Towards Water Systems Security and Sustainability Using Deep Learning

Chhayly Sreng[1*], Justice Lin[1*], Dong Sam Ha[1], Sook S. Ha[1], A. Lynn Abbott[1], Feras A. Batarseh[2†]

[1]Bradley Department of Electrical and Computer Engineering, Virginia Tech, Arlington, VA, USA
[2]Department of Biological Systems Engineering, Virginia Tech, Blacksburg, VA, USA
{chhaylysreng, justicelin5403, ha, sook, abbott, batarseh}@vt.edu

Abstract

Wastewater treatment plants (WWTPs) face significant challenges due to varying influent conditions, multiple operational constraints, and a constant lack of reliable datasets to manage and monitor water quality and flow using automated approaches. This paper introduces a novel framework showcasing soft sensors that are aimed at enhancing the sustainability and security of wastewater quality indicators using deep learning. We develop a trustworthy soft sensor that utilizes artificial intelligence (AI) approaches to provide nitrate NO3 predictions at the WWTP, as well as context-based evaluations to estimate overall predictive uncertainty. Contextual elements are injected into the model to allow for more accurate and relevant water quality monitoring, especially in different conditions (such as rain and snow). In addition, in this paper, we present a time-series Generative Adversarial Network (GAN), namely H2OGAN to address data scarcity and to improve model training by generating synthetic data that mirrors the statistical properties of water datasets from both controlled and real-world environments. Data in turn also train against data poisoning attacks on water supply systems, rendering these systems more secure. Our results indicate the potential uses of the integration of soft sensors and H2OGAN to significantly improve the operational efficiency of WWTPs by providing robust AI-driven tools for offering secure and sustainable water monitoring solutions.

Introduction

Today, about 25% of the global population (or more) lack access to clean water, 50% lacks access to sanitation services, and 30% lack access to hygiene facilities (Richards et al. 2023). In response to these issues, one of the tools applied to address these challenges is artificial intelligence (AI). AI is increasingly proposed to aid with evaluating water system deficiencies and other relevant issues that we present in this paper. According to a study by Alam et al., the water industry is actively investing in AI, with market research forecasting investments to reach an estimated $6.3 billion by 2030. Moreover, AI is expected to save 20% to 30% of operational expenditures by decreasing the cost and optimizing the usage of chemicals in water treatment (if applied correctly). These statistics indicate a promising future for AI in the water sector (Batarseh and Kulkarni 2023). The primary reason contributing to this optimism is that AI applications in water treatment have made the process easy due to their modest implementation, flexibility, generalization, and design simplicity (Alam et al. 2022).

Sensors in Water Systems

The use of sensors is constantly on the rise in many industrial and scientific applications, making accurate sensor measurements crucial. Instruments such as nitrate (NO_3) sensors are subject to environmental conditions, calibration drift uncertainties, are costly, difficult to set up, and unreliable in many cases(Dewhurst and Tian 2008). Researchers have turned to advanced computational methods, including mathematical modeling, statistical analysis, and machine learning, to overcome these limitations (Sun and Ge 2021). Deep learning techniques (presented in the next section) have shown promise in outperforming traditional methods in many applications by achieving higher accuracy, but they are often criticized as 'black-box' models due to their lack of transparency (Batarseh and Freeman 2022). Over time, these regular sensors can get less accurate due to aging, while in contrast, data-driven models such as soft sensors get better over time due to accumulating data, consequently leading to better accuracy. In this paper, we present a novel soft sensor for measuring nitrate in water systems. The current state of the art provides evidence that AI is the leading approach to such defenses due to its ability to adequately identify unwarranted pattern shifts in networks and datasets, a feature that is not achievable using traditional approaches (Batarseh and Kulkarni 2023). Deep learning models, such as Artificial Neural Networks (ANNs), are indeed known for requiring large amounts of data to train effectively. This data-hungry nature can pose challenges, particularly in domains where collecting data is difficult or expensive, or when dealing with rare events. Several research papers and works address issues related to data availability, offering insight, methodologies, and innovations to mitigate these challenges. For instance, Elbasi et al. discusses the application of AI in agriculture, which directly impacts water management, such as via irrigation. They highlight the use of sensors and soil sampling for data collection, which is crucial for managing water resources in agriculture. The paper emphasizes the role of

[*]These authors contributed equally.
[†]Corresponding author.

AI in improving farmers' profitability and the overall economy, which is closely tied to efficient water use (Elbasi et al. 2023). Additionally, Mueller et al. identifies significant gaps in datasets for water assessment, particularly in the measurement and reporting of geographic water shortfalls. They underscore the need for comprehensive datasets to enable effective water management decisions in businesses (Mueller et al. 2015). Suchetana et al. introduces AI techniques on water data to promote sustainable usage. They discuss AI's insights for both short-term and long-term water policy decisions, highlighting the potential and challenges of using AI in water management (Suchetana et al. 2023). Li et al. details AI's role in optimizing drinking water treatment processes. They discuss AI's potential in water quality diagnosis, decision-making, and operation process optimization, although they also note challenges in data availability and quality (Li et al. 2021).

