# Water quality anomaly detection research based on GRU-PINN model

Xinyu Zhao*

Faculty of Information Science and Engineering, Ocean University of China, 266005, Qingdao, China

**Abstract.** Maintaining high-quality water resources is essential for sustainable urban water resource management and public health. This paper introduces the GRU-PINN model, developed based on the Gated Recurrent Unit (GRU) network and integrated with a Physics-Informed Neural Network (PINN), to analyze real-world monitoring data from a water treatment company. By embedding domain-specific physical constraints into the loss function, the model enhances interpretability and reduces false alarms. The proposed approach begins with feature engineering on the raw dataset, including missing value imputation, normalization, trend feature extraction, and rolling feature computation. Feature selection is then performed based on feature importance ranking and mutual information analysis. The GRU-PINN model is subsequently employed for anomaly detection in the dataset. The performance of the model is evaluated using the F1-score, precision, and recall. The F1-score, which represents the harmonic mean of precision and recall, is particularly suitable for imbalanced datasets, as the dataset used in this paper contains very few anomaly instances. Experimental results demonstrate that the proposed GRU-PINN model outperforms traditional models by achieving a higher F1-score, thereby improving anomaly detection performance.

## 1 Introduction

Water quality is essential for life, health, and well-being, but it is threatened by various natural and anthropogenic factors, such as pollution, climate change, population growth, urbanization, and industrialization. These factors can cause water quality issues, such as eutrophication, acidification, contamination, and infection, which can affect the safety and efficiency of water treatment systems, as well as the protection of aquatic ecosystems and human health [1-2]. Therefore, water quality monitoring is crucial for detecting and addressing these issues. Water quality monitoring involves the measurement and analysis of various physical, chemical, and biological parameters of water using different methods. Traditionally, these methods are manual sampling and laboratory analysis, which are costly, time-consuming, labor-intensive, and prone to errors [3]. Moreover, they provide limited spatial and temporal coverage of water quality data, which may not capture the dynamic and complex nature of water systems [4].

Machine learning is a branch of artificial intelligence that can provide powerful and flexible methods for anomaly detection in water quality data streams. Machine learning methods can learn from data without explicit rules or assumptions, and can handle complex and nonlinear relationships between variables. Machine learning methods can be broadly classified into three categories: supervised learning, unsupervised learning, and semi-supervised learning [5]. In previous studies, traditional machine learning techniques have been widely employed for anomaly detection in water quality data. Methods such as Support Vector Machines (SVM), Decision Trees (DT), One-Class Support Vector Machines (OCSVM), and Autoencoders (AE) have shown promising results in identifying abnormal patterns within complex datasets [6-8]. These approaches have contributed significantly to the early stages of intelligent water quality monitoring by enabling automated detection of deviations from normal conditions. However, while these models have demonstrated notable progress, they often rely heavily on well-balanced datasets and handcrafted features, which may limit their adaptability and robustness in real-world scenarios characterized by noise, non-stationarity, and data imbalance.water quality data is inherently temporal in nature, feature engineering is essential to effectively capture its dynamic characteristics.

With the rapid advancement of deep learning techniques in recent years, an increasing number of researchers have explored the application of deep learning models to the task of anomaly detection in water quality data. Fehst [9] investigated the impact of various feature engineering techniques on model performance. Their findings demonstrated that automatic feature learning based on Long Short-Term Memory (LSTM) networks yielded superior results, highlighting the potential of deep learning methods in capturing the complex temporal dependencies inherent in water quality data. Similarly, Chen [10] proposed an approach that employs Convolutional Neural Networks (CNN) as feature extractors, which are subsequently integrated with

* Corresponding author: zhaoxinyu@stu.ouc.edu.cn

bidirectional LSTM networks. Additionally, they incorporated cost-sensitive learning strategies to enhance the model's performance on highly imbalanced datasets.

In this paper, a Gated Recurrent Unit (GRU) network is employed, where the Physics-Informed Neural Network (PINN) is integrated into the loss function of the neural network. By combining this approach with feature-engineered data, real-time anomaly detection is achieved within a sliding window framework.

## 2 Data processing

### 2.1 Data sources

The dataset utilized in this paper originates from the GECCO 2018 Industrial Challenge, specifically provided by Thüringer Fernwasserversorgung (TFW) [11]. TFW conducted measurements at critical points throughout the entire water distribution system, particularly at the outlets of water treatment plants and the inlets and outlets of water towers. The data supplied by TFW was measured at various stations near the outlet of a water treatment plant.

