WILEY | Hindawi

Research Article

Relationships among Seepage, Water Level, and Rainfall of a Fill Dam by Decision Tree Analysis

Seong-Kyu Yun [ID],[1] Jiseong Kim [ID],[2] Eun-Sang Im [ID],[3] and Gichun Kang [ID][4]

[1]*Engineering Research Institute, Gyeongsang National University, Jinju 52828, Republic of Korea*
[2]*Department of Cadastre & Civil Engineering, Vision College of Jeonju, 235 Cheonjam-ro, Wansan-gu, Jeonju, Jeollabuk-do 55069, Republic of Korea*
[3]*Water Energy & Infrastructure Research Center, K-water, 1689 Yuseongde-ro, Yuseong-gu, Daejeon 34045, Republic of Korea*
[4]*Department of Civil Engineering, College of Engineering, Gyeongsang National University, 501 Jinjudero, Jinju, Gyeongsangnam-do 52828, Republic of Korea*

Correspondence should be addressed to Gichun Kang; gkang@gnu.ac.kr

This study is to present acceptable measurement values through decision tree analysis for the seepage, which is an important measuring item of the fill dam. The seepage of the dam under study increases rapidly when rainfall occurs, when the direct inflow of rainfall from the downstream slope and both sides occur. Therefore, the allowable seepage of fill dam considering rainfall and water level is required. Decision tree analysis was conducted for one domestic fill dam by setting the seepage as a response variable and setting rainfall and water level as explanatory variables. At this time, in order to analyze the effects of rainfall on the seepage more closely, the data subject was classified into two groups consisting of a rainfall-free group and a rainfall-occurring group. Group A, which is a rainfall-free group, had 97.7% of the seepage data distributed under the conditions of 98.50 mm/day of the antecedent 5-day rainfall selected as the first explanatory variable. The average seepage of the group was between 12.01 L/min and 26.35 L/min. Group B, which is a rainfall-occurring group, had 85.7% of the water leakage data distributed under conditions of 38.50 mm/day of daily rainfall selected as the first explanatory variable, with an average of 23.70 L/min.

1. Introduction

In Korea, about 20,000 dams serve an important function of industrial infrastructure. However, recently, the aging of dams, the increasing frequency of earthquakes, and weather fluctuations have threatened the stability of dams, and the development and improvement of design, construction, and maintenance technologies to secure the stability of dams is urgently required. In particular, the need for repair and reinforcement is emerging as multipurpose dams, which have been constructed in Korea since the 1960s and are gradually aging. Among the 37 multipurpose dams and water supply dams in Korea, about 30% have passed more than 30 years, and more than 50% of the 14 water supply dams have passed about 30 to 60 years. Problems with these

dams are thought to not only cause water disasters but also negatively affect the nation's water supply. Therefore, various kinds of measuring instruments such as water seepage meter, pore pressure meter, seismometer, earth pressure gauge, clinometer, and settlement gauge are buried and installed in the dam and used for stability analysis through real-time or manual monitoring [1]. And the data obtained from these various instruments is critical as basic data for maintenance or research of dams [2]. In particular, the amount of seepage flowing through the dam is important for understanding the motion and mysterious symptoms of the dam, and accurately identifying them is a prerequisite for dam management [3]. As shown in Figure 1, the dam is referred to as stage 1 until it reaches normal high-water level for the first time after the commencement of

impoundment, and it is referred to as stage 2 from the first normal high-water level to the time when the dam's movement reaches a safe state. And after the dam's movement has reached a stable state, it is called stage 3. As shown in Figure 1, the seepage is typically at its peak in stage 1 and stabilizes over time with a gradual decrease in the seepage, resulting in a constant value in stage 3.

According to the US Committee on Large Dams [5], 77 cases of collapse of fill dams in the US (accidents that occurred until the 1980s) were analyzed, and the destruction caused by leaks and piping through dam body or foundation ground accounted for 44%. Seepage water meter of the fill dam is an important measuring item that can monitor this type of destruction. In addition, it is necessary for seepage measurement management to establish an acceptable measurement, which means a safe level for the observed values, and to evaluate the stability of the dam by comparing the current measurement with the acceptable measurement. It is important to establish the decision of the acceptable measurement considering many assumptions or environmental conditions included in the design [6]. However, there is an actual difference in the predicted seepage in the design and the measurement of the seepage of the dam considering the surrounding environmental conditions. Therefore, the dam shall analyze test impoundment and the measurement data of operation period to check the normal range of the dam and establish the acceptable measurement based on it. However, it is common to establish acceptable measurements of seepage by using measurements obtained during operation of the dam due to no implementation of test impoundment, absence of instruments, and deterioration of reliability due to the inflow of rainfall.

For the dam measurements, Kuperman et al. [7] considered the behavior of the dam to be normal if the measurements measured from the same instrument are within a certain range under conditions such as water levels similar to the past. Lee [8] calculated the upper and lower limits by methods of the Shewhart control chart method and linear regression analysis according to the aging characteristics within the normal range of dam behavior and presented it as the management criteria. Ryan [9], Lewis et al. [10], and Myers and Montgomery [11] conducted studies on many measurements that did not follow a normal distribution and found that measured values by dam measuring instruments often exhibited asymmetric distributions that did not follow a normal distribution. In addition, Park and Park [12] said that using Shewhart control chart based on normality assumption for control chart of measured values following asymmetric distributions is less efficient in managing measure variability and increases the probability of making errors as asymmetry increases, proposing a quartile control chart as an alternative to this problem. In fact, however, many fill dams have limitations in introducing the method above as the water catchment wall is located at the bottom of the downstream slope and the seepage increases due to the direct inflow of rainfall through the downstream slope.

Decision tree analysis is easy to understand the classification structure of the data and can explain the reasons for the decision-making so that the effect of water levels and
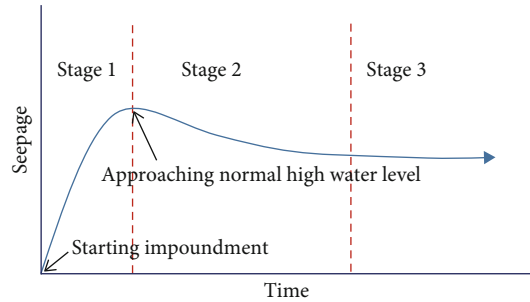

Figure 1: Management stage of fill dams [4].

Table 1: Summary of the dam for study.

Dam type	Center-cored rockfill dam
Dam crest elevation (m)	44
Dam crest wide (m)	8
Upstream slope	$1:2.2$
Downstream slope	$1:1.8$
Dam crest length (m)	108
Dam height (m)	35.2
Dam volume (m^3)	153,000

rainfall, which are highly correlated with the water seepage of the dam, can be considered [4]. In this study, the seepage is determined as a response variable for one fill dam, while rainfall and water level are set as explanatory variables. At this time, in order to analyze the effects of rainfall on the seepage more closely, the decision tree analysis was conducted by classifying measured data into two groups consisting of a rainfall-free group and a rainfall-occurring group, using the daily rainfall and antecedent 5-day rainfall as explanatory.