Synthetic Data Generation for Water Systems

The GANs framework, introduced by Goodfellow et al. in 2014, represents an approach for estimating generative models to synthesize data. The core idea involves training two models simultaneously: a generative model (*G*) and a discriminative model (*D*). *G* is trained to capture the data distribution, while *D* is trained to distinguish between actual data and the data generated by *G*. This setup forms a minimax two-player game, where *G* tries to maximize the probability of *D* making a mistake. GANs are unique in that they do not require Markov chains or unrolled approximate inference networks during training or generation of samples, and they can be trained using backpropagation. The effectiveness of GANs is demonstrated through both qualitative and quantitative evaluations of the generated samples. GAN models are necessary for many DL applications, such as security, data augmentation, and privacy preservation (Wang et al. 2017). They work by understanding and replicating data distributions, generating new data based on learned parameters. These models, when applied to various benchmark datasets, show high variance in outputs but remain competitive with other generative models without needing manual intervention during learning (Goodfellow et al. 2014). GANs have also been utilized to enhance intrusion detection systems against CPS attacks by generating synthetic samples (Shahriar et al. 2020). While GANs have shown promising results, their application in water systems is still a relatively new field, and further research is needed to fully realize their potential. The methods presented in this paper aim to fill that gap.

Our Contribution

Addressing these challenges, the research framework developed and presented here consists of two main components: a trustworthy soft sensor (for measuring and evaluating water-related values - i.e., a sustainability goal); and H_2OGAN (for synthetic data generation - i.e., a data security goal against anomalies in water parameters), both presented in the next section. This paper is structured as follows: the next section presents our methodology and testing in detail, section 3 discusses the experimental results, and the final section provides discussions and conclusions.

Methodology

Each of the two components presented in this paper (soft sensor and H_2OGAN) is designed to enhance water management at wastewater treatment plants (WWTPs) data security and overall sustainability. Central to this framework is the quality of data and the selection of appropriate input variables, crucial for the effective operation of soft sensors in estimating water quality variables. It is achieved by estimating predictive uncertainty and conducting context-based evaluations for NO_3, taking into account various factors such as water conditions, weather factors, and anomaly events. Challenges such as poor data quality and difficulties in selecting the right variables can significantly impair sensor accuracy, highlighting the need for a robust soft sensing approach capable of handling the variability and complexity of data typical in WWTP environments. This component employs deep learning to enable non-AI experts to estimate predictive uncertainty and perform context-based evaluations for variables like NO_3, taking into account varying water conditions, weather, and anomaly events. In parallel, H_2OGAN, a time-series GAN-based model, enhances the security and efficiency of water systems by supporting functions such as data augmentation, anomaly detection, risk assessment, and predictive model optimization. It generates and analyzes realistic water data within the expected constraints of water parameter characteristics, thereby providing enriched data inputs and refined predictive capabilities essential for advanced water system monitoring and decision-making. H_2OGAN builds upon the principles and architecture of TimeGAN (Yoon, Jarrett, and van der Schaar 2019), drawing motivation from the time-series characteristics of water data as discussed in existing literature (Teramoto, Crioni, and Chang 2021; Zanotti et al. 2023), and from experimental and real-world environment. This inheritance enables a robust approach for understanding and synthesizing the complex dynamics in sequential data. Notably, H_2OGAN not only adopts the TimeGAN framework, but also extends it.

Data Used in the Study

In this work, we utilize three datasets: the AI & Cyber for Water & Ag (ACWA) lab dataset, the Alexandria Renew Enterprises (AlexRenew) dataset, and the National Oceanic and Atmospheric Administration (NOAA) dataset (for weather and contextual data). The ACWA lab dataset (Batarseh et al. 2023a), a water testbed with computational resources, sensors, and water systems, simulates real-world conditions to address data quality and availability issues in AI water system research. We also present the AlexRenew dataset, reflecting external factors such as weather changes and events. Both datasets are used in the experiments (see Figure 1).

Trustworthy Soft Sensor

Researchers have explored evaluating their models across various timelines or categories (Yan et al. 2020; Li et al.

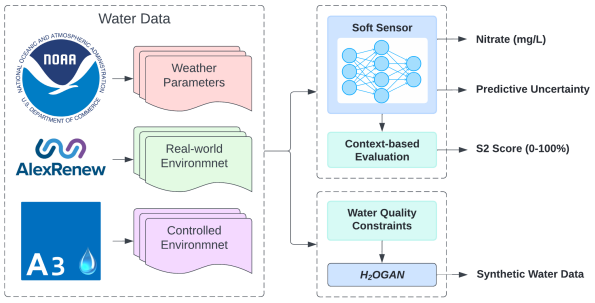

Figure 1: The proposed pipeline

2024), yet it remains a challenge due to models' context-specific performance. Understanding this enables us to leverage models more effectively by concentrating on their strengths in specific contexts.

In the biological processes of the treatment plants, NO_3 levels are largely influenced by water quality attributes such as pH, DO, NH_3, and temperature (Wilczak et al. 1996; Nitrification 2000). In addition, due to physical limitations in collecting NO_3 levels, they can be influenced by external factors such as weather and extreme events, including overflow, flooding, rain, and snow, making accurate prediction and monitoring challenging.

In this study, we develop a scoring system that incorporates historical data to assess the likelihood that the soft sensor performs well in sensor readings based on context. The experiment is divided into two parts: initially, it focuses on a controlled environment dataset (i.e., ACWA), influenced solely by water quality parameters and flow dynamics. Subsequently, the experiment extends to real-world water system dataset (i.e., AlexRenew), incorporating additional factors such as weather changes and extreme events such as rain, snow and flooding, allowing a thorough evaluation of the soft sensor under varied conditions.

