ENHANCING FLOOD ROUTING SIMULATION OF DAM SAFETY REMOTE MONITORING SYSTEM BY MACHINE LEARNING TECHNIQUE

[1]KANOKVATE TUNGPIMOLRUT, [2]RANGSARIT VANIJJIRATTIKHAN, [3]JITTIWUT SUWATTHIKUL, [4]MANUT TANTICHAROEN, [5]CHIRAPOT VUTHIPRAPHAN, [6]ARNAN JOMTARAX, [7]SEUBSUANG KACHAPORNKUL, [8]SHINICHI MIYAKAWA, [9]SHINSUKE MIWA, [10]TOSHIYUKI MIYACHI, [11]JESSADDA KARNJANA

[1,2,3,6,7,11]National Electronics and Computer Technology Center, Thailand
[4,5] Sirindhorn International Institute of Technology, Thailand
[8,9,10] National Institute of Information and Communications Technology
E-mail: [1]kanokvate.tungpimolrut@nectec.or.th, [2]rangsarit.vanijjirattikhan@nectec.or.th, [3]jittiwut.suwatthikul@nectec.or.th, [4]6522780616@g.siit.tu.ac.th, [5]6522781598@g.siit.tu.ac.th, [6]arnan.jomtarax@nectec.or.th, [7]seubsuang.kachapornkul@nectec.or.th, [8]miyakawa@nict.go.jp, [9]danna@nict.go.jp, [10]miyachi@nict.go.jp, [11]jessada.karnjana@nectec.or.th

Abstract - In this paper, a concept for integrating machine learning–based inflow forecasting into the flood routing simulation of a dam safety remote monitoring system is proposed. A Fully Connected Neural Network (FCNN) is selected as the forecasting model for reservoir inflow at the Ubol Ratana Dam, used here as a case study. Based on nine years of recorded daily historical environmental parameters at the dam site, preliminary results indicate that the model achieves improved performance, with a correlation coefficient ranging from 0.60452 to 0.82921. This is higher than the results obtained from a model based on monthly and global parameters, which produced a correlation coefficient of approximately 0.40. The 15-day-ahead forecasted inflow aligns well with the actual inflow. When integrated into the flood routing simulation, a similar pattern of spillway gate operations is observed, and the reservoir water level can be effectively managed. These findings validate that the proposed concept can enhance flood routing simulations by expediting the process of identifying suitable hydrographs that align with realworld conditions—traditionally done manually and randomly—thereby making the system more adaptive.

Keywords - Dam Safety Management, Flood Routing Simulation, Machine Learning, Reservoir Operation.

I. INTRODUCTION

Flooding is one of the major natural disasters worldwide, including in ASEAN countries. With the increasing impact of severe climate change, floods now significantly affect both human lives and various forms of social infrastructure. In predominantly agricultural regions such as those in ASEAN, dams are considered essential tools for water management, serving to mitigate the impacts of both floods and droughts. Effective dam operation has therefore become crucial for ensuring water security and preventing severe flooding. Some hydraulic and hydrologic analysis methods and software platforms are fundamental tools used by dam operators for the calculation of inundation maps, the planning of dam operations, and the development of emergency action plans, including evacuation routes during crises. In conventional approaches, flow rates and water levels at various upstream and downstream locations are calculated using physical simulation methods such as finite element analysis and other numerical modeling techniques [1]. These simulations require detailed 2D or 3D
geographical information and a wide range of parameters - including soil absorption properties, water evaporation rates, and surrounding environmental conditions. These parameters vary over time and are influenced by many factors, making them difficult to measure accurately across large, open areas. This directly affects the precision of simulation results.

