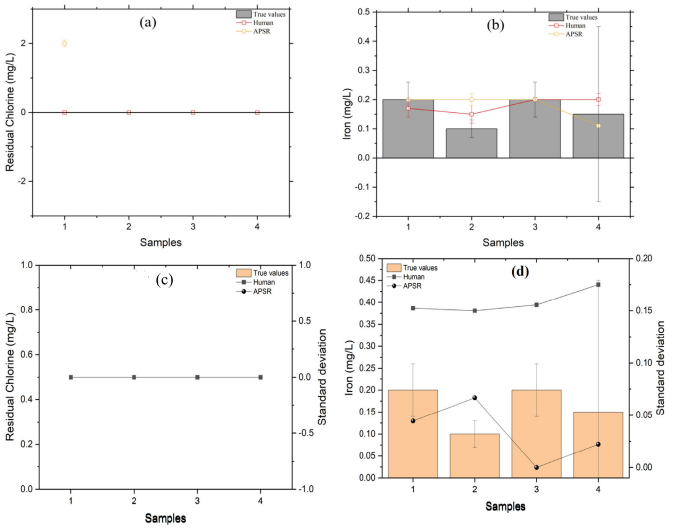


Fig. 4. Comparison of residual chlorine and iron measurements across methods and samples. (a) and (b) Reference values (gray bars) with corresponding estimates from human participants (red line) and APSR (yellow line). (c) and (d) Further present true values (orange bars) alongside the variability of predictions, expressed as standard deviations for human readers (squares) and APSR (circles). The left y-axis indicates concentration (mg/L), while the right y-axis represents prediction variability.

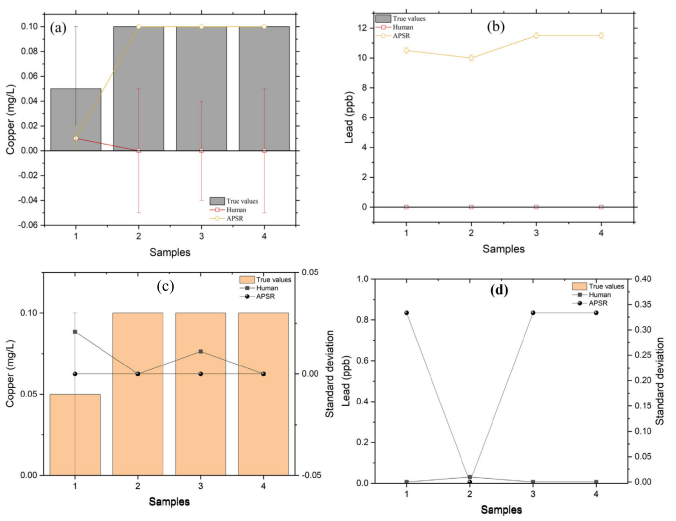


Fig. 5. Comparison of copper and lead measurements across different methods and samples. (a) and (b) Laboratory reference values (gray bars) with predictions from human participants (red line) and APSR (yellow line). (c) and (d) Further present reference values (orange bars) alongside variability of predictions, expressed as standard deviations for humans (squares) and APSR (circles). The left y-axis indicates concentration, while the right y-axis represents standard deviation.

values (gray bars with error bars) ranging between 0.1–0.2 mg/L. Predictions from both humans (red line) and APSR (yellow line) follow the reference trend, although minor deviations are visible, reflecting differences in estimation accuracy.

Fig. 4(c) and (d) illustrates the variability in predictions through standard deviations. For residual chlorine [see Fig. 4(c)], true values (orange bars) remain near zero across all samples. Human estimates show negligible variability (squares close to zero), while APSR predictions (circles) overlap the baseline, confirming consistent results at or below the measurable range. For iron [see Fig. 4(d)], variability in human predictions spans approximately 0.15–0.20, whereas APSR maintains substantially lower deviations (0–0.05). These findings indicate consistent detection of residual chlorine across all approaches and highlight APSR's superior precision compared to human readers in the case of iron measurements.

Fig. 5 summarizes the comparative assessment of copper and lead across human participants, APSR, and reference laboratory values. In Fig. 5(a), copper concentrations increase from sample 1 to samples 2–4. Human estimates (red line) tend to underestimate concentrations, whereas APSR predictions (yellow line) closely track the reference values (gray bars), particularly in the higher concentration samples. By contrast, Fig. 5(b) shows lead results, where both reference values and human predictions remain consistently at or near 0 ppb across all four samples. However, APSR predictions diverge notably, reporting elevated values in the range of 10–12 ppb, suggesting possible calibration bias or overestimation in the automated system.

