Imaginex International Journal of Interdisciplinary Research and Reviews
Volume 1, Issue 3, November – December 2026
ISSN: 3140-0094 (Online)
www.iijirr.com

GW-EdgeTwin: An Edge-Intelligent IoT and Digital-Twin Framework for Real-Time Greywater Quality Forecasting and Treatment Control in Circular Campus Water Systems

Dr. Muhammadu Sathik Raja[1], Dr. Karthiga B[2], Ms. Sri Harini S[3], Dr. Surya[4]

[1]Research Coordinator, Dhanalakshmi Srinivasan Institute of Research and Technology. *scewsadik@gmail.com*
[2]Associate Professor, Department of Electronics and Communication Engineering, Dhanalakshmi Srinivasan Institute of Research and Technology. *karthiga.jaya15@gmail.com*
[3]Assistant Professor, Department of Bio-Medical Engineering, Dhanalakshmi Srinivasan Institute of Research and Technology. *sriharinisaravanan07@gmail.com*
[4]Associate Professor, Department of Biochemistry, Dhanalakshmi Srinivasan Institute of Research and Technology. *chinnaasurya@gmail.com*

Abstract

Greywater can be reused on site only if its quality is known, yet its organic and microbial load is still measured in the laboratory, and the results usually come too late for a treatment unit to respond to changes in the incoming water. We built GW-EdgeTwin to remove this delay. It links a fifteen-node ESP32–LoRa sensor network, NVIDIA Jetson Orin Nano edge processors and a process digital twin of a 2 KLD circular recycling pilot plant. A soft sensor that combines temporal convolution with a Transformer estimates chemical oxygen demand (COD), five-day biochemical oxygen demand (BOD_5) and *E. coli* load from pH, turbidity, total dissolved solids (TDS), temperature and flow. Before each estimate, a multivariate anomaly gate screens out faulty sensor readings. The digital twin then turns these forecasts into set-points for the bio-reactor, filtration and ultraviolet advanced oxidation (UV-AOP) stages. On the PJ-GWQ-2026 dataset, the framework reached a reuse-class accuracy of 94.6% and an R^2 of 0.93 for COD with an edge latency of 38 ms, and it outperformed LSTM, GRU, Transformer and random forest baselines.

Keywords: Greywater reuse, Edge intelligence, Internet of Things, Digital twin, Soft sensor, Temporal convolution–Transformer, Circular water systems

1. Introduction

Water stress in fast-growing cities has made domestic greywater, the lightly polluted water from showers, wash basins, laundries and kitchens, a useful second source of water for institutional campuses. Its quality, however, changes a great deal over a single day. Detergents, personal-care products and food residues reach the drain in irregular bursts, and a fixed treatment schedule cannot guarantee safe reuse when the load keeps shifting [1].

Quality checks still depend on grab samples tested for chemical oxygen demand (COD), five-day biochemical oxygen demand (BOD_5) and faecal indicators. These tests take hours or even days, so by the time a result arrives the water has often been reused already. Data-driven methods have begun to narrow this gap: hybrid learning pipelines can now direct greywater samples to suitable treatment and irrigation routes [2]. Digital twins offer another option, because they mirror plant behaviour in software and let operators adjust conditions before a deviation reaches the outlet [3].

Edge–cloud collaboration has been shown to shorten the response time of water-quality analytics in distribution networks [4], and LoRaWAN nodes can send multiparameter readings from remote sites while using very little power [5]. Even so, the work in this area is scattered. Sensing studies usually end with a visual display. Prediction studies work on historical data stored centrally. Digital-twin studies deal with large municipal plants, not with small greywater units. We found no study that joins low-cost sensing, on-site prediction and automatic treatment adjustment for decentralised greywater in a single loop.

GW-EdgeTwin brings three new elements together. The first is a cross-node anomaly gate, which removes drifting or fouled sensor readings before they reach the predictor but lets real pollution pulses through. The second is a soft sensor in which a temporal convolution branch and a Transformer branch are joined by gated fusion; it runs on Jetson Orin Nano edge processors and infers laboratory-grade parameters from low-cost probes. The third is a process digital twin of the 2 KLD (2 m^3/day) pilot plant, which turns the forecasts into set-points for the bio-reactor, the sand–carbon filter and the ultraviolet advanced oxidation (UV-AOP) unit.