Data Preprocessing and Model Development The development stage includes four primary steps: (1) data cleaning, imputation, and normalization, (2) neural network model development, and (3) hyperparamater tuning via grid search, and (4) model testing. For instance, in the first phase, we used data imputation, specifically through linear interpolation, to address gaps in the dataset, a common issue in time-series data in water systems (Gnauck 2004). Despite its simplicity, linear interpolation is favored for its effectiveness in bridging these gaps, thereby maintaining the continuity and integrity of the dataset. To ensure uniformity in the data range and distribution, the standardization scaling technique is applied to datasets. The performance of a deep learning model (developed in Python) proved to be highly dependent on the selection of hyperparameters (Yu and Zhu 2020). As noted by Yu and Zhu (2020), random search is highly effective in most cases when compared to the other algorithms such as grid search (Liashchynskyi and Liashchynskyi 2019). This technique involves defining a hyperparameter space, randomly sampling from it, training and evaluating models with different hyperparameter combinations, and selecting the best-performing model.

Evaluation Metrics In our approach, we use a predictive uncertainty estimation method as follows: an input sample x^* is fed into multiple neural network ensembles; each ensemble produces a prediction y_i^* for the input sample. These individual predictions are then aggregated to obtain a final prediction y^* by calculating the mean of the predictions. Additionally, the variance σ^* of the predictions is computed to provide a measure of model uncertainty. This method ensures that both the prediction and the associated uncertainty are accounted for, enhancing the robustness and reliability of the model's output. We use the following evaluation metrics in this study (limits and ranges shown in Table 1): Root Mean Square Error (RMSE), Nash-Sutcliffe Efficiency (NSE), and Relative Standard Error (RSR).

Context Clustering Experimentation This section describes the approach used to create context classes based on each of the water quality parameters such as pH, DO, water temperature, air temperature, and NH_3 and so on, by using K-means clustering (Likas, Vlassis, and J. Verbeek 2003). The objective is to segment water quality data into distinct groups that reflect similar water quality characteristics. This segmentation helps to identify patterns or anomalies within the data, enabling context-specific evaluation.

S2 Score Methodology The evaluation of NO_3 soft sensing (S2) is conducted using a composite score derived from multiple metrics. According to the suggested guidelines in Table 1 for evaluating the performance of hydrological modeling, each metric is assigned a point. Each method is then scored as meeting criteria (1) or not (0).

Sat. Rating	References
NSE>0.5	Moriasi et al. (2007); Duda et al. (2012)
MAE<20	Shyu et al. (2023)
MAPE<25%	Shyu et al. (2023)
PBIAS<25%	Moriasi et al. (2007); Duda et al. (2012)
RSR<0.6	Moriasi et al. (2007); Duda et al. (2012)

Table 1: Guidelines for evaluating the performance of hydrological modelling

Let's denote the S2 score as $S2$, and let $x_1, x_2, x_3, ..., x_n$ represent the different parameters, which could include pH, DO, temperature, NH_3, precipitation, events whether it is operating under flooding, overflowing, raining, or snowing. The normalized score S that ranges from 0 to 100% (or equivalently, 0 to 1 for a proportion) can be calculated using the following formula:

$$S2 = \frac{\sum_{i=1}^{n} w_i \cdot f_i(x_i)}{\sum_{i=1}^{n} w_i M_i} \quad (1)$$

Where: w_i are the weights assigned to each parameter x_i, indicating its importance in the overall score. These weights allow for the flexibility to prioritize certain environmental factors over others based on the context or objectives of the evaluation. For simplicity, we will assume that all $w_i = 1$

which means all context are equally important. $f_i(x_i)$, is the number of criteria met by x_i in a particular context class, contributing into a score to the overall score. For each parameter x_i, the function f_i evaluates the following conditions in Table 1 and sums the number of true statements. M_i is the maximum possible score or value that $f_i(x_i)$ can yield.

Scoring is based on a testing set from historical records under diverse contextual scenarios Lin et al. (2023). This approach provides a holistic view for water utility operators and AI practitioners.

H_2OGAN

Utilizing boundary-based constraints, H_2OGAN creates three distinct y-axis regions based on specific contexts for each parameter, defined by four adjustable boundaries. The middle region represents the ideal range for each water parameter, while the low and high regions are considered acceptable yet cautionary ranges. This segmentation of data points across different regions ensures that the data are appropriately scaled for H_2OGAN. The model is designed to exploit data points in the low and high regions, which are the anomalies within the boundary constraints, optimizing the synthesizing of those data points. The model also uses a modified generator loss function, which includes an additional calculation to ensure similar anomaly ratios between the generated and actual data.

Boundary-Based Constraints The flexibility of boundary setting in H_2OGAN is a crucial aspect that allows it to adapt to different water contexts (as shown in Table 2). Its ability to adjust these boundaries is particularly useful when dealing with datasets from different sources that have distinct characteristics and requirements.

