A Study on Anomaly Detection and Hybrid Forecasting Model Using Underground Stormwater Pipe Water Level Time-Series Data for Urban Flood Prediction

Jang-won Lee, Nak Hyun Jung

Abstract

As urban flooding risks intensify due to climate change, developing accurate and real-time predictive systems has become critical for disaster preparedness and response. This study proposes an AI-based flood prediction model that leverages real-time rainfall and inflow water level time-series data. The system consists of two core modules: an anomaly detection and auto-correction module to ensure data reliability and quality, and a deep learning-based time-series forecasting module for predicting inflow levels based on rainfall trends.

The anomaly detection module employs a reconstruction error method based on Long Short-Term Memory (LSTM) networks to identify abnormal data points that deviate from normal time-series patterns. Detected anomalies are corrected via linear interpolation, where the anomalous data points are removed and replaced with new values estimated from surrounding normal data points.

The forecasting module, built upon the LSTM architecture, captures complex rainfall-runoff relationships and temporal dynamics to provide accurate water level predictions. Additionally, hybrid models including ConvLSTM and LSTM-Transformer were designed and evaluated for comparative performance.

Experimental results demonstrated that the proposed system achieved over 95% accuracy in anomaly detection and correction. On real-world datasets, the LSTM model outperformed baseline methods, achieving a Mean Squared Error (MSE) of 0.000089, a Mean Absolute Error (MAE) of 0.006778, and a high R2 score of 0.972614. These results suggest the model's potential to enhance the reliability and efficiency of urban flood response decision-making and to contribute to the strengthening of disaster management capabilities.

☞ Keywords: Time Series Prediction, Anomaly Detection, Water Level Prediction, Hybrid Model, LSTM, ConvLSTM, Transformer

1. Introduction

Extreme weather events, driven by global climate change, are increasing in frequency and intensity, causing severe socioeconomic problems in urban areas. These problems include significant loss of life and property, as well as the disruption of critical social infrastructure due to flooding.

The rise of impermeable surfaces from industrialization and urbanization has led to a sharp increase in rainwater runoff, which, combined with the limitations of existing urban drainage systems, exacerbates the risk of unpredictable urban flooding. To effectively respond to these disasters, it is essential to develop a reliable and accurate prediction model for flood situations, one that can integrate real-time rainfall and water level data collected from hard-to-access areas such as underground storm sewers, which are often prone to power and communication issues.

1.1 Research Background

Historically, urban flood prediction has relied primarily on hydrological and hydraulic physics-based models. While these models are grounded in physical principles to describe system operations, they face significant challenges in accurately reflecting the heterogeneity, nonlinearity, and dynamic changes of complex urban environments in real time. The calibration and validation of model parameters demand enormous amounts of time and specialized expertise. Furthermore, data quality issues, such as missing values and outliers from field sensors, can severely degrade the predictive performance of these models. This data quality problem is particularly critical for disaster prediction systems that require high real-time accuracy and reliability.

1.2 Need for Research

In recent years, artificial intelligence (AI) technology, and deep learning in particular, has demonstrated exceptional performance in learning and predicting complex patterns in large-scale time-series data, leading to its application in various engineering fields. This study aims to leverage the potential of AI technology to improve the accuracy and reliability of urban flood prediction. Specifically, the research proposes an end-to-end framework for predicting inflow water levels in underground sewer and stormwater pipes. This framework utilizes real-time rainfall and water level data and incorporates correlation-based feature engineering, as well as a robust anomaly detection and automatic correction model based on LSTM networks.

The study also investigates LSTM, ConvLSTM, and LSTM-Transformer hybrid models for prediction. The goal is to overcome the limitations of existing models and provide a robust framework for proactive and systematic urban flood response. The research focuses on integrating anomaly detection and time-series prediction technology to ensure data integrity and predictive accuracy, and to support decision-making with visualized data.

1.3 Research Objectives

The ultimate objective of this study is to develop an AI-based urban flood situation prediction model that utilizes real-time rainfall and water level time-series data. This model is intended as a technological solution to effectively address the growing damage from urban flooding caused by climate change. Urban flooding is a complex phenomenon that can involve sewer/stormwater overflow, river flooding, underground facility inundation, and drainage system degradation due to short-term heavy rainfall. The difficulty in predicting and responding to these events makes them a major issue with direct this, the research is structured around three specific objectives

To enhance the accuracy of flood prediction based on real-world sensor data, a time-series-based multi-anomaly detection technique and an automatic correction algorithm were designed and developed. This aims to minimize data loss and noise issues caused by communication errors, battery depletion, and environmental interference, thereby ensuring the reliability and quality of the input data.

To accurately reflect the nonlinear relationship between rainfall and inflow water levels, deep learning-based time-series prediction models were applied. The research specifically focuses on hybrid models from the LSTM, CNN, and Transformer families to learn water level change patterns. The goal is to improve the prediction model's performance by enabling it to detect and predict rapid water level increases in urban flood scenarios.

The developed system is intended to be a tool for enhancing the reliability and efficiency of urban disaster response decision-making. By doing so, it can contribute to a tangible strengthening of the flood response capabilities of local governments and disaster response agencies. This research is expected to function as a robust model for an AI-based disaster response system, contributing to smart city-based disaster prevention and the enhancement of urban resilience.