Our aim is a greywater recycling chain that checks itself all the time, adjusts itself, and can be copied to other campuses and housing communities. Replacing slow laboratory checks with real-time edge intelligence supports safer non-potable reuse, reduces energy use and backwash-water demand, and builds up the records that circular

water management depends on. More broadly, the work shows that low-cost hardware combined with process-aware learning can give decentralised reuse schemes the reliability they need before institutions and the public will accept them.

2. Literature Review

Continuous water-quality monitoring now usually starts with low-cost sensors. Point-to-point LoRa links have carried temperature, pH, conductivity and turbidity data across river catchments of several square kilometres on a small energy budget [6]. Printable housings and simpler wiring have cut the effort needed to install real-time water-level sensors in deep wells [7]. Distributed wireless networks have also been used to protect groundwater quality in cross-border wetlands, with attention to scalability and low power draw [8]. These studies show that the sensing layer is no longer expensive. Most of them, however, stop once the data have been delivered.

Raw sensor data suffer from drift, biofouling and transmission faults. One multivariate network, built from several convolutional branches and long short-term memory (LSTM) units, picked out anomalous readings in real time with high precision [9]. A hybrid convolution–LSTM model estimated dissolved oxygen from other sensor signals when no dedicated probe was available [10], and an LSTM soft sensor has been designed to estimate hard-to-measure variables in activated-sludge plants [11]. Learned models can therefore both clean sensor data and fill in missing measurements. All of these models, though, were built for municipal plants or rivers, not for greywater.

The next step is forecasting effluent quality when the load varies. Shyu et al. compared four machine learning algorithms for predicting COD, TSS and *E. coli* in a field-tested onsite treatment system with a small dataset [12]. Another study tested LSTM, GRU and Transformer soft sensors under dry-weather, rain and storm scenarios and found that the best model changed with the hydraulic regime [13]. In a full-scale plant, a probabilistic encoder–decoder LSTM gave better total-nitrogen predictions during shock-load events [14]. Because the results depend so much on the scenario, a hybrid model that can follow both short and long temporal patterns is worth testing.

Operators will only trust a model if they can see why it gives a certain output. Li et al. applied Shapley additive explanations to an LSTM effluent model to separate the contribution of each feature and time lag [15]. Nasir and Li compared local and global explanation tools for gradient boosting, random forest and neural network predictors of treatment-plant variables [16], and later tested the same approach on small and large plants to check how well it generalises [17]. These studies show that explanations can be added without losing accuracy. None of them, however, placed the explanation step inside an edge controller.

Optimisation and distributed learning have also been explored. Bayesian-optimised gradient boosting, combined with a multi-objective search, balanced effluent quality against energy use in a treatment process [18]. A federated framework fused data from several sensors to predict industrial pollutants without moving raw data off site [19], and a purpose-built compiler made contaminant detection possible on low-resource edge hardware [20]. Even with this progress, no study has combined low-cost sensing, anomaly screening, edge inference and twin-based control for decentralised greywater. GW-EdgeTwin was designed to provide that combination.

3. Methodology

GW-EdgeTwin works as a closed loop of sensing, prediction and control, organised in four layers (Figure 1). In the perception layer, fifteen IP65 ESP32 nodes carry pH, turbidity, TDS, temperature and flow probes and connect through LoRa to a campus gateway. The edge layer has four NVIDIA Jetson Orin Nano units, placed at the greywater collection sump, the bio-reactor outlet, the filter outlet and the UV-AOP outlet. A GPU workstation hosts the analytics layer, where the soft sensor is trained and the digital twin is maintained. The application layer shows the dashboards and sends control commands to the pumps, the dosing units and the UV lamp driver.

