M.S. Thesis Proposal

A Study on the Improvement of Flood Season in Korea Considering the 21 Century Observations

21세기 관측자료를 고려한 홍수기 개선에 관한 연구

2022년 6월

Seoul National University

Civil and Environmental Engineering

Jaehwang Lee

Advisor: Professor, Kim Young-Oh

A Study on the Improvement of Flood Season in Korea Considering the 21 Century Observations

21세기 관측자료를 고려한 홍수기 개선에 관한 연구

지도 교수 김 영 오

이 논문을 공학석사 학위논문으로 제출함

2022년 **6**월

서울대학교 대학원

건설환경공학부

이 재 황

이재황의 공학석사 학위논문을 인준함

2022년 **8**월

위 원 장 _____ **(인)**

부위원장 _____ **(인)**

위 원 _____ **(인)**

Abstract

According to the sixth report of the Intergovernmental Panel on Climate Change, extreme weather events such as heavy rains and floods are predicted to become more frequent and severe owing to the rise in the global temperature. In Korea, abnormal climates such as rapid increase in the frequency of typhoons and the longest rainy season in the history has been reported, and the damage caused by them has been severe. However, the flood season, which has been occurring for over 50 years, is still established nationwide without considering climate characteristics and changes. This reveals the limitations of nonstructural countermeasures against flooding in Korea and highlights the need to improve the flood season establishment considering climate change.

Therefore, in this study, the problems of the current flood season were analyzed in terms of period and space using statistical techniques. Subsequently, the basis of the establishment of the current flood season was examined, and seven new flood seasons were proposed using extension and shift methods based on the analogical results. The Yongdam dam was selected as the study area because it met the four criteria of this study, and it was simulated and evaluated by predicting the inflow using a long short-term memory optimal model to generate an inflow hydrologic curve. This curve was employed to determine the discharge amount by a simulation method established by applying the basic dam operation rules and the rigid reservoir operation method. The optimal flood season for the study area was identified by

evaluating the flood reduction effect using both the method with nondamage and dam design release established in this study and method with river design flood and dam design relese, which is adopted in practice for deriving the discharge amount.

Keywords: climate change, regional rainfall characteristics, flood season, LSTM, Rigid ROM, Evaluation methods
***Student Number*: 2020-26297**

CONTENTS

List of Tables

List of Figures

CHAPTER 1. INTRODUCTION

1.1Problem statement

Korea has a unique precipitation pattern, accounting for 54% of the annual precipitation with 710.9 mm of precipitation in summer. Considering these characteristics, Korea designates the period from June 21 to September 20 as the so-called flood season, and prepares for flooding by operating dams differently during this period (Dam Management Regulations, 2015). However, due to the abnormal climate that has occurred in the 21st century, despite operational changes, the country is experiencing enormous flood damage (Meteorological Agency, 2011; Cha Eun-jung, 2006). In particular, the typhoon in October attacked the Korean Peninsula four times in the past 10 years (Jung et al., 2018), and in 2020, the longest rainy season ever recorded since the meteorological observation of 54 days (Ministry of Public Administration and Security, 2020) .

Looking at a study based on changes in precipitation characteristics in the recent flood season in Korea, Lee and Kwon (2004) divided Korea into four regions and compared the increase and decrease in precipitation during the flood season in the middle and end of the 20th century. Through this, precipitation during the flood season in Korea shows an increasing trend, and a clear increase trend was confirmed in August. Koh et al. (2005) showed that the precipitation from late July to early August has increased significantly in recent years, increasing regional differences. It

was also found that the cause of this increase in precipitation in August was meteorologically, the location of the North Pacific high pressure was extended to the west (Ha et al., 2007).

Considering these changes in precipitation characteristics, studies on preparation for flooding during the flood season have mainly aimed to find the optimal water level and capacity of the dam. As a representative example, Sim et al. (1995) proposed the application of a variable limiting water level for optimal reservoir operation at the end of the flood season. However, in the current situation where the uncertainty of climate change is increasing, the limiting water level adjustment in the limited storage capacity has a limitation in that it is disadvantageous in terms of water supply. In addition, research on the lack of flood control capacity of existing dams is being conducted (Jang et al., 2014), and there is also a study on improving the target water level for dam flood control by using the inflow-to-reservation volume ratio (Kwak Jae-won, 2021).

However, there is still no review of the problem and related studies on the period of the flood season. The document that first specified the legal flood period has not changed since it was enacted in 1974 as 「Soyanggang Dam Management Regulations」. This is interpreted as failing to take into account changes in rainfall patterns in the 21st century. Also, considering that one flood season was applied to the whole country, the strong regional characteristics of recent precipitation are not taken into account. In order to improve these limitations, it is necessary to discuss the improvement of the current flood season.

Therefore, this study intends to review the validity of the current flood season through 21st century observation data. In addition, instead of finding flood countermeasures through changes in water level and capacity of existing dams, we propose a methodology to improve the period of the current flood season to prepare for flooding.

1.2Research Objectives

The ultimate purpose of this study is to propose a new flood season considering climate change in the target watershed. To this end, first, the limit of the current flood season is analyzed by comparing the precipitation of the 20th and 21st century flood seasons through statistical techniques. This is because it is necessary to check the precipitation trend due to climate change and the limitations that the current flood season did not take into account for flood preparation. Second, this study proposes a new flood season in the target watershed. This proposed a new flood season in consideration of the expert's advice and the statistical change of the current flood season establishment methodology. Lastly, this study intends to select a new flood season most suitable for the target watershed through evaluation and analysis. to determine the discharge amount. The LSTM model of deep learning was used for the predicted inflow, which is the input data of Rigid ROM. Finally, by applying the method based on nondamage and dam design release and the method based on river design flood and 200-year frequency dam design release, a new flood season suitable for the target watershed is proposed.

1.3Thesis Organization

Chapter 2 of this paper summarizes the inference results of the current flood season establishment methodology and the flood season dam operation method. In addition, four representative ROMs of the simulated operation method were investigated. In this case, previous studies on the method of deriving the predicted inflow to be used as input data were reviewed. In Chapter 3, the limitations of the current flood season were analyzed, and basic statistics and hypothesis tests were applied. In Chapter 4, a study area was selected according to four criteria and 7 new flood seasons suitable for the target watershed were proposed. The proposed flood season was simulated and evaluated, and a flood season suitable for the target watershed was finally proposed by analyzing the evaluation results. Finally, Chapter 5 describes the conclusion and future research plans.

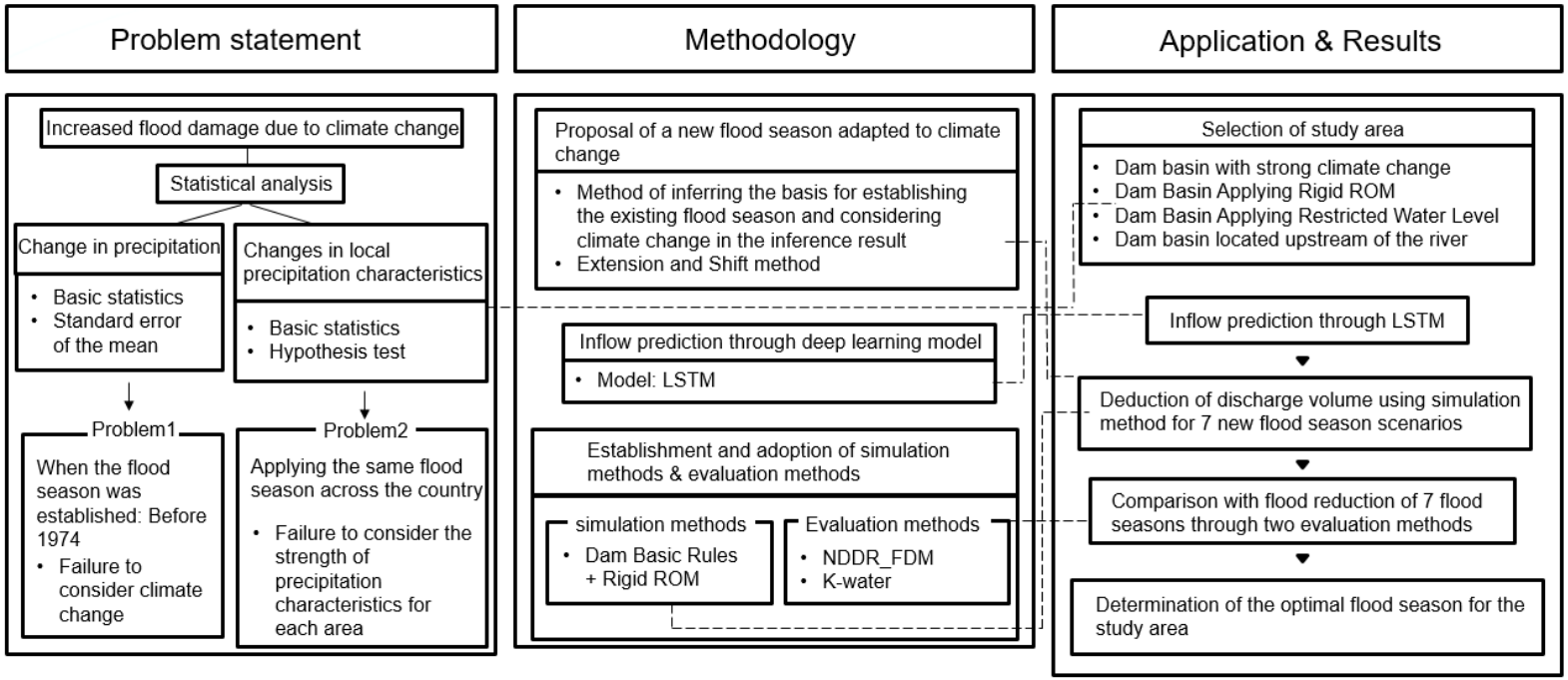


Figure 1.1 Thesis flowchart

CHAPTER 2. THEORETICAL BACKGROUNDS

2.1Flood Season in Korea

In Korea, many dams are being built for efficient water resource management. Since 2/3 of the rainfall in Korea is concentrated in the rainy season, it is the most important to efficiently operate dams during the rainy season to secure the maximum amount of water after the rainy season and to prepare for flooding by using the flood control capacity during the rainy season. That is, during the flood season, the reservoir flood control capacity is secured to store the inflow of the dam caused by the torrential rains, while the proper reservoir storage volume must be secured even after the flood season by proper dam discharge during the flood season in preparation for the water supply after the flood season. It is basic goal of dam operation.

Reservoir operation can be explained as a long-term operation plan centered on the dry season, which places importance on water supply and power generation, and as a short-term operation plan during the flood season, where dimensional aspects such as flood control are important. Here, the short-term reservoir operation aims at optimal flood control to minimize flood damage downstream of the dam by using most of the flood control capacity of the dam.

2.1.1 Logics behind the Current Flood Season

Literature search

For the current flood season, related prior studies and reports are insufficient. In addition, the establishment methodology of the current flood season is unreported. Therefore, in this study, the existing flood season establishment methodology was inferred through various literature surveys.

A domestic flood period-related literature study was conducted, such as on dam management regulations, multipurpose dam operation manuals, and dam and weir linkage operation regulations. The rationale for the establishment of the flood period can be found in its definition. “Flood period” is the period from June 21 to September 20 during which flood damage is probable to occur. Thus, it can be inferred that the existing flood period was determined using variables related to the possibility of flood damage.