1.4 Thesis Structure

This thesis is structured as follows: Chapter 1, Introduction, presents the research background, necessity, objectives, and thesis structure. Chapter 2, Theoretical Background, reviews existing literature on time-series data anomaly detection and prediction and summarizes the principles and features of relevant models. Chapter 3, Research Methodology, defines the proposed model and its methodology, providing a detailed explanation of the system's architecture, data collection, anomaly detection and correction, and the deep learning-based water level prediction process. Chapter 4, Implementation and Verification, presents the research process and results, including the model's prediction accuracy using various metrics. Chapter 5, Conclusion and Future Plans, summarizes the main research findings and suggests directions for future research.

2. Theoretical Background

Time-series data-based anomaly detection and prediction are core topics actively addressed in a variety of fields, including smart cities, industrial control, environmental monitoring, and disaster prediction. This study emphasizes the importance of a deep learning-based approach and meticulous data quality management.

2.1 Prior Research on Time-Series Data Anomaly Detection

2.1.1 Definition and Importance of Anomalies

An anomaly (or outlier) is a data point or sequence that deviates from the general pattern of a dataset. Effectively detecting anomalies in time-series data is crucial for ensuring data reliability and enhancing the performance of subsequent analysis and prediction models.

According to a study by Freeman, Merriman, Beaver, & Mueen (2021), anomalies can arise from various sources such as sensor malfunctions, communication errors, or genuine abnormal phenomena. They can be broadly classified into three types: Point Anomalies, where a single data point is significantly different from others; Contextual Anomalies, where a data point is considered abnormal only within a specific context or time period and Collective Anomalies, where a series of data points, individually normal, collectively form an abnormal pattern.[5]

2.1.2 Research on Statistical-Based Anomaly Detection Trends

In the early days of time-series anomaly detection, statistical approaches were predominantly used. A study by Braei & Wagner (2020) discusses statistical-based anomaly detection as a key category, comparing it with machine learning and deep learning methods. Statistical methods use a dataset's statistical properties to identify observations that deviate from normal patterns, often by assuming a specific statistical distribution or defining a normal range based on metrics like mean, variance, or quartiles.[3]

While these methods have played a significant historical role, they have limitations in capturing the complex temporal dependencies inherent in time-series data. This is because they may not adequately consider unique time-series characteristics such as trends, seasonality, or periodicity. Consequently, recent research has seen a shift toward machine learning and deep learning-based methods for anomaly detection.

2.1.3 Anomaly Detection Research by Core Time-Series Characteristics

The study by Freeman et al. (2021) evaluated twelve anomaly detection techniques across different time-series characteristics, including seasonality, trend, concept drift, and missing values. The researchers compiled ten time-series corpus datasets for this purpose, with some from the Numenta Anomaly Benchmark and others labeled manually. The study provided criteria for determining which technique is most suitable for a given situation, using metrics such as AUC (Area Under ROC Curve), windowed F-score for time-interval precision/recall, and NAB (Numenta Anomaly Benchmark) scores for rewarding early detection.[5]

2.1.4 Ensemble-Based Anomaly Detection Research

Single anomaly detection models, while strong in detecting specific anomaly types, are limited in their ability to effectively identify complex anomalies from diverse causes. To overcome this limitation and enhance the robustness of anomaly detection, ensemble methods that combine multiple single models have received significant attention.

According to a study by Boateng & Bruce (2023), a comparison of five ensemble configurations—four variations based on weighted voting and a stacking approach—was conducted on various PLC-based ICS datasets.[2] The stacking-based model achieved the most superior performance across key metrics such as precision, recall, and F1-score. This model's detection rate was consistently better, and its false positive rate was lower than both single-anomaly detectors and voting-based models. This performance difference was statistically significant, as confirmed by statistical hypothesis testing, which adds to the credibility of the findings. This approach demonstrates how combining the strengths of individual models can effectively identify a wider range of anomaly types.

2.1.5 IoT Time-Series Data Anomaly Detection Research

In a paper on IoT time-series data anomaly detection, Shahanas, Akthar, Anand, Rakshitha, & Amirthavalli (2024) proposed a hybrid model that combines a Bi-directional Long Short-Term Memory (Bi-LSTM) autoencoder with an One-Class Support Vector Machine (OC-SVM). The Bi-LSTM autoencoder learns the temporal patterns of the IoT time-series data and is used to reconstruct the data. Normal data is accurately reconstructed, but anomalous data generates a large reconstruction error. The OC-SVM then uses this reconstruction error to determine if a data point is an anomaly. The model achieved a high accuracy of 96.34%, proving its effectiveness in leveraging the complex temporal characteristics of IoT time-series data for anomaly detection. The same study also evaluated a hybrid model combining an LSTM autoencoder with Isolation Forest. This model, which achieved 90.8% accuracy, performed less effectively, possibly due to the limitations of Isolation Forest on certain anomaly types or a less powerful synergistic effect compared to the Bi-LSTM+OC-SVM combination.[8]

2.2 Time-Series Data Prediction

Time-series data prediction, which involves learning patterns from past data to infer future values, plays a crucial role in disaster prediction and early warning systems like those for urban flooding.