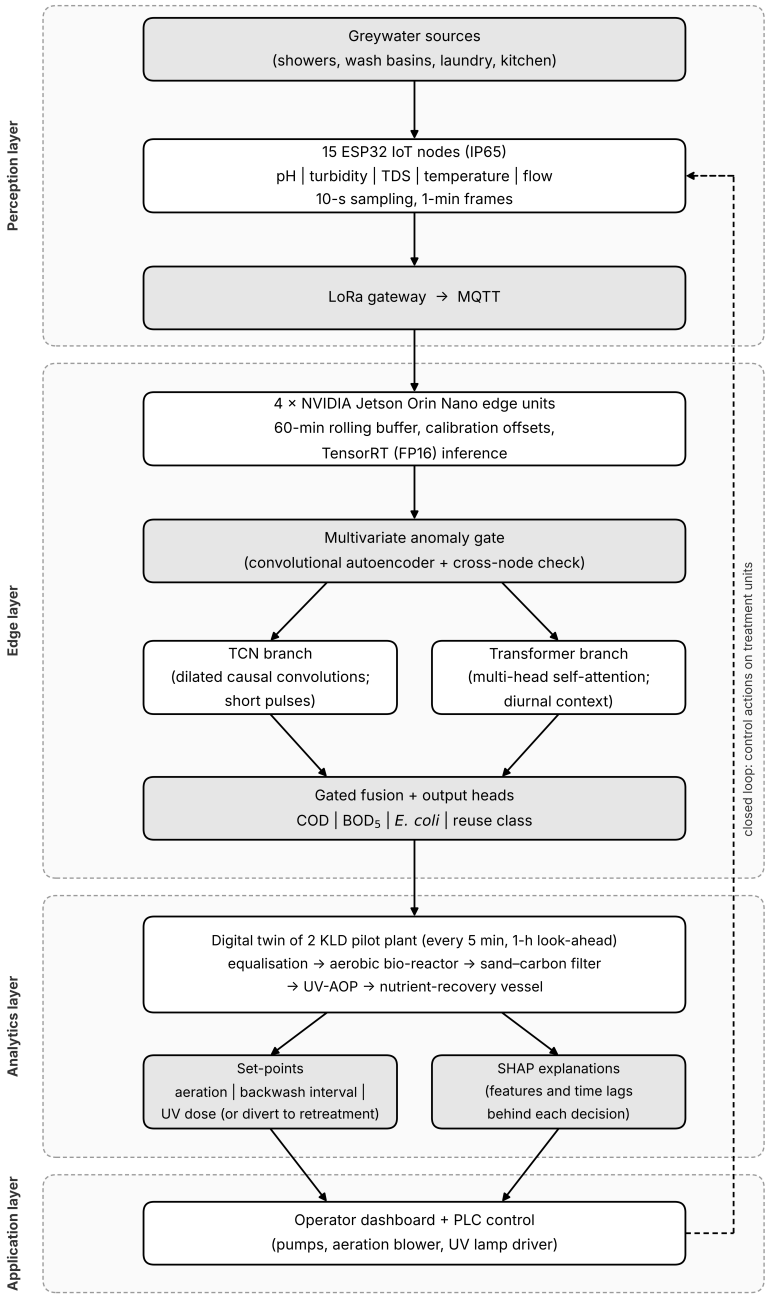

Figure 1. Proposed architecture of the GW-EdgeTwin framework

A. Dataset Description and Preparation

We used the PUNARJAL Greywater Quality Telemetry Dataset (PJ-GWQ-2026). It was generated and curated by the Smart Water Innovation Laboratory under the PUNARJAL circular water programme. The dataset pairs one-minute telemetry from the fifteen sensor nodes with laboratory reference values for COD, BOD_5 and *E. coli*, which were obtained with a multiparameter analyser, a COD digester, a BOD incubator, a UV–visible spectrophotometer and membrane-filtration microbiology. Every record contains a timestamp, the node identifier, five sensed attributes, four derived features (hour of day, day type, cumulative flow and turbidity gradient), three reference targets and a reuse class: toilet flushing, landscape irrigation or retreatment. To prepare the data, we synchronised the clocks, removed records from calibration periods, forward-filled gaps shorter than five minutes and applied min–max scaling fitted on the training data only.

B. Sensor Network and Edge Acquisition

Each node reads its probes every ten seconds, averages the readings into one-minute frames and sends them over LoRa, with the spreading factor adjusted to the link quality. The gateway passes the frames over MQTT to the nearest Jetson Orin Nano, which keeps a rolling sixty-minute window for each node. Every two weeks the low-cost probes are calibrated against the reference analyser, and the offsets are stored on the edge device, so every

later stage works with corrected values. Packet loss, battery voltage and signal strength are logged with each frame; this lets the system tell a communication fault from a real change in water quality. Because the data are buffered at the edge, the treatment loop keeps running even when the link to the analytics server drops.

C. Multivariate Anomaly Gate

Every one-minute window goes through an anomaly gate before prediction. The gate is a convolutional autoencoder trained only on data from verified normal operation: the encoder compresses the multivariate window into a short latent vector and the decoder rebuilds it. If the reconstruction error is above a threshold set on the validation data, the window is flagged. A flag does not mean the data are thrown away. The gate first compares the error pattern with that of neighbouring nodes, since a single faulty probe looks different from a real pollution pulse, which shows up at several nodes at once. Faulty channels are filled with values from the last reliable state, and a maintenance ticket is raised for that node. Real pulses pass through unchanged, with an alert flag attached.