Recently, techniques for more accurately forecasting Reservoir Water Level (RWL), which is determined by the
balance between total inflow and outflow, have been proposed but this remains a highly challenging problem. Various approaches have been introduced, including smoothing techniques [2], averaging techniques [3], statistical learning methods [4], and machine learning techniques [5–6]. Scenario-based RWL forecasting and flood routing simulations using historical hydrographs have also been proposed [7]. Inflow forecasting using machine learning has shown promising potential. For example, one study [8] used 20 years of monthly global indices—such as those from the North Atlantic region—to forecast inflows for three dams located in South Korea. While the results demonstrated the potential of machine learning methods, the overall performance was still limited. Flood routing simulation remains a widely used method to help dam operators determine the most effective operational strategies under various scenarios. This method uses historical inflow data, known as inflow hydrographs, to estimate the optimal spillway gate openings in order to balance downstream outflow and reservoir water levels. Traditionally, these hydrographs are compiled

manually from historical records spanning 5, 10, or even 100 years. Based on this data, upper and lower rule curves are defined for each dam, serving as guidelines for flood routing simulations. However, due to the recent impacts of climate change, inflow patterns have increasingly deviated from historical trends in unpredictable ways. As a result, it has become more difficult to regulate reservoir water levels using conventional methods, and in some cases, reservoir levels have exceeded critical thresholds.

In this paper, we propose a new approach: integrating machine learning based inflow forecasting into flood routing simulations. The proposed inflow forecasting model leverages recent datasets collected from the dam area and enables more dynamic adaptation, rather than relying solely on static historical hydrographs. This approach aims to improve prediction accuracy and support more adaptive and rapid dam-operation strategies in response to rapidly changing climate conditions.

Furthermore, the simulations are performed on the CyReal platform [9-10], which is built on the StarBED testbed [11], developed by the National Institute of Information and Communications Technology (NICT) in Japan. This facility offers a large-scale environment for emulating complex network topologies using both physical and virtual modes. It also supports technologies such as IoT, Cyber-Physical Systems, and digital twins, enabling comprehensive investigations of ICT systems through simulation, emulation, and real-world implementation in an integrated manner. The remainer of this paper is organized as follows: Section II describes the implementation of a dam safety management and water governance system in Thailand, designed to collect the dataset required for machine learning based inflow forecasting and flood routing simulation. Section III outlines the principles of the selected machine learning model used for inflow forecasting in the flood routing simulation. Section IV presents the preliminary simulation results used to validate the effectiveness of the proposed concept.

II. DAM SAFETY MANAGEMENT AND WATER GOVERNANCE SYSTEM – CASE STUDY IN THAILAND

A. Dam Safety Remote Monitoring System

For large dams, many sensors have been installed in the dam body and surrounding areas to monitor and assess safety conditions starting from the construction phase. Monitoring and assessment criteria generally follow the guidelines issued by the International Commission on Large Dams (ICOLD) [12]. Traditionally, data from each sensor, installed at different parts of the dam structure, are collected manually by staff. This process is time-consuming and prone to human error. In addition, data processing and analysis often become bottlenecks, hindering accurate and timely decision-making. These issues are even more critical during crisis situations, such as heavy rainfall accompanied by thunderstorms. To address these challenges, the Dam Safety Remote Monitoring System (DS-RMS) has been developed [13]. The system has been installed at 14 large dams across Thailand, including Ubol Ratana Dam. Both hardware and software components were developed locally, and key features of the system are briefly described below.

a) Hardware components: The Remote Terminal Unit (RTU) is a key component of the DS-RMS system. It is a microcontroller-based device developed to collect data from various sensors with different output signal interfaces, including 4–20 mA, 0–12 V, serial communication, and vibrating wire. The RTU is also equipped with communication modules that support fiber optic and 4G mobile signals. The RTU module, surge protection circuit, and power supply are housed in a waterproof steel enclosure, as shown in Fig. 1. The RTU box and some sensorssuch as the weather station, which collects environmental parameters including rainfall, temperature, humidity, wind speed, and air pressure-are installed both inside and outside a concrete cabinet. Since the entire system is used outdoors 24/7, all hardware components are designed to comply with and pass IEC standards for EMI/EMC and surge protection, ensuring high reliability.

b) Software Components: The DS-RMS system consists of two main software components: one for data collection/visualization, and the other for decision-making and warning systems. All data collected from the RTUs is transmitted to the database (Microsoft SQL Server) for further analysis to support dam safety management and water governance. Environmental parameters are collected at predefined intervals-for example, every four hours-while some operational data such as reservoir water levels and outflows (through generators and spillways) are recorded daily. The inflow values are estimated using a simple relation between the current and previous water levels and outflows, as described in [7]. Since its deployment, the system has recorded approximately 3,000 days of estimated inflow values flowing into the reservoir.