Parameter	Inner Bounds	Outer Bounds
ACWA_DO	[6.5, 8]	[5, 14]
ACWA_TDS	[150, 300]	[0, 600]
ACWA_NO_3	[0, 10]	[0, 40]
ACWA_pH	[6.5, 8.5]	[5.5, 9]
ACWA_Temp	[20, 22]	[18, 25]
AlexRenew_DO	[0.3, 1.8]	[0.0, 5.2]
AlexRenew_NH_3	[112.8, 185.6]	[0.0, 1036.4]
AlexRenew_NO_3	[114.4, 168.4]	[0.0, 1036.4]
AlexRenew_pH	[6.8, 7.1]	[0.0, 10.2]
AlexRenew_Temp	[79.6, 88.4]	[43.6, 102.7]

Table 2: Outer Boundaries for each water parameter

For ACWA, the outer and inner boundaries are set based on guidelines from literature (Patel and Vashi 2015; Summers 2020) and from the U.S. Environmental Protection Agency (EPA) (Environmental Protection Agency (EPA) 2023). In the case of AlexRenew however, the determination of outer boundaries is derived from the minimum and maximum measurements across all data points for a given water parameter (given its a real-world dataset). The inner boundaries, on the other hand, are computed based on the average of all minimum and maximum measurements. Using the DO data from AlexRenew as an example, as illustrated in Figure 2, the inner boundaries are indicated by the dashed green lines, while the outer boundaries are indicated by the dashed red lines, collectively defining the low, middle, and high regions of the data range. These constraints capture hydrological concepts and deem the proposed GAN water-specific.

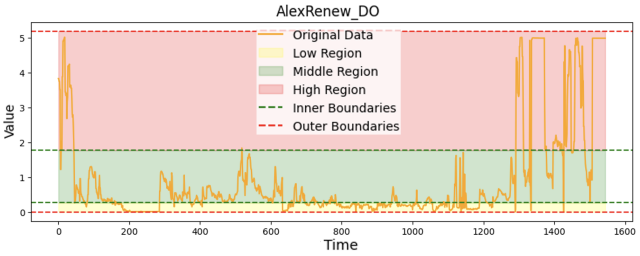

Figure 2: Demonstration of boundaries and regions using AlexRenew's DO variable

The noise vectors (created for each water parameter) are stacked together to form the complete set of noise vector sequences for a batch - which is a part of the data generation process. This is represented as $noise^{D\times B}$, where D is the dimension of a noise vector. Specifically, D is the product of the timestamp length and the number of water parameters.

***H_2OGAN*-Specific Loss Calculation** The generator loss is adjusted with an extra computation to maintain similarity in the anomaly ratios between the synthetic and real data. This extra computation guarantees that the synthetic data preserves the distribution of the real data, implying that the proportions of data points in the low, middle, and high regions stay consistent. However, data points in the low and high regions are exploited and can appear anywhere within their respective regions. The objective is to optimize the use of data points in the low and high regions as much as possible, while avoiding detection by the discriminator. In order to compare the distribution between the original and synthetic sets, each $noise_{i,j}$ in $noise^{D\times B}$ is utilized. This is for determining $\hat{P}_{L,i,j}$ and $\hat{P}_{H,i,j}$, which are shown in Equation 2 and 3.

$$F_{L,i}(a) = \begin{cases} 1, & \text{if } a < \min_{\text{inner},i} \\ 0, & \text{otherwise} \end{cases}$$
$$\hat{P}_{L,i,j} = \frac{1}{B}\sum_{k=1}^{B} F_{L,i}(noise_{i,j,k}) \tag{2}$$

$$F_{H,i}(a) = \begin{cases} 1, & \text{if } a > \max_{\text{inner},i} \\ 0, & \text{otherwise} \end{cases}$$
$$\hat{P}_{H,i,j} = \frac{1}{B}\sum_{k=1}^{B} F_{H,i}(noise_{i,j,k}) \tag{3}$$

As shown in Equation 4, the sum of all absolute differences between $P_{L,i,j}$ as well as between $\hat{P}_{L,i,j}$, $P_{H,i,j}$, and

$\hat{P}_{H,i,j}$ is for every possible pair of the i-th water parameter and for the j-th sequence column. This loss is then incorporated as a parameter in the computation of the generator loss.

$$anomaly\_dist\_loss = \frac{1}{D_1 \times D_2} \sum_{j=1}^{D_1} \sum_{i=1}^{D_2} \Big(|P_{L,i,j} - \hat{P}_{L,i,j}| + |P_{H,i,j} - \hat{P}_{H,i,j}| \Big) \tag{4}$$

The $anomaly\_dist\_loss$ ensures that the synthetic data maintains a similar anomaly ratio to the real data. This is particularly important as it allows the model to exploit data points in the low and high regions of the distribution, optimizing their use while avoiding detection by the discriminator. This balance between exploitation and detection evasion is what allows the H_2OGAN model to generate synthetic data that preserves the distribution of the real data.