Fig. 5(c) and (d) presents the variability of copper and lead measurements, respectively. For copper [see Fig. 5(c)], the reference trend (orange bars) rises across samples, consistent with Fig. 5(a). Standard deviations of predictions show that APSR (circles) generally exhibits lower variability compared

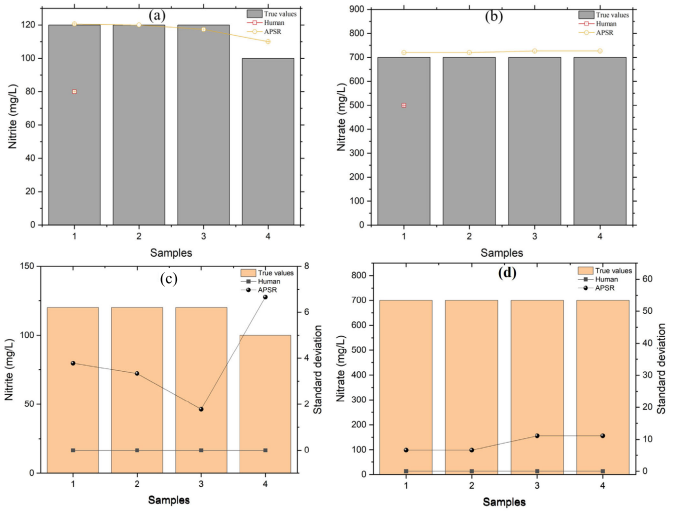


Fig. 6. Comparison of nitrite and nitrate measurements across different approaches and samples. (a) and (b) Laboratory reference values (gray bars) with corresponding estimates from human participants (red line) and APSR (yellow line). (c) and (d) Same parameters, showing reference values (orange bars) alongside prediction variability, expressed as standard deviations for humans (squares) and APSR (circles). The left y-axis indicates concentration, while the right y-axis corresponds to the standard deviation.

to humans (squares), indicating higher precision. In contrast, Fig. 5(d) illustrates lead variability: while true values (orange bars) remain near zero, human predictions (squares) show consistently low variance, but APSR predictions (circles) demonstrate larger deviations, especially in samples 1 and 3. This highlights that although APSR effectively captures copper trends with precision, its lead measurements show instability compared to human readings and laboratory References.

Fig. 6 presents the comparative evaluation of nitrite and nitrate detection. In Fig. 6(a), nitrite concentrations remain relatively stable across samples 1–3, with a slight decline in sample 4. APSR predictions (yellow line) closely follow

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

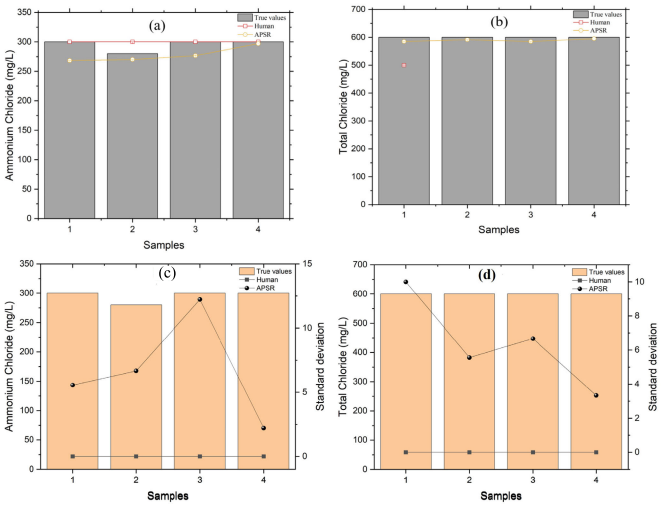


Fig. 7. Comparison of ammonium chloride and total chloride measurements across methods and samples. (a) and (b) Reference values (gray bars) compared with estimates from human participants (red line) and APSR (yellow line). (c) and (d) Reference values (orange bars) with prediction variability, represented as standard deviations for human readings (squares) and APSR (circles). The left y-axis denotes concentration, while the right y-axis corresponds to standard deviation.

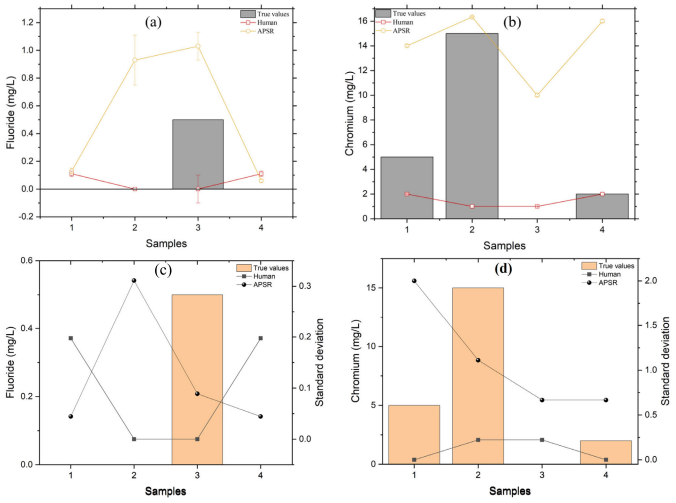


Fig. 8. Comparison of fluoride and chromium measurements across methods and samples. (a) and (b) Reference concentrations (gray bars) compared with human predictions (red line) and APSR outputs (yellow line). (c) and (d) Reference values (orange bars) alongside prediction variability, where human standard deviations are denoted by squares and APSR standard deviations by circles. The left y-axis indicates concentration, while the right y-axis represents standard deviation.