All data stored on the dam server is replicated to a central server, as shown in Fig. 2. Dam site staff can monitor the stored data via a web-based application for day-to-day operations and maintenance. With the help of the developed expert system, all stored data is automatically interpreted to detect any warning conditions. In case of an alert, SMS and email notifications are sent to field operators for immediate inspection and resolution. Additionally, dam safety committee members and operators at headquarters

can access all necessary data and information through the central server. In some cases, warning messages are also sent to selected representatives in downstream communities via a mobile application.

Fig. 3 illustrates the data flow of the DS-RMS system. Raw data is processed and converted into engineering data, which can be visualized and analyzed to evaluate dam behavior. Moreover, this engineering data serves as input to the expert system for safety evaluation and warning generation. The automated data processing saves time and enables multi-dimensional data analysis, which simplifies dam safety assessments. The system also supports flood routing simulations, which can be performed via the DS-RMS web interface from any location, including both the dam site and the headquarters.

B. Flood routing simulation

Fig. 4 shows an example of dam operation during a crisis



Fig. 1. RTU module installed in RTU box.

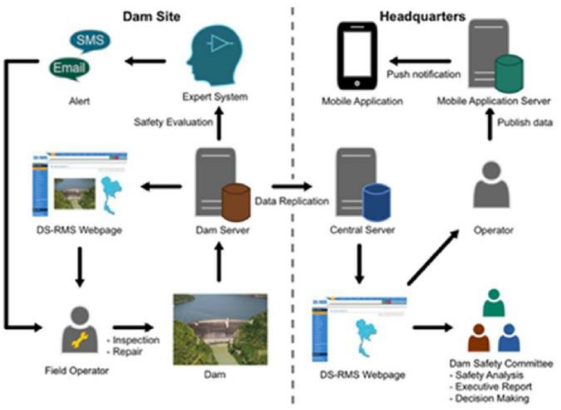


Fig. 2. DS-RMS system diagram.

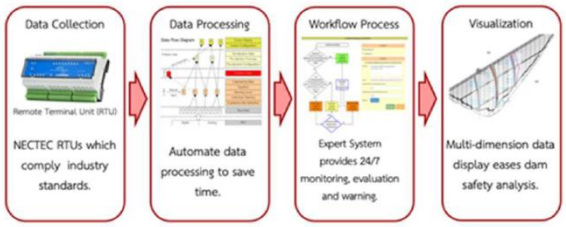


Fig. 3. Dataflow of DS-RMS.

and the need for flood routing simulation at Ubol Ratana Dam in 2017. Typically, dam operators maintain the

Reservoir Water Level (RWL) between the Lower Rule Curve (LRC) and the Upper Rule Curve (URC), which are determined based on historical data. However, when unexpected or unforecastable storms impact the dam area, the resulting increase in RWL due to excessive inflow can push the level beyond the critical threshold of the Normal High Water Level. In such cases, flood routing simulation becomes essential to balance the risk of dam failure with the potential flooding impacts downstream. Flood routing simulation is a core component of the DSRMS. It involves estimating changes in reservoir water levels and outflows in response to varying inflows, spillway operations, and hydrological conditions. The simulation is implemented using hydrologic methods such as the storage– indication routing technique. This approach uses the conservation of mass while simplifying the conservation of



Fig. 4. Water situation at Ubol Ratana Dam in 2017.