2.2.1. Research on Statistical Time-Series Prediction Models

Widely used statistical models in time-series forecasting, such as ARIMA and Prophet, are important research subjects due to their characteristics and real-world limitations. ARIMA is effective at predicting future values using past data and past prediction errors, while Prophet excels at flexibly modeling seasonality, trends, and holiday effects, making it a strong choice for business time-series forecasting.

However, according to a study by Taylor & Letham (2017), applying these statistical models to complex, nonlinear time-series patterns, such as those in Facebook event data, reveals their inherent limitations. Automated methods like auto.arima often struggle to effectively capture trend changes and multiple seasonalities. This suggests that statistical models have limitations when dealing with highly complex and unpredictable time-series patterns, such as those found in rainfall and water level data.[7]

2.2.2 Research on Runoff Prediction Models Using ML-based SVR and RF

Bargam, Boudhar, Kinnard, Bouamri, Nifa, & Chehbouni (2024) proposed a model for daily runoff prediction in a data-scarce Moroccan river basin using Support Vector Regression (SVR) and Random Forest (RF). The predictive variables selected were rainfall, previous day's flow, and previous day's snow cover area. The models were calibrated using a yearly cross-validation method. SVR, using a linear kernel, was optimized for hyperparameters via GridSearchCV, while RF used 1,300 trees to prevent overfitting.[1]

The results showed that SVR, with an NSE of 0.59, achieved higher accuracy during the verification phase than RF (NSE 0.53) and Multiple Linear Regression (MLR, NSE 0.54). Both models demonstrated a strength in predicting rapid flow fluctuations. However, the performance showed significant variability in yearly cross-validation (NSE 0.05~0.87), which was attributed to uneven hydrological conditions and the quality of the input data. This study serves as a case study demonstrating that machine learning techniques, particularly SVR's nonlinear predictive ability, can be effective for hydrological time-series prediction.

2.2.3 Research on LSTM-GRU Hybrid Models for Channel Basin Water Level Prediction

The study by Widiasari & Efendi (2024) proposed a deep learning framework for predicting water levels in the lower section of a flood channel in Semarang, Indonesia. The study used IoT-based multi-variable rainfall time-series data (from upstream and downstream points). The researchers compared LSTM, GRU, and a hybrid LSTM-GRU model, evaluating their performance with both single-variable (downstream rainfall only) and multi-variable (upstream and downstream rainfall) inputs.[9]

The LSTM model, which excels at capturing long-term dependencies, showed high prediction reliability and accuracy (approximately 88%). GRU, with its simpler gate structure, proved more computationally efficient and stable with volatile time series, achieving around 92% accuracy and lower error metrics (MSE, RMSE) when using both upstream and downstream rainfall. The hybrid LSTM-GRU model, which combines the strengths of both, showed the most significant performance improvement when sufficient input information was available, reaching an accuracy of about 92.5%. This suggests that the hybrid approach, by combining LSTM's memory retention with GRU's efficiency, can effectively model complex rainfall-water level relationships.

2.2.4 ConvLSTM Hybrid Prediction Model Research

The ConvLSTM architecture, which combines the strengths of CNN and LSTM, has recently gained attention in the field of hydrology. CNNs can extract spatial features from time-series inputs, while LSTMs can learn the temporal dependencies between these features.

A study by Oddo, Bolten, Kumar, & Cleary (2024) used ConvLSTM to predict stream stage height with multi-modal inputs, including remote sensing and on-site rainfall/runoff data. This hybrid model reduced the prediction error (RMSE) by approximately 26% compared to existing LSTM-based models, significantly improving the sensitivity of flood forecasting.[6]

2.2.5 LSTM Transformer Hybrid Prediction Model Research

To overcome the limitations of single deep learning models and maximize prediction accuracy, hybrid approaches that combine multiple models are being actively researched. A study by Fan, Zhang, Chen, & Xu (2025) proposed a hybrid prediction model that integrates a Transformer, LSTM, and the Sparrow Search Algorithm (SSA).

This model is designed to effectively learn the complex structure of time-series data by combining the Transformer's ability to extract global features with LSTM's capacity for processing temporal dependencies. The SSA serves to automatically optimize the model's key hyperparameters, maximizing prediction performance while overcoming the limitations of manual tuning. The model was experimentally validated using rainfall and water level data from Zhengzhou, China. The results showed that the Transformer-LSTM-SSA hybrid model significantly outperformed standalone Transformer or LSTM models. It recorded up to 10% higher Nash-Sutcliffe Efficiency (NSE) and reduced Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) by more than 30% and 10%, respectively. These findings demonstrate the model's effectiveness in handling the nonlinearity and long-term dependencies of time-series data, and its potential as a practical approach for urban flood prediction systems.[4]

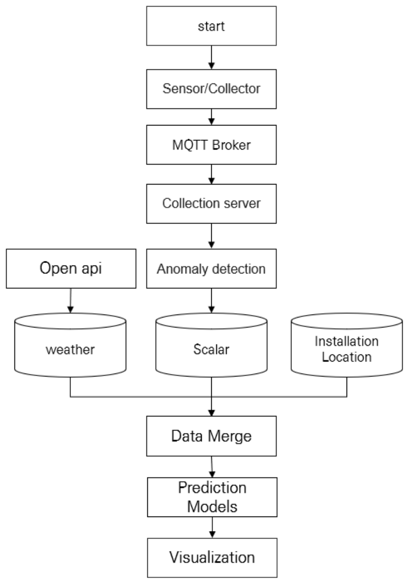
3. Research Methodology

This study proposes an AI-based flood situation prediction model and a real-time architecture for urban flood response. The system is designed as a pipeline encompassing IoT data collection, anomaly detection and correction, feature extraction and preprocessing, deep learning-based water level prediction, and the utilization of prediction results to support decision-making.