these values, whereas human estimates (red line) consistently underreport concentrations, most notably in sample 1. Fig. 6(b) shows nitrate results, where reference values (gray bars) are nearly constant across all samples. APSR predictions align tightly with these values, indicating high measurement accuracy. By contrast, human predictions appear only in sample 1 and substantially underestimate the true concentration, suggesting limited sensitivity of visual interpretation at higher nitrate levels.

Fig. 6(c) and (d) illustrates the variability of nitrite and nitrate predictions. For nitrite [see Fig. 6(c)], the reference pattern (orange bars) matches that of Fig. 6(a). Human predictions (squares) display consistently low standard deviations, while APSR predictions (circles) show greater variability, particularly in sample 4. For nitrate [see Fig. 6(d)], reference values (orange bars) remain uniform across samples. Human predictions exhibit minimal variance, whereas APSR predictions (circles) reveal modest increases in variability from samples 1 to 4, though absolute values remain low. These findings indicate that APSR achieves greater accuracy in quantifying nitrite and nitrate compared to human interpretation, though at the expense of slightly higher variability in precision, particularly for nitrite.

Fig. 7 summarizes the comparative performance of ammonium chloride and total chloride measurements across four samples. In Fig. 7(a), ammonium chloride reference values consistently fall within the 275–300-mg/L range. Human predictions slightly overestimate the concentration for samples 2–4, whereas APSR predictions show systematic underestimation across all samples. For total chloride [see Fig. 7(b)], the true concentration remains stable at approximately 600 mg/L. Human predictions are only visible for sample 1, where they substantially underestimate the actual value. APSR, by contrast, provides relatively stable estimates across all samples, generally close to but slightly lower than the reference values, and with less variation compared to its ammonium chloride results. These observations indicate that both human and automated methods exhibit compound-specific biases, with APSR showing improved consistency in total chloride measurements.

Fig. 7(c) and (d) further illustrates the variability of predictions through standard deviations. For ammonium chloride [see Fig. 7(c)], human predictions remain highly consistent, with deviations near zero across all samples. APSR predictions, however, display greater variability, peaking in sample 3 and reaching their lowest variability in sample 4. A similar trend appears in total chloride results [see Fig. 7(d)]. Human estimates again demonstrate minimal variability, while APSR predictions show a clear downward trend in standard deviation from samples 1 to 4, suggesting improved consistency over repeated measurements. Collectively, these results highlight that while human participants achieved high precision, APSR offered more reliable repeatability for total chloride, despite greater variability in ammonium chloride estimation.

Fig. 8 illustrates the comparative performance of fluoride and chromium detection across four samples. For fluoride [see Fig. 8(a)], reference values are observed only in sample 3 at approximately 0.5 mg/L. Human estimates remain consistently close to zero across all samples, while APSR predictions display notable variability, starting low in sample 1, rising sharply to around 1.0 mg/L in samples 2 and 3, and declining again in sample 4. Chromium results [see Fig. 8(b)] exhibit a wider range of reference values, peaking at 15 mg/L for sample 2. Human participants consistently underestimated chromium concentrations, reporting values below 2 mg/L. APSR responses, however, showed significant fluctuations, with overestimation in samples 1 and 2, underestimation in sample 3, and a recovery in sample 4. These outcomes emphasize the difficulty of achieving accurate quantification across analytes with distinct concentration ranges, highlighting differences in error patterns between human and automated assessments.