momentum, making it well-suited for reservoirs with relatively level water surfaces. Inputs such as elevation–storage data, elevation– discharge relationships, and forecasted inflow hydrographs are fed into the simulation, allowing dam operators to predict reservoir water levels at hourly or sub-hourly intervals. A distinctive feature of modern flood routing simulations is the integration of configurable, humanreadable spillway operation rules stored in databases and represented in tree-based structures. These rules link conditions—such as reservoir water level ranges and gate opening states—to specific operational actions. The outputs, delivered as CSV files, provide time series data for inflows, outflows, reservoir water levels, and gate openings. Validated against dam construction manuals and historical events, these simulations have demonstrated high accuracy, enabling more reliable decision-making during both flood and drought conditions.

III. MACHINE LEARNING MODEL FOR INFLOW FORECASTING

The inflow into the reservoir directly affects the reservoir water level and it is one of the main causes of dam failure, particularly overtopping. Overtopping occurs when the amount of accumulated water in the reservoir exceeds its designed capacity, resulting in excessive pressure on the dam structure, which can ultimately lead to structural failure from the top. The primary source of inflow is rainfall in upstream areas. Due to the wide area coverage, varying terrain, and differing soil conditions, direct measurement of inflow is challenging.

Therefore, inflow is typically estimated indirectly using the reservoir water level and recorded outflow data. During the rainy season, a massive volume of inflow over consecutive days can cause the reservoir water level to rise excessively, increasing the risk of dam failure. Conversely, in the dry season, maintaining a high reservoir water level is essential to prevent drought conditions. Thus, more accurate long-term inflow forecasting, for example over a 15-day period, is crucial for effective dam safety management and water governance.

Various approaches to inflow forecasting have been proposed. In this paper, we select a simple machine learning model, the Fully Connected Neural Network (FCNN), to evaluate its performance in inflow forecasting compared to the previous Long Short-Term Memory (LSTM) model [8] that uses global indices, as well as its effectiveness when integrated into flood routing simulations.

A. Fully Connected Neural Network

Inflow is a type of time series data that can be forecasted using machine learning approaches. Various models, such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, have been proposed for time series forecasting tasks. However, these models can be overly complex and time-consuming to train. In this paper, a Fully Connected Neural Network (FCNN) is used for inflow forecasting.

The inflow is forecasted using six environmental variables collected from the dam site area. As the FCNN does not maintain hidden state memory like RNN or LSTM models, its structure is simpler. While this simplicity improves training efficiency, the lack of memory for long sequences may reduce the effectiveness of the model in capturing long-term dependencies. The mathematical formulas for hidden state and output of FCNN can be expressed as follows:

$$h = \sigma(W_{xh}x + b_h) \quad (1)$$

$$y = \sigma(W_{hy}h + b_y) \quad (2)$$

where σ is the activation function Rectified Linear Unit (ReLU) and x is input vector. W_{xh} and W_{hy} are weight matrices from input to hidden layer and from hidden layer to output layer, respectively. x_h being input to hidden and h_y being hidden to output. b_h and b_y are hidden bias and output bias, respectively.

The ReLU Activation Function can be expressed as follows:

$$f(x) = \begin{cases} 0 \; for \; x < 0 \\ x \; for \; x \geq 0 \end{cases} \quad (3)$$

Since inflow is a real-valued quantity, the output must not be negative. ReLU (Rectified Linear Unit) converts all negative values to zero. It is not a linear function and is simple to compute. The simplicity of using only positive values and zeros makes it efficient for this type of data.

In this paper, Huber loss is used for loss evaluation during the training phase. Huber loss combines Mean Squared Error (MSE) and Mean Absolute Error (MAE), making it less sensitive to outliers. It is defined by the following equation:

$$L_\delta(a) = \begin{cases} \frac{1}{2}a^2 & for \; |a| \leq \delta, \\ \delta.\left(|a| - \frac{1}{2}\delta\right), & otherwise \end{cases} \quad (4)$$

where a is the error, i.e., the difference between a predicted value and the target value, and δ is the adjustable threshold parameter, defaulted to 1. MSE is used if $|a|$ is smaller or equal to δ, and MAE is used otherwise.