D. Hybrid TCN–Transformer Soft Sensor

The cleaned window then enters a soft sensor with two branches. The temporal convolution branch uses stacked dilated causal convolutions [21] to pick up short pulses, such as detergent surges that last only a few minutes. The Transformer branch applies multi-head self-attention [22] to the whole sixty-minute window, so it can follow slower daily patterns that depend on building occupancy. The two outputs are concatenated and fed to a gated fusion layer, which learns, sample by sample, how much weight each branch should get. Three regression heads estimate the COD, BOD_5 and *E. coli* concentrations, and a classification head assigns the reuse class. All heads are trained together with a weighted multi-task loss. The trained model is then quantised to half precision with TensorRT and deployed on the Jetson devices.

In equation form, the TCN branch applies dilated causal convolutions to the input sequence x, using a filter f of size k and a dilation factor d (Eq. 1). Each head of the Transformer branch applies scaled dot-product self-attention (Eq. 2):

$$F(t) = \Sigma_{i=0}^{k-1} f(i) \cdot x(t - d \cdot i) \tag{1}$$

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^{T} / \sqrt{d_k})\, V \tag{2}$$

For every sample, the gate g decides how much weight goes to the TCN output h_{TCN} and how much to the Transformer output h_{TF} (Eq. 3). The four output heads are trained jointly with a weighted multi-task loss (Eq. 4):

$$g = \sigma(W_g[h_{TCN}; h_{TF}] + b_g), \quad h = g \odot h_{TCN} + (1 - g) \odot h_{TF} \tag{3}$$

$$L = \lambda_1 L_{COD} + \lambda_2 L_{BOD} + \lambda_3 L_{EC} + \lambda_4 L_{CE} \tag{4}$$

where σ is the logistic sigmoid function and $\odot$ denotes element-wise multiplication. The terms L_{COD}, L_{BOD} and L_{EC} are the regression losses for COD, BOD_5 and *E. coli*, L_{CE} is the cross-entropy loss of the reuse-class head, and λ_1–λ_4 are the task weights.

E. Digital Twin and Treatment Control

The digital twin models the 2 KLD pilot plant as a chain of linked units: equalisation, aerobic bio-reactor, sand–activated carbon filter, UV-AOP chamber and nutrient-recovery vessel. Each unit model joins first-principle mass balances with data-driven correction terms learned from the plant's operating history. Every five minutes the twin takes the latest soft-sensor forecasts and simulates the next hour of treatment under a set of candidate settings. It then picks the aeration rate, filter backwash interval and UV dose that reach the target reuse class with the least energy. These settings go to the programmable controller through the edge node. If the twin predicts that no setting will meet the reuse limits, the flow is sent to the retreatment loop.

F. Explainability, Dashboard and Evaluation Protocol

For each prediction, Shapley additive explanations [23] are computed on the analytics server. They show which sensors and time lags drove the estimate, and a summary is attached to every control decision. Operators see node status, predicted quality, twin recommendations and these explanations on a web dashboard served from the storage and hosting server. We assessed the regression targets with the coefficient of determination, root mean square error and mean absolute error, and the reuse classification with accuracy, precision, recall and F1-score. Edge latency and power draw were measured on the Jetson devices. The baselines were LSTM, GRU, standalone Transformer and random forest models trained on the same splits.

4. Results and Discussion

We built and tested the framework on a testbed that combined a workstation with edge hardware. Model training, digital-twin simulation and explanation analysis ran on a GPU workstation with an Intel Core i9-13900K processor, 64 GB RAM, a single NVIDIA RTX-class GPU with 24 GB memory and 4 TB solid-state storage,

under Ubuntu 22.04 LTS. The models were written in Python with PyTorch; the baselines used scikit-learn, the explanations used SHAP, and TensorRT was used to optimise the models. Edge inference ran on four NVIDIA Jetson Orin Nano kits with JetPack, each in an enclosure with a 4G/LoRa module. The fifteen ESP32 nodes ran firmware written in the Arduino framework, and the dashboard was served from the network-attached storage and hosting unit.