2. Target Dam for Research

The dam subject to the research is a central core rockfill dam, and the cross section of the dam is constructed in order of filter and rockfill, with the core installed on the axis of the dam and directed upstream and downstream. The main dimensions of the dam are shown in Table 1. The dam was completed in 2007 and has been around 13 years since impoundment (Figure 2).

2.1. Installation Status of Water Seepage Measuring Instruments. The water seepage measurement instrument of the dam under research was installed to identify the changes in the amount of water penetrating through the dam body and the foundation ground and to understand the soundness of the barrier function of the dam. First of all, a barrier water catchment wall is installed at the lower fore-end of the dam, and the wall is connected by a seepage measurement room and an induction pipe. The water seepage of the fill dam to be studied is measured in real time (1 time/hr). Figure 3 shows V-notch, a seepage measurement
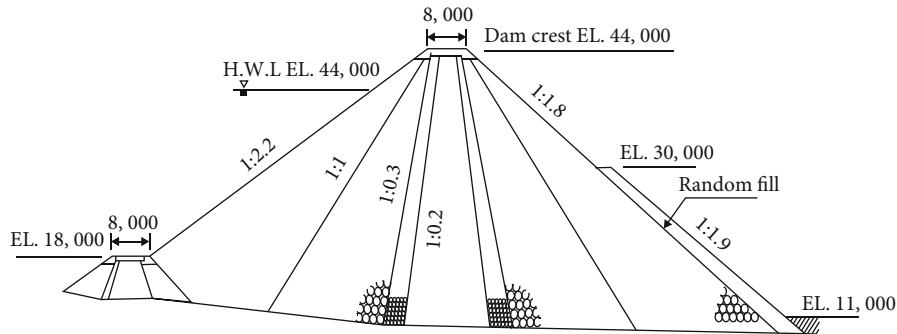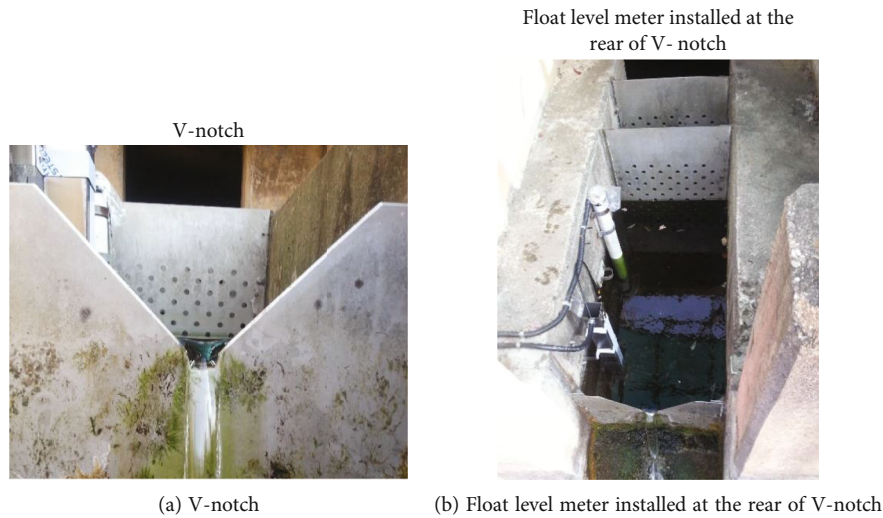
FIGURE 2: Cross-sectional plan (unit: mm).

(a) V-notch

(b) Float level meter installed at the rear of V-notch

FIGURE 3: Instruments for measuring seepage of a fill dam.

instrument installed in the dam, and a water gauge. The dam has one seepage water meter installed, with bottom width of waterway (B) of 0.5 m and height from the bottom of the waterway to the bottom of the V-notch (D) of 0.3 m. The depth (h) and angle (a) of the V-notch are 0.2 m and 90 degrees, respectively. To measure the seepage, a float type water gauge is installed at the entrance of seepage water to automatically measure the height flowing over the V-notch (h) and calculate it by converting it to the rate of flow. In the case of the dam studied, it was installed in the specifications suggested by the International Organization for Standardization ISO [13] so that its accuracy has already been verified [14].

3. Measurement Status

3.1. Storage Level and Rainfall. Figure 4 shows changes in water level and rainfall over time. The water level, which is an important factor in dam management and operation, has an excellent data management with missing rate of 0% within the data collection period. As shown in the figure, a surge in rainfall indicates a rise in the water level. The average water level is EL.35.68 m, and the standard deviation is

2.97 m. For statistical analysis of water levels and rainfall, the measured results from June 1st, 2009 to June 10th, 2019 were used. The average daily rainfall of the reservoir in the dam was 3.27 mm/day, and the maximum daily rainfall was approximately 234 mm/day, which occurred on October 05, 2016.

Another explanatory variable, the antecedent 5-day rainfall, is shown in Figure 5. The reason for the application of the antecedent 5-day rainfall is that the inflow of rainfall usually lasts for several days after rainfall, and the rainfall, which is the standard of the Antecedent Soil Moisture Content (AMC) applied in hydrological flood simulations, is the antecedent 5-day rainfall [15]. The average of the antecedent 5-day rainfall was 16.31 mm/day, and the maximum antecedent 5-day rainfall was approximately 354 mm/day, which occurred on September 18th, 2012.

3.2. Seepage of Water. For statistical analysis of seepage quantities, water levels and rainfall measurement results at the same time were utilized. The average missing rate of seepage during the data collection period was 21.6% (792 days/3,662 days). Also, since the barrier wall is located in the fleet of the downstream, the seepage of water of the fill
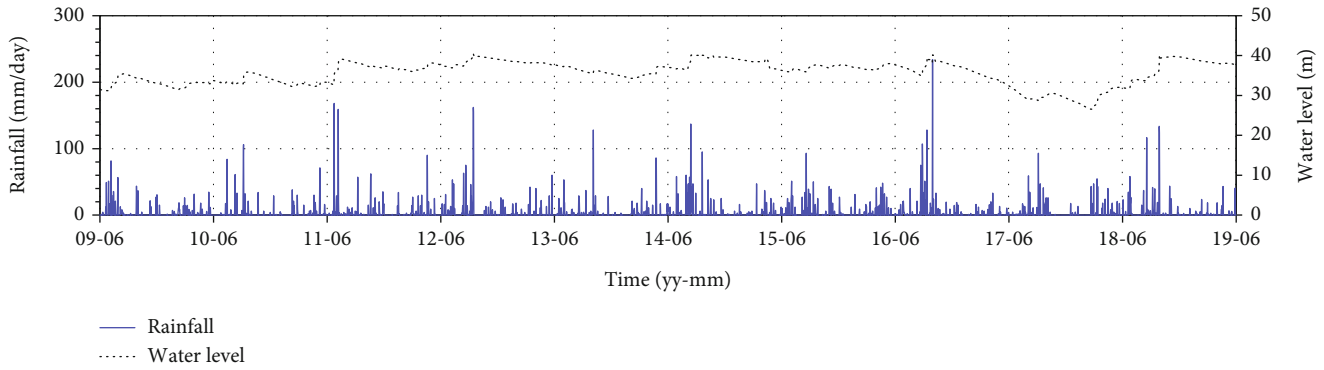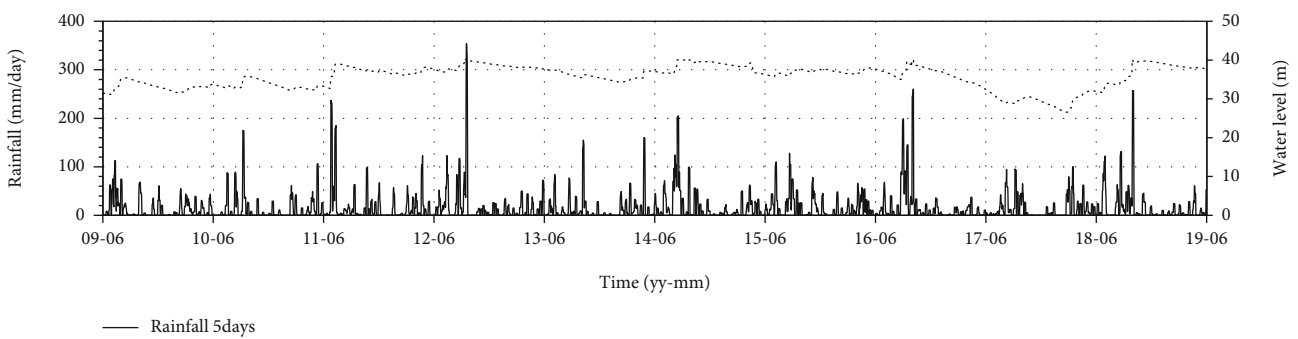
Figure 4: Water level and daily rainfall.