This dataset comprises eleven key features: Time, Tp, Cl, pH, Redox, Leit, Trueb, Cl_2, Fm, Fm_2, and Event. Table 1 provides a detailed description of these features. The dataset spans four months, capturing water quality metrics with timestamps.Figure 1 presents a logarithmic-scale visualization of the pronounced class imbalance within the dataset, employing a dual-color bar chart to contrast the distribution of anomalous ("True") versus normal ("False") instances. It is evident that the data set exhibits an extreme imbalance, with only approximately 0.012% of the data rows labeled as "True" and the remaining 99.988% labeled as "False". This level of extreme imbalance poses a significant challenge for machine learning algorithms, as the models tend to perform better in predicting the majority class while neglecting the minority class. Consequently, this results in a high overall accuracy but poor performance in anomaly detection, as the models rarely identify anomalies.

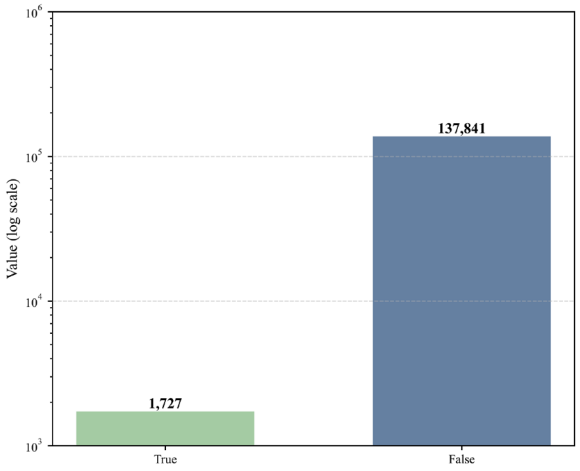


**Figure 1.** Imbalanced binary label distribution in the original dataset

**Table 1.** Description of water quality parameters in the GECCO 2018 dataset.

| No. | Variable | Description |
|---|---|---|
| 1 | **Time** | Time of measurement, given in following format: yyyy-mm-dd HH:MM:SS. |
| 2 | **Tp** | The temperature of the water, given in °C. |
| 3 | **Cl** | Amount of chlorine dioxide in the water, given in mg/L ($MS_1$). |
| 4 | **pH** | PH value of the water. |
| 5 | **Redox** | Redox potential, given in mV. |
| 6 | **Leit** | Electric conductivity of the water, given in µS/cm. |
| 7 | **Trueb** | Turbidity of the water, given in NTU. |
| 8 | **Cl_2** | Amount of chlorine dioxide in the water, given in mg/L ($MS_2$). |
| 9 | **Fm** | Flow rate at water line 1, given in $m^3$/h. |
| 10 | **Fm_2** | Flow rate at water line 2, given in $m^3$/h. |
| 11 | **Event** | Binary event flag, where 'True' indicates an anomaly and 'False' denotes normal conditions. |

### 2.2 Feature engineering

The raw dataset first requires preprocessing to impute missing values and perform normalization. In this paper, we employ KNN imputation with 5-fold cross-validation, evaluating values of $k$ ranging from 3 to 10, ultimately determining the optimal neighbor count as $k = 5$. Following imputation, min-max scaling is applied to normalize the data to a bounded interval [0,1], which helps prevent overfitting.

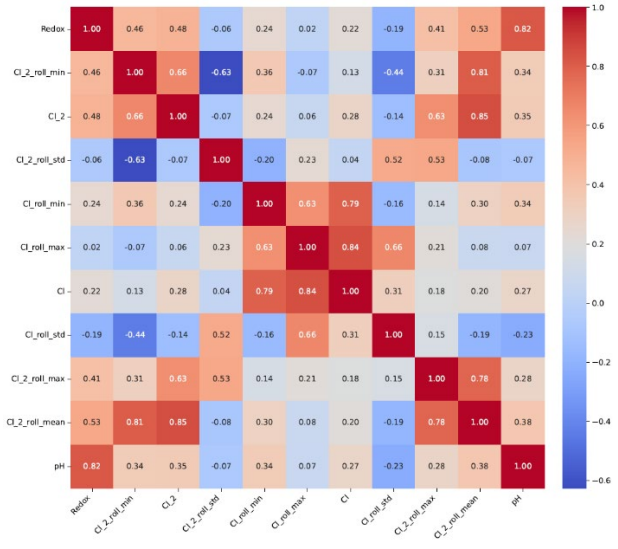


**Figure 2.** Mutual information heatmap of top features highlighting redundancy.

Since water quality data is inherently temporal in nature, feature engineering is essential to effectively capture its dynamic characteristics. For each water quality parameter p ∈{Tp, Cl, pH, Redox, Leit, Trueb, Cl_2, Fm, Fm_2}, we compute trend features (Trend) and rolling-window features with a window size of w. The rolling features include moving average (MA), rolling standard deviation (RSD), and rolling extrema (RMax/RMin).