The PJ-GWQ-2026 records were split by time, not at random, so that the test data always came after the training data. The first 70% of the timeline was used for training. The next 15% formed the validation set, which we used for early stopping, for choosing the anomaly threshold and for tuning hyperparameters. The last 15% was kept as an unseen test set. Scaling parameters were fitted on the training portion only, and windows that crossed a split boundary were dropped to avoid information leakage. Every baseline used the same splits, input windows and target definitions, so any difference in performance comes from the model and not from the data.

Table 1 compares GW-EdgeTwin with the four baselines on the test set. For COD, the proposed soft sensor reached an R^2 of 0.93 and an RMSE of 11.4 mg/L. The strongest baseline, the standalone Transformer, reached 0.88 and 14.7 mg/L, and the LSTM 0.86 and 16.2 mg/L. For BOD_5, the R^2 was 0.91 with an RMSE of 5.6 mg/L, and the *E. coli* error fell to 0.21 log CFU/100 mL. Reuse-class accuracy was 94.6% (F1-score 0.945), compared with 91.2% for the Transformer and 84.1% for the random forest. The largest gains were for COD and BOD_5. Both respond to sudden detergent pulses as well as to slower occupancy cycles: the dilated convolution branch captures the pulses, the attention branch captures the cycles, and the gated fusion balances the two. The recurrent baselines followed smooth changes well but lagged behind sharp pulses. The standalone Transformer reacted to pulses but tended to over-smooth the morning peak, and the random forest, which has no temporal memory, gave the widest error band. *E. coli* is still the hardest target, because microbial counts depend on factors such as temperature history and biofilm sloughing that the probes see only indirectly. Even so, the framework beat every baseline on this parameter, which matters for the toilet-flushing and irrigation classes, where pathogen limits decide compliance.

Table 1. Prediction performance of GW-EdgeTwin and baseline models on the test set

Model	COD R^2	COD RMSE (mg/L)	BOD_5 R^2	BOD_5 RMSE (mg/L)	*E. coli* MAE (log CFU/100 mL)	Reuse-class accuracy (%)	F1-score
Random forest	0.79	21.8	0.75	9.4	0.38	84.1	0.836
LSTM	0.86	16.2	0.83	7.8	0.30	89.3	0.890
GRU	0.85	16.9	0.82	8.1	0.31	88.7	0.884
Transformer	0.88	14.7	0.85	7.2	0.27	91.2	0.910
GW-EdgeTwin (proposed)	0.93	11.4	0.91	5.6	0.21	94.6	0.945

Figure 2 plots the soft-sensor COD estimates against the laboratory values for a 240-hour period of the test set. Over this window, the predicted curve stays close to the measured series, including the two high-load episodes near hours 55 and 167 and the shorter peaks near hours 88 and 195. Short spikes from kitchen and laundry discharges are picked up with only a small delay, and the estimates fall back to baseline at the same rate as the measured values. On days when occupancy dropped, the model adjusted to the lower load without systematic bias. This level of agreement suggests that the soft sensor could reduce how often routine COD grab samples are needed for operational decisions, with laboratory analysis kept for periodic checks. Close tracking at the peaks matters most, because these are the periods when fixed treatment schedules most often fail to meet reuse limits.