Results and Discussions

The results of hyperparameters' tuning in the NO_3 soft sensor, for two models (ACWA and AlexRenew) are as follows: a learning rate of 0.001; The ACWA model has 14 hidden layers with the following distribution of neurons across the layers: 66, 672, 800, 32, 672 in the first five layers, 224, 160, 160, 32, 320 in layers six to ten, and 512, 576, 928, 512 in layers eleven to fourteen. The AlexRenew model has 18 hidden layers, with 770, 672, 672, 928, 608 in the first five layers, 704, 416, 160, 224, 672 in the next five layers, and 192, 992, 32, 32 in the subsequent four layers, finishing with 32, 32, 32, 32 in the last four layers. All layers in both models utilize the ReLU activation function. The differences in the best configurations of the ACWA lab (Batarseh et al. 2023b) and AlexRenew datasets are due to the varying network capacities and complexity requirements of the models. Both models are trained for approximately 200 to 300 epochs, with batch sizes of 64 and 128, respectively, to optimize performance based on the mean square error (MSE) criterion. According to the results, the soft sensor performs very well with the ACWA dataset, reflected by an NSE of 0.948 and a low MAE of 3.183. The unexpectedly high MAPE of 450.323 with a low MAE may suggest the presence of skewed data, particularly if the actual values in the dataset are very small or close to zero. In this case, even minor absolute errors can produce large percentage errors. For the AlexRenew dataset, the sensor shows good but lesser performance with an NSE of 0.719 and a much higher MAE of 17.686. MAPE for AlexRenew is significantly lower than in ACWA's, i.e., at 0.537; indicating more consistent percentage errors. RSR is acceptable for both datasets, with ACWA at 0.22 and AlexRenew at 0.529, implying satisfactory performance. The hyperparameter configurations may contribute to the model's tendency to slightly underestimate NO_3 levels in both datasets. A more pronounced underestimation is observed in the AlexRenew dataset, as indicated by a PBIAS of -1.938, in contrast to the -0.911 recorded for ACWA.

Context-Based Evaluations This section presents evaluations of different contexts or ranges, with performance metrics in the ACWA dataset and AlexRenew datasets. For DO, the NSE is satisfied at 0.709 and 0.963 for a range of 10.74 to 338.36 and 5.46 to 10.72, suggesting that the model performs well. The MAE and MAPE indicate that the model's output can be quite off, with the lowest error at a higher DO range. We first begin with a summary of results for ACWA data. In ACWA, for EC, the model shows high performance in one context (0.0 to 360.26) with an NSE of 0.878 and low MAE/MAPE, but it's less accurate at higher ranges of conductivity, indicated by a lower NSE of 0.858. For pH, the model has a high NSE across all pH ranges, indicating accurate predictions, with the highest accuracy in the range of 5.73 to 6.48. The negative NSE in the range of 6.49 to 7.11 suggests the model predictions are worse than an average-based prediction. For air temperature, the model's performance is poor in the range of 21.28 to 22.43, as indicated by negative NSE values. Conversely, in the range of 22.51 to 23.77, despite the high MAPE, the MAE remains low, and the NSE exceeds the satisfactory benchmark. These results suggest that the model can capture the dynamics of the system; however, the high MAPE with the low MAE is likely due to actual NO_3 values being close to zero. The model tends to overestimate NO_3 levels (positive PBIAS) except at the highest EC range. The RSR values are generally low, suggesting that the model errors are modest relative to the variability of the observed data.

We plot the physical sensor against the soft sensor with predictive uncertainty using our model and present the S2 score, as shown in Figure 3. Both sensors' readings show NO_3 levels fluctuating over time with some periods of increased variability and spikes. These could represent natural variations in NO_3 concentration or response to specific environmental events. Additionally, the results for the AlexRenew data are also captured, they are presented next. For DO (in AlexRenew data), the soft sensor's accuracy is moderate-to-low in predicting NO_3 levels based on DO. It has an acceptable NSE of 0.771 for the lowest range but drops significantly for the mid-range (1.65 to 5.01), and slightly improves for higher DO levels. For ammonia NH_3, the NSE indicates that the sensor performs poorly for lower and higher ranges, with the best performance at a mid-range (134.07 to 384.32) with an NSE of 0.815. For temperature, the model performs poorly at predicting NO_3 levels for the lower temperature range (60.85 to 79.2), as suggested by the negative NSE. The performance improves significantly in higher temperature ranges, indicated by NSE values of 0.796 and 0.768 for the mid and higher temperature ranges, respectively. For pH, the sensor shows good accuracy in a lower pH range (3.71 to 7.17) with an NSE of 0.762, but a negative NSE (-0.215) for the higher pH range (7.18 to 10.0), indicating poor model predictions in more alkaline conditions. In the case of air temperature, the sensor's accuracy improves with higher temperatures, with moderate performance at lower temperatures (NSE of 0.413) and better performance at higher temperatures (NSE of 0.782). Prediction error, as indicated by MAE and MAPE, decreases in the higher temperature range, and PBIAS indicates a trend from overestimation to slight