Figure 2 presents the mutual information correlation heatmap of features after feature engineering. This analysis enables the reduction of feature dimensionality from both mathematical and statistical perspectives, as redundant features may compromise computational efficiency and adversely affect model performance.

# 3 Methods

## 3.1 Fundamentals of Gated Recurrent Unitns

The GRU was first proposed by Cho et al. [12] as a simplified variant of the Recurrent Neural Network (RNN). While achieving performance comparable to that of the LSTM network, the GRU requires fewer trainable parameters, making it a more computationally efficient alternative. This architectural efficiency stems from GRU's simplified gating mechanism, which maintains temporal modeling capabilities through strategic information flow control while reducing computational complexity. The operational dynamics of GRU are governed by the following equations:

$$\begin{aligned} \mathbf{z}_t &= \sigma(\mathbf{W}_z\mathbf{x}_t + \mathbf{U}_z\mathbf{h}_{t-1} + \mathbf{b}_z) \\ \mathbf{r}_t &= \sigma(\mathbf{W}_r\mathbf{x}_t + \mathbf{U}_r\mathbf{h}_{t-1} + \mathbf{b}_r) \\ \tilde{\mathbf{h}}_t &= \tanh(\mathbf{W}_h\mathbf{x}_t + \mathbf{U}_h(\mathbf{r}_t \odot \mathbf{h}_{t-1}) + \mathbf{b}_h) \\ \mathbf{h}_t &= \mathbf{z}_t \odot \mathbf{h}_{t-1} + (1 - \mathbf{z}_t) \odot \tilde{\mathbf{h}}_t \end{aligned} \quad (1)$$

where $\mathbf{z}_t \in \mathbb{R}^n$ denotes the update gate vector controlling historical information retention, $\mathbf{r}_t \in \mathbb{R}^n$ represents the reset gate vector regulating past state influence, and $\tilde{\mathbf{h}}_t \in \mathbb{R}^n$ is the candidate hidden state combining current inputs with gated historical context; $\mathbf{W}_z, \mathbf{W}_r, \mathbf{W}_h \in \mathbb{R}^{n\times d}$ are input transformation weight matrices, $\mathbf{U}_z, \mathbf{U}_r, \mathbf{U}_h \in \mathbb{R}^{n\times n}$ are recurrent connection weight matrices, $\mathbf{b}_z, \mathbf{b}_r, \mathbf{b}_h \in \mathbb{R}^n$ denote corresponding bias vectors, $\sigma(\cdot)$ is the sigmoid activation function compressing values to [0,1], and $\odot$ signifies the Hadamard (element-wise) product operation.

## 3.2 Loss function design

A critical component of our methodology is a specialized loss function that integrates data-driven learning with physics-based constraints. This composite loss function combines the mean squared error (MSE) loss with a physics-informed loss term:

$$L_{\text{total}} = L_{\text{MSE}} + \lambda_1 \cdot L_{\text{phy}} \quad (2)$$

The MSE loss minimizes the discrepancy between predicted and observed values:

$$L_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N}\left(y_{\text{true},i} - y_{\text{pred},i}\right)^2 \quad (3)$$

where $y_{\text{true,i}}$ represents the ground truth value, $y_{\text{pred,i}}$ denotes the predicted value, and N is the total number of data points.

The physics-informed loss enforces adherence to the governing convection-diffusion equation:

$$L_{\text{phy}} = \frac{1}{N}\sum_{i=1}^{N}\left(\text{CDM}_i - y_{\text{pred},i}\right)^2 \quad (4)$$

where $\text{CDM}_i$ corresponds to the solution derived from the convection-diffusion model for the i-th data point.

The physical principles are derived from the convection-diffusion equation:

$$\frac{\partial C}{\partial t} + \vec{u}\cdot\nabla C = D\nabla^2 C + S \quad (5)$$

where C is the concentration of the substance, $\vec{u}$ represents the velocity field, D is the diffusion coefficient, and S denotes the source term.