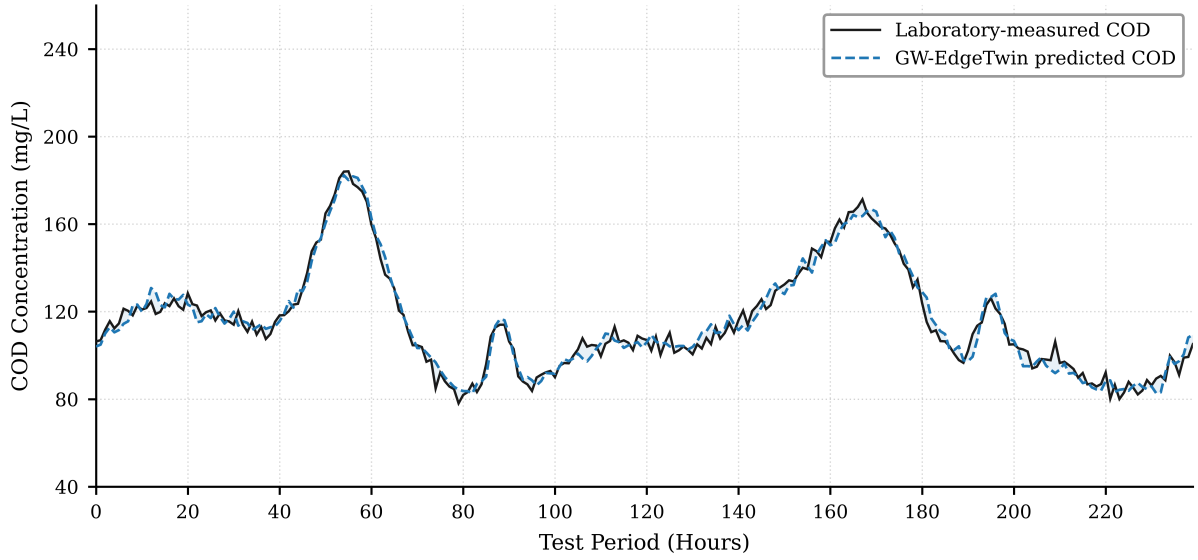

Figure 2. Soft-sensor COD estimates versus laboratory-measured COD over a 240-hour period of the test set

Table 2 summarises the edge deployment and closed-loop treatment results and compares GW-EdgeTwin control with the fixed operating schedule previously used at the pilot plant. After TensorRT optimisation, inference on a sixty-minute window took 38 ms on the Jetson Orin Nano at an average board power of 9.4 W, which keeps the edge layer within the energy budget of the treatment skid. The anomaly gate flagged faulty channels with a precision of 96.2% and a recall of 93.5%, so biased inputs did not reach the predictor. Under twin-guided control, the share of treated water meeting its target reuse class rose from 78.5% to 94.6%, and backwash water fell from 310 to 205 L/day. Energy use in the UV-AOP stage and across the plant was also lower than under the fixed schedule. These savings came from the twin lowering the UV dose and spacing out backwashes during low-load periods, and raising them only when the forecasts showed more organic or microbial load coming. Accurate prediction is therefore useful not just for monitoring: once it is linked to a process model, it also brings operational savings.

Table 2. Edge deployment and treatment-control performance

Metric	Fixed schedule	GW-EdgeTwin
Inference latency per window (ms)	–	38
Average edge board power (W)	–	9.4
Anomaly-gate precision (%)	–	96.2
Anomaly-gate recall (%)	–	93.5
Treated water meeting target reuse class (%)	78.5	94.6
Backwash water (L/day)	310	205

Figure 3 shows the mean absolute Shapley values of the input features for the COD prediction. Turbidity and its short-term gradient have the largest influence, followed by total dissolved solids and the hour-of-day feature, which stands in for occupancy. Flow rate matters most at peak hours, when dilution and loading change quickly, while temperature has a smaller but steady effect. This ranking fits what is known about the process: suspended and colloidal matter make up most of the organic load in campus greywater, and dissolved solids reflect the detergent and salt input. Because the ranking makes physical sense, operators have a reason to trust the predictions, and maintenance staff know to calibrate the turbidity and TDS probes first, as errors in these channels have the strongest effect on the output.