Traditional hydrological and hydraulic models, while accurate, are computationally expensive and limited in their real-time application, particularly when faced with missing values or anomalies in sensor data. This study addresses these limitations by using an AI-based model and leveraging real-world water level data combined with Korea Meteorological Administration rainfall data to construct a robust preprocessing pipeline for feature extraction, anomaly detection, and correction.

3.1 Research Design and Procedure

The research focuses on using time-series data collected from the extremely challenging and unstable environment of underground storm sewers. The goal is to solve data quality issues (anomalies) and develop an accurate inflow water level prediction model. Recognizing that traditional time-series analysis techniques are ill-suited for handling the complexity and noise of such environmental data, the study integrates deep learning-based anomaly detection and correction methods into the prediction modeling pipeline. A comparative analysis of various deep learning models' predictive performance was also conducted. The experimental process is outlined in Figure 1.



(Figure 1) Urban Flood Artificial Intelligence Response System Sewage/Stormwater Level Prediction Experiment.

3.2. Real-time Data Collection and Preprocessing

The time-series data used in this study was collected from the underground sewer and stormwater pipes in a specific area that experienced severe flooding during recent summer heavy rainfall. The data includes water level (in cm) measured at 1-minute intervals over the month of April 2025, merged with nearby Korea Meteorological Administration rainfall data based on the datetime.

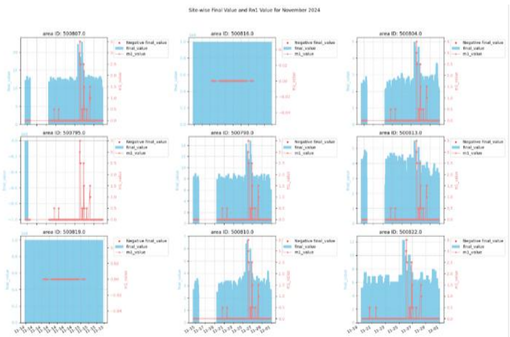
The specifications of the dataset used in the study are provided in Table 1.

(Table 1 Water Level and Rainfall Data Information)

Category	Measurement Area	Period	Number of Areas	Daily Measurement Frequency	Measured Data Points

Water Level	Seoul	2025/04/01~2025/04/30	9	1,440 times/min	388,200
Rainfall	Seoul	2025/04/01~2025/04/30	1	48 times/30 min	1,440

As is characteristic of underground sewer and stormwater pipes, the data collection process frequently encounters anomalies or missing values due to various environmental factors, such as sensor errors, communication failures, and battery replacements. These issues are major factors that degrade the accuracy of prediction models. Figure 2 illustrates the extent of incorrect missing values that occurred in November 2024 from sensors installed in nine different areas due to communication failures and sensor defects.

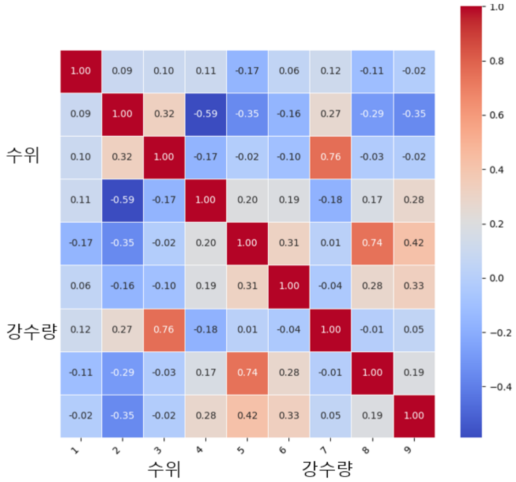


(Figure 2) 2024 November

According to the research design, the raw data underwent a series of preprocessing steps to be refined into a format suitable for prediction model training. After loading the data, key columns containing water level, rainfall, and time information were selected. Time-related columns were converted to the proper format and sorted chronologically to meet the basic requirements of time-series analysis.

3.2.1 Correlation Analysis

To analyze the relationship between rainfall and the water level in the sewer/stormwater pipes, Korea Meteorological Administration data (including temperature, humidity, and wind) was interpreted. A correlation heatmap, shown in Figure 3, revealed a very strong positive correlation of 0.76 between water level (variable 3) and rainfall (variable 7), indicating that the two variables move in the same direction almost simultaneously.