Given that our water quality sensors provide temporal but limited spatial data, we adopt a one-dimensional simplification:

$$\frac{\partial C}{\partial t} + \vec{u}\frac{\partial C}{\partial x} = D\frac{\partial^2 C}{\partial x^2} + S(x,t) \quad (6)$$

The source term $S(x,t)$ is estimated from temporal variations in water turbidity:

$$S(x,t) = \frac{\partial T}{\partial t} \approx \frac{T_{i+1} - T_i}{t_{i+1} - t_i} \quad (7)$$

This dual-component loss function ensures that the model not only minimizes prediction errors but also adheres to the fundamental principles of fluid dynamics, thereby improving its predictive accuracy and generalizability in water quality modeling applications.

## 3.3 GRU-PINN architecture

The convection-diffusion model provides a theoretical foundation for modeling transport, mixing, and dispersion of substances in fluid environments. We leverage this physical framework by implementing a PINN approach that enhances standard deep learning architectures with physical constraints. As illustrated in Figure 3, our GRU-PINN architecture integrates recurrent neural networks with physical principles to improve predictive accuracy for water quality parameters. This architecture combines the strengths of GRU networks in capturing temporal dependencies with the physical consistency enforced by PINN. The GRU component processes the sequential water quality data:

$$\mathbf{h}_t = \text{GRU}(\mathbf{x}_t, \mathbf{h}_{t-1}; \theta_{\text{GRU}}) \quad (8)$$

where $\mathbf{h}_t \in \mathbb{R}^{64}$ is the hidden state at time t, summarizing temporal information from the input sequence $\mathbf{x}_t$ and the previous hidden state $\mathbf{h}_{t-1}$. $\theta_{\text{GRU}}$ represents the trainable parameters of the GRU network, which includes update and reset gates that enable efficient learning of temporal patterns in the water quality data. We selected a GRU with 64 hidden units as a trade-off between model expressiveness and computational efficiency, a configuration widely adopted in time series modeling due to its ability to mitigate vanishing gradients while maintaining simplicity compared to LSTM.

The physics-informed aspect is introduced through a separate neural network F, which takes the GRU hidden state $\mathbf{h}_t$ as input and outputs a prediction of the water quality parameter concentration $\hat{C}_t$:

$$\hat{C}_t = F(\mathbf{h}_t; \theta_{\text{phy}}) \quad (9)$$

The network F is a fully connected feedforward neural network with 3 hidden layers, each containing 32 neurons, and parameters $\theta_{\text{phy}}$. This architecture was empirically determined through ablation studies, balancing the capacity to approximate the concentration field with regularization against overfitting. Crucially, F ensures that predictions adhere to the underlying convection-

diffusion equation, integrating domain knowledge with data-driven learning.

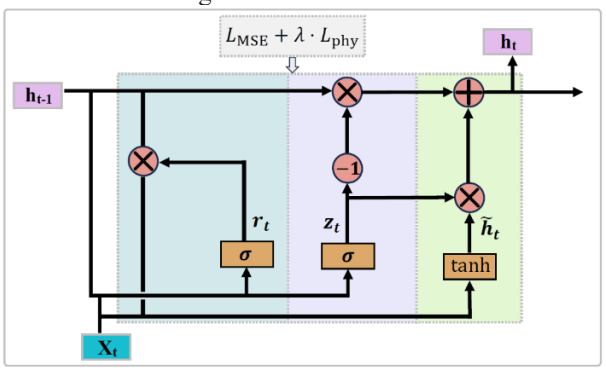


**Figure 3.** Structure of GRU-PINN.

### 3.4 Performance evaluation metrics

In this paper, we employ the F1 score as the primary evaluation metric to assess the model's anomaly detection performance, as it provides a balanced measure between precision and recall.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{10}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{11}$$

$$\text{F1-Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

where TP, TN, FP, and FN represent the counts of true positives, true negatives, false positives, and false negatives, respectively. The F1-Score, being the harmonic mean of precision and recall, is particularly suitable for imbalanced datasets where anomaly instances are typically rare compared to normal observations.

## 4 Experimental results and analysis

The experimental evaluation was conducted on a computational platform equipped with an NVIDIA GeForce RTX 3060 Ti GPU (12GB VRAM), utilizing PyTorch 2.60 and Python 3.7. This configuration was selected to ensure efficient training of deep learning models while maintaining reproducibility. The GPU's tensor cores and mixed-precision capabilities were leveraged to accelerate matrix operations inherent to recurrent neural network architectures. All models were implemented using PyTorch's native CUDA backend, with deterministic algorithms enabled to guarantee consistent results across multiple runs.

Table 2 presents a comprehensive comparison of model performance across three key metrics: F1-Score, Precision, and Recall. These metrics were chosen as they collectively provide a robust assessment of anomaly detection capabilities, particularly important given the typically imbalanced nature of anomaly detection datasets. The results reveal several critical observations.