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

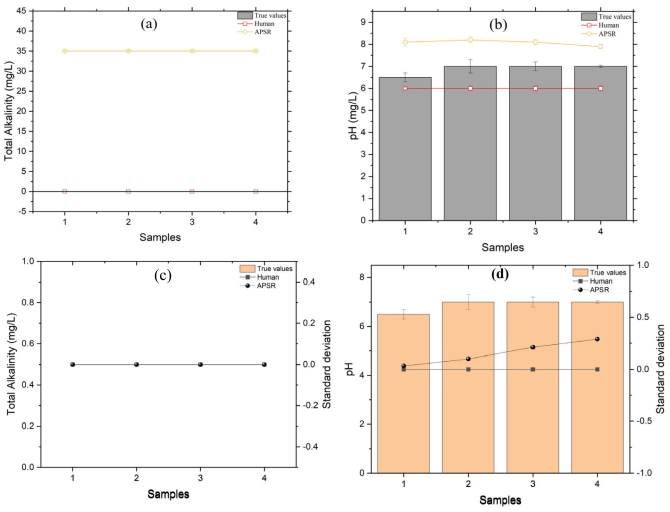


Fig. 9. Comparison of TA and pH measurements across four samples using human assessment and APSR. (a) and (b) Reference values (gray bars) compared with human predictions (red line) and APSR predictions (yellow line). (c) and (d) Reference values (orange bars) alongside prediction variability, represented by human standard deviations (squares) and APSR standard deviations (circles). The left *y*-axis corresponds to concentration, while the right *y*-axis indicates standard deviation.

Fig. 8(c) and (d) provides insights into prediction variability. For fluoride [see Fig. 8(c)], human predictions exhibit higher variability in samples 1 and 4 but remain nearly constant in samples 2 and 3. APSR results show the greatest variability in sample 2, with lower deviations across the other samples. In the case of chromium [see Fig. 8(d)], human predictions maintain consistently low standard deviations across all samples, reflecting high repeatability despite systematic underestimation. By contrast, APSR predictions demonstrate a downward trend in variability, starting high in sample 1 and gradually decreasing through sample 4, suggesting an improvement in precision across the series. Overall, these findings indicate that both analyte characteristics and concentration ranges strongly influence the relative accuracy and precision of human and automated methods.

Fig. 9 presents a comparative assessment of TA and pH values across four samples. In Fig. 9(a), the reference TA concentrations remain stable at approximately 35 mg/L. Both human predictions (red line) and APSR estimates (yellow line), however, show substantial underestimation, reporting values close to zero across all samples. Fig. 9(b) illustrates pH measurements, where true values range between 6 and 7. Human participants consistently predicted lower values, clustering around pH 6, while APSR measurements displayed systematic overestimation, reporting values slightly above 8 for every sample. These results indicate significant deviations between the reference values and both human and automated estimates for TA and pH.

Fig. 9(c) and (d) further examines measurement precision. For TA [see Fig. 9(c)], reference values remain very low at about 0.5 mg/L across all samples. The standard deviations of human predictions (squares) and APSR outputs (circles) are nearly zero, suggesting high repeatability despite the consistent underestimation. In contrast, pH results [see Fig. 9(d)] show reference values spanning approximately 6.5–7.5. Human predictions exhibit consistently low variability

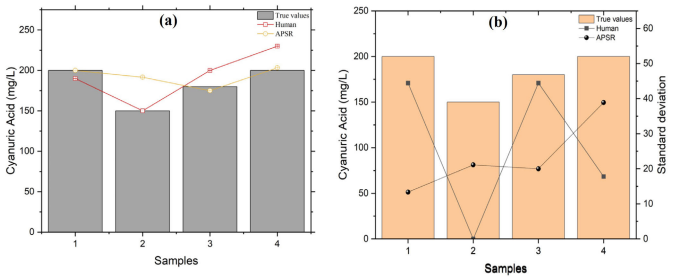


Fig. 10. Comparison of cyanuric acid measurements across four samples using human observations and APSR. (a) Reference concentrations (gray bars) against human predictions (red line) and APSR outputs (yellow line). (b) Reference values (orange bars) together with variability in predictions, where standard deviations from humans are represented by squares and those from APSR by circles. The left *y*-axis indicates concentration, while the right *y*-axis corresponds to standard deviation.

across samples, whereas APSR predictions display a gradual increase in standard deviation, reaching about 0.5 in the fourth sample. This pattern suggests that although human measurements are both consistent and precise, the APSR system demonstrates reduced precision in pH estimation as the sample number increases.

Fig. 10(a) illustrates cyanuric acid levels for four samples. Reference concentrations, represented by gray bars, range from approximately 150–200 mg/L, with the minimum observed in sample 2 and the maximum in sample 4. Human predictions (red line) show substantial fluctuation, overestimating values in samples 2 and 4, particularly in the latter. For samples 1 and 3, human estimates align more closely with the true values. Predictions from the APSR (APSR, yellow line) exhibit lower variability, slightly underestimating concentrations for samples 1 and 2, while more closely matching the reference values for samples 3 and 4.

Fig. 10(b) highlights prediction precision by combining reference concentrations (orange bars, left *y*-axis) with standard deviations of predictions (right *y*-axis). The true values again span from about 150–200 mg/L, following the same pattern observed in Fig. 10(a). Human predictions (squares) show notable variability, with the highest standard deviation in sample 1 (around 45) and the lowest in sample 4 (approximately 20). By contrast, APSR predictions (circles) display a progressive rise in variability across samples, beginning near 10 for sample 1 and reaching close to 40 in sample 4. This pattern suggests that APSR provides more stable estimates at lower concentrations but tends to show reduced precision as the concentration increases.