(Figure 3) Precipitation and sewage/storm water level correlation heat map

3.2.2 Feature Engineering

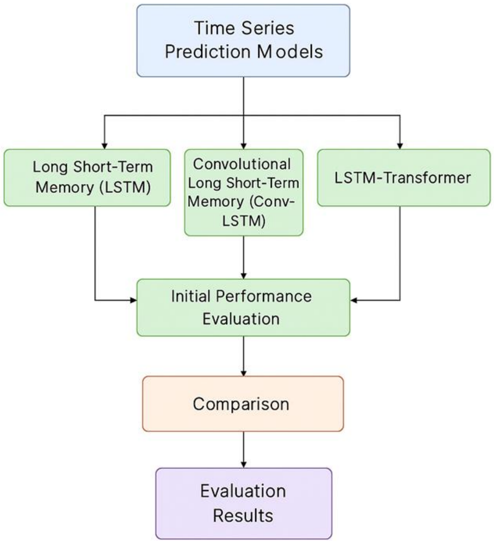
To better reflect the relationship between rainfall and water level, a new feature was created by calculating the difference between the two variables. This new feature helps the model learn the relative change in water level in response to changes in rainfall, which contributes to capturing the nonlinear and complex impact of rainfall on water level.

3.2.3 Time Series Anomaly Detection and Correction

The study used an LSTM-based reconstruction error method to detect anomalous data that deviates from the normal time-series pattern. This technique is advantageous for effectively identifying abnormal fluctuations within time-series data. For the segments identified as anomalies, the values were corrected using linear interpolation. This process involves removing the anomalous data points and estimating new values based on surrounding normal data points, thereby making the model less susceptible to noise.

3.3 Prediction Model Configuration and Training

After anomaly correction and feature engineering, the time-series data was used to train and configure three deep learning-based time-series prediction models. All models were designed to predict all feature values, including the water level, at the next time step based on an input sequence. These models were selected because they are suitable deep learning architectures for learning complex and nonlinear time-series patterns like those found in underground sewer and stormwater pipe dat



<Figure 4> Constructing a Time Series Water Level Prediction Model

3.3.1 LSTM (Long Short-Term Memory) Model

The LSTM model, specialized for learning the temporal dependencies of time-series data, was configured with two LSTM layers and a dropout layer. The first LSTM layer has 128 units and uses a Rectified Linear Unit (ReLU) activation function. To maintain the input data in a sequence format, return_sequences=True was set. The input shape was defined as (seq_length, n_features). The second LSTM layer has 64 units and is configured with return_sequences=False to pass only the output of the final time step to the next layer, summarizing the entire time-series data. A dropout layer with a rate of 0.2 was inserted between the LSTM layers to prevent overfitting. The final output layer is a Dense layer with n_features units and a linear activation function, which is suitable for predicting continuous numerical values.

3.3.2 ConvLSTM (Convolutional LSTM) Model

This model was designed to learn both temporal patterns and the spatio-temporal correlations that arise in multivariate sensor data. The core of this model is the ConvLSTM2D layer, which combines convolutional operations with an LSTM structure. The model consists of two ConvLSTM2D layers, a Flatten layer, and a Dense output layer.

The first ConvLSTM2D layer uses 64 filters and a kernel size of (1, 3), with a ReLU activation function. It is set to return_sequences=True to allow the learning of spatial patterns at each time step and connect them temporally. The second ConvLSTM2D layer has 32 filters and is set to return_sequences=False to summarize the information from the final time step. The Flatten layer converts the spatio-temporal information into a one-dimensional vector. Dropout layers with a rate of 0.2 are used to prevent overfitting. The final Dense output layer has n_features units and a linear activation function for regression.

3.3.3 LSTM-Transformer Model

This hybrid model combines LSTM and a Transformer

Encoder to learn both the temporal patterns and the important relationships between time points. This architecture was designed to effectively capture long-term dependencies and crucial time-point information in complex time-series data like that from underground storm sewers.

The initial input sequence is processed by an LSTM layer with 64 units, a ReLU activation function, and return_sequences=True. The LSTM output is then passed to a Transformer Encoder block, which is applied twice. Each block consists of Multi-Head Attention and a Feed Forward network. The attention mechanism is configured with head_size=64 and num_heads=4, allowing for parallel analysis of interactions between time points from different perspectives. LayerNormalization and Dropout (0.2) are included in each block to ensure stable learning and prevent overfitting. The Transformer blocks emphasize important time points for prediction based on learned weights. The final output from the Transformer Encoder is passed to a Dense output layer with n_features units and a linear activation function.

3.4. Research Environment

The research environment for the AI-based flood prediction model was configured on a cloud-based platform to ensure scalability, flexibility, and stability. This setup enables efficient research and platform operation. The hardware and software components used are detailed in Table 2.