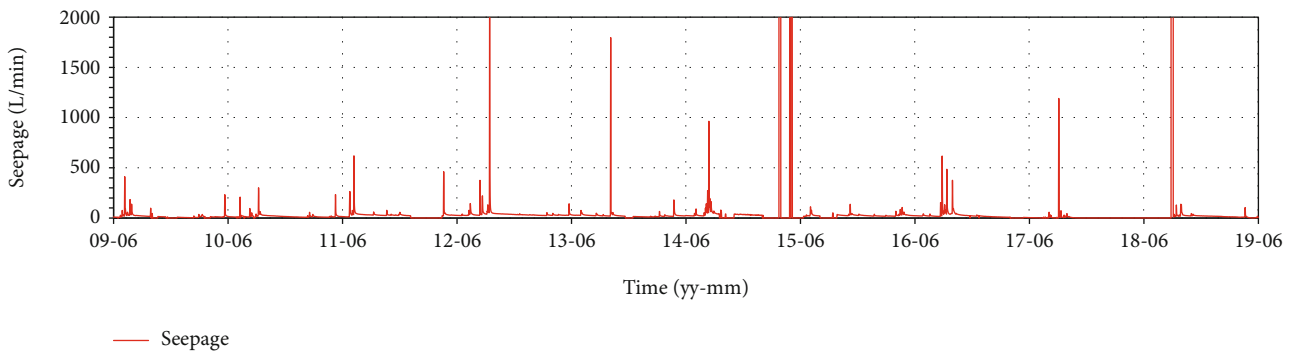Figure 5: Water level and antecedent 5-day rainfall.

Figure 6: Seepage data.

dam increases rapidly when rainfall occurs as the direct inflow of rainfall from the downstream slope and the left and right sides occurs. As shown in Figure 6, potential upper bound outliers are distributed throughout the data, and seepage quantities above 1000 L/min may be measured during rainfall.

3.3. Removal of Outliers of Seepage. Statistical analysis of seepage shows an extreme asymmetric distribution with an average observed value of 178.74 L/min, a median of 18.4 L/min, and a standard deviation of 2,950 L/min. In the case of seepage water meters, it is deemed impossible to remove outliers effectively by searching for outliers based on univariate

such as z-score, since they are measured in one place. Therefore, in this study, a data analysis-based outlier removal method using rainfall and water level data is applied for quantitative analysis of seepage data. The maximum capacity of the V-notch measuring seepage is applied by the Kindsvater-Shen equation presented by the International Organization for Standardization ISO [13], and the calculated maximum possible observation was found to be approximately 1,500 L/min [16]. As illustrated in Figure 7(a), it can be seen that seepage data exceeding the maximum possible observations are distributed throughout the time series data. In addition, Figures 7(b)–7(d), which illustrate the relationship between seepage, rainfall, and water level, show that no extreme values

(a) Raw data for seepage

(b) Raw data for seepage with daily rainfall

(c) Raw data for seepage with antecedent 5-day rainfall

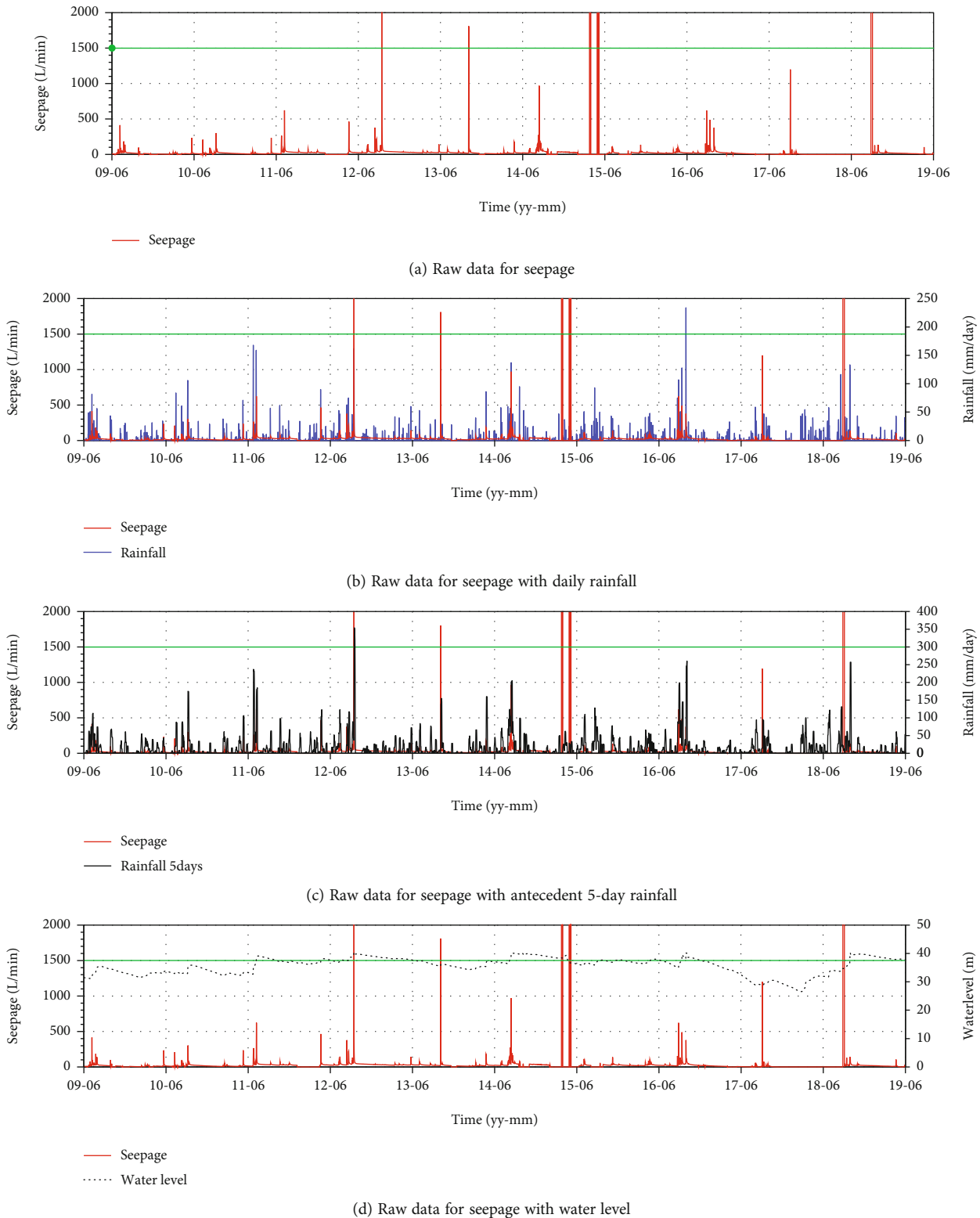(d) Raw data for seepage with water level