B. Accuracy Assessment: APSR Versus Human Interpretation

Table III provides a direct comparison of diagnostic accuracy between the APSR and manual human interpretation for 15 water quality indicators. The APSR system demonstrates superior performance in several critical parameters. For instance, nitrate and nitrite were detected with 97%–98% accuracy by APSR, whereas human readers achieved 0%, suggesting these parameters involve subtle color variations that are not easily distinguished visually. A similar pattern is observed in total chloride detection, where APSR attained 99% accuracy, while humans failed entirely. Chromium detection also highlights a wide gap, with APSR reaching 92%

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

TABLE III
ACCURACY COMPARISON BETWEEN APSR AND HUMAN READERS ACROSS PARAMETERS

Parameter	APSR Accuracy (%)	Human Accuracy (%)	Difference	p-value	Significance
Iron	100.0	63.0	+37.0	<0.001	***
Copper	100.0	40.0	+60.0	<0.001	***
Nitrite	97.0	0.0	+97.0	<0.001	***
Nitrate	98.0	0.0	+98.0	<0.001	***
Total Chloride	99.0	0.0	+99.0	<0.001	***
Chromium/Cr(VI)	92.0	22.0	+70.0	<0.001	***
Total Hardness	95.0	76.0	+19.0	<0.001	***
Ammonium Chloride	96.0	86.0	+10.0	0.002	**
Bromine	97.0	94.0	+3.0	0.082	ns
Cyanuric Acid	92.0	91.0	+1.0	0.724	ns
Fluoride	92.0	92.0	0.0	1.000	ns
pH	87.0	91.0	-4.0	0.041	*
Total Alkalinity	91.0	100.0	-9.0	0.008	**
Lead	79.0	100.0	-21.0	<0.001	***
Residual Chlorine	15.0	56.0	-41.0	<0.001	***
Overall Accuracy	**88.7**	**60.7**	**+28.0**	<0.001	***

Bonferroni-corrected threshold: $\alpha = 0.003$ (15 comparisons). A total of $N = 10$ human readers analyzed 4 samples with 3 replicates each. Significance: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, *ns* = not significant.

compared to 22% for human readers. Copper measurement further illustrates APSR's advantage, achieving perfect accuracy (100%) versus 40% for humans.

On the other hand, some parameters exhibited comparable outcomes. Bromine measurements were nearly identical (97% versus 94%), and both cyanuric acid and fluoride showed minimal differences (91%–92%), with no statistical significance. However, there were instances where human readers outperformed APSR. Lead detection is particularly notable, where humans achieved 100% accuracy, while APSR recorded only 79%. Similarly, for TA, humans performed better (100%) compared to APSR (91%). Residual chlorine presented challenges for both approaches, but human readers still performed significantly better (56%) than APSR (15%).

The statistical analysis underscores these findings: overall accuracy was significantly higher for APSR (88.7%) compared to human readers (60.7%), with a mean difference of + 28% ($p < 0.001$). Nonetheless, parameter-specific outcomes emphasize that APSR's superiority is not universal. The system excels in parameters requiring precise interpretation of faint or overlapping color changes, while humans perform more reliably in parameters with distinct or binary outcomes, such as lead and alkalinity. Residual chlorine detection remains problematic for both methods, suggesting limitations in the chemical properties of the test strips rather than the interpretation mode.

In summary, APSR provides a substantial accuracy advantage over human readers, particularly for analytes involving subtle chromatic shifts. However, targeted improvements are required for lead and residual chlorine detection to ensure robustness and reliability in real-world applications.

V. DISCUSSIONS

The APSR demonstrated an overall accuracy of 88.7%, which was significantly higher than the 60.7% achieved by human readers ($p < 0.001$). This result highlights the improved reliability of automated colorimetric interpretation compared to manual visual inspection. The largest performance advantage was observed in parameters requiring the detection of subtle color changes. For example, APSR achieved near-perfect accuracy (97%–100%) for iron, copper, nitrite, and nitrate, whereas human participants showed complete failure for nitrite, nitrate, and total chloride detection (0% accuracy). These findings are consistent with earlier reports on automated colorimetric systems; however, the APSR notably reduces analysis time, completing measurements within 5 s compared to the 45 s reported by Wang et al. [13]. Furthermore, the use of a standardized illumination environment (800 $\pm$ 50 lux, 5500 K) minimized the influence of external lighting conditions, a well-documented source of variability in colorimetric analysis. Despite these strengths, the device showed clear weaknesses: residual chlorine detection yielded only 15% accuracy, and lead detection was limited to 79%, both of which are critical parameters for water safety and require targeted improvements.