(Table 2) Platform and Research Environment Configuration

Category	Item	Specification/Version
Hardware	Processor	Intel CPU 8-core / Colab support
	Memory	16GB / High-capacity RAM
	Graphics Accelerator	T4 GPU
Software	Programming Language	Python 3.*
	Data Analysis	Pandas, NumPy
	Time-series Modeling	LSTM, ConvLSTM, LSTM Transformer
	Visualization Tools	Matplotlib, Plotly
	Front end/UI	React, Next.js, Chart.js, Redis
	Back end	Spring Boot, Gradle
	Database	MySQL
	Repository & Build	GitLab / Jenkins

This cloud-based configuration allows for flexible resource allocation and is optimized for the high-speed storage and retrieval of large-scale time-series data. A cloud IoT platform was used for real-time data collection from sensors via lightweight communication protocols like MQTT. The preprocessing and anomaly correction logic, implemented using Python-based data analysis libraries, was executed using serverless functions or container-based services to ensure data quality. For visualization and user interaction, the front-end (React, Next.js, Chart.js) and back-end (Spring Boot) were deployed on cloud hosting services, with the back-end managing the prediction model API and database access. The use of CI/CD services (GitLab, Jenkins) streamlined the development and deployment process, ensuring stable system operation. This robust, cloud-native environment provides a solid foundation for expanding the AI-based flood prediction model into a real-world urban disaster response system.

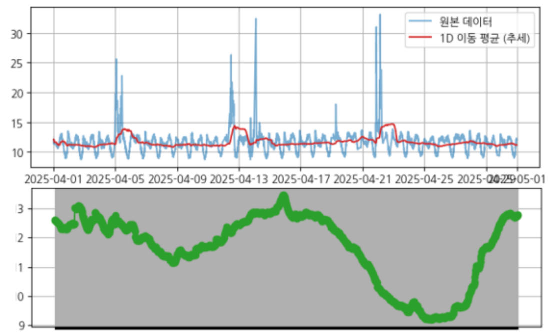
4. Implementation and Verification

The proposed platform was implemented and verified by applying it to an area at risk of urban flooding.

4.1 Implementation Settings

The time-series data used in this study was collected from the underground sewer/stormwater pipes of a specific area that experienced flooding during the summer of 2023. The data, measured at 1-minute intervals for one month in April 2025, is prone to frequent anomalies due to noise, sensor errors, and communication failures. These data quality issues are a major cause of performance degradation in prediction models. To address this, the collected raw data underwent a series of preprocessing steps to prepare it for model training.

First, columns containing time information were converted to the datetime format and the data was sorted chronologically. This ensured the fulfillment of basic requirements for time-series analysis, as shown by the trend, seasonality, and overall progression of the water level information in Figure 5.



<Figure 5> Water level information Time series trends, trends, and seasonality

During the preprocessing stage, an LSTM-based reconstruction error method was applied to detect anomalies, a technique well-suited for identifying abnormal time-series patterns. The detected anomalies were corrected using linear interpolation, which involved removing the anomalous data points and replacing them with values estimated from surrounding normal data points. This correction process helped the prediction model become less sensitive to noise and better learn the true data patterns. Finally, the numeric feature columns were normalized to the range using a Min-Max Scaler, a step performed after anomaly correction to improve the deep learning model's performance.

The verification process involved collecting water level data from IoT measurement devices and ultrasonic sensors installed on-site, along with nearby rainfall data from the Korea Meteorological Administration. This data, sampled at a fixed interval of 1 minute, was transmitted to a collection server via the MQTT communication protocol. The collected raw data was processed and corrected before being stored in a time-series database, enabling efficient storage, retrieval, and real-time monitoring. Figure 6 displays a visualization screen of the platform, showing collected rainfall and water level data over time on the left, and a visual representation of the water level inside the sewer pipe on the right.

<Figure 6> Visualization of water level information in April 2025

4.2. Results Verification

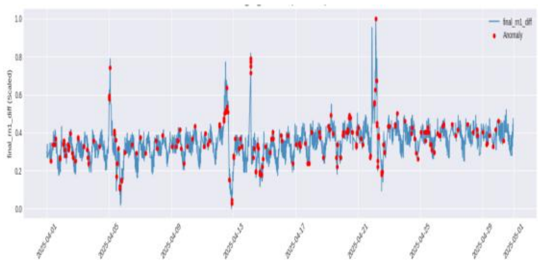
The improved performance from the anomaly correction and the model stacking approach was confirmed to be statistically significant, indicating that the results were not merely coincidental or specific to a particular data sample. The data used in this study was real-world time-series data measuring water levels in underground storm sewers, which are inherently susceptible to various factors and missing values.

In real-world environments, factors like communication failures, sensor inactivity during battery changes, changes in internal/external temperature and humidity, and sensor malfunctions due to nearby construction are frequent. These issues degrade water level data quality and increase the likelihood of false anomaly detections. Considering these variables, the experiment did not just analyze detection accuracy but also quantitative prediction error metrics such as Mean Squared Error (MSE). The results showed that prediction models trained on the anomaly-corrected time-series data achieved a remarkable 59.17% average decrease in MSE compared to models without correction. This indicates that the anomaly detection and correction process made a substantial contribution to improving prediction accuracy. Furthermore, the model's reproducibility was also a key evaluation factor. The stacking-based anomaly detection model maintained stable performance with an MSE variation of only ±1.7% in repeated experiments under identical conditions and parameter settings. This supports the notion that the proposed technique is less sensitive to external environmental changes or temporary anomalies and can be operated with high reliability in a real-time system.