FIGURE 7: Comparisons among seepage, rainfall, and water level.

of rainfall and water level occur when the maximum possible observations of seepage are exceeded. Therefore, considering the relationship between the maximum possible observations of seepage, rainfall, and water level, observations exceeding 1,500 L/min were judged as simple outliers and eliminated.

The seepage data contains approximately 22% (missing rate) of observations recorded as "0(zero)" indicating the missing. It was compared to rainfall and water level data to distinguish whether the observation was missing or actual. Figure 8(a) below illustrates the relationship between the
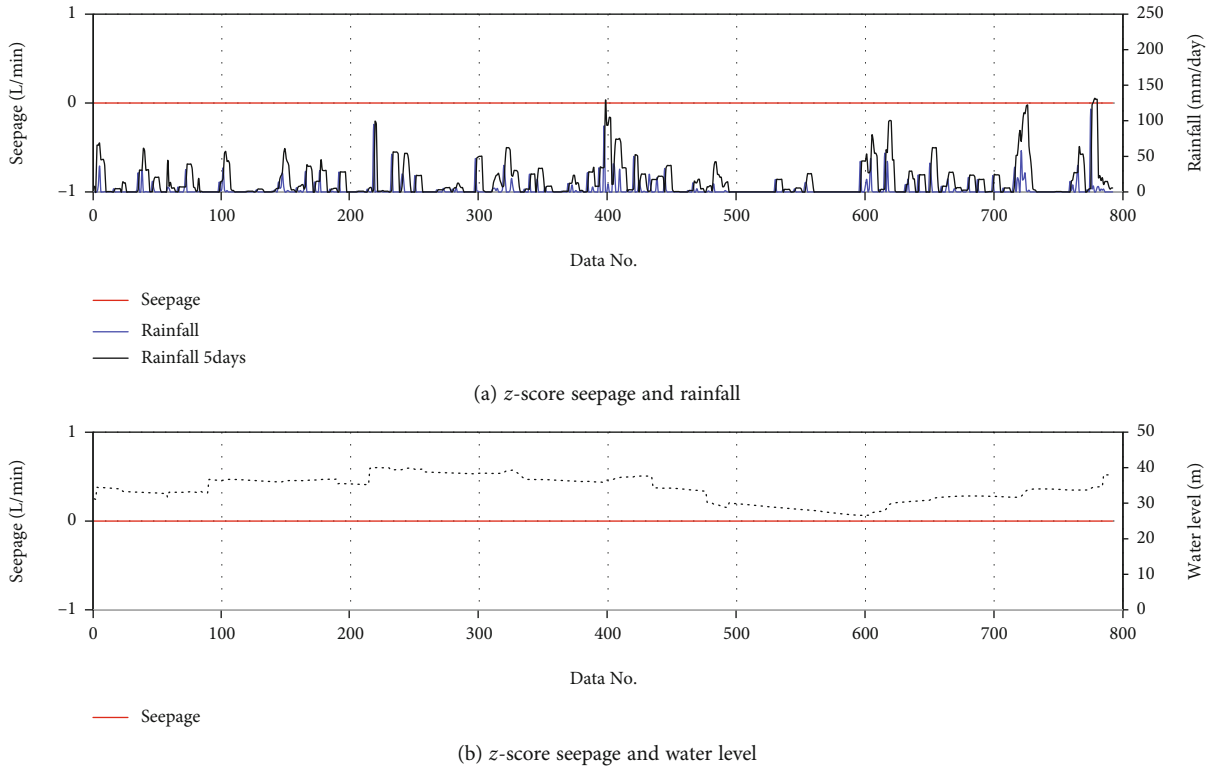
(a) *z*-score seepage and rainfall

(b) *z*-score seepage and water level

FIGURE 8: Time histories for no data.

time recorded as "0," and Figure 8(b) illustrates the relationship with the water level. The observation was eliminated by considering that the actual rainfall occurred when the seepage was recorded as "0(zero)" in Figure 8, and that the change in the water level was distributed above and below the average.

Figure 9 compares the seepage-rainfall and seepage-water level after removing the measurement of seepage water measurements determined to be outliers. In Figures 9(a) and 9(b), which illustrate the seepage and rainfall, the seepage is shown to increase when the rainfall increases significantly. In Figure 9(c), which illustrates the relationship between seepage and water level, it can be seen that seepage increases and decreases with changes in water level as well as rainfall. In other words, the seepage of the dam subject to research is shown to be sensitive to rainfall and water level as response variables.

3.4. Removal of Outliers of Seepage. To closely analyze the effects of rainfall on seepage, the data to be analyzed were classified into two groups consisting of a rainfall-free group (group A) and a rainfall-occurring group (group B). In the case of rainfall, the daily rainfall and antecedent 5-day rainfall in net unit were generated and applied in consideration of the hydrologic response time of the dam basin. The detailed AMC conditions according to the antecedent 5-day rainfall are as shown in Table 2 below [15].

Figures 10 and 11 show the classification into rainfall-occurring group and rainfall-free group, respectively, after removing the seepage outliers and missing values from the

raw data. Of 2854 data excluding 792 missing data and 15 outliers from total measured data of 3,662 data (June 1st, 2009~June 10th, 2019), group A, a rainfall-free group, has 2133 data, while group B, a rainfall-occurring group, has 721 data.

4. Decision Tree Analysis

4.1. Purpose and Process of Decision Tree Analysis. Decision tree analysis is more applicable because it is easy to understand the classification structure of the data and can explain the reasons for the decision, unlike the Neural Network Analysis which is a similar type. Algorithms for tree structure formation in decision tree analysis are currently being developed in various ways. Decision trees start from the roots and are formed by dividing the segmented joints until each branch becomes the end joint. Like this, in order to complete the decision tree, several steps must be performed on the selection of a splitting rule, the selection of a stopping rule to stop the splits, the selection of a pruning method, and, if there is a defect within the input variable value, imputation method [16]. Mainly known algorithms include CART (Classification and Regression Trees), CHAID (Chi-squared Automatic Interaction Detection), C5.0 [17], and C4.5 [18], and this study conducted the seepage analysis using the most commonly used CART algorithm [19].