B. Layers and parameters

The model has three FCNN layers consisting of 64 units with ReLU activation function, 32 units with ReLU activation function, and 1 unit with linear activation function. Six environmental parameters—rainfall, wind speed, humidity, temperature, reservoir water level, and air pressure—were collected by RTUs from 2016 to 2024 and used as inputs to the model. A total of 3,000 daily data samples were used. Since the reservoir water level is directly related to inflow, only one previous sample of the reservoir water level is used as input. The dataset was split into 70% training data, 20% validation data, and 10% test data. The number of epochs and batch size were initially set to 100 and 32, respectively. The min-max function was applied for data normalization. The Huber loss function was used. During experimentation, some settings and parameters were adjusted to identify the best-performing model.

IV. SIMULATION RESULTS

A. Fully Connected Neural Network

In this test, each model was retrained three times to assess variation between iterations. The model with 500 epochs, a batch size of 32, and no shuffling, produced the highest Correlation Coefficient (CC) between the predicted inflow and the ground truth. The values of Root Mean Square Error (RMSE) and CC for the training, validation, and test sets in each

iteration are shown in Table 1. The plots of the loss and CC for each epoch are shown in Figs. 5. The plot comparing the forecasted inflow with the actual inflow is shown in Fig. 6.

B. Integrated Flood Routing Simulation

The integrated flood routing simulations were conducted under two conditions: one using actual inflow data and the other using forecasted inflow data from the inflow forecasting module.

The forecasted inflow exhibited peak values that aligned well with the actual inflow, resulting in a similar gate-opening pattern. Consequently, the long-term reservoir water level could be managed more effectively and in a timely manner during crisis conditions based on the long-term inflow forecasting. In the actual inflow case, as shown in Fig. 7(a), the inflow hydrograph displayed multiple sharp peaks with magnitudes exceeding 1,600 cms.

These rapid inflow surges caused steep increases in reservoir water level (RWL), beyond the 182 m MSL (Mean Sea Level) normal pool level. Due to reservoir storage effects, the RWL peaks lagged several hours behind the inflow peaks. To control the situation, the spillway gates were opened substantially, up to about 4.75 m. This scenario illustrates the necessity for aggressive, real-time operational decisions when faced with sudden extreme inflows to ensure dam safety.

In the forecasted inflow case, as shown in Fig. 9(b), the hydrograph was smoother, with peak inflows of around 1,200-1,350 cms, considerably lower than those recorded in the actual case. The resulting RWL increases were more gradual and gate openings were limited to 1.75 m. While the reservoir response exhibited a similar lag of several hours between inflow and RWL, the smoother profile of the predicted inflow represents a limitation that should be considered.

However, the strong similarity between the forecasted and actual inflows can serve as a guideline for expediting the

	Iteration 1	Iteration 2	Iteration 3
Train-RMSE	11.02	10.49	11.05
Train-Last CC	0.74995	0.76756	0.75172
Train-Best CC	0.75508	0.76756	0.75517
Validate-RMSE	23.39	23.62	23.23
Validate-Last CC	0.66386	0.65031	0.67160
Validate-Best CC	0.66790	0.67848	0.67272
Test-RMSE	12.82	13.28	13.90
Test -Last CC	0.64332	0.63293	0.60452
Test-Best CC	0.81794	0.82921	0.70910