In addition, we conducted a literature survey on overseas flood periods, such as data from the Ministry of Land, Infrastructure, and Transport in Japan. Although it was impossible to confirm the actual basis for the establishment of the flood period in Japan, it the commonality of establishment of a flood period and dam management with Korea was confirmed. However, in Japan, the establishment and operation of the flood season are different for each dam.

Table 2.1 key findings of literature review

	Key findings
Definition	· The period from June 21st to September 20th during which flood damage is likely to occur · "Restricted water level" refers to the water level set to secure flood control capacity during the flood season, and the highest water level t hat must be maintained when there is no flood inflow
Priority by application	· During flood season, flood control takes precedence over other uses · Power generation can be done during the flood season
Flood alert system	· The water level is expected to rise above the restricted water level · The Korea Meteorological Administration issues a heavy rain advisory or warning in the upstream area of the dam
Flood season Task	· Collecting meteorological and hydrological data upstream and downstream of the dam · Collection and analysis of hydrological data such as water level and rainfall upstream of the dam · Establishment of flood control plan and determination of preliminary release
Flood control	· In the case of multi-purpose dams, flood control is carried out by maximizing the flood control capacity during the flood season.
The first specified flood season in Korea	· Soyanggang dam management regulations · March 4th , 1974
Overseas flood season operation cases	· Japan has a flood season and dam operation using the restricted water level · Flood season vary by region

Expert interview

The conducted literature surveys on the legal flood period revealed the limitations of insufficient flood period-related information; therefore, experts were interviewed for further investigation. These interviews were conducted with experts from the time of flood period establishment and experts working at the Water Resources Information Center of the Flood Control Center.

Information on the flood period, summarized through the interviews, is as follows. The flood season is defined in the dam weir and dam management regulations of the Korea Water Resources Corporation, and it is proposed in terms of dam flood level management. In 1970, the flood period was assigned as "June 21 to September 20" in the Disaster Prevention Work Manual (SOP booklet). Thus, it can be inferred that this period was actually proposed before 1970. Moreover, the Flood Countermeasures Act Enforcement Decree specifies that the head of the Disaster Response Headquarters has the right to operate the floodgates of a multipurpose dam during a joint work period of the Central Disaster Response Headquarters (dispatched to the central government). The flood period is presumed to be set as the period from the rainy season in South Korea to the end of typhoons, based on rainfall statistics (30 years). In addition, it is determined that it will be useful to adjust the flood season considering recent precipitation patterns, because research on the flood season has revealed the necessity to maintain the water level limit during the flood season. Moreover, it will be beneficial to determine the flood season for each watershed unit individually. In addition, a flood management plan should be formed by

quantitatively determining the effect of climate change by simulating dam operation according to the existing dam operation rules considering the precipitation changes. The above is summarized in Table 2.2.

Table 2.2 Summary of expert interviews

Expert interview summary
• The flood season is judged as a term proposed in terms of dam flood level management, and is defined in the dam management r egulations
• 1970s, the flood season was already marked in the Disaster Prevention Work Manual (SOP booklet)
• It is thought that the period was set when the head of the Disaster Response Headquarters had the right to operate during the joint work period of the Central Disaster Response Headquarters
• The basis for setting the period is estimated to be the period from the rainy season in South Korea to the end of the typhoon based on rainfall statistics (30 years)

Results of inferring the rationale for establishing the flood season

The basis for the establishment of the current flood period was inferred by combining data obtained from literature search and information acquired through expert interviews. Statistics were verified by considering the entry periods of the rainy season fronts in Korea from 1961 to 1973, which were the average data at the time of the establishment of the flood season, as the starting points of the flood seasons (The Meteorological Agency, 1995). Based on the average values, the starting point is June 23, which is close to the starting point of the current legal flood period, June 21. Thus, it was determined that the basis for establishing the starting point was the entry of the rainy season front in Korea. In addition, the statistics of the last points of the last typhoons that affected Korea from 1941 to 1970 (Meteorological Agency, 2011) were the average data at the time of the establishment of the flood season. Based on them, mid-September and September 20 were found as the ending points of the current legal flood season. Thus, the last point of the last typhoon, which was close to the day and had an impact on Korea, was the basis for the ending point of the legal flood period.

Table 2.3 Results of applying 30-year data at time of establishment to inferred method

	Start point	End point
	Data, at the time of enactment (1961 ~ 1973)	Data, at the time of enactment (1941 ~ 1970)
Minimum	June 14	Early August
1st quartile	June 24	Late August
Median	June 24	Early September
3rd quartile	June 25	Late September
Maximum	July 1st	Mid-October
Mean	**June 23**	**Mid-September**
Standard deviation	4.13	1.9

2.1.2 Dam Operations in Flood Season in Korea

In Korea, to prepare for flood damage, the flood season is designated. "Flood season" refers to the period from June 21 to September 20 when flood damage is probable to occur, and during this period, the operation of a dam is different from that in the dry season.

Dam Operation in Flood Season in Korea

In Korea, many dams are being built for efficient water resource management. Two-thirds of the rainfall in Korea occurs in the rainy season. Therefore, efficiently operating dams during it to secure the maximum amount of water after it and prepare for flooding using the flood control capacity during the rainy season are most important. Specifically, during the flood season, the reservoir flood control capacity is secured to store the inflow of dams caused by torrential rains. Appropriate reservoir

storage volume must be secured even after the flood season by suitable dam discharge during it in preparation for the water supply after it. This is basic objective of dam operation.

Reservoir operation can be explained as a long-term operation plan centered on the dry season, which places importance on water supply and power generation, and as a short-term operation plan during the flood season, in which dimensional aspects such as flood control are important. Short-term reservoir operation aims at the optimal flood control to minimize the flood damage downstream of a dam using most flood control capacity of the dam.

According to the dam management regulations in Korea, a multipurpose dam has the highest water level in the part used for flooding called “normal high water level (NHWL),” and has a “restricted water level (RWL),” which is the water level set to secure the flood control capacity during the flood season. The highest water level to be maintained is selected and operated (Figure 2.1). In the case of dams without a flood water level (FWL), the NHWL is set as the limiting water level and flexibly operated according to the hydrological conditions.

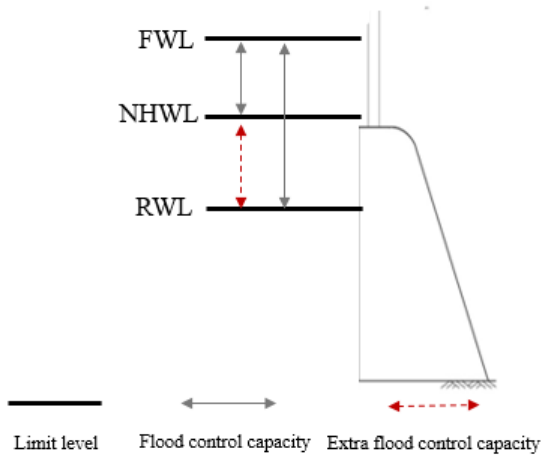


Figure 2.1 Method of securing flood control capacity during flood season

In addition, dam reservoirs are operated according to the priority of each use specified in the dam management regulations according to the hydrological conditions. During the flood season, flood control takes precedence over other uses. The priorities are summarized in Table 2.4.

Table 2.4 Priority of water use during flood season

Water use purpose	Priority	Detailed usage
Flood control	1	Preliminary release
		Flood control
Supply	2	Supply of water for domestic & Industrial and Agricultural use
		River management use
		River improvement use
Power generation	1	Power generation use
Other	3	Other unavoidable reasons (Improvement of dam utility, maintenance, etc.)

“Flood control” refers to the storage of all or a part of the flood volume flowing into a dam using its flood control capacity to minimize the discharge size, and the technique used for this purpose is the ROM. However, power generation can be done even during the flood season.

To limit the discharge amount, the discharge of the water stored in a dam should not exceed the design discharge amount for it. However, if there is a risk of dam collapse owing to high water level, all water gates are opened and the maximum discharge exceeding the design discharge is released.

In other scenarios, the dam manager must maintain the water level in the dam appropriately by flexible determination of the fluctuating hydrological conditions. At this time, meteorological and hydrological data upstream and downstream of the dam, hydrological data such as water level and precipitation upstream of the dam, and the results of flood hydrological analysis should be considered.

Dam Operation in Flood Season in Abroad

Overseas, countries with similar precipitation characteristics to Korea also operate dams in a period to prepare for flooding. Representatively, in the case of Japan, the flood season is designated identically to in Korea, and dams are operated in preparation for flooding. The method is the same as in Korea in that the NHWL, which is maintained during the dry season, is lowered to the RWL during the flood season to secure and operate the flood control capacity (Figure 2.2). In Japan, the NHWL and the RWL are divided by a ratio based on the amount of water stored up

to the design flood level. For the NHWL and the RHL, 80% and 60% standard water levels are designated, respectively.

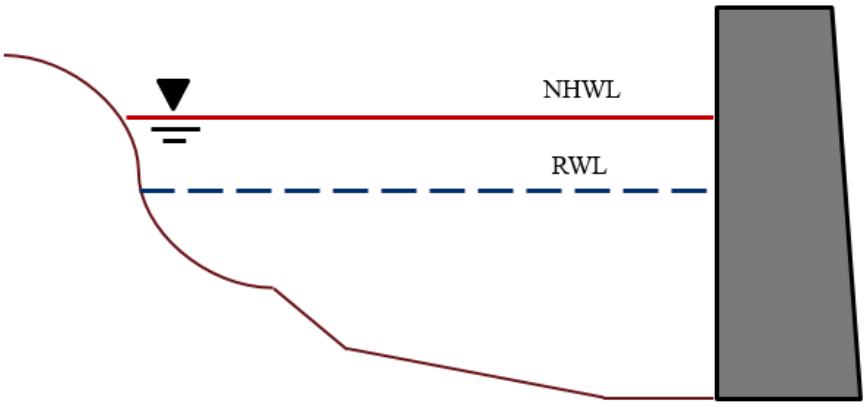


Figure 2.2 Example of water level limit at Kanagawa dam

However, Japan is different in that it adopts a method that considers regional precipitation characteristics by applying different flood seasons to each river. The flood seasons of the Seongsan, Miho, Tonegawa, and Kanagawa dams are from June 1 to October 15, from June 15 to October 15, from July 1 to September 30, and from July 1 to October 1, respectively (Figure 2.3).

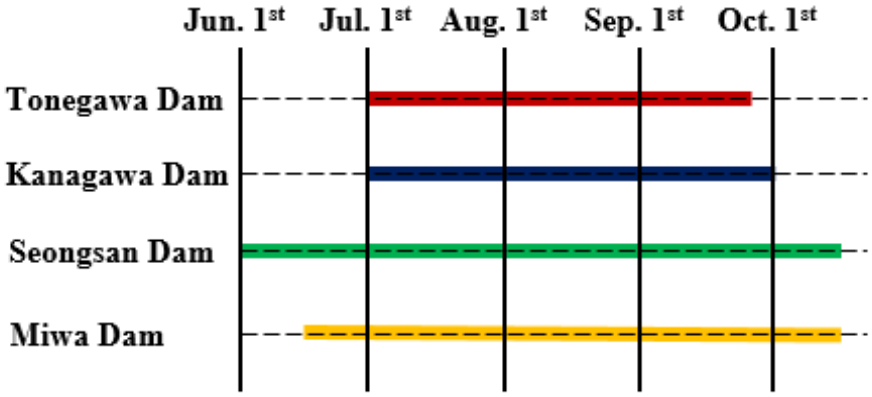


Figure 2.3 Examples of different flood seasons for different dam basins in Japan

2.2Reservoir Operation Method

The primary principles of reservoir operation during a flood are to manage a flood utilizing its storage space to prevent flood damage downstream while ensuring a sufficient volume of water in the reservoir for various water uses after the flood. To achieve this, first, the reservoir capacity is filled below the typical NHWL feasibly, to prepare for diverse water demands throughout the non-flood season. Finally, during the flood period, it should be possible to assure the dam safety by keeping the water level below the FWL.

The operation of a dam/reservoir in case of a flood is based on the reservoir operation rate stipulated in the dam management regulations.

The ROM is classified into simulation and optimization methods.

An optimization method seeks to optimize the operation of a reservoir in the event of a flood by optimizing various explanatory conditions added using methods such as linear programming and dynamic programming. Owing to the diversity of state variables according to the operation, many calculation processes are required, and practical access is difficult. Therefore, an optimization method is hardly adopted for short-term real-time reservoir operation for flood control.

Although a simulation method generally has a limitation in that it is difficult to obtain the optimal solution, it is suitable for dealing with detailed and complex problems, and its theory is simple and practically accessible. The rigid, auto, technical, and spillway rule curve (SRC) ROMs employed in this study are available as single ROM using a simulation operation method. In addition, linked operation of these

reservoir methods has been studied. Table 2.5 summarizes the ROMs.

Table 2.5 ROM types and characteristics

Approach	Method	characteristics
Optimization	· Linear programming (LP) · Dynamic programming (DP)	· Requires a lot of calculations · Practical approach is hard
Simulation	· Auto ROM · SRC ROM · Rigid ROM · Technical ROM	· Theory is simple · Practical approach is easy · Often used during flood season

2.2.1 Auto ROM

The auto ROM is the simplest method to operate a reservoir, and it can only ensure securing of water and the safety of a dam. If the water level of a dam is lower than the target water level (NHWL or RWL), the water gate does not discharge until the target water level is reached. If the water level exceeds the target water level, the spillway discharge rating curve can be used to determine the amount of discharge. This is a method to maintain the target water level by discharging the entire amount of inflow. However, if there is a power generation facility, the discharge of the power generation is continuous. Therefore, in principle, this method ensures that the water level of the reservoir does not rise above the target water level, which prevents the flood control space of the reservoir to be fully utilized.

The operation of a reservoir by the auto ROM can be divided into three types as follows:

(1) If the water level of the reservoir is below the target water level, the water gate is kept closed.

(2) When the water level of the reservoir reaches the target water level, the water gate is partially opened to keep the water level of the reservoir at the target water level, and the inflow flood is discharged.

(3) If the water level of the reservoir exceeds the target water level, the water gate is completely opened and discharged. Subsequently, when the water level drops to the target water level, step (2) is repeated.

It is not recommended to use the auto ROM when the floor elevation of the water gate is below the NHWL, such as for medium-sized and large dams, or when the RWL is set, such as for multipurpose dams.

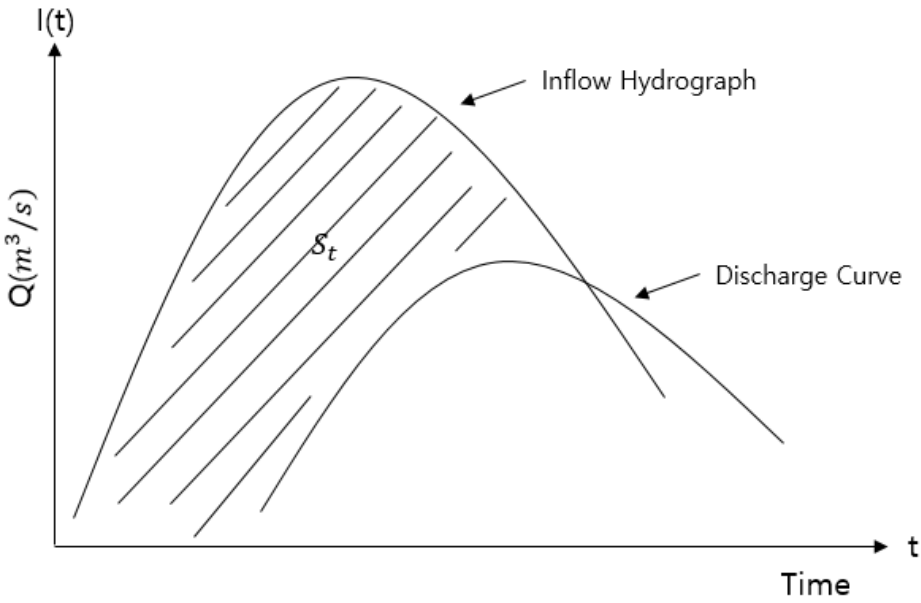


Figure 2.5 Auto ROM operating method graph

2.2.2 SRC ROM

The SRC ROM is similar to auto ROM in that it is used regardless of the prediction of the hydrologic curve of the inflow flood into the reservoir. It determines the discharge amount to downstream of a dam using an SRC.

Because this method discharges the flood at a predetermined discharge amount according to the reservoir level, it can be an appropriate flood control plan when a flood volume similar to the planned flood volume flows in. In addition, flood control is easy and the flood control capacity is highly usable. However, if the inflow flood amount is significantly different from the planned flood amount, it has the disadvantage that the flood progress cannot be appropriately reflected.

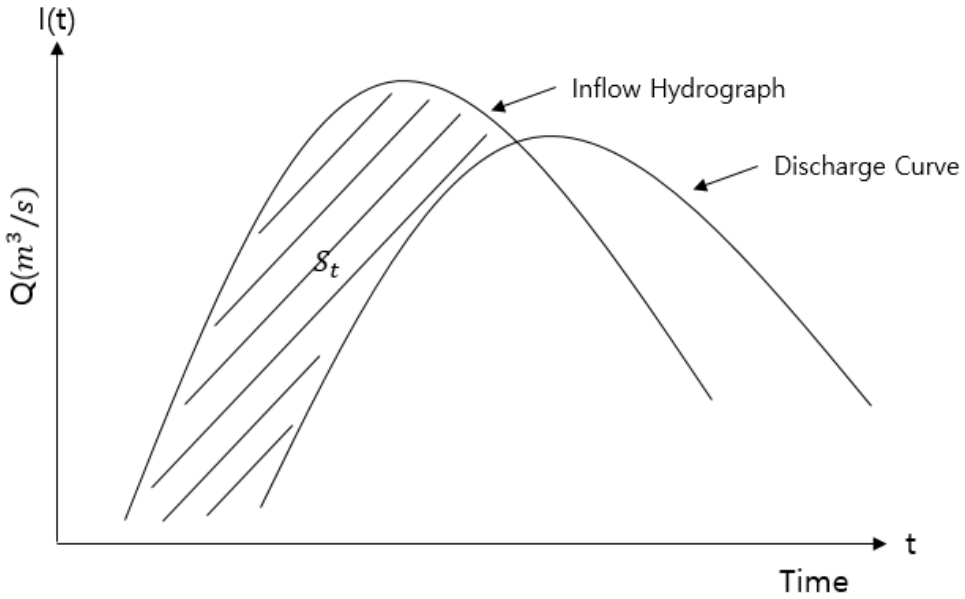


Figure 2.6 SRC ROM operating method graph

2.2.3 Technical ROM

The technical ROM determines the discharge amount based on the predicted inflow curve, and it can be used in a flood control model combined with a flood prediction model. Specifically, a certain amount of water is discharged with O_t obtained from the following equation when the water storage, S_t, between the dam level at the predicted time, t_1, and the FWL is stored by the flood.

$$S_t = \int_{t_1}^{t_2} [I_t - O_t] dt$$

In this equation, it is the predicted inflow curve, $t2$ is the time at which It descends and becomes equal to the constant discharge, and Ot is the constant discharge.

This operation method determines the discharge amount that matches the target flood control capacity with the reservoir discharge for storing the predicted inflow after the current time in a dam. Therefore, it is the most effective method among simulation operation techniques for flood control in reservoirs. Because the discharge amount determined changes according to the predicted hydrologic curve, during the flood season, when real-time analysis is performed, the discharge amount also changes based on the analysis time and the measured hydrologic curve. The former is related to the error of the predicted hydrologic curve and the latter to the uncertainty of the rainfall prediction.

To appropriately use the technical ROM, the complete inflow flood hydrologic curve during the duration of the flood must be accurately predicted; therefore, the accuracy of the outflow calculation model is the key. However, when an actual flood event occurs, the applicability of the runoff calculation model to temporally changing rainfall events is limited. Therefore, accurately predicting the complete inflow flood hydrologic curve is difficult, and many errors are bound to occur. Therefore, the practical applicability of the technical ROM is lower than that of the rigid ROM.

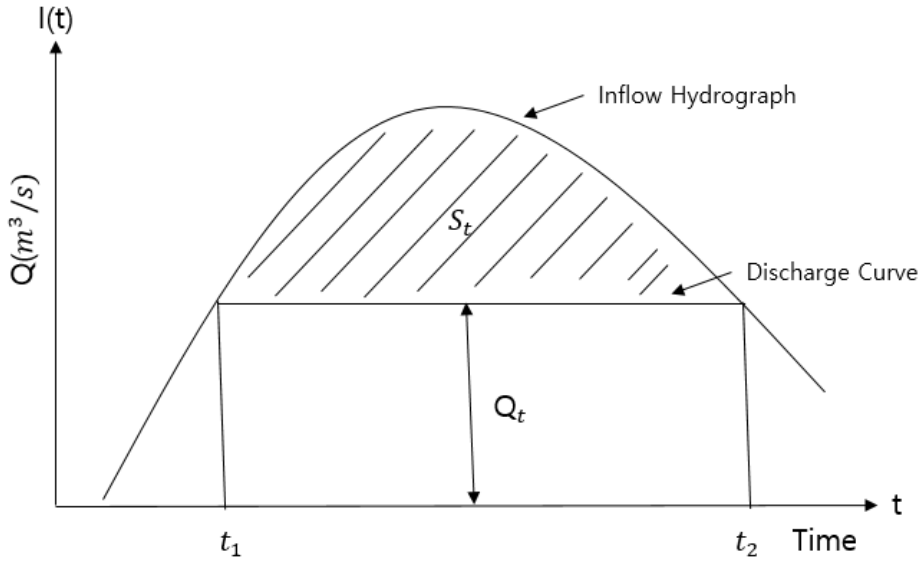


Figure 2.7 Technical ROM operating method graph

2.2.4 Rigid ROM

The rigid ROM is a method of operating a reservoir by discharging at a constant rate a fixed amount determined based on the observation value of the inflow at the time of the hydrological operation, expected inflow, hydrologic curve, and water level. In the case of flood control, the total storage is computed by estimating the input hydrologic curve, and discharging is performed at a constant rate when the inflow reaches its maximum. The reservoir is operated to match the regulated capacity. If the discharge becomes the ratio to the inflow until the inflow reaches the expected maximum inflow, the ratio is determined using the following equation:

$$V(t) = (1-\alpha)\int_{t_0}^{t_P} I(t)\,dt + \int_{t_0}^{t_P} I(t)\,dt - I(t_e)(t_e - t_P)$$

Here, V(t) is the flood control capacity at the time of flood prediction, t_o is the time of flood prediction, t_p is the time of maximum flood inflow, t_e is the time at which the flood inflow and the discharge coincide, and α is the ratio of the inflow and the discharge.

The rigid ROM cannot easily perform hydrological manipulation and is difficult to use with complex hydrologic curves with multiple peaks. However, it is extensively used in practice because it has the advantages of maximizing the flood control capacity and reducing the flood damage during low-frequency flooding.

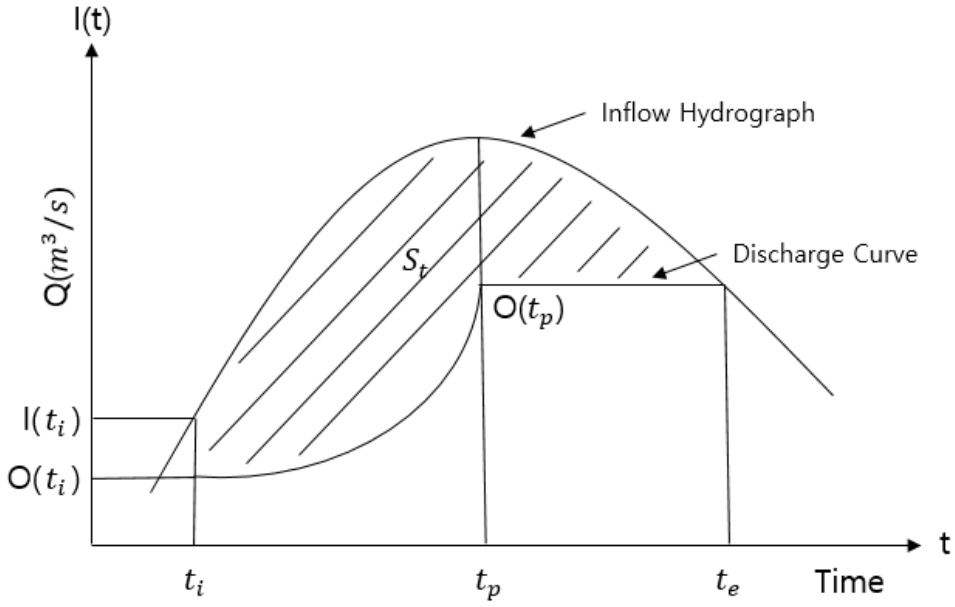


Figure 2.8 Rigid ROM operating method graph

Each of the previous four reservoir management techniques has advantages and disadvantages. The advantages and disadvantages of flood control and practical applicability, which are the main parts of this study, are explained as follows (Table 2.6). The auto ROM is the most vulnerable to flood control, as it is only used for the safety of dams and securing water. The SRC, technical, and rigid ROMs have excellent flood control ability. Among them, in the case of flood preparation for climate change, the SRC ROM, which is operated after determining the amount of discharge according to the water level in advance, is unsuitable. Therefore, in the area of flood control for climate change, the technical and rigid ROMs are excellent. In terms of practical applicability, the technical ROM, which is considerably affected by the accuracy of the predicted inflow hydrograph and discharges in a certain amount, is relatively inferior to the rigid ROM.

Table 2.6 Advantages and disadvantages of ROMs

	Advantages	Disadvantages
Auto ROM	· Simple operation method · Securing water and ensuring the safety of dams	· In case of flood, causing damage to the downstream because of la rge amount of release · Unable to utilize flood control capacity · **Difficult to use in practice** in the case of multipurpose dams
SRC ROM	· No need to predict the flood hydrologic curve · Easy flood control by operating at the w ater level of the reservoir	· If there is large difference between inflow and design flood, **the fl ood situation is not properly reflected**
Technical ROM	· Maximize flood control capacity · The most effective method	· Accurate prediction of flood hydrologic curve is required · **Less practical applicability than Rigid ROM**
Rigid ROM	· **Flood control capacity can be utilized to the maximum, so it is widely used i n practice**	· Difficulty in operation · Difficult to apply to complex hydrographs with multiple peaks

2.3 Flood Forecasting

2.3.1 Fundamentals on Flood Forecasting

Hydrological data are critical for comprehending the hydrological process and recognizing its characteristics to protect people and property from future natural disasters. While establishing water resource planning and developing large-scale water structures in the future, accurately comprehending the design hydrologic volume, such as rainfall and runoff, which are hydrological design criteria, is important. However, for observatories that obtain hydrological data in Korea, the recording period is short because most of them have been installed recently. In addition, understanding the characteristics of the time series data using given data is very important, because the reliability of the data is frequently poor and there are many missing values. (Kim et al., 1997)

Forecasting refers to the estimation of the state of a variable at a specific time or in a specific time range. Such actionable and accurate predictions are essential for decision-makers to identify trends in an environment in which rapid climate changes and fluctuations of various variables constantly occur.

2.3.2 Machine Learning for Flood Forecasting

Machine learning is a branch of artificial intelligence that is used to automatically and intuitively recognize patterns in datasets without requiring explicit programming. In traditional programming, data are input, and output data are obtained using a function, as shown in Figure 2.9. However, in machine learning, input and output data are input

to obtain a function. Machine learning has high performance and relatively lesser complexity than existing models, making it easier to solve complex problems (Mosavi et al., 2018; Wagenaar et al., 2020).

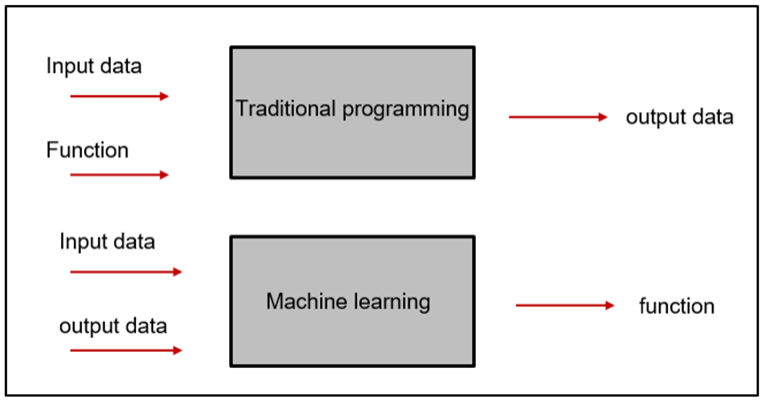


Figure 2.9 Difference between machine learning and traditional programming

Natural disasters such as floods are very complex to model. In the existing case, traditional programming contributes to minimizing the damage to properties and human life due to flooding. However, machine learning methods have considerably contributed to the development of prediction systems that provide better performance and efficient solutions than traditional programming. Therefore, in this study, machine learning methods were adopted to obtain high-accuracy and efficient prediction models (Mosavi et al., 2018).

Over the past two decades, machine learning methods have been continuously

evolving, demonstrating a predictive power that outperforms those of conventional approaches. When the performance of many existing models and machine learning prediction models was compared, the accuracies of the latter were higher and they were proven to be suitable for flood prediction (Abbot et al., 2014; Fox et al., 2005; Merz, B et al., 2010).

The overall flowchart of a machine learning method is shown in Figure 2.10. When the input and output data are set, the data are divided into datasets for training and performance test, respectively. In this study, the data were generally divided in an 80:20 ratio. In addition, in the training stage, K-fold cross-validation was performed using a total of five pictures, and the optimal hyperparameters were obtained. The test was conducted with a model using the hyperparameters, and the accuracies of several models obtained accordingly were compared for the final model selection (Gizaw et al., 2016; Campolo, m et al., 1999).

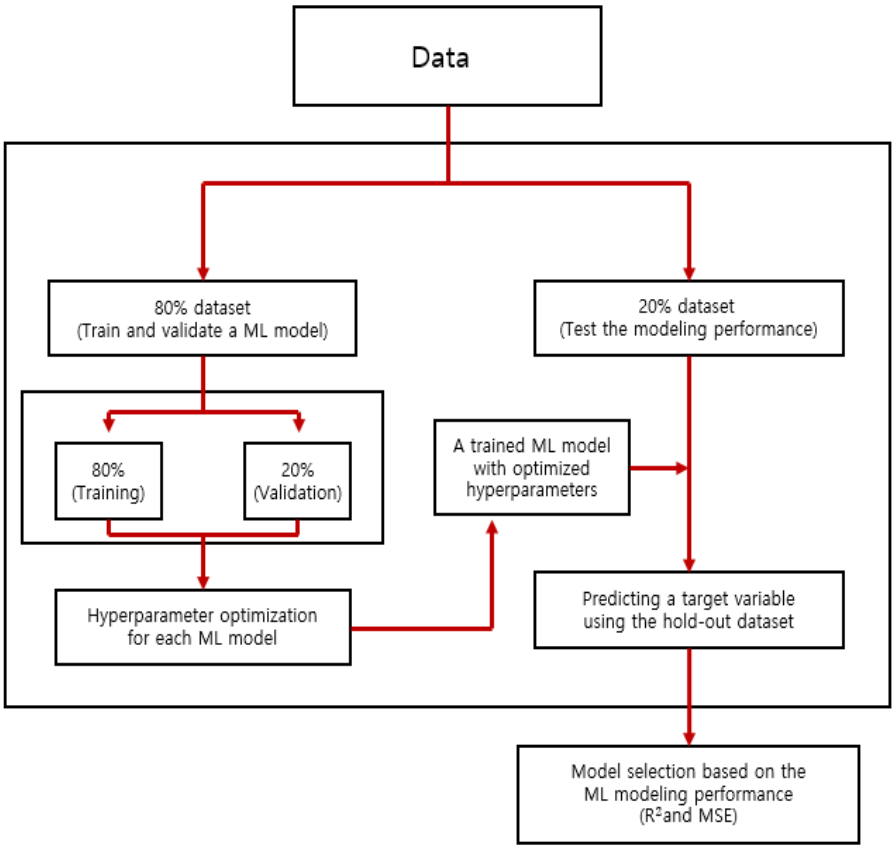


Figure 2.10 Machine learning flowchart (ML: machine learning, MSE: mean square error, R^2: coefficient of determination)

2.3.2 Long Short-Term Memory (LSTM)

Since Lapedes and Farber (1987) used a multilayer neural network model for time series prediction, such research using artificial neural networks became after the 1990s. Weigend (1990) proved the accuracy of multilayer neural networks using sunspots data compared to a threshold autoregressive model. Since then, research on time series prediction using this model has been actively conducted. Mozer presented

modified multilayer neural network models at the Santa Fe Institute workshop in 1992, which used an RNN. This neural network type is classified into an Elman network (Elman, 1990) and a Jordan network (Jordan, 1990) according to criteria. An Elman network is a neural network model in which the output value of the hidden layer is fed back to the input layer, and a Jordan network has a structure in which the output value of the output layer is fed back to the input layer. Since then, research on RNNs, e.g., predicting stock price change patterns on time series data, is being actively conducted.

In the field of hydrology, various studies on the application of neural networks are being actively conducted. Abroad, studies on flow rate prediction have been steadily progressing. In general, various hydrologic models such as physical and empirical models have been used for flow rate prediction. however, in early research, studies using physical models were predominant (Bicknell et al., 1996; Kim et al., 2007; Neitsch et al., 2011; Kang et al., 2013; Devia et al., 2015; Noh et al., 2016). However, there were difficulties in considering various variables, and Hsu et al. (1995) showed that artificial neural networks can be used in areas where explaining the physical process in rainfall-runoff modeling is difficult. Dawson and Wilby (1998) had suggested that an artificial neural network model can be used as a flow rate prediction model by learning about rainfall-runoff. Presently, studies on the application of artificial neural networks to the field of hydrology are being conducting (Kim, 2020). Imrie et al. (2000) applied an artificial neural network model to river flow prediction and proposed a method to improve the performance. Recently, studies using deep

learning-based models based on artificial neural networks have emerged. that mimic the human brain neural network structure and outperform existing machine learning-based models (Chen et al., 2018; Shoaib et al., 2016; Assem et al., 2017). Moreover, research has shown that the introduction of LSTM to an existing RNN model improves the prediction performance (Tian et al., 2018; Kratzert et al., 2018; Hu et al., 2018).

In Korea, machine learning is being actively employed in the field of hydrology prediction. The inflow of the Yongdam multipurpose dam located upstream of the Geumgang was predicted using the LSTM technique (Mok et al., 2020). Moreover, the real-time prediction of the inflow of the dam was reversed using the average rainfall of the dam basin, measured dam inflow, and predicted dam inflow. A previous study used a propagation neural network model for predictions (Kang et al., 2004). In addition, Lee et al. (2020) evaluated the prediction performance of the LSTM method according to the time interval of observation data. They compared it with the water level of the Oesong water level station located in the Namgang dam basin. Hwang (2021) predicted the inflow of the Sapkyo lake by adjusting the sequence length for the applicability of the LSTM model. Heo and Bae (2021) also used the LSTM method to estimate the inflow amount at a watershed upstream of the dam by the preceding time.

An artificial neural network model has a basic structure compared with many deep learning models, and it can solve problems by changing the binding

strength with nodes (neurons) composed of synaptic bonds (Heo and Bae, 2021). Based on the theory, it can provide a generalized optimal output for a given input by finding a pattern for a given input value and target value through learning (Goodfellow et al., 2016). The model is as shown in the figure below. It consists of an input layer, a hidden layer, and an output layer. In case of a linear combination that multiplies the input value by a weight and subsequently transforms it nonlinearly by an activation function. It has a structure that transmits or outputs to a layer (Fig. 2.11).

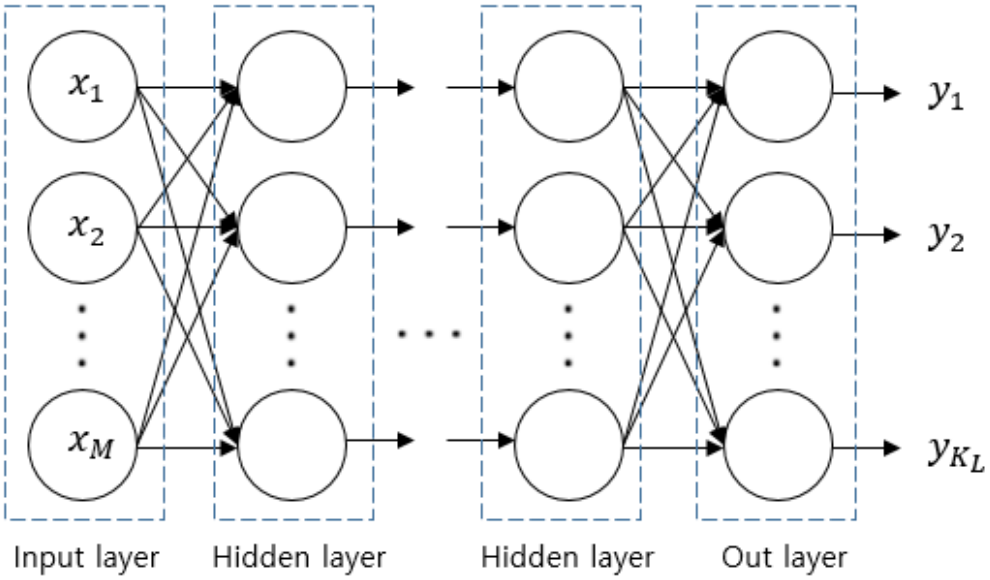


Figure 2.11 Structure of artificial neural network

An RNN, which is belongs to deep learning, is an effective deep learning technique for learning time series data from a structure in which a specific part is repeated (Lee,

2017). RNNs are multilayer perceptrons. The structure of an RNN is simply composed of three levels: an input layer, a hidden layer, and an output layer, similar to an artificial neural network. However, the part that is different from the basic artificial neural network is that the input and output layer of the neural network are influenced. However, RNNs have the disadvantage of long-term dependence, which prevents effective learning owing to the gradient loss when processing the current node and the distant past.

To solve the long-term dependency problem of RNNs, Hochreiter and Schmidhuber (1997) developed the LSTM model. This model is known to be more advantageous in predicting time series data because it can solve the problem of gradient loss that causes long-term dependence, which is a disadvantage of conventional RNNs (Q.-K. Tran and Song, 2017). The main flows in an LSTM model are a memory cell that can maintain state over time and three nonlinear gates that regulate the flow of data into and out of the cell (Figure 2.12).

Each gate of the LSTM is as follows. In the first stage, the forget gate (ft) receives the previous state of h_{t-1} and the new input, x_t, in the cell state, and decides what information to discard. This step of selecting information to be maintained through the cell state is expressed as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Here, σ is the activation function, W_f is the weight, and b_f is the bias.

In second step, the input gate (it) decides which new information to store in the input cell state. First, a sigmoid function determines the value to be updated, and subsequently, a new cell state C_t is created using a hypertangent function, which is expressed as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\widetilde{C_t} = \sigma(W_c \cdot [h_{t-1}, x_t] + b_c)$$

Here, $\tilde{C}t$ is the state of the newly updated cell, Wi and Wc are the weights of the function, and bi and bc are the biases of the function.

In the last step, the output gate (ot) decides what to output using an activation function. In addition, the output value, ht, of the current time is updated using the hypertangent function.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \cdot \tanh(C_t)$$

Here, ht denotes the current output value that is input to the next step.

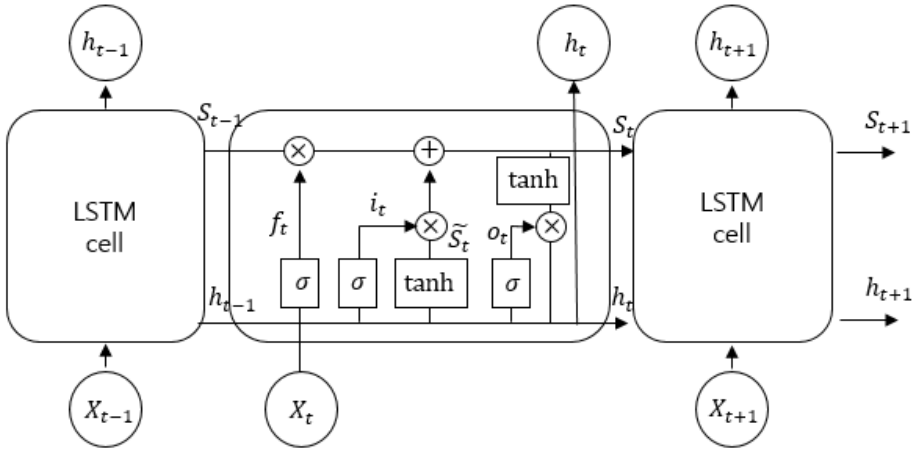


Figure 2.12 Structure of LSTM

CHAPTER 3. STATISTICAL ANALYSIS FOR FLOOD SEASON

3.1 Are Rainfall Patterns for the Korean Peninsula Changed in the 21 Century?

According to the sixth report of the Intergovernmental Panel on Climate Change, extreme weather such as heavy rains and flooding are predicted to become more frequent and severe in the future owing to the increase in the global temperature. Korea has also experienced abnormal climate such as the longest rainy season in the history and an increase in the frequency of typhoons in 2020, which caused severe social and economic damage. This flood damage has revealed the limitations of countermeasures against floods in Korea, and the designated flood season does not appropriately reflect the characteristics of climate change. In this chapter, the problems of the current flood season are identified by analyzing parts about climate change and regional characteristics using statistical techniques on two data groups: 20th and 21st century groups.

Prior Research

Precipitation in Korea has been steadily increasing, particularly in summer. As a preliminary study, to analyze the pattern changes in the flood season in the 20th

century (Journal of the Korean Geographical Society, 2004), four regions in Korea were examined (midwestern, Honam, Yeongdong and Yeongnam, and Jeju island regions). For the evaluation, the data were divided into two groups: from 1941 to 1970 and from 1971 to 2000. The results of the study confirmed that the precipitation in the four regions increased by 2.8 mm, 57.7 mm, 28.3 mm, and 38.9 mm, respectively. In addition, when comparing regionally, the standard deviation was 19.84, which is quite high. It was confirmed that the precipitation characteristics of each region in Korea are strong, and the difference between them has been gradually increasing even from the mid-20th century. Table 3.1 summarizes the abovementioned research.

Table 3.1 Thirty-year average monthly precipitation increase and decrease by region

Area / M Y	Midwest			Honam			Yeongnam & Yeongdong			Jeju island		
	41-70	71-00	Increase & Decrease	41-70	71-00	Increase & Decrease	41-70	71-00	Increase & Decrease	41-70	71-00	Increase & Decrease
6	156.9	122.1	-34.8	161.0	189.8	+28.8	142.3	152.7	+10.4	162.3	189.8	+27.5
7	333.4	294.7	-38.7	236.9	255.0	+18.1	222.5	202.4	-20.1	216.0	232.3	+16.3
8	211.5	318.5	+107.0	191.3	238.5	+47.2	166.6	221.5	+54.9	223.1	258.0	+34.9
9	162.7	132.1	-30.6	171.8	135.5	-36.6	174.4	157.6	-16.8	228.0	188.2	-39.8
Sum	864.5	867.3	+2.8	761.1	818.8	+57.7	705.8	734.2	+28.4	829.4	868.3	+38.9

3.1.1 Basic Statistics

In this study, to examine whether climate change is considered in the current flood season, the flood season rainfall patterns of the 20th and 21st centuries were quantified and compared by statistical techniques. First, basic statistical analysis was performed. This is the most important step for data analysis and provides the most basic characteristics of data.

The data used for the basic statistical analysis were the average annual data of accumulated precipitation (mm) by year from 1971 to 2000 as the data of the 20th century and the accumulated precipitation (mm) by year from 2001 to 2020 as the 21st century. Additionally, the cumulative precipitation during the flood period was quantified by comparing the 20th and 21st century results for the same period.

The results of the basic statistical analysis for the 20th and 21st centuries are shown as boxplots in Figure 3.1. The increasing trend can be intuitively confirmed. Numerically, the annual cumulative precipitation in the 21st century increased by 45.47 mm from 1181.8 mm to 1227.27 mm on average compared to that in the 20th century. In addition, the annual cumulative precipitation during the flood period increased by 105.91 mm from 717.27 mm to 818.18 mm.

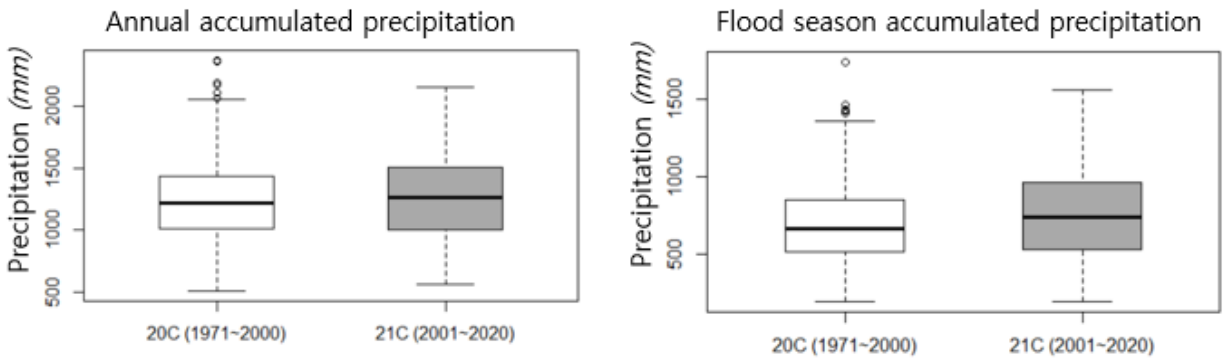


Figure 3.1 a) Boxplots showing annual cumulative precipitation changes in 20th and 21st centuries, b) Boxplots showing annual cumulative precipitation during the flood season changes in 20th and 21st centuries

3.1.2 Hypothesis Test

In this study, an increase in precipitation was additionally confirmed through hypothesis testing. For the changes in precipitation in the 20th and 21st centuries of the flood season, 18 multi-purpose dam basins in Korea were conducted. The null hypothesis of the hypothesis test is the $\mu_0 = \mu_1$, and the μ_0 is the average of the accumulated precipitation during the flood period of the 20th century, and μ_1 was designated as the average of the accumulated precipitation during the flood period of the 21st century, and a one-sided test was conducted.

When the 17 dams except for Gunwi Dam, which did not meet the assumption conditions, were carried out, it can be seen that the change in precipitation shows an increasing trend in a total of 7 dam basins including the Seomjingang Dam. Through this, almost half of the dam basins show an increasing trend in precipitation, confirming the increasing trend in precipitation.

3.2 Are Regional Rainfall Patterns Changed in the 21 Century?

3.2.1 Basic Statistics

In this study, the 21st century rainfall patterns were quantified by region using statistical techniques. For the statistical comparison, the analysis was conducted by dividing the data into two groups—20th and 21st centuries—identical to in the basic statistics analysis described in Chapter 3.1.1. The accumulated precipitation (mm) by year from 1971 to 2000 of each of the 18 domestic multipurpose dams was used as the 20th century data, and that from 2011 to 2020 composed the 21st century data. In addition, the change in the accumulated precipitation (mm) during each flood season was analyzed for the same period.

A comparison of the rainfall data of each dam basin is shown as boxplots in Figure 3.3. Examining the overall increase/decrease trends, the annual accumulated precipitations of all dams except the Gunwi, Miryang, Buan, and Boryeong dams showed increasing trends, confirming that the annual accumulated precipitations in the dam basins increased by approximately 78%. Based on the median, which is a basic statistic that is lesser affected by outliers than other indexes, the Juam dam showed an increase by approximately 400 mm, whereas the Boryeong dam present a decrease by approximately 80 mm.

Comparison of the flood season cumulative data showed similarity to trend discussed above (Figure 3.4). Overall, the precipitation in the flood seasons of all 18 dams except the Boryeong dam, i.e., approximately 94% dam basins, showed

increasing trends compared to the those in the 20th century. Among them, the difference of Juam dam increased the most, whereas that of the Boryeong dam decreased the most, by approximately 290 mm based on the median value.

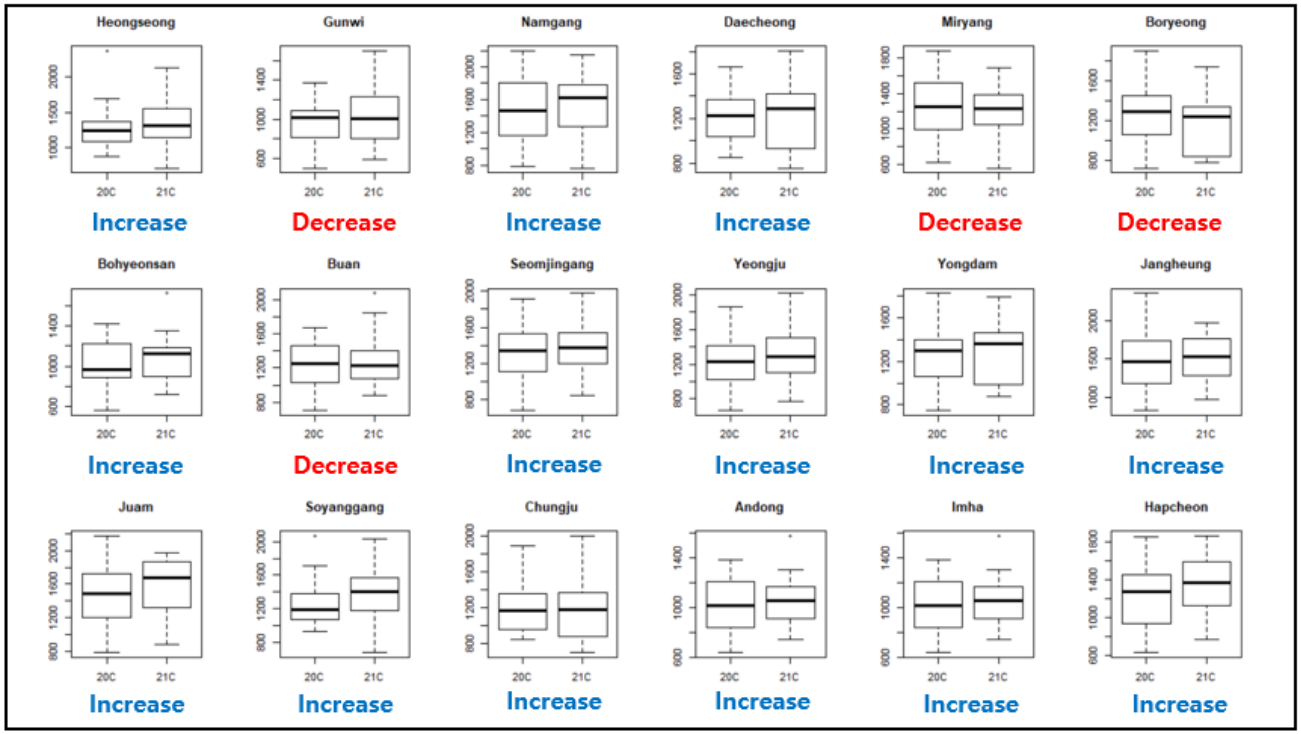


Figure 3.2 Boxplots showing annual cumulative precipitation changes in 20th and 21st centuries (20C and 21C, respectively) of multipurpose dams across country

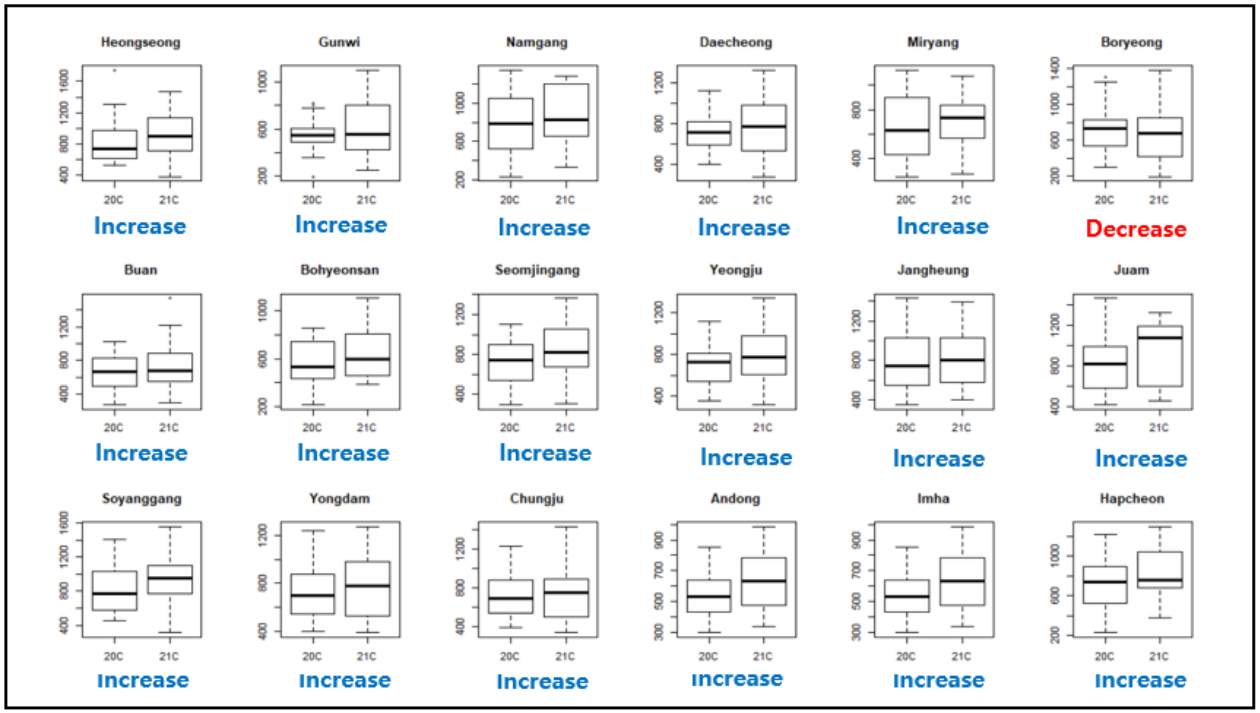


Figure 3.3 Boxplots showing annual cumulative precipitation during flood season changes in 20th and 21st centuries (20C and 21C, respectively) of multipurpose dams across country

3.2.2 Hypothesis Test

In this study, the 21st century rainfall patterns were quantified by region using statistical techniques. Statistical comparisons and hypothesis testing confirmed increasing trends for the changes in the precipitations in the 20th and 21st centuries. However, in Japan, the flood seasons for all dams are different, and the degrees of changes in the precipitation for all dams are compared considering the strong regional characteristics of the recent precipitation.

Therefore, in this study, analysis of the 21st century rainfall pattern change was conducted for each multipurpose dam in Korea by hypothesis tests. The accumulated precipitation data during the flood period from 1971 to 2000 were taken as the 20th century data and those from 2001 to 2020 as the 21st century data.

First, the normality was confirmed by Shapiro–Wilk test and QQ-plots. Moreover, the equality of variance was confirmed by F-tests. Based on the results, all the multipurpose dams except the Gunwi dam confirmed the establishment of the condition for using the independent sample T test.

The null hypothesis in this study is that “the average of the accumulated precipitation of a dam during the flood period does not change,” and the alternative hypothesis is that “the average of the accumulated precipitation of a dam during the flood periods of the two periods increases.” For this, a T-test of a one-sided test was performed.

As the significance level, the most commonly used value of 0.05 was used,

and the results were ranked according to the p-values from the hypothesis tests. Table 3.2 compares the degrees of changes in the accumulated precipitations during the flood periods of 17 domestic multipurpose dams.

It shows that the Seomjingang dam has a p-value of 0.038, whereas the Daecheong dam has a p-value of 0.3. Consequently, the p-value range shows a large difference. Thus, it is determined as additional evidence that the difference in the precipitation characteristics for each dam basin, from the basic statistical analysis discussed in Chapter 3.2.1, is significant.

Table 3.2 Ranking based on the p-values from hypothesis tests of multipurpose dams across country

Rank	Dam	Null hypothesis reject p-value	Used data (yr)
1	Seomjingang	0.038	1973~2000, 2001~2020
2	Hapcheon	0.059	1973~2000, 2001~2020
3	Bohyeonsan	0.083	1973~2000, 2001~2020
4	Soyanggang	0.096	1971~2000, 2001~2020
5	Andong, Imha	0.101	1973~2000, 2001~2020
7	Yongdam	0.130	1973~2000, 2001~2020
8	Juam	0.133	1973~2000, 2001~2020
9	Buan	0.138	1973~2000, 2001~2020
10	Yeongju	0.144	1973~2000, 2001~2020
11	Namgang	0.200	1973~2000, 2001~2020
12	Jangheung	0.206	1973~2000, 2001~2020
13	Hoengseong	0.215	1973~2000, 2001~2020
14	Boryeong	0.221	1973~2000, 2001~2020
15	Chungju	0.281	1973~2000, 2001~2020
16	Miryang	0.289	1973~2000, 2001~2020
17	Daecheong	0.300	1973~2000, 2001~2020

Based on analysis using the previous four statistical techniques, this study confirmed two problems of the current flood season.

First, it does not consider the increase in the precipitation in the 21st century compared to that in the 20th century. The first report describing the current flood season is the Soyanggang Dam Management Regulations of 1974. Thus, it can be inferred that the current flood season was designated before 1974, i.e., the same period has been adopted for more than 49 years. It is determined that the current flood season does not consider the climate change, as discussed in Chapter 3.1.1.

Second, strengthening of the regional precipitation characteristics is not considered. The same period is adopted for the current flood season across the country, which is interpreted as not considering the differences in the size and trend of increasing precipitation in each region, as confirmed from the results in Chapter 3.1.2. Thus, in Korea, a new flood season should be established in the same direction as in Japan, by considering the difference in the precipitation characteristics for each water system.

CHAPTER 4. TESTING FOR FLOOD SEASON ADJUSTMENT

4.1 Study Basin: Yongdam Multipurpose Dam

4.1.1 Hydrological Characteristics

In this study, the Yongdam dam basin of the Geum river was finally selected by considering the four criteria defined in Chapter 4.1.1. The Yongdam dam has storage capacities of 695.8 million m^3, 762.6 million m^3, and 833.3 million m^3 under the RWL, NHWL, and FWL, respectively. Because this study was conducted by hour, by scaling, the design discharge amount becomes 11.56 million m^3/h and the nondamage discharge becomes 1.08 million m^3/h.

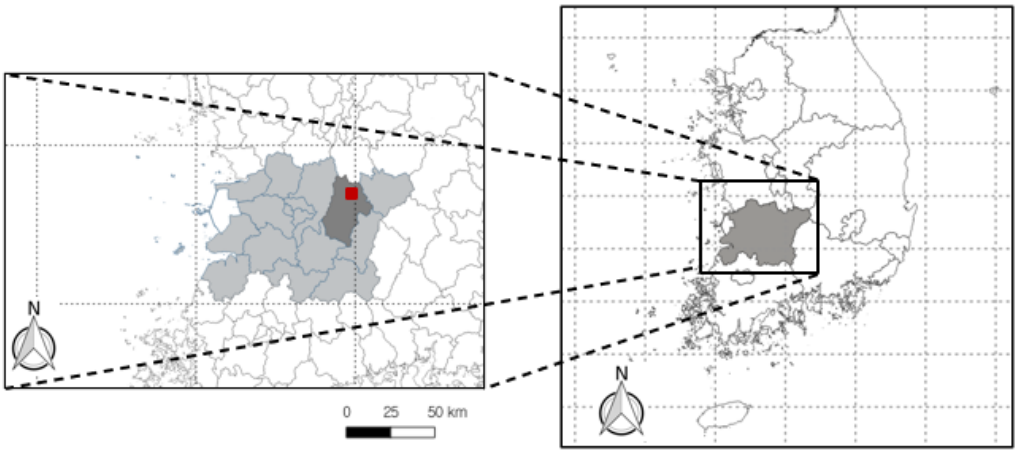


Figure 4.1 Map of study area

4.1.2 Reservoir Operation Principle

The study area was selected considering the results of the hypothesis test presented in Chapter 3. Moreover, the watershed operating dams were those examined by the simulation method established in this study, whether RWL was applied, and the dam basin located upstream of the river.

Dam Basin Using Rigid ROM

Multipurpose dams in Korea are operated by reservoir operation techniques. The types and characteristics of the ROMs are summarized in Chapter 2.2. Among them, in this study, a dam using the rigid ROM was adopted considering that a simulation method applying the rigid ROM to the basic rules of dam operation was established to determine the amount of discharge. Table 4.1 lists the ROMs for all multipurpose dams in Korea.

Table 4.1 ROMs used by multipurpose dams in Korea

Number	Dam	Basin	Application operation method
1	Soyanggang	Han River	Rigid+Technical ROM
2	Chungju	Han River	SRC ROM
3	Hoengseong	Han River	Rigid ROM
4	Andong	Nakdong River	Technical ROM
5	Imha	Nakdong River	Rigid ROM
6	Seongdeok	Nakdong River	Technical ROM
7	Yeongju	Nakdong River	Technical ROM
8	Gunwi	Nakdong River	Rigid ROM
9	Gimcheonbuhang	Nakdong River	Rigid ROM
10	Bohyeonsan	Nakdong River	Technical ROM
11	Hapcheon	Nakdong River	Rigid ROM
12	Namgang	Nakdong River	Technical ROM
13	Miryang	Nakdong River	Rigid ROM
14	Yongdam	Geum River	Rigid ROM
15	Daecheong	Geum River	Rigid ROM
16	Seomjingang	Seomjin River	Technical ROM
17	Juam	Seomjin River	Rigid ROM
18	Boryeong	etc	Rigid ROM
19	Jangheung	etc	Technical ROM

Comparison of Degree of Climate Change by Dam Basin

A watershed with a small p-value was selected from the results of the hypothesis tests. This is because, compared to other watersheds, its rainfall pattern in the 21st century is relatively larger than that in the 20th century. In Chapter 3, it is confirmed that there is a large difference in the increase and decrease in the precipitation by region. Therefore, to compare the magnitude of the flood reduction effect, a pilot watershed was selected as an area with a large variation in the precipitation due to climate change. Table 4.2 compares the effect of climate change by river for the multipurpose dams in Korea.

Table 4.2 Comparison of effect of climate change by river for multipurpose dams in Korea

Number	Dam	Basin	Reject null hypothesis p-value
1	Soyanggang	Han River	0.096
2	Chungju	Han River	0.281
3	Hoengseong	Han River	0.215
4	Andong	Nakdong River	0.101
5	Imha	Nakdong River	0.101
6	Seongdeok	Nakdong River	-
7	Yeongju	Nakdong River	0.144
8	Gunwi	Nakdong River	-
9	Gimcheonbuhang	Nakdong River	-
10	Bohyeonsan	Nakdong River	0.083
11	Hapcheon	Nakdong River	0.059
12	Namgang	Nakdong River	0.200
13	Miryang	Nakdong River	0.289
14	Yongdam	Geum River	0.130
15	Daecheong	Geum River	0.300
16	Seomjingang	Seomjin River	0.038
17	Juam	Seomjin River	0.133
18	Boryeong	etc	0.221
19	Jangheung	etc	0.206

Dam Basin Applying RWL

For the multipurpose dams in Korea, the NHWL is applied as the limiting water level during the dry season except during the flood season. In addition, during the flood season, for the dam operation, the water limit is changed to the flood season limit water level. However, in some dams, the NHWL is maintained even during the flood season because of the determination of no requirement to lower the water level limit. In this study, the simulation method determined lowering the flood season limit water level as the basic operation rule of a dam. Therefore, application of the flood season limit water level was designated as the pilot watershed selection criterion. Table 4.3 summarizes this.

Table 4.3 RWL application by river for multipurpose dams in Korea

Number	Dam	Basin	Use of restricted water level
1	Soyanggang	Han River	○
2	Chungju	Han River	○
3	Hoengseong	Han River	○
4	Andong	Nakdong River	-
5	Imha	Nakdong River	○
6	Seongdeok	Nakdong River	○
7	Yeongju	Nakdong River	○
8	Gunwi	Nakdong River	-
9	Gimcheonbuhang	Nakdong River	○
10	Bohyeonsan	Nakdong River	-
11	Hapcheon	Nakdong River	-
12	Namgang	Nakdong River	-
13	Miryang	Nakdong River	-
14	Yongdam	Geum River	○
15	Daecheong	Geum River	-
16	Seomjingang	Seomjin River	-
17	Juam	Seomjin River	-
18	Boryeong	etc	-
19	Jangheung	etc	○

Dam Basin Upstream of River

In addition, for the accuracy of the inflow prediction, which is presented in Chapter 4.2, a dam located upstream of a river that is least affected by all variables except the inflow and the precipitation was selected. Accordingly, the Soyang river dam in the Han river, Imha dam in the Nakdong river, and Yongdam dam in the Geum river were selected. Among them, the Yongdam dam, which is the most upstream, was finally selected.

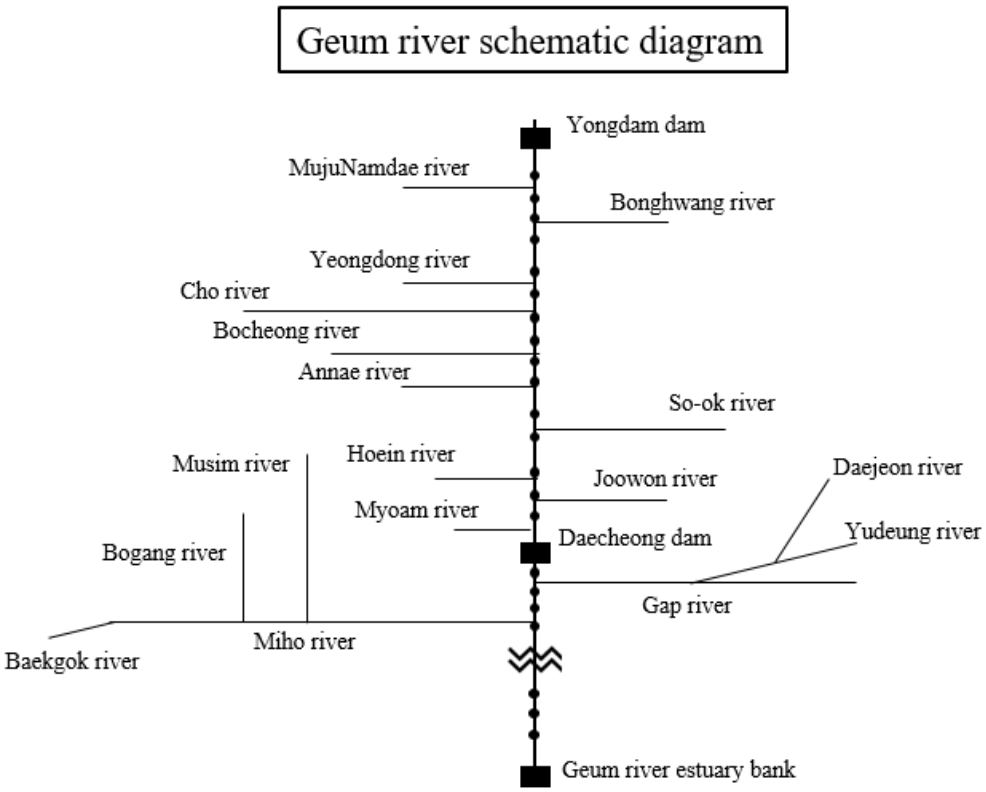


Figure 4.2 Schematic of Geum river

4.2 Flood Season Adjustment Proposal

4.2.1 Analysis of Flood Season Beginning and Ending

A new flood season, Ex1, was determined using the starting date of the Korean rainy season and the last date of the last typhoon affecting Korea, which were the basis for establishing the current flood season as described in Chapter 3.1.3. From Chapter 3.1.3, the basis for establishing the current flood season is that the starting point is the entry point of the rainy season front into Korea and the ending point is the time of the last typhoon that affected Korea. Accordingly, the extent of change in the basic statistics of each variable due to climate change was confirmed using the data of the most recent period available from white papers and applied to propose a new flood season. For the rainy season, using the data from 1991 to 2020, the entry point was confirmed as 5 days earlier than that based on the 30-year data at the time of flood period establishment. The last typhoon affecting Korea was confirmed to be delayed by approximately 10 days using data from 1981 to 2010. Considering this, the current flood season from June 21 to September 20 was changed by 15 days, and a new period from June 16 to September 30 was proposed.

Table 4.4 a) Comparison of changes in starting point of rainy season by applying 30-year data at time of flood period establishment and most recent 30-year data to inferred method, b) Comparison of changes in ending point of last typhoon affecting Korea by applying 30-year data at time of flood period establishment and most recent 30-year data to inferred method

a)

	Date at the time of enactment (1961 ~ 1973)	Recent date (1991 ~ 2020)
Minimum	June 14	June 10
Median	June 24	June 18
Maximum	July 1st	June 26
Mean	June 23	June 18
Standard deviation	4.13	4.15

b)

	Date at the time of enactment (1941 ~ 1970)	Recent date (1981 ~ 2010)
Minimum	Early August	Early August
Median	Early September	Mid-September
Maximum	Mid-October	Mid-October
Mean	Mid-September	Late September
Standard deviation	1.9	1.6

4.2.2 Proposal of Adjustment Candidates

Considering climate change, six additional flood seasons were proposed based on expert opinions. For Ex2–4, considering the relatively insignificant amount of precipitation at the starting point of the flood season, only the ending point was changed by 1 day, 5 days, and 10 days, respectively, without any change in the starting point. Extension 5 was set as a case in which the change between the starting and ending points was large, and the flood season was considerably extended from June 1 to September 30. Regarding the case, the total number of days of the current flood season was not changed, and only the period was shifted. Sh1 was set as from June 16 to September 15, advancing the flood period by five days considering only the change in the rainy season entry time. Sh2 was established as from July 1 to September 30, delaying by ten days considering only the change in the last typhoon affecting Korea (Figure 4.3).

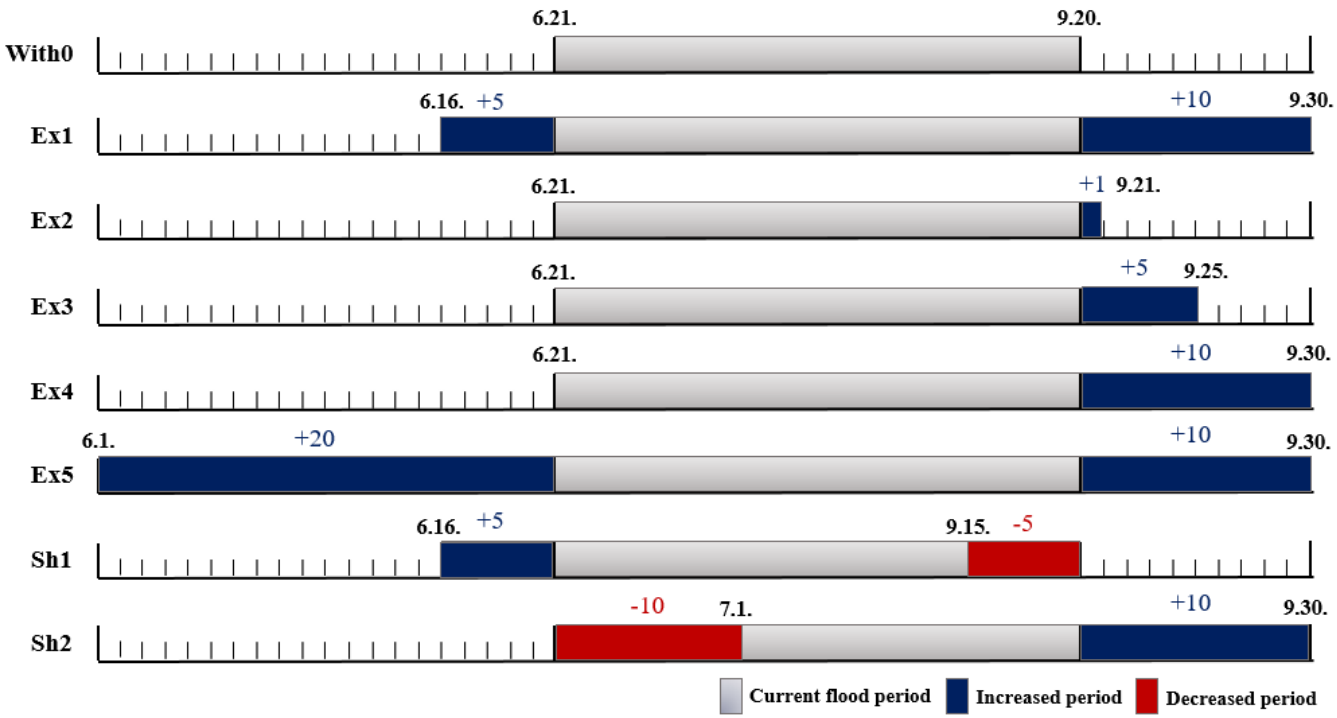


Figure 4.3 Current flood season and seven additional flood seasons proposed in this study

4.3 Hourly Inflow Forecasting with LSTM

The inflow hydrologic curve predicted during the current flood period was input into a simulation program to determine the amount of discharge and proceed with the operation of a dam. In this study, the inflow data to be used with the simulation method were predicted and derived by deep learning. In this study, the optimal model was found by adjusting the input data and parameters of the LSTM model, and by adjusting the lead time, the most suitable value was selected.

4.3.1 Input Data for LSTM

Input data

This study used two input data to predict the hourly inflow. The first were the observed inflow data, for which the data from K-water were used, and the equation is expressed as follows:

$$\hat{Q}(t+l) = f[Q(t), Q(t-1), \cdots, Q(t-r)]$$

Here, t denotes the current time and l is the lead time. r is the sequence length and represents the time consumed for forecasting. (t) denotes the observed inflow at time t.

The second input data used were the observed inflow and precipitation data. According to Mok et al. (2020), high-accuracy results are obtained when observed precipitation data are used for peak inflow forecasting.

Considering this, the data from K-water were obtained and used as input data. This can be expressed as

$$\hat{Q}(t+l) = f[R(t), R(t-1), \cdots, R(t-r); Q(t), Q(t-1), \cdots, Q(t-r)]$$

As in the previous equation, t denotes the current time, l denotes the lead time, and r is the sequence length. (t) is the observed precipitation at time t, and (t) is the observed inflow at time t.

Data preprocessing

When a neural network model learns a wide range of data, the function values diverge and degrade the prediction performance; therefore, the data are processed into useful information by several methods. This process is called data preprocessing, and typical preprocessing methods include normalization and standardization. In addition, outliers and missing values are removed and used, respectively; however, this process was not performed in this study because the outliers were the main data and not many missing values were found. Because the hourly inflow data of the Yongdam dam had a very wide range, from a minimum value of 0.1 $m3/s$ to a maximum of 3,373.2 $m3/s$, a preprocessing process was absolutely necessary. Scaling was performed using the following regularization formula:

$$Y_i = \frac{X_i - X_{min}}{X_{max} - X_{min}}$$

Here, Y_i is the normalized variable value, X_i is the actual variable value, X_{min} is the minimum value of the variable, and X_{max} is the maximum value of the variable.

Lead time

In the hourly inflow forecasting, the lead time of the number of hours to predict is very important in determining the discharge amount. Specifically, a long lead time is good for preparing for the future; however, if the lead time is extremely long, the accuracy is lowered, which may lead to insufficient flood countermeasures owing to the incorrect predicted inflow hydrologic curve. Therefore, obtaining the appropriate lead time is important. At this stage, determining the minimum time required for K-water to instruct a multipurpose dam management office to release the dam is necessary, and this is done in the order as shown in Figure 4.4. K-water reports the dam discharge plan to the Han River Flood Control Center from 16:00 of the previous day to 16:00 of the same day. After the dam discharge is approved, the multipurpose dam management office is instructed to release the water. At this time, the minimum time required to take precautionary measures before floodgate operation, such as downstream patrol, alarm broadcasting, and inspection, is 3 h. Therefore, the minimum lead time in this study was set as 3 h. The lead time was determined using

the average ROM application time in the simulation performed with the current flood season. Because the ROM application time is the length of the flood event, it is important for determining the amount of discharge. The average length of flood events in the simulation method of this study is approximately 7 h. Therefore, a second lead time was selected as 7 h. Finally, for a third lead time, setting the maximum reporting time as 27 hours would be optimal; however, it will cause the accuracy to sharply drop. Therefore, the third lead time was set as 12 h, which is half the reporting time, and prediction is made over a longer time according to the accuracy of the lead time of 12 h. It was carried out in a way to determine whether to proceed.

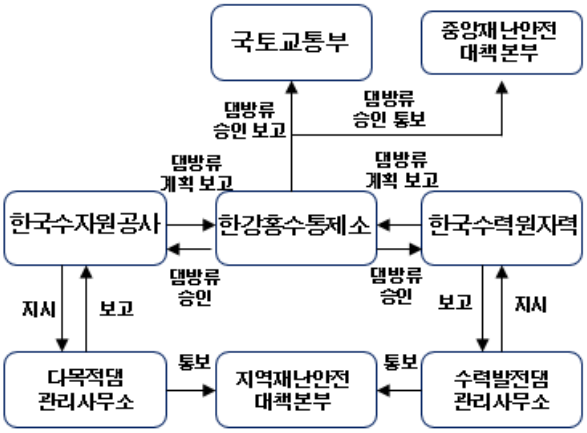


Figure 4.4 Flowchart of dam discharge decision process

4.3.2 LSTM Model Calibration

Parameter

1) Sequecne length

The sequence length determines the number of hours (or units) of data of the past that will be used in the model to predict data at a certain time.

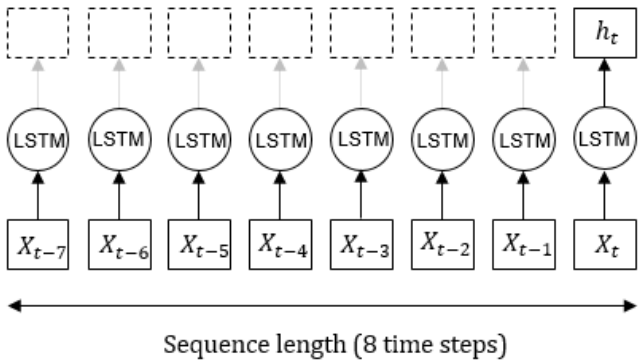


Figure 4.5 Sequence length

In this study, the period during which the ROM was used in the simulation method adopted in the last ten years based on the study area was considered. The evaluation was conducted in three steps: an average period of 7 h, a maximum time of 24 h, and an intermediate period of 16 h.

2) Hidden layer

A hidden layer is a layer between the input and output layers. In detail, it is a layer with a perceptron added between the input layer composing the input features and the output layer composing the output values. Consequently, many hidden layers imply a deep network and high performance. However, if there are excessive hidden layers, there is a risk of overfitting. The method for determining the number of hidden layers is not established, and in general, the optimal hidden layers is obtained based on the

number used in previous studies of the same system or by a trial-and-error method. In this study, ten commonly used hidden layers were fixed and applied based on previous studies.

3) Learning rate

The learning rate is an indicator that enables fast learning when it is high; however, if the minimum value is not determined, problem of overshooting can occur, as shown in Figure 4.6; therefore, adjustment is required. Kingma and Ba (2014) stated that the optimal learning rates in Adam's technique are 0.001 and 0.002, from which 0.001 was selected, fixed, and used in this study.