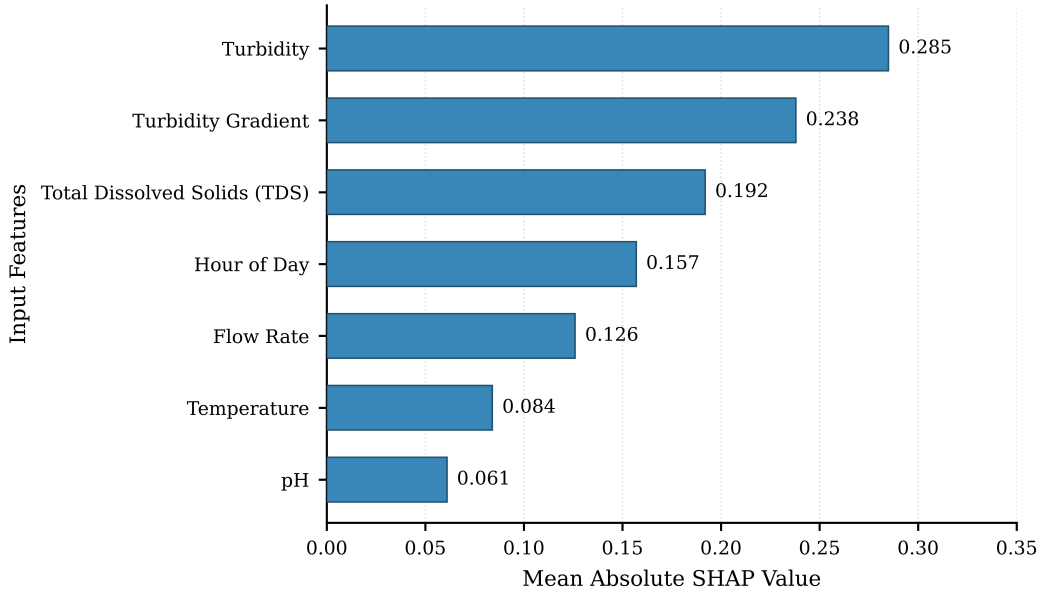

Figure 3. Feature contribution ranking from Shapley additive explanations

The operator dashboard of the implemented system combines a campus map with the live status of every sensor node, real-time and forecast curves for COD, BOD_5 and *E. coli*, the current reuse class of each treated-water tank, and the set-points that the digital twin recommends for aeration, backwash and UV dose. A side panel lists the anomaly alerts and maintenance tickets raised by the anomaly gate, and each control recommendation can be opened to show its explanation summary. During operation, staff used the dashboard to confirm the twin's decisions, override them when needed and export daily compliance reports for the institutional water committee.

The three parts of GW-EdgeTwin support each other. The anomaly gate protects the quality of the predictions, the hybrid soft sensor gives laboratory-grade estimates within the time limits of the edge hardware, and the digital twin turns those estimates into treatment actions that save energy and improve compliance. Every component runs on hardware bought within a modest project budget and uses low-cost probes, so the framework can be repeated in hostels, academic blocks and residential communities. The time-based evaluation also shows that the model keeps its performance on data recorded after the training period, which is essential for long-term use in the field.

5. Conclusion and Future Work

We have presented GW-EdgeTwin, an edge-intelligent framework for decentralised greywater recycling that combines low-cost IoT sensing, anomaly screening, a hybrid temporal convolution–Transformer soft sensor and a process digital twin. Instead of waiting for delayed laboratory results, the framework estimates organic and microbial load on site, continuously, and uses these estimates to adjust treatment in real time. Its main contributions are the cross-node anomaly gate, the gated fusion of short- and long-range temporal features, and the link between edge prediction and twin-based set-point selection. In practice, a campus can raise reuse compliance and reduce energy use and backwash-water demand with modest hardware. So far the system has been deployed at one campus pilot, and the next step is to validate it in other climates and building types. Future work will add federated learning across sites, extra probes for nutrients and emerging contaminants, and links to aquifer recharge planning.

Declarations

Conflict of interest. The authors declare that they have no conflict of interest.

Funding. This research was funded by Dhanalakshmi Srinivasan Institute of Research and Technology (DSIRT).

Data availability. The data that support the findings of this study are available from the authors on reasonable request.

Use of generative AI. Generative AI tools were used to a limited extent in preparing this manuscript. The authors reviewed and edited all such content and take full responsibility for the published article.

References

[1] Mustafa, M., Epelle, E. I., Macfarlane, A., Cusack, M., Burns, A., & Yaseen, M. (2025). Innovative approaches to greywater micropollutant removal: AI-driven solutions and future outlook. *RSC Advances, 15*(16), 12125. https://doi.org/10.1039/D5RA00489F

[2] Sujadi, H., Koswara, E., Gaffar, A. A., Subandi, D., Nur, T. I., Ain, A. N., & Badrujaman, J. (2026). A hybrid artificial intelligence pipeline for greywater classification and irrigation suitability assessment. *Ingénierie des Systèmes d'Information, 31*(1), 67–76. https://doi.org/10.18280/isi.310107

[3] Wang, A.-J., Li, H., He, Z., Tao, Y., Wang, H., Yang, M., Savic, D., Daigger, G. T., & Ren, N. (2024). Digital twins for wastewater treatment: A technical review. *Engineering, 36*, 21–35. https://doi.org/10.1016/j.eng.2024.04.012

[4] Shahra, E. Q., Wu, W., Basurra, S., & Aneiba, A. (2024). Intelligent edge-cloud framework for water quality monitoring in water distribution system. *Water, 16*(2), 196. https://doi.org/10.3390/w16020196

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

[6] Pires, L. M., & Gomes, J. (2024). River water quality monitoring using LoRa-based IoT. *Designs, 8*(6), 127. https://doi.org/10.3390/designs8060127

[7] Kondapuram, A. R., Treytl, A., Bigler, T., & Sauter, T. (2024). Design and evaluation of an IoT-based real-time underground water-level sensor system. In *2024 IEEE Applied Sensing Conference (APSCON)*. IEEE. https://doi.org/10.1109/APSCON60364.2024.10466069