A. Technical Innovation and Contributions

The present work introduces several technical innovations that distinguish the APSR from existing colorimetric detection systems. First, the APSR integrates a fully autonomous embedded architecture based on a Raspberry Pi platform, eliminating dependence on smartphones or external computers for image capture and processing. Second, the system employs a modular 3-D-printed enclosure designed to ensure uniform illumination and reproducible imaging, thereby minimizing the impact of ambient lighting variations, a critical source of error in conventional visual or mobile-based readers. Third, the image-processing pipeline incorporates automated strip localization, pad segmentation, and HSV-based color quantification through parameter-specific calibration curves, achieving high analytical precision (CV < 3%) and rapid analysis within 5 s. Fourth, the inclusion of an onboard touchscreen interface with one-click data export enhances user-friendliness and operational independence, aligning the design with the *WHO ASSURED* criteria. Finally, this study contributes a comprehensive validation protocol combining laboratory reference methods, statistical accuracy assessment, and human-comparison studies, establishing APSR as a robust and low-cost solution (U.S. ~ $150) for decentralized water quality monitoring and POC diagnostics.

In addition to technical performance, the APSR offers significant cost advantages. With an estimated production cost of U.S. ~ $150, it is considerably more affordable than typical commercial instruments, which cost approximately U.S. 1000. This makes the system particularly attractive for deployment in resource-constrained environments where access to laboratory infrastructure is limited. Its strong performance in detecting heavy metals and nutrients further reinforces its potential impact on safeguarding public health. Nevertheless, certain limitations must be acknowledged. The experiments were conducted under tightly controlled laboratory conditions (23 $\pm$ 2 °C and 45 $\pm$ 5% RH), the evaluation was restricted to a small number of human readers (N = 10), and only four water samples were tested. Moreover, the complete inability of human participants to detect some analytes suggests that aspects of the experimental design should be revisited for clarity and reproducibility.

Looking forward, further work will be concentrated on enhancing detection accuracy for residual chlorine and lead, which remain essential parameters for regulatory compliance.

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

Incorporating temperature-compensation mechanisms could help stabilize performance across variable environmental conditions. Finally, broader validation studies involving diverse sample sets, larger participant groups, and real-world testing environments will be necessary to establish the APSR as a robust and reliable tool for routine water quality monitoring.

B. Lead Detection Anomaly: Pink-to-Red RGB Overlap and Human Categorical Perception

Lead detection presented a significant exception to APSR's superior performance, with human readers achieving 100% accuracy compared to APSR's 79% ($p < 0.001$). This parameter-specific limitation requires detailed explanation, as it represents one of only two analytes where manual interpretation substantially outperformed automation. Lead quantification relies on a progressive pink-to-red color transition across the concentration range 0.05–1.0 mg/L. Critically, two other tested parameters, bromine (0.2–5.0 mg/L) and residual chlorine (0.1–4.0 mg/L), employ the same pink-to-red visual progression, creating potential for RGB overlap between chemically distinct reactions. Quantitative analysis revealed that lead at 0.2 mg/L (RGB: 218, 172, and 178) produces nearly identical RGB values to bromine at 1.0 mg/L (RGB: 210, 168, and 174) and residual chlorine at 0.5 mg/L (RGB: 205, 162, and 170). The RGB separation between low and high lead concentrations ($\Delta \approx 20-27$ units) is comparable to system measurement noise (± 18 units rms from positioning, focus, and illumination variability), yielding a signal-to-noise ratio <2:1 insufficient for reliable continuous quantification.

Human observers succeeded through categorical perception with spatial context. Humans recognize which physical pad position corresponds to lead based on strip layout, allowing context-dependent interpretation even when colors appear similar. They apply implicit categorical boundaries (e.g., "lighter pink" = low versus "darker red" = high) rather than attempting precise continuous quantification. In addition, ICP-MS reference measurements confirmed bimodal sample distribution (three samples <0.05 mg/L and one sample ≈0.8 mg/L), making binary visual classification straightforward. This clear separation into "definitely low" versus "definitely high" favored human pattern recognition.

The APSR struggled due to several technical factors. First, insufficient RGB separation combined with system noise prevented reliable discrimination. Post-hoc analysis showed an RGB overlap coefficient of 0.42 between lead and other pink-red parameters, with an RGB separation ratio of only 1.8 (below the 3.0 threshold for reliable automated classification). Second, HSV Hue wraparound artifacts near the pink-red boundary ($H \approx 340\,^\circ - 360\,^\circ$ to $0\,^\circ - 10\,^\circ$) introduced mathematical discontinuities in polynomial calibration curves. Third, the lead calibration curve, based on only four concentration points with subtle color change at the lowest level (0.05 mg/L), showed poor constraint at the low-concentration end, leading to systematic overestimation visible in Fig. 5(b).