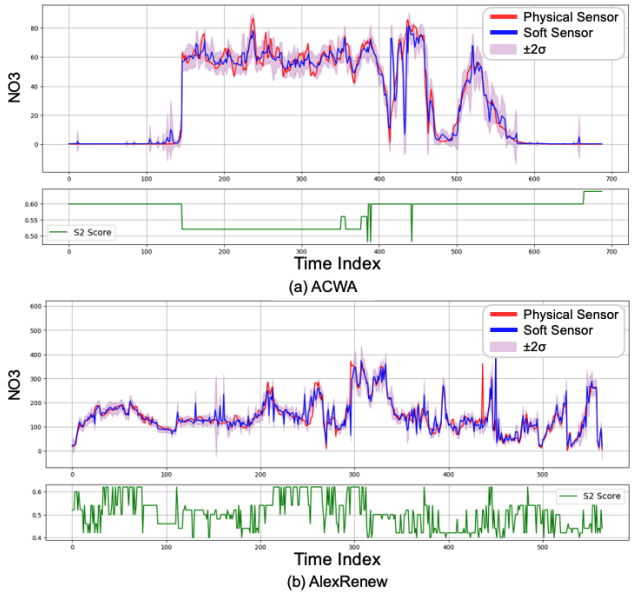

Figure 3: Plot of physical sensor vs soft sensor for NO_3

underestimation as temperature increases. In normal conditions, the sensor has an NSE of 0.66, indicating a good fit to the observed data. The MAE and MAPE are relatively low, suggesting acceptable accuracy in normal weather. PBIAS is slightly negative, showing a minor tendency for underestimation. The soft sensor performs best under overflowing conditions with an NSE of 0.902, which denotes a high level of accuracy. The MAE score is the lowest among all events; but the MAPE score is also low, which suggests accurate estimations during overflowing. However, the PBIAS is negative (-8.42), indicating significant underestimation of NO_3 levels during overflowing. During rainfall, the model's performance drops (NSE of 0.483), with higher MAE and MAPE values compared to normal and flooding conditions. The PBIAS is very positive, indicating a substantial overestimation of NO_3 levels. For snowy conditions, the model has a relatively high NSE of 0.673, similar to normal conditions. MAE and MAPE are modest, compared to all events and in the guidelines Table 1, and the PBIAS is close to zero, indicating minimal bias.

Validation and Visualization For validation, we plot the physical sensor against the soft sensor with predictive uncertainty using our model and present the S2 score, as shown in Figure 3. It seems that the soft sensor line closely follows the physical sensor line, providing a good estimate of the NO_3 levels. The confidence interval ($\pm 2\sigma$) shown in purple around the soft sensor readings is crucial to understanding the uncertainty in the soft sensor estimates. The width of the confidence interval indicates the level of confidence we can have in the soft sensor's predictions.

H_2OGAN Outcomes

We evaluated the effectiveness of the synthetic data generation process by comparing the original and synthetic datasets for ACWA (as shown in Table 3) and AlexRenew (as shown in Table 4). The results confirm that the synthetic data accurately reflects the statistical characteristics of the original data within predetermined boundaries. Detailed comparative analyses, including visualizations and summary statistics, are presented in Table 3 and shown in Figure 4.

(a) ACWA_DO

	mean	std	Q1	Q2	Q3
Orig	8.61	2.12	6.29	9.42	10.60
Synth	8.07	2.18	6.23	6.88	10.34

(b) ACWA_TDS

	mean	std	Q1	Q2	Q3
Orig	281.41	247.14	14.56	191.66	579.67
Synth	282.27	248.68	34.34	191.34	581.72

(c) ACWA_NO_3

	mean	std	Q1	Q2	Q3
Orig	14.68	10.67	3.13	17.32	23.84
Synth	13.89	10.76	3.09	13.62	22.60

(d) ACWA_pH

	mean	std	Q1	Q2	Q3
Orig	6.22	0.49	5.88	6.02	6.69
Synth	6.24	0.46	5.92	6.04	6.49

(e) ACWA_Temp

	mean	std	Q1	Q2	Q3
Orig	22.21	0.99	21.97	22.07	22.30
Synth	22.22	0.72	21.94	22.11	22.46

Table 3: Summary statistics comparison - ACWA

In the ACWA experiment, the synthetic data for DO presents a slightly lower mean (8.073) than the original data (8.61), indicating a tendency towards conservative estimates in synthetic generation. The associated time series plots reveal a smoothing of extremes by the synthetic data. TDS shows a similar mean for both the original (281.414) and synthetic data (282.265), highlighting the effectiveness of the synthetic modeling process. NO_3 concentration in the synthetic data is modestly diminished in both mean (13.888) and standard deviation, reflecting the intention to minimize extreme values and variability compared to the original mean of 14.676. The pH measurements demonstrate exceptional consistency between datasets, with the synthetic data replicating the mean (6.236) and quartile values with high precision, closely matching the original mean of 6.224. Lastly, the synthetic water temperature closely mirrors the mean of the original data (22.206) but with reduced variability (standard deviation of 0.721), portraying a more consistent synthetic data profile. For AlexRenew, synthetic data indicates lower average DO levels (0.618) compared to the original (0.874), pointing to a tendency for fewer fluctuations and a more balanced representation of high and low values. For NH_3 and NO_3, the synthetic data shows lower average concentrations (118.928 and 118.423 respectively) and a narrower range of values, which suggests the model does not replicate the wider variety of unusual or extreme values found in the original data (138.397 and 129.718 respectively). The pH levels match closely between

(f) AlexRenew_DO

	mean	std	Q1	Q2	Q3
Orig	0.87	1.30	0.19	0.36	0.85
Synth	0.62	1.02	0.20	0.42	0.59

(g) AlexRenew_NH_3

	mean	std	Q1	Q2	Q3
Orig	138.40	66.60	92.58	135.81	181.15
Synth	118.93	49.72	80.42	116.67	159.69

(h) AlexRenew_NO_3

	mean	std	Q1	Q2	Q3
Orig	129.72	82.44	82.43	122.71	157.80
Synth	118.42	56.99	83.50	121.15	148.23