**Table 2**. Performance comparison of anomaly detection models (F1-Score, Precision, Recall).

| Model | F1-Score | Precision | Recall |
|---|---|---|---|
| **SVM** | 0.81 | 1.00 | 0.69 |
| **DT** | 0.79 | 0.76 | 0.88 |
| **AE** | 0.78 | 0.70 | 0.88 |
| **LSTM** | 0.86 | 0.75 | 1.00 |
| **GRU** | 0.83 | 1.00 | 0.71 |
| **GRU-PINN** | 0.91 | 1.00 | 0.83 |

This paper systematically compares the performance of traditional machine learning models and deep learning architectures, with a particular focus on GRU, LSTM, and the proposed GRU-PINN model. The dataset, following rigorous feature engineering, was partitioned using a 70:30 train-test split. A sliding window approach with a fixed window size of 60 time steps was employed to evaluate all models uniformly. Traditional machine learning models, such as SVM, DT, and AE, demonstrated only moderate performance. These outcomes suggest that conventional methods struggle to capture the complex temporal dynamics often present in water quality data. Among the deep learning models, the GRU architecture achieved perfect precision (1.00), indicating its strong ability to minimize false positives. However, its relatively lower recall (0.71) reflects a conservative detection strategy, where some true anomalies are potentially missed. In contrast, the LSTM model exhibited an inverse performance pattern, achieving perfect recall (1.00) but at the expense of reduced precision (0.75). This suggests a more aggressive detection approach that successfully captures all anomalies but is prone to generating more false alarms.

The proposed GRU-PINN architecture strikes an optimal balance, achieving perfect precision (1.00) and substantially improved recall (0.83) compared to the baseline GRU. This results in the highest overall F1-score (0.91) among all models evaluated. Specifically, it represents a 0.5 improvement over the LSTM model and an 0.8 improvement over the standard GRU. These findings underscore the effectiveness of incorporating physical constraints into the learning framework, which enhances the model's capacity to accurately detect anomalies in temporally dynamic datasets.

## 5 Conclusion

This paper introduces the GRU-PINN model, specifically designed for anomaly detection in urban water quality monitoring systems. The proposed approach begins with feature engineering on the raw dataset, incorporating both trend and rolling features. Subsequently, a feature selection process is conducted based on feature importance ranking and mutual information correlation among features to identify the most relevant attributes. The GRU-PINN model integrates domain-specific physical constraints into a data-driven GRU architecture, enabling the model to better capture the underlying physical dynamics of water quality parameters. This integration enhances both the accuracy and robustness of

anomaly detection. Compared to conventional standalone models, the GRU-PINN model achieves a higher F1 score, improving its capability to detect anomalies while reducing false alarms, thereby enhancing the system's reliability in real-world applications.

## References

1. M.G. Uddin, A.I. Olbert, S. Nash, CERAI 594-599 (2020).
2. van Vliet, M.T.H., Thorslund, J., Strokal, M. *et al.*, Nat Rev Earth Environ **4**, 687–702 (2023).
3. J. Liu, D. Wu, H. Mohammed, R. Seidu, *Water* **16**, 1238 (2024)
4. U. Ahmed, R. Mumtaz, H. Anwar, S. Mumtaz, A.M. Qamar, Water Supply **20**, 28 (2020)
5. M.N. Kanyama, F.B. Shava, A.M. Gamundani, A. Hartmann, Phys. Chem. Earth A/B/C **134**, 103558 (2024)
6. Z.X. Tian, J.P. Jiang, L. Guo, P. Wang, *Proc. ACAI* 518 (2012)
7. Y. Yuan, K. Jia, *Proc. ICSPCC* 1 (2015)
8. S. Fang, W.Z. Sun, L. Huang, *J. Phys.: Conf. Ser.* **1345**, 022054 (2019)
9. V. Fehst, H.C. La, T.-D. Nghiem, B.E. Mayer, P. Englert, K.-H. Fiebig, *Proc. GECCO* 5 (2018)
10. X.G. Chen, F. Feng, J.K. Wu, W.Y. Liu, *Proc. GECCO* 3 (2018)
11. F. Rehbach, S. Moritz, S. Chandrasekaran, M. Rebolledo, M. Friese, T. Bartz-Beielstein, Accessed: Feb **19**, 2019 (2018)
12. K. Cho, B. van Merrienboer, D. Bahdanau, Y. Bengio, arXiv:1409.1259 (2014)