The CART algorithm is a methodology for generating multiple subset trees of the data and finding the optimal subset tree among them. The CART algorithm is applicable to nominal, ordinal, and continuous variables and is
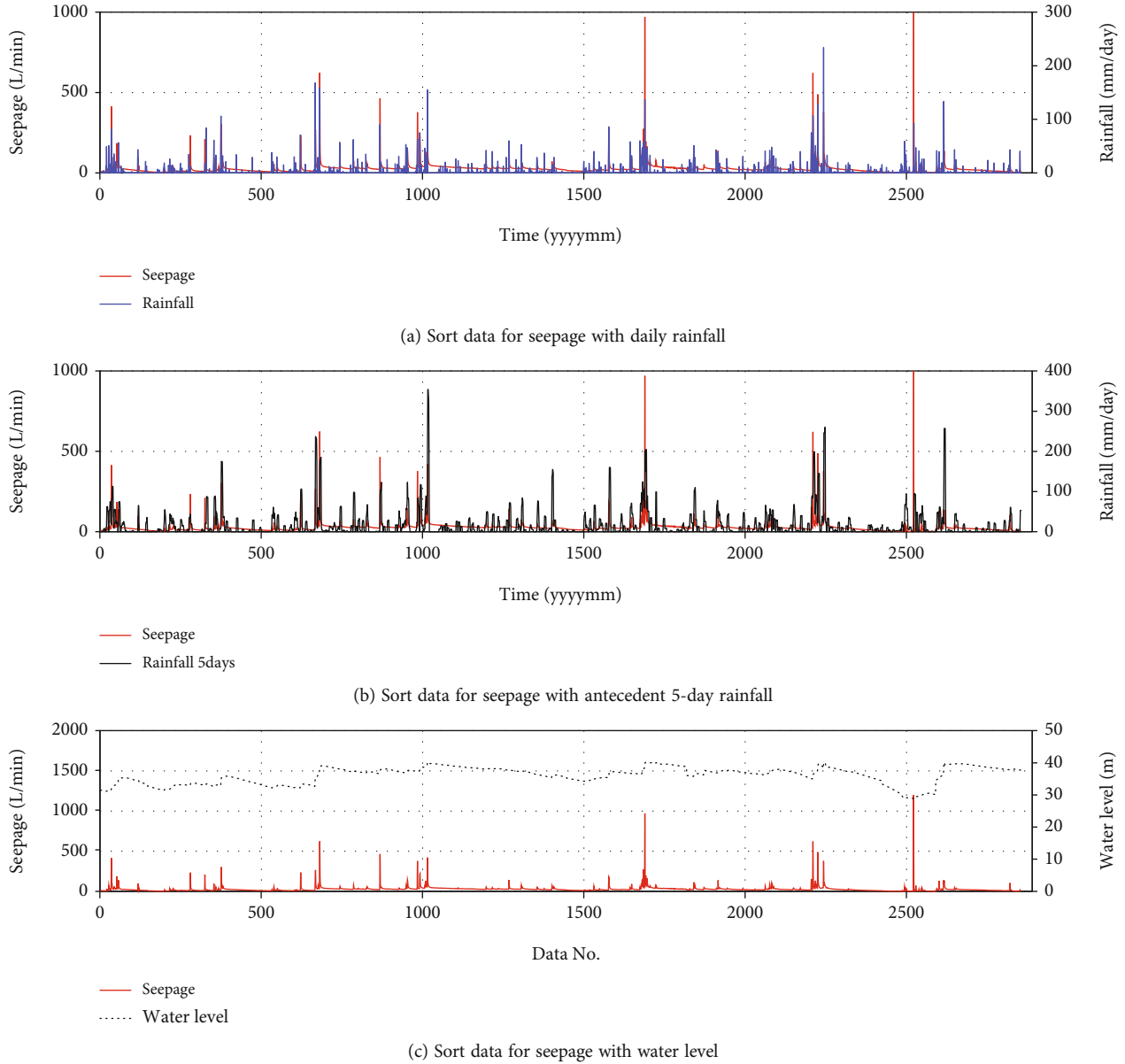
(a) Sort data for seepage with daily rainfall

(b) Sort data for seepage with antecedent 5-day rainfall

(c) Sort data for seepage with water level

Figure 9: Relationship after removal of outliers and missing data.

Table 2: AMC conditions according to antecedent 5-day rainfall [15].

| AMC condition | Antecedent 5-day rainfall P_5 (mm) | |
	Dormant season	Growing season
I	$P_5 < 12.70$	$P_5 < 35.56$
II	$12.70 < P_5 < 27.94$	$35.56 < P_5 < 53.34$
III	$P_5 > 27.94$	$P_5 > 53.34$

characterized by structuring the model's composition according to the conditions of explanatory variables in the order of root nodes, child nodes, and branches, as shown in Figure 12. The root node has the most influential variable among the explanatory variables describing the change in the response variable, and it constructs a binary branch according to the conditions of the corresponding variable. The branch node includes the first explanatory variable or other explanatory variables, and the leaf node is the final-stage node divaricated from the root and each joint, with one leaf node representing clusters according to classified rules. Then, the regression equations estimated from classified cluster units are aggregated to produce a model that can predict the behavior of the response variable.