(i) AlexRenew_pH

	mean	std	Q1	Q2	Q3
Orig	6.93	0.47	6.76	6.79	7.08
Synth	6.95	0.46	6.72	6.91	7.20

(j) AlexRenew_Temp

	mean	std	Q1	Q2	Q3
Orig	87.60	7.98	84.50	88.59	94.40
Synth	88.86	6.73	84.70	89.58	94.57

Table 4: Summary statistics comparison - AlexRenew

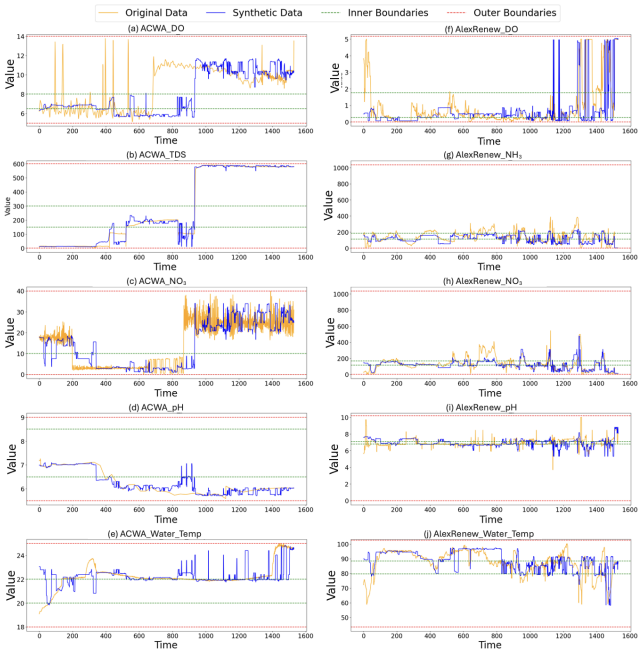

Figure 4: Water parameters: original vs synthetic

the synthetic and original data, showing the precision of the model in mirroring the acidity or alkalinity of water (synthetic mean of 6.945 as compared to the original mean of 6.93). Interestingly, H_2OGAN appears to capture fluctuations across all five parameters effectively, demonstrating its capability to handle real-world data complexities.

Conclusions

The context-based evaluation methodology and scoring developed in this study offer a novel framework for assessing soft sensor trustworthiness. By segmenting the dataset into different contexts—ranging from water quality parameters to weather conditions and anomaly events—the study quantifies the sensor's performance in each scenario, providing different results based on context. The S2 score, crafted from the context-based evaluation, plays an important role in translating complex performance metrics into understandable and actionable insights. By assigning scores to different contexts, the study illustrates the sensor's strengths and areas for improvement. The S2 scoring system is rooted in the assessment against established hydrological modeling guidelines (NSE > 0.5, MAE $< 20\%$, MAPE $< 25\%$, PBIAS $< 25\%$, RSR < 0.6), as shown in Table 1, and offers a clear and concise method for evaluating soft sensor performance across a spectrum of real-world conditions. Addressing the potential security threats, the entities an attacker may compromise include the sensor nodes, PLCs, and the communication channels within the water monitoring system. The attacker could alter the sensor measurements by injecting noise or false data, leading to a degree of change that might range from minor fluctuations to significant deviations, depending on the attack's sophistication. Whether the attack is stealthy or detectable depends on the anomaly detection methods in place. Additionally, the deployment of H_2OGAN highlights its impact on improving the availability, quality, and poisoning of water data. Through these applications, H_2OGAN has demonstrated its capability to generate synthetic data that can seamlessly blend with real data, even when sophisticated security anomalies are included. The performances of H_2OGAN with both ACWA and AlexRenew data further emphasize its effectiveness in creating data that reflects the environmental context. This synthetic water data can be useful in training robust and secure AI solutions intended for prediction, protection, optimization, and various other applications. H_2OGAN offers system robustness via more data, and aids in developing detection methods for anomaly and security detection. This is achieved by capturing complex patterns in sequential data representing changes in water parameters over time. These combined advantages position H_2OGAN as a promising approach for improving the efficiency and data security of water supply systems.

References

Alam, G.; Ihsanullah, I.; Naushad, M.; and Sillanpää, M. 2022. Applications of artificial intelligence in water treatment for optimization and automation of adsorption processes: Recent advances and prospects. *Chemical Engineering Journal*, 427: 130011.

Batarseh, F. A.; and Freeman, L. 2022. *AI Assurance: Towards Trustworthy, Explainable, Safe, and Ethical AI*. Academic Press.

Batarseh, F. A.; and Kulkarni, A. 2023. AI for Water. *Computer*, 56(3): 109–113.

Batarseh, F. A.; Kulkarni, A.; Sreng, C.; Lin, J.; and Maksud, S. 2023a. ACWA: an AI-driven cyber-physical testbed

for intelligent water systems. *Water Practice & Technology*, 18(12): 3399–3418. Publisher: IWA Publishing.

Batarseh, F. A.; Kulkarni, A.; Sreng, C.; Lin, J.; and Maksud, S. 2023b. ACWA: an AI-driven cyber-physical testbed for intelligent water systems. *Water Practice and Technology*, 18(12): 3399–3418.

Dewhurst, R.; and Tian, G. 2008. Sensors and sensing systems. *Measurement Science and Technology*, 19: 020101.