Table. 1 Evaluation comparison

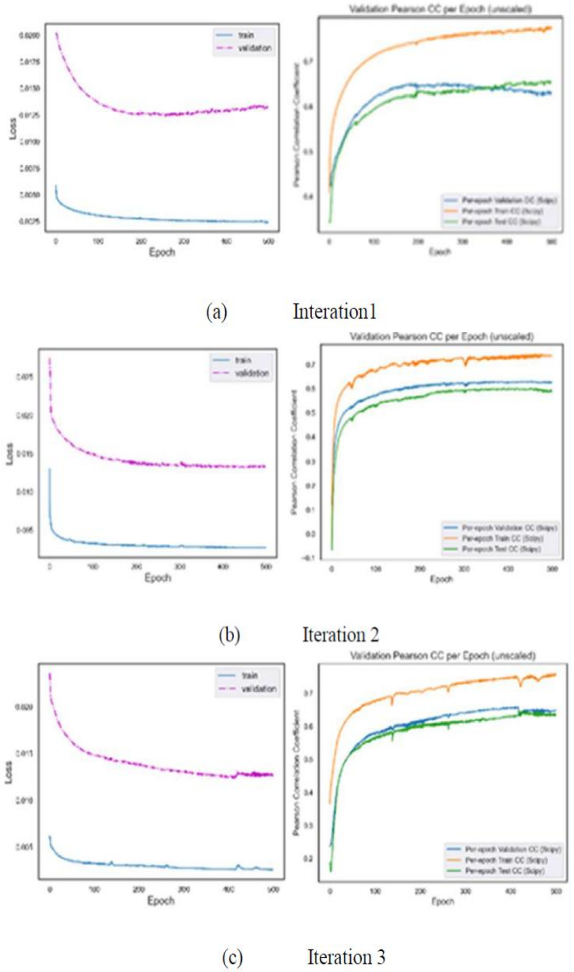


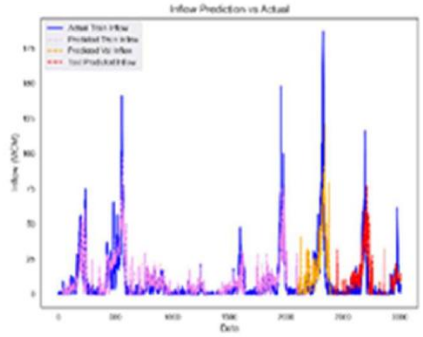
Fig. 5. Loss and correlation coefficient at each epoch.

selection of the most suitable recorded hydrograph for flood routing simulations, a process that is typically performed manually and somewhat randomly by dam operators.

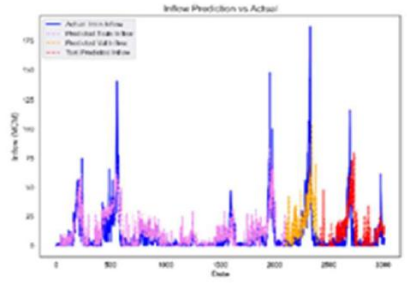
Consequently, the proposed adaptive real-time integrated flood routing simulation, combined with real-world collected data, can effectively reduce the underestimation of extreme inflow magnitudes and timing. This enables dam operators to prepare proactive response plans that closely align with scenarios observed during actual inflow peaks.

V. CONCLUSION

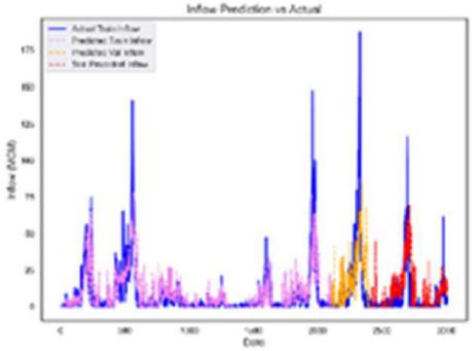
The necessity and implementation of a dam safety remote monitoring system—including flood routing simulation as a main module for balancing dam safety and water governance—have been described. The concept of



(a) Iteration 1



(b) Iteration 2

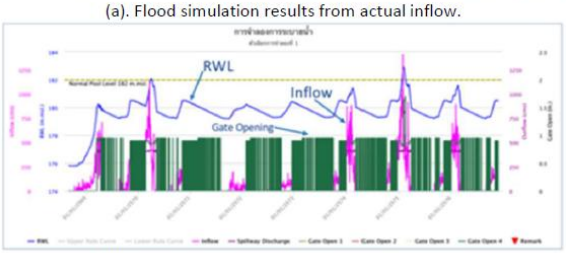


© Iteration 3

Fig. 6. Actual inflow and forecasted inflow.