In particular, the CART algorithm can be proposed as an alternative to regression analysis for the cases such as the presence of interactions of independent variables or the presence of multicollinearity problems. In the CART algorithm, a separation criterion is a criterion for determining the choice of explanatory variables and the merging of
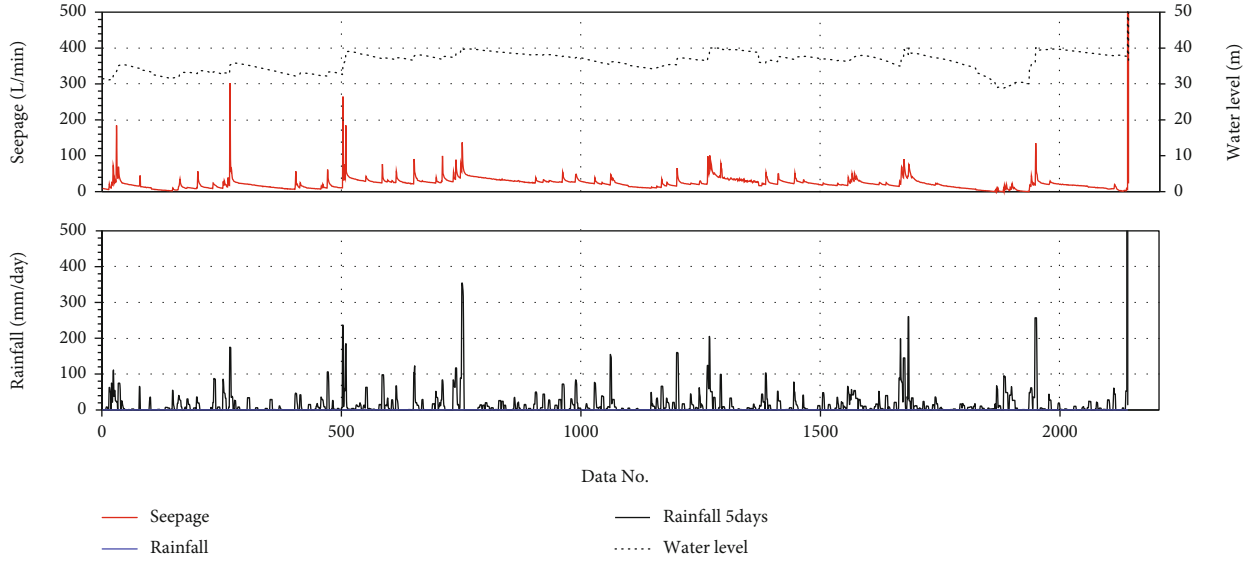
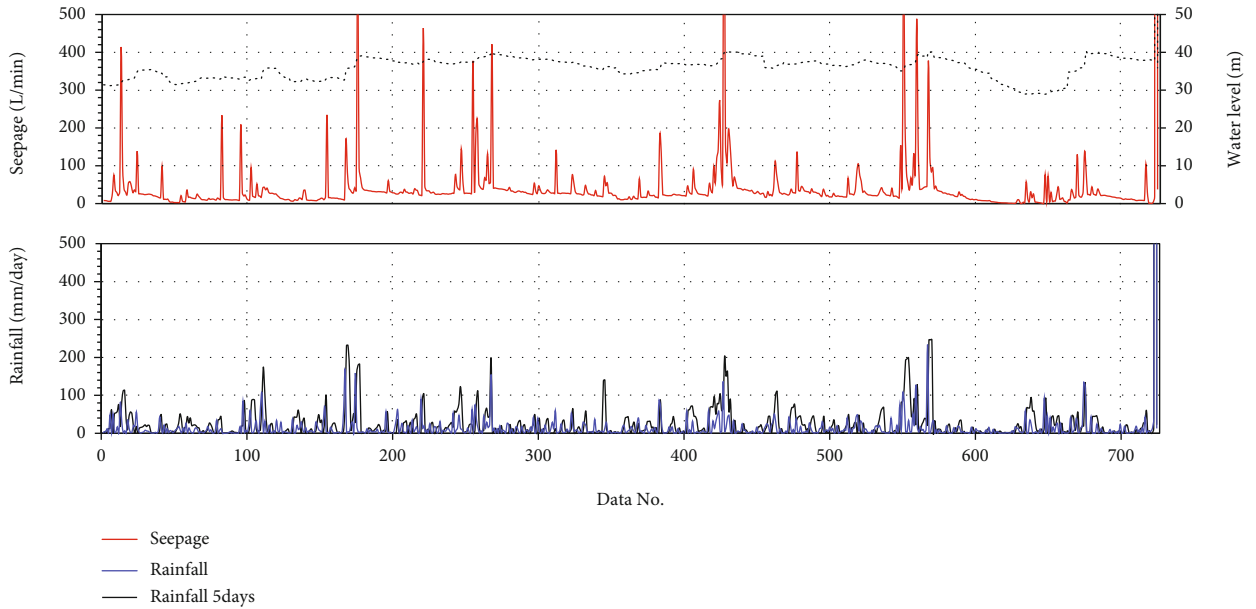FIGURE 10: Seepage, rainfall, and water level for group A.

FIGURE 11: Seepage, rainfall, and water level for group B.

categories when child nodes are formed from parent nodes. Therefore, quantifying the separation criteria requires quantification of explanatory variable selection and separation conditions that best distinguish the distribution of the response variables. For discrete data, the separation occurs based on the frequency of each category of response variables, while the separation of joints occurs based on the mean and standard deviation of the target variables for continuous data [20]. Detailed separation criteria and conditions for each data type are summarized in Table 3.

In this paper, Matlab R2020a was utilized as a tool for decision tree analysis, and the variance reduction was applied as a separation criterion as the target data is contin-uous. Variable reduction, $R(t)$, is defined as shown in equation (1) for each separation node for the data set of N-explanatory variable $x_1 = (x_1 \cdots x_d)$ and response variable $y_1 = (y_1 \cdots y_n)$.

$$R(t) = \frac{1}{n} \sum_{y_i, x_i \in t} (y_i - \bar{y}(t))^2, \tag{1}$$

Here, $\bar{y} = 1/n(t) \sum_{y_i, x_i \in t} y_i$.

When the set of all possible branch conditions in each separation step is called S, and each branch condition is
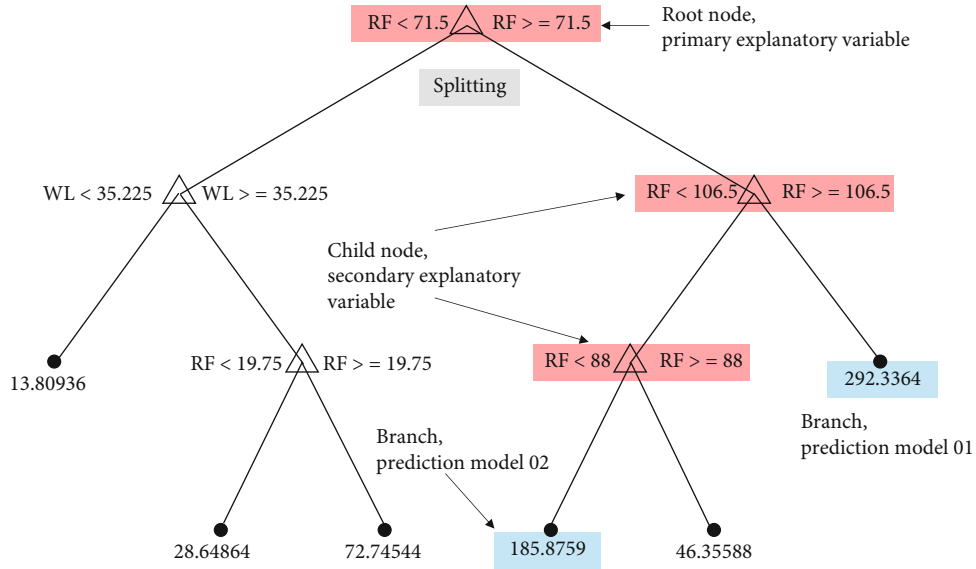
Figure 12: CART algorithm.

Table 3: Splitting rules for decision trees.

| Categories	Splitting rules	Splitting conditions
	p value for χ^2	Under the condition that p value is minimized
Discrete distribution	Gini index	Under the conditions that most reduce the Gini index
	Entropy index	Under conditions that reduce the most entropy index
Continuous distribution	p value for F	Under conditions where p value is minimized
	Decrements of variance	Under conditions that maximize the amount of variance reduction
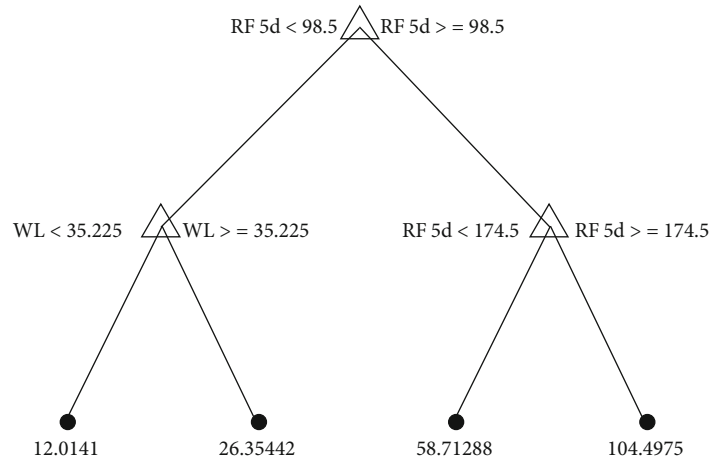
Figure 13: Decision tree for group A (seepage, antecedent 5-day rainfall, and water level).

called s, the optimal branch condition is s^* as shown in the following equation.

$$\Delta R(s, t) = R(t) - R(t_L) - R(t_R),$$
$$\Delta R(s^*, t) = \max \Delta R(s, t) \, ; s \in S. \tag{2}$$

Here, $R(t_L)$ and $R(t_R)$ are the variance reduction of the left/right branches in each joint.