Duda, P. B.; Hummel, P. R.; Donigian Jr, A. S.; and Imhoff, J. C. 2012. BASINS/HSPF: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4): 1523–1547. Publisher: American Society of Agricultural and Biological Engineers.

Elbasi, E.; Mostafa, N.; AlArnaout, Z.; Zreikat, A.; Cina, E.; Varghese, G.; Shdefat, A.; Topcu, A.; Abdelbaki, W.; Mathew, S.; and Zaki, C. 2023. Artificial Intelligence Technology in the Agricultural Sector: A Systematic Literature Review. *IEEE Access*, 11: 171–202.

Environmental Protection Agency (EPA). 2023. Estimated Nitrate Concentrations in Groundwater Used for Drinking. Retrieved April 4, 2024.

Gnauck, A. 2004. Interpolation and approximation of water quality time series and process identification. *Analytical and Bioanalytical Chemistry*, 380(3): 484–492.

Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative Adversarial Nets. *Advances in Neural Information Processing Systems 27*, 26722680.

Li, D.; Yang, C.; Li, Y.; Zhou, C.; Huang, D.; and Liu, Y. 2024. A deep semi-supervised learning framework towards multi-output soft sensors development and applications in wastewater treatment processes. *Journal of Water Process Engineering*, 57: 104654.

Li, L.; Rong, S.; Wang, R.; and Yu, S. 2021. Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: A review. *Chemical Engineering Journal*, 405: 126673.

Liashchynskyi, P.; and Liashchynskyi, P. 2019. Grid search, random search, genetic algorithm: a big comparison for NAS. *arXiv preprint arXiv:1912.06059*.

Likas, A.; Vlassis, N.; and J. Verbeek, J. 2003. The global k-means clustering algorithm. *Pattern Recognition*, 36(2): 451–461.

Lin, J.; Sreng, C.; Oare, E.; and Batarseh, F. A. 2023. NeuralFlood: an AI-driven flood susceptibility index. *Frontiers in Water*, 5: 1291305.

Moriasi, D. N.; Arnold, J. G.; Van Liew, M. W.; Bingner, R. L.; Harmel, R. D.; and Veith, T. L. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3): 885–900. Publisher: American society of agricultural and biological engineers.

Mueller, S. A.; Carlile, A.; Bras, B.; Niemann, T.; Rokosz, S. M.; McKenzie, H. L.; Kim, H.; and Wallington, T. 2015. Requirements for water assessment tools: An automotive industry perspective. *Water Resources and Industry*, 9: 30–44.

Nitrification, T. F. 2000. Technology Fact Sheet.

Patel, H.; and Vashi, R. 2015. Acknowledgment. In *Characterization and Treatment of Textile Wastewater*.

Richards, C. E.; Tzachor, A.; Avin, S.; and Fenner, R. 2023. Rewards, risks and responsible deployment of artificial intelligence in water systems. *Nature Water*, 1(5): 422–432.

Shahriar, M. H.; Haque, N. I.; Rahman, M. A.; and Alonso, M. 2020. G-ids: Generative Adversarial Networks assisted Intrusion Detection System. *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, 376–385.

Shyu, H.-Y.; Castro, C. J.; Bair, R. A.; Lu, Q.; and 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.

Suchetana, B.; Srivastava, B.; Gupta, H. P.; and Saharia, M. 2023. Promoting Sustainable Water Usage and Management With Water Data, AI and Policy. *Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD)*.

Summers, K. 2020. *Water Quality - Science, Assessments and Policy*.

Sun, Q.; and Ge, Z. 2021. A survey on deep learning for data-driven soft sensors. *IEEE Transactions on Industrial Informatics*, 17(9): 5853–5866.

Teramoto, E. H.; Crioni, P. L.; and Chang, H. K. 2021. Daily time series of groundwater recharge derived from temporal variation of water level. *Sustainable Water Resources Management*, 7(4).

Wang, K.; Gou, C.; Duan, Y.; Lin, Y.; Zheng, X.; and Wang, F.-Y. 2017. Generative Adversarial Networks: Introduction and Outlook. *IEEE/CAA Journal of Automatica Sinica*, 4(4): 588–598.

Wilczak, A.; Jacangelo, J. G.; Marcinko, J. P.; Odell, L. H.; and Kirmeyer, G. J. 1996. Occurrence of nitrification in chloraminated distribution systems. *Journal AWWA*, 88(7): 74–85.

Yan, W.; Xu, R.; Wang, K.; Di, T.; and Jiang, Z. 2020. Soft sensor modeling method based on semisupervised deep learning and its application to wastewater treatment plant. *Industrial & Engineering Chemistry Research*, 59(10): 4589–4601.

Yoon, J.; Jarrett, D.; and van der Schaar, M. 2019. Time-series Generative Adversarial Networks. In *Advances in Neural Information Processing Systems*, volume 32.

Yu, T.; and Zhu, H. 2020. Hyper-Parameter Optimization: A Review of Algorithms and Applications.

Zanotti, C.; Rotiroti, M.; Redaelli, A.; Caschetto, M.; Fumagalli, L.; Stano, C.; Sartirana, D.; and Bonomi, T. 2023. Multivariate Time Series Clustering of Groundwater Quality Data to Develop Data-Driven Monitoring Strategies in a Historically Contaminated Urban Area. *Water*, 15(1): 148.