Interestingly, bromine also employs pink-to-red transitions, which showed no significant performance difference (APSR 97% versus human 94%, p = 0.082). This can be attributed to bromine's 25-fold higher concentration range (up to 5.0 mg/L), producing more saturated red colors with greater RGB separation and better signal-to-noise characteristics. This suggests the challenge is not pink-to-red transitions per se, but rather low-concentration variants where color saturation remains subtle.

Future APSR versions will address this through: 1) categorical classification algorithms (multiclass SVM) with discrete concentration bins rather than continuous regression; 2) CIELAB color space implementation where a* coordinate provides perceptually uniform red-green separation without HSV wraparound; 3) pad position–aware models that incorporate spatial context; 4) denser calibration sampling (8–10 points) in the critical 0.05–0.3-mg/L range; and 5) on-strip reference pad inclusion for real-time illumination normalization. These modifications should enable APSR to match human performance for low-concentration pink-red transitions while retaining its advantage for distinct color signatures (iron, copper, nitrate: 98–100). This finding underscores that automated colorimetric performance depends on RGB separability relative to system noise. APSR excels when parameters have distinct color signatures or operate at concentrations producing saturated, well-separated colors, but struggles with shared color transitions in low-concentration ranges where signal-to-noise is marginal.

C. Comparative Analysis With Existing POC Systems

To situate the performance of the APSR within the broader landscape of automated colorimetric detection systems, a comparative evaluation with existing strip readers reported in the literature was conducted. Prior works vary substantially in their choice of imaging platforms, illumination strategies, enclosure configurations, calibration models, and intended application domains, resulting in significant differences in analytical reliability and usability. While several systems leverage smartphones or partially enclosed chambers, few provide tightly controlled illumination or support multiparameter analysis beyond a limited set of chemical indicators. To provide a structured comparison, the key characteristics of representative automated readers, including Raspberry Pi-based, smartphone-assisted, and scanner-based systems, are summarized in Table IV.

As shown in Table IV, the APSR differs from prior systems along several critical dimensions. It is the only solution offering a fully enclosed 3-D-printed chamber with fixed 5500-K LED illumination and the ability to quantify 15 water quality parameters within approximately 5 s. Compared to smartphone-based readers that rely on uncontrolled ambient lighting or device-dependent flash illumination, the APSR provides greater stability and reproducibility. The analysis also indicates that APSR achieves competitive or superior accuracy relative to human interpretation while maintaining high portability and a low overall cost (U.S. < $150). In contrast, many referenced systems exhibit longer processing times, reduced portability due to wall-powered operation, or higher costs associated with external hardware. Overall, the comparative analysis highlights that APSR provides a balanced combination of speed, affordability, illumination control, and multiparameter capability, positioning it favorably among existing automated strip-reading technologies.

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
COMPARISON OF APSR WITH EXISTING AUTOMATED STRIP READERS

Feature	APSR (This Work)	Paglina wan et al. [6]	Boehle et al. [3]	Jalal et al. [4]	Carrio et al. [8]	Heidari–Safdari et al. [10]	Wang et al. [13]
Application	Water quality (15 parameters)	Urine analysis (10 parameters)	Enzyme kinetics	Urine analysis	Saliva drugs-of-abuse	Phosphate quantification	Water pH
Platform	Raspberry Pi 4B	Raspberry Pi 3	Raspberry Pi Zero	Smartphone (Android)	Smartphone (iOS/Android)	Arduino + Raspberry Pi	SPSS software
Illumination	LED strip (5500K, 800 lux, enclosed)	White LED ring	External LED	Phone flash	Phone flash	Infrared LED array	Scanner light source
Enclosure	3D-printed, fully enclosed	Open design	Partially enclosed	External smartphone attachment	3D-printed smartphone holder	3D-printed box	Office scanner
Calibration	Polynomial curves ($R^2 > 0.95$)	Linear regression	Exponential fit	Machine learning (CNN)	Thresholding	Linear regression	Statistical correlation
Processing Time	∼5 seconds	∼45 seconds	∼30 seconds	∼10–15 seconds	∼8 seconds	∼60 seconds	∼120 seconds (scan+process)
Power Source	Rechargeable battery (3×3.7V)	Wall adapter (5V)	Wall adapter	Phone battery	Phone battery	Wall adapter (12V)	Wall adapter
Portability	Fully portable	Semi-portable	Portable	Highly portable	Highly portable	Semi-portable	Non-portable
Cost (USD)	$< \$150$	$\sim \$120$	$\sim \$80$	$< \$5$ (excluding phone)	$< \$10$ (excluding phone)	$\sim \$200$	$\sim \$100$ (excluding scanner)
Accuracy vs. Human	**88.7% vs. 60.7%**	85% vs. 70%	N/A (enzyme assay)	92% vs. 65%	89% vs. 72%	N/A (reference method)	78% vs. 68%