The predictor importance (PI) of the explanatory variables can be calculated by equation (3) through the explanatory variable (x_i) selected from the optimal branch condition of each joint, s^*, and variance reduction, $\Delta R(s^*, t)_{x_i}$. The

TABLE 4: Results of decision tress analysis for group A.

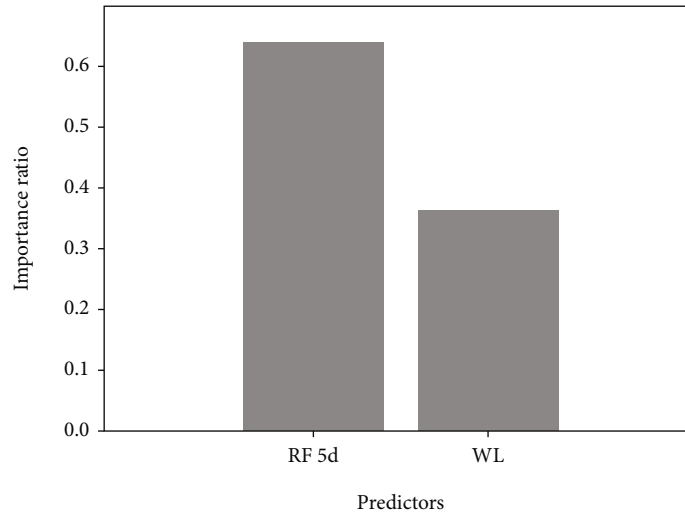| Primary explanatory variable	Second explanatory variable	Response variable Seepage (LQ, L/min)	
Antecedent 5-day rainfall (RF_5 d, mm/day)	Water level (WL, m)	Mean	Subset n
RF_5 d ≥ 174.5		104.5	18 (0.8%)
98.5 ≤ RF_5 d < 174.5		58.72	32 (1.5%)
RF_5 d < 98.5		21.8	2,083 (97.7%)
	WL ≥ 35.23	26.35	1,421 (66.6%)
	WL < 35.23	12.01	662 (31.0%)

FIGURE 14: Predictor importance estimates for group A.

explanatory variable with larger PI means better explanation for the response variable.

$$PI_{x_i} = \frac{1}{N_t} \left(\Delta R(s^*, t)_{x_i} \right). \tag{3}$$

4.2. Decision Tree Analysis of Seepage of Seepage. The analytical tree model is organized in the order of root (the primary explanatory variable), branch (other explanatory variables) division, and leaf (classification group/predictive model), which facilitates the analysis of multivariate factors with interactions. Therefore, an analytical tree analysis using the CART algorithm to analyze the effects of rainfall and water level on the causes of changes in seepage was conducted.

Figure 13 and Table 4 are the results of decision tree analysis of group A, representing the analytical tree results of the rainfall-free group performed by setting the seepage (LQ) as a response variable and antecedent 5-day rainfall (RF 5 d) and water level (WL) as explanatory variables. Figure 14 shows the dimensionless variable importance PI for the antecedent 5-day rainfall and water level applied as explanatory variables. Here, dimensionless PI presents the size of PI of each explanatory variable as a ratio to the total PI. As the explanatory variable antecedent 5-day rainfall was 0.64 and water level was 0.36, the antecedent 5-day rainfall
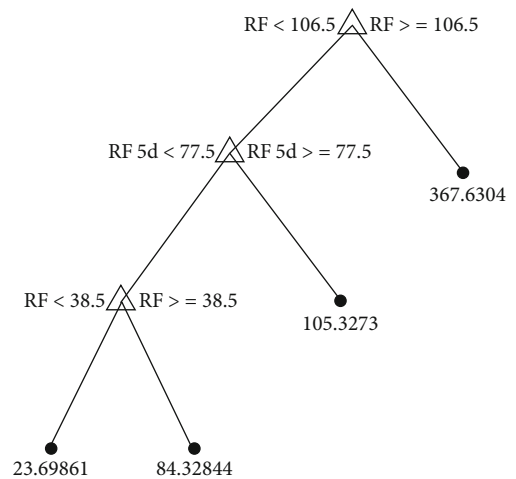

FIGURE 15: Decision tree for group B (seepage, rainfall, antecedent 5-day rainfall, and water level).

was selected as the primary explanatory variable. The statistics of the seepage group were calculated according to classified conditions to analyze the changes in seepage according to changes in rainfall and water level. As summarized in Table 4, the branch conditions of the antecedent 5-day

TABLE 5: Results of decision tress analysis for group B.

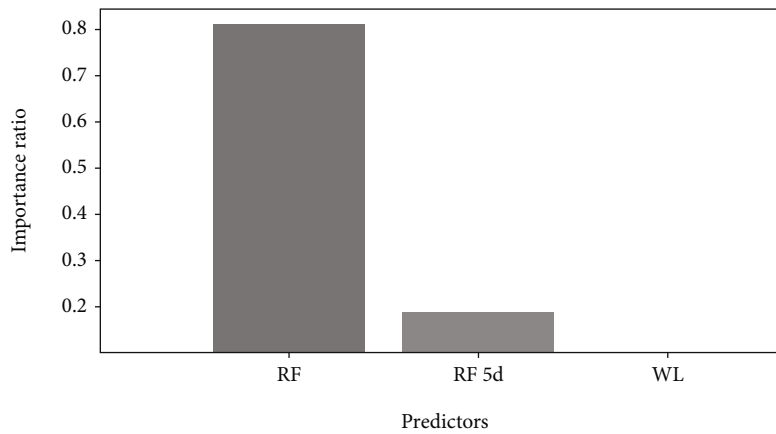| Primary explanatory variable	Second explanatory variable	Response variable Seepage (LQ, L/min)	
Rainfall (RF, mm/day)	Antecedent 5-day rainfall (RF_5 d, mm/day)	Mean	Subset n
RF ≥ 106.5		367.63	9 (1.2%)
	RF_5 d ≥ 77.5	105.33	48 (6.6%)
	RF_5 d < 77.5	27.9	664 (92.1%)
RF > 38.5		84.33	46 (6.4%)
RF < 38.5		23.7	618 (85.7%)

FIGURE 16: Predictor importance estimates for group B.

rainfall selected as the primary variable were 98.5 mm/day and 174.5 mm/day, with the water level acting as the explanatory variable under the conditions of antecedent 5-day rainfall below 98.5 mm/day. The branch condition of the water level was analyzed to be 35.23 m. In particular, 97.7% of the water leakage data are distributed under conditions of antecedent 5-day rainfall below 98.5 mm/day, with the average value of the group being 12.01 L/min~26.35 L/min, which is significantly lower than the average value of the relative group (58.72 L/min to 104.50 L/min). Therefore, it is determined that the antecedent 5-day rainfall can be considered as a major influence factor on the leakage at the point of no rainfall, and antecedent 5-day rainfall of 98.5 mm/day can be presented as the allowable seepage value for the increase in the amount of water leakage.