[8] Radić, J., Brkić, M., Keser, T., Obrovski, B., Mihajlović, I., & Vasiljević Toskić, M. (2024). Distributed wireless IoT based sensing and quality monitoring system in protection of wetlands groundwater areas. *Measurement Science and Technology, 35*(12), 125110. https://doi.org/10.1088/1361-6501/ad774c

[9] El-Shafeiy, E., Alsabaan, M., Ibrahem, M. I., & Elwahsh, H. (2023). Real-time anomaly detection for water quality sensor monitoring based on multivariate deep learning technique. *Sensors, 23*(20), 8613. https://doi.org/10.3390/s23208613

[10] Bagheri, M., Bagheri, K., Farshforoush, N., Velazquez, A., & Liu, Y. (2024). A novel hybrid deep learning model for real-time monitoring of water pollution using sensor data. *Journal of Water Process Engineering, 68*, 106595. https://doi.org/10.1016/j.jwpe.2024.106595

[11] Recio-Colmenares, R., León Becerril, E., Gurubel Tun, K. J., & Conchas, R. F. (2023). Design of a soft sensor based on long short-term memory artificial neural network (LSTM) for wastewater treatment plants. *Sensors, 23*(22), 9236. https://doi.org/10.3390/s23229236

[12] Shyu, H.-Y., Castro, C. J., Bair, R. A., Lu, Q., & Yeh, D. H. (2023). Development of a soft sensor using machine learning algorithms for predicting the water quality of an onsite wastewater treatment system. *ACS Environmental Au, 3*(5), 308–318. https://doi.org/10.1021/acsenvironau.2c00072

[13] Voipan, D., Voipan, A. E., & Barbu, M. (2025). Evaluating machine learning-based soft sensors for effluent quality prediction in wastewater treatment under variable weather conditions. *Sensors, 25*(6), 1692. https://doi.org/10.3390/s25061692

[14] Yin, H., Chen, Y., Zhou, J., Xie, Y., Wei, Q., & Xu, Z. (2025). A probabilistic deep learning approach to enhance the prediction of wastewater treatment plant effluent quality under shocking load events. *Water Research X, 26*, 100291. https://doi.org/10.1016/j.wroa.2024.100291

[15] Li, R., Feng, K., An, T., Cheng, P., Wei, L., Zhao, Z., Xu, X., & Zhu, L. (2024). Enhanced insights into effluent prediction in wastewater treatment plants: Comprehensive deep learning model explanation based on SHAP. *ACS ES&T Water, 4*. https://doi.org/10.1021/acsestwater.4c00040

[16] Nasir, F. B., & Li, J. (2024). Comparative analysis of machine learning models and explainable artificial intelligence for predicting wastewater treatment plant variables. *Advances in Environmental and Engineering Research, 5*(4), 020. https://doi.org/10.21926/aeer.2404020

[17] Nasir, F. B., & Li, J. (2026). Application of explainable artificial intelligence and machine learning in predicting wastewater treatment plant variables: A comparative study of small- and large-scale treatment plants. *Environmental Science: Advances, 5*, 1174. https://doi.org/10.1039/D5VA00408J

[18] Liu, T., Zhang, H., Wu, J., Liu, W., & Fang, Y. (2024). Wastewater treatment process enhancement based on multi-objective optimization and interpretable machine learning. *Journal of Environmental Management, 364*, 121430. https://doi.org/10.1016/j.jenvman.2024.121430

[19] Ramadan, M. N. A., Ali, M. A. H., Khoo, S. Y., & Alkhedher, M. (2025). SecureIoT-FL: A federated learning framework for privacy-preserving real-time environmental monitoring in industrial IoT applications. *Alexandria Engineering Journal, 114*, 681–701. https://doi.org/10.1016/j.aej.2024.11.069

[20] Cerioli, A., Petrosino, L., Sasso, D., Laroche, C., Piechowiak, T., Pezzarossa, L., Merone, M., Vollero, L., & Sabatini, A. (2025). Efficient detection of microplastics on edge devices with tailored compiler for TinyML applications. *IEEE Access, 13*, 90970–90982. https://doi.org/10.1109/ACCESS.2025.3567816

[21] Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint*, arXiv:1803.01271. https://doi.org/10.48550/arXiv.1803.01271

[22] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*. https://doi.org/10.48550/arXiv.1706.03762

[23] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*. https://doi.org/10.48550/arXiv.1705.07874