4.2.1 Anomaly Detection and Correction Results

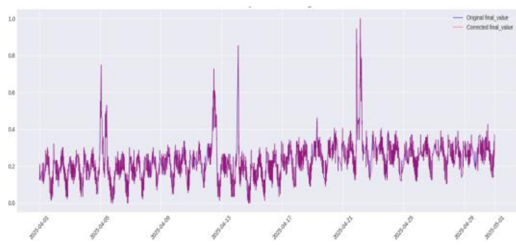
Anomaly detection and correction experiments were performed on the collected water level time-series data. The data, measured at 1-minute intervals over one month in April 2025, contained frequent abnormal patterns caused by sensor malfunctions, communication failures, and extreme rainfall, which are characteristic of underground sewer systems in low-lying urban areas.

The LSTM-based time-series anomaly detection model identified sections with high reconstruction errors, particularly focusing on periods with sudden increases or decreases in water level. Sudden water level changes that did not correlate with rainfall were also detected, which were presumed to be abnormal noise or external disturbances. As shown in Figure 7, the feature engineering metric derived from the difference between rainfall and water level also reflected similar anomaly patterns.



(Figure 7) Features Engineering Results Between Precipitation and Water Level

For the data identified as anomalous, linear interpolation was applied for correction. The interpolation technique replaced the values at the anomalous time points with a linear function connecting the preceding and succeeding normal values. This method was effective in preserving the continuity and pattern of the time series while mitigating the anomalous signals. As shown in Figure 8, during the mid-April period with numerous anomalies, the corrected time series showed a smoother change compared to the raw data, with the sudden change patterns being mitigated. The correction process also had a consistent effect on the derived feature engineering metrics, where the abrupt amplitude of the difference metric was smoothed out. This contributed to the overall stability and interpretability of the time series, ensuring a data quality that reduced the potential for distortion in subsequent analysis and model training.



(Figure 8)Correction results of data from LSTM model abnormalities

It should be noted that linear interpolation is only effective when a sufficient normal data segment is available. If the correction period is long or nonlinear patterns exist, it may risk distorting the original dynamic pattern. Therefore, future research should include a performance comparison of different correction techniques and a quantitative evaluation of the impact of the corrected data on prediction models and decision-making processes. Overall, linear interpolation-based anomaly correction proved to be an effective method for improving the quality of water level data and derived metrics and can be used as a core foundational step for ensuring the reliability of time-series-based applications such as flood prediction and urban flood response systems.

4.2.2 Water Level Time-Series Data Prediction Results

Three deep learning models—LSTM, ConvLSTM, and LSTM-Transformer—were trained and evaluated using the anomaly-corrected time-series data. The models' performance was assessed on a test dataset by comparing their predictions to actual water level values using key metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and R2 Score.

The evaluation results for the test dataset, summarized in Table 3, show that the LSTM model recorded the lowest MSE (0.000089) and MAE (0.006778) compared to the other models. This indicates that the LSTM model's predictions had the smallest average error relative to the actual values, suggesting it performed the most accurate predictions on the test dataset.

(Table 3)Comparison of Prediction Performance by Model

Model Name	MSE	MAE	R2 Score	RMSE
LSTM	0.000089	0.006778	0.972614	0.009434
ConvLSTM	0.000218	0.011817	0.932757	0.014765
LSTM Transformer	0.000498	0.018368	0.846364	0.022316

The R2 Score, which measures how well the model explains the variance of the target variable (water level), was highest for the LSTM model at 0.972614. This means the LSTM model was the most successful at capturing and explaining the variability in the underground water level time series. Overall, the LSTM model demonstrated superior predictive performance compared to both the ConvLSTM and LSTM-Transformer models. While the ConvLSTM model performed slightly worse than the LSTM, it still surpassed the LSTM-Transformer model, which recorded the lowest performance among the three.

The results suggest that for the anomaly-corrected time-series water level data from the target region, the LSTM model is the most suitable for prediction. However, to enhance model generalization and ensure stable predictions under various conditions, further efforts such as hyperparameter tuning, exploring different feature engineering techniques, and validating with an expanded dataset are necessary. For the ConvLSTM and LSTM-Transformer models, their more complex architectures suggest they have the potential for higher performance depending on the data's characteristics and quantity. Additional architectural adjustments and improved training strategies could lead to performance enhancements for these models.

5. Conclusion and Future Plans

5.1. Conclusion

To expand into a proactive response system for minimizing urban flood damage, this study proposed a system that utilizes time-series-based anomaly detection and correction alongside a deep learning-based LSTM network for inflow water level prediction. By ensuring the reliability of real-time rainfall and inflow water level data and predicting flood situations with high accuracy, the system supports efficient, data-driven decision-making. Experiments on a real-world dataset confirmed that the proposed model achieved over 95% accuracy in both anomaly detection/correction and prediction, representing an improvement of over 95% compared to existing prediction models. These findings suggest that the system can contribute to the systematization of urban flood response and, ultimately, to minimizing casualties and property damage.

5.2 Thesis Limitations and Solutions

While this study proposed an effective AI-based system for urban flood prediction, it contains several limitations that should be addressed in future work.