Figure 15 and Table 5 are the results of decision tree analysis of group B, representing the analytical tree results of the rainfall-occurring group performed by setting the seepage (LQ) as a response variable and antecedent 5-day rainfall (RF 5 d) and water level (WL) as explanatory variables. Figure 16 shows the dimensionless PI. As the explanatory variable daily rainfall (RF) was 0.81, the antecedent 5-day rainfall (RF 5 d) was 0.19, the water level (WL) was 0.00, and daily rainfall was selected as the primary explanatory variable. Group B, a rainfall-occurring group, was analyzed to have no effect of water level as an explanatory variable. To analyze the changes in seepage, the statistics of the seepage group were calculated according to the classified condi-

tions. As summarized in Table 5, the branch conditions of daily rainfall selected as the first variable were 106.5 mm/day and 38.5 mm/day, while the antecedent 5-day rainfall only worked as an explanatory variable in the range of 38.5 mm/day to 106.5 mm/day. In other words, the effect of antecedent 5-day rainfall on the amount of leakage is determined to be significant only at the time of occurrence of daily rainfall bigger than medium-size. In addition, in group B, 85.7% of the water leakage data are distributed under the conditions of daily rainfall below 38.5 mm/day, and the average value of the group was 23.70 L/min, which is significantly lower than the average value of the relative group (84.33 L/min~367.63 L/min). Therefore, it is determined that the daily rainfall (RF) can be considered as a major influence factor on the amount of water leakage at the time of rainfall, and that the daily rainfall of 38.5 mm/day can be presented as the allowable seepage value for the increase in the amount of water leakage.

5. Conclusions

In this study, the following results were obtained by classifying the group into two groups consisting of a rainfall-free group and a rainfall-occurring group to conduct decision tree analysis considering the effects of dam water level, daily rainfall, and antecedent 5-day rainfall on the seepage, which is the primary measured item for prediction of leakage of fill dam and piping..

As a result of the decision-making tree analysis on rainfall-free group (group A), there were the most data (66.6%) with the antecedent 5-day rainfall, selected as the primary explanatory variable, of less than 98.5 mm/day and water level higher than EL.35.225 m, and the average seepage at this time was 26.35 L/min.

As a result of the decision-making tree analysis on rainfall-occurring group (group B), the branch conditions of the daily rainfall, selected as the primary explanatory variable, were 106.5 mm/day and 38.5 mm/day. In addition, it was analyzed that the change in seepage during rainfall is not related to the water level. 85.75% of seepage data was distributed under the conditions with the antecedent 5-day rainfall, the primary explanatory variable, of less than 77.5 mm/day and daily rainfall less than 38.5 mm/day, and the average seepage at this time was 23.70 L/min. Also, when daily rainfall was more than 38.5 mm/day under the same conditions, the average seepage was 84.33 L/min.

Therefore, the seepage of the dam subject to research was found to be more directly affected by rainfall than by water level. Rather than presenting a single value as the acceptable seepage of the fill dam, the acceptable seepage according to the explanatory variables determined by the decision tree analysis can be presented, respectively.

Data Availability

The data set used in this study is available through Water Energy & Infrastructure Research Center, K-water (http://www.kwater.or.kr/kiwe/main.do).

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

References

[1] K-Water, *A Study on the Combined Analysis Method of Measurement Data for Dam Safety and Maintenance(2)*, (K-Water)16, 2019.

[2] korean geotechnical society, "Design, Construction and Safety Management for the Dam and Embankment," in *Goomi book*, KGS, 2010.

[3] H. K. Park and D. H. Shin, *Safety management of dams*, CIR, 2013.

[4] J. Lee, S. Yoon, E.-S. Im, and G. Kang, "A study on management criteria of seepage for fill dams considering rainfall effect," *Journal of the Korean Geotechnical Society*, vol. 36, no. 5, pp. 5–16, 2020.

[5] United States Cold Storage, *Lessons of Dam Incidents*, USCOLD, USA, 1975.

[6] B. D. Noh, *A study 314 on road slope safety investigation plan for disaster prevention*, 7th Seminar for Disaster Prevention Administration, 2002.

[7] S. Kuperman, M. Moretti, J. Pinfari, and E. Carneiro, "Placing limit values on instrument readings," *Hydro Review*, vol. 15, no. 3, pp. 24–31, 2007.

[8] J. Lee, "Application of statistical method for establishment of alert level for dam instrumentation," *Journal of the Korean Geotechnical Society, KSCE*, vol. 36, no. 5, pp. 283-284, 2014.

[9] P. T. Ryan, *Statistical Methods for Quality Improvement*, John Wiley and Sons, second edition, 2000.

[10] L. S. Lewis, C. D. Montgomery, and H. R. Myers, "Examples of designed experiments with nonnormal responses," *Journal of Quality Technology*, vol. 33, no. 3, pp. 265–278, 2001.

[11] R. Myers and D. Montgomery, *Response Surface Methodology: Product and Process Op-Timization Using Designed Experiments*, John Wiley & Sons, second edition edition, 2002.

[12] S. Park and H. Park, "A control chart method using quartiles for asymmetric distributed processes," *Journal of the Korean Statistical Society*, vol. 19, no. 1, pp. 81–96, 2006.

[13] International Organization for Standardization, *Water Flow Measurement in Open Channels Using Weirs and Venturi Flumes- Part1: Thin-Plate Weirs*, ISO, 1998.

[14] C. Jang and J. Lee, "Development of national standard technology and status of v-notch for measuring seepage of dam in Korean, water for future," *The KoreaWater Resources Association*, vol. 51, no. 5, pp. 57–63, 2018.

[15] Soil Conservation Service, *National Engineering Handbook, Section 4, Hydrology*, SCS, 1972.

[16] K-water, *(Field investigations and improvement plansfor for existing seepage instruments of 25 dams in Korea)*, (K-Water), 2017.

[17] Z. Guo, Y. Shi, F. Huang, X. Fan, and J. Huang, "Landslide susceptibility zonation method based on C5.0 decision tree and K-means cluster algorithms to improve the efficiency of risk management," *Geoscience Frontiers*, vol. 12, no. 6, article 101249, 2021.

[18] N. S. Park, Y. Lee, S. Chae, and S. Yoon, "A study on the statistical predictability of drinking water qualities for contamination warning system," *Journal of Korean Society of Water and Wastewater*, vol. 29, no. 4, pp. 469–479, 2015.

[19] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Egression Trees*, Taylor & Francis, 1984.

[20] H. Kang, S. Han, J. Choi, E. Kim, and I. Um, *Data Mining Methodology for Big Data Analysis*, Free Academy, 2014.