(a). Flood simulation results from actual inflow.



(b). Flood simulation results from predicted inflow.

Fig. 7. Flood routing simulation results.

integrating machine learning based inflow forecasting into the flood routing simulation to enhance its adaptability and timeliness has also been presented. Preliminary results show improved inflow forecasting performance and demonstrate the effectiveness of the proposed concept in managing reservoir water levels. Longer-term and more accurate inflow forecasting—such as by using regional data shared among ASEAN countries via the CyReal testbed—should be further investigated to enable better dam operations across the region.

VI. ACKNOWLEDGEMENT

This work was produced as part of the ASEAN IVO project (http://www.nict.go.jp/en/aseanivo/index.html),titled ‘Cyber to Real World Integrated Testbed for Dam Safety Management and Water Governance System,’ with financially support from NICT http://www.nict.go.jp/en/index.html).

REFERENCES

[1] S. Sarwari, Y. Qiao and M. Abubakar Rafi, “A Short Review Note on Finite Element Method for Hydraulic Structure Engineering,” Journal of Engineering, Scince and Technology Trends, vol 1, No. 2, August 2024.

[2] Lydiane F. Souza, “Forecast Modeling a Time Series of Water Reservoir Levels using Exponential Smoothing Method.” In https://doi.org/10,48550/arXiv.2306.05199.

[3] A. S. Azad, R. Sokkalingam, H. Daud and S. K. Adhikary, “Application Water Level Prediction Through Seasonal Autoregressive Integrated Moving Average: Red Hills Reservoir Case Study,” in Proc. IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, 2022.

[4] R. Obringer and R. Nateghi, “Predicting Urban Reservoir Levels using Statistical Learning Techniques,” in Scientific Reports, 2018

[5] S. Rani and F. Parekh, “Predicting Reservoir Water Level using Artificial Neural Network,” International Journal of Innovative Research in Science, Engineering and Technology, vol. 3, issue 7, 2014.

[6] F. Unes, M. Demirci, B. Tasar, et al., “Estimating Dam Reservoir Level Fluctuations using Data-driven Techniques,” Pol. J. Environ. Std., vol. 28, issue 5, 2019.

[7] R. Vanijjirattikhan, C. Thongthamchart, P. Rakcheep, et al., “Reservoir Flood Routing Simulation for Dam Safety Management in Thailand,” vol.16, issue 4, June 2021.

[8] H. Han, D. Kim, W. Wang and H. S. Kim, “Dam Inflow Prediction using Large-scale Climate Variability and Deep Learning Approach: a Case Study in South Korea.” Water Supply, vol. 23, No. 2, 2023

[9] M. Toshiyuki, I. Yota, M. Shinichi, et al., “CyReal: A Unified Framework for Continuous Evaluation of Simulations, Emulations, and Real Systems,” in Proc. EAI International Conference on Simulation Tools and Techniques (SIMUtools 2024), Bratislava, Slovakia, 2024

[10] K. Hiroi, T. Inoue, K. Akashi, et al., “ARIA: Interactive Damage Prediction System for Urban Flood using Simulation and Emulation Federation Platform,” in Proc. ACM International Joint Conference on Pervasive and Ubiquitous Computing, London, 2019, pp. 284-287.

[11] https://testbed.nict.go.jp/jgn/english/index.html

[12] ICOLD dam safety guideline homepage, https://www.scribd.com/document/661664292/ICOLD-

DAMSAFETY- GUIDELINES-ENGLISH-PAGES-ASSORTED, last accessed 2025/09/08.

[13] J. Suwatthikul, R. Vanijjirattikhan, U. Supakchukul, et al., "Development of Dam Safety Remote Monitoring and Evaluation System," Journal of Disaster Research, vol. 16, Issue 4, Jun 2021.

★★★