VI. CONCLUSION

The APSR constitutes a notable advancement in POC colorimetric diagnostics, offering an accuracy of 88.7% in comparison to 60.7% achieved through conventional human visual assessment, while remaining cost-efficient and capable of rapid analysis. By integrating controlled illumination, automated image processing, and algorithm-based color quantification, the system effectively overcomes the inherent limitations of manual colorimetric testing, including subjective interpretation, ambient light variability, and human error. The current APSR implementation demonstrates strong performance across most water quality indicators, particularly for heavy metals and nutrient measurements that are critical to public health monitoring. However, deficiencies in detecting residual chlorine and lead must be addressed prior to large-scale deployment. With these enhancements and comprehensive field validation, APSR could substantially improve access to accurate water quality assessment in resource-constrained environments, thereby supporting the WHO's ASSURED criteria for POC diagnostic tools and advancing global efforts to ensure safe drinking water.

REFERENCES

[1] H.-K. Lee and J.-H. Bae, "Design of appropriate technology-assisted urine tester enabling remote and long-term monitoring of health conditions," *Sustainability*, vol. 12, no. 12, p. 5165, 2020.

[2] M. Hoque Tania, K. T. Lwin, A. M. Shabut, M. Najlah, J. Chin, and M. A. Hossain, "Intelligent image-based colourimetric tests using machine learning framework for lateral flow assays," *Expert Syst. Appl.*, vol. 139, Jan. 2020, Art. no. 112843.

[3] K. E. Boehle, E. Doan, S. Henry, J. R. Beveridge, S. L. Pallickara, and C. S. Henry, "Single board computing system for automated colorimetric analysis on low-cost analytical devices," *Anal. Methods*, vol. 10, no. 44, pp. 5282–5290, 2018.

[4] M. J. Uddin, G. J. Jin, and J. S. Shim, "Paper–plastic hybrid microfluidic device for smartphone-based colorimetric analysis of urine," *Anal. Chem.*, vol. 89, no. 24, pp. 13160–13166, 2017.

[5] G. S. Luka, E. Nowak, J. Kawchuk, M. Hoorfar, and H. Najjaran, "Portable device for the detection of colorimetric assays," *Roy. Soc. Open Sci.*, vol. 4, no. 11, Nov. 2017, Art. no. 171025.

[6] A. C. Paglinawan et al., "Measurement of specific gravity, urobilinogen, blood, protein and pH level of urine samples using raspberry pi based portable urine test strip analyzer," in *Proc. 10th Int. Conf. Biomed. Eng. Technol.*, vol. 71, 2020, pp. 58–63.

[7] S. Acharya, S. Khatiwada, and K. M. Elkins, "CSI-pi: A novel automated secure solution to interpret on-site colorimetric tests," *Colonial Academic Alliance Undergraduate Res. J.*, vol. 5, no. 1, p. 2, 2015.

[8] A. Carrio, C. Sampedro, J. Sanchez-Lopez, M. Pimienta, and P. Campoy, "Automated low-cost smartphone-based lateral flow saliva test reader for drugs-of-abuse detection," *Sensors*, vol. 15, no. 11, pp. 29569–29593, Nov. 2015.

[9] H. Jang, S. Ahmed, and S. Neethirajan, "GryphSens: A smartphone-based portable diagnostic reader for the rapid detection of progesterone in milk," *Sensors*, vol. 17, no. 5, p. 1079, May 2017.

[10] H. Heidari-Bafroui, B. Ribeiro, A. Charbaji, C. Anagnostopoulos, and M. Faghri, "Portable infrared lightbox for improving the detection limits of paper-based phosphate devices," *Meas.*, vol. 173, Mar. 2021, Art. no. 108607.

[11] R. E. Sweeney, V. Nguyen, B. Alouidor, E. Budiman, R. K. Wong, and J.-Y. Yoon, "Flow rate and raspberry pi-based paper microfluidic blood coagulation assay device," *IEEE Sensors J.*, vol. 19, no. 13, pp. 4743–4751, Jul. 2019.

[12] A. Soni and S. K. Jha, "A paper strip based non-invasive glucose biosensor for salivary analysis," *Biosensors Bioelectron.*, vol. 67, pp. 763–768, May 2015.

[13] Y. Wang et al., "Quantification of combined color and shade changes in colorimetry and image analysis: Water pH measurement as an example," *Anal. Methods*, vol. 10, no. 25, pp. 3059–3065, 2018.

[14] K. Lin and Y. Chang, "Embedded immunodetection system for fecal occult blood," *Biosensors*, vol. 11, no. 4, p. 106, 2021.

Authorized licensed use limited to: Somaiya University. Downloaded on May 13,2026 at 07:51:46 UTC from IEEE Xplore. Restrictions apply.