First, obtaining a sufficient quantity of high-quality time-series rainfall and water level data, particularly for various flood scenarios, remains a challenge in real urban environments. To overcome this, the study proposes a solution of using simulation results from physics-based hydrological/hydraulic models (e.g., HEC-RAS 2D) as synthetic data for training the AI model. Additionally, data augmentation techniques or transfer learning from similar regions could be explored to mitigate the data scarcity problem.

Second, a model trained on data from a specific region may have limited generalization capabilities when applied to other areas with different topographical and hydrological characteristics. A solution would be to explicitly include unique geographical and environmental features of each region, such as watershed area, slope, land use type, and soil characteristics, as model inputs. Adopting meta-learning techniques to develop a model that can quickly adapt to diverse environments is another promising approach.

Third, the current model provides a single prediction value without explicitly offering information on prediction uncertainty (e.g., a confidence interval). To address this, uncertainty quantification techniques such as Monte Carlo Dropout or Bayesian Deep Learning could be integrated to provide reliability intervals, which would more effectively support the decision-making process.

5.3 Future Plans

Future research will aim to deepen and expand the study of urban flood prediction in a multi-faceted manner to maximize accuracy and efficiency.

First, a new hybrid model structure that combines a Large Language Model (LLM) with the existing architecture will be designed. This will involve exploring optimized hyperparameter tuning and training strategies to maximize prediction accuracy and efficiency.

Second, research will focus on enhancing prediction accuracy through multi-sensor fusion and spatio-temporal data analysis. The goal is to fuse diverse heterogeneous sensor data, such as sewer pressure, soil moisture, and IoT-based CCTV video. Advanced deep learning models like Graph Neural Networks (GNNs), which can account for spatio-temporal correlations, will be explored to model the complex flow of water within the urban network. Furthermore, virtual sensor research will be introduced in areas where physical sensors are difficult to install, a step that will address the data scarcity problem and expand the prediction range.

Third, the study plans to link the prediction results to an automated control and decision-making support system. This would allow for the automatic execution of practical flood prevention measures or the proposal of optimal response strategies to decision-makers. This integration could be optimized through the learning of control policies based on Reinforcement Learning.

Finally, to ensure the continuous operation of a real-time prediction system, energy-efficient model implementation will be explored. This will involve researching model lightweighting techniques (e.g., Model Quantization, Pruning, Knowledge Distillation) and implementing a system based on Edge Computing to enable high-performance prediction even in resource-constrained environments.

References

[1] Bargam, B., Boudhar, A., Kinnard, C., Bouamri, H., Nifa, K., & Chehbouni, A. (2024). Evaluation of the support vector regression (SVR) and the random forest (RF) models accuracy for streamflow prediction under a data-scarce basin in Morocco. Discover Applied Sciences, 6(6), Article 306. https://doi.org/10.1007/s42452-024-05994-z

[2] Boateng, E. A., & Bruce, J. W. (2023). Unsupervised ensemble methods for anomaly detection in PLC-based process control. arXiv preprint arXiv:2302.02097.

[3] Braei, M., & Wagner, S. (2020). Anomaly detection in univariate time-series: A survey on the state-of-the-art. arXiv preprint arXiv:2004.00433. https://arxiv.org/abs/2004.00433

[4] Fan, Z., Zhang, J., Chen, Y., & Xu, H. (2025). Urban flood prediction model based on Transformer-LSTM Sparrow Search Algorithm. Water, 17(3), 385. https://doi.org/10.3390/w17030385

[5] Freeman, C., Merriman, J., Beaver, I., & Mueen, A. (2021). Experimental comparison and survey of twelve time series anomaly detection algorithms. arXiv preprint arXiv:2109.08055. https://arxiv.org/abs/2109.08055

[6] Oddo, P. C., Bolten, J. D., Kumar, S. V., & Cleary, B. (2024). Deep convolutional LSTM for improved flash flood prediction. Journal of Hydrology, 627, 130260. https://doi.org/10.1016/j.jhydrol.2023.130260

[7] Taylor, S. J., & Letham, B. (2017). Forecasting at scale. The American Statistician, 72(1), 37-45. https://doi.org/10.1080/00031305.2017.1380080

[8] Shahanas, S., Akthar, A., Anand, S., Rakshitha, & Amirthavalli, M. (2024). Anomaly detection in time series data in IoT environment. International Journal of Innovative Research in Technology, 11(5).

[9] Widiasari, I. R., & Efendi, R. (2024). Utilizing LSTM-GRU for IoT-based water level prediction using multi-variable rainfall time series data. Informatics, 11(4), 73. https://doi.org/10.3390/informatics11040073

Author Biography

Jang-won Lee

Graduated from Seoul School of Integrated Sciencs & Technologies(aSSIST), Graduate School of AI, Department of AI Advanced Studies, Major in AI Big Data(Master of Engineering)

Research Interests : AI, Time Series Data Forecasting, Digital Transformation, Generative, Data Analysis.

E-mail : lessle23277@gmail.com

Nak Hyun Jung

Research Professor,The Institute for Industrial Policy Studies Visting Professor, Seoul School of Integrated Sciences & Technologies(aSSIST)
Head of Process Management Team, Mirae Asset Securities
Ph.D. in Business Administration , aSSIST University
Research Interests : AI, Finance, Time Series, Business Strategy
E-mail : nhjung.phd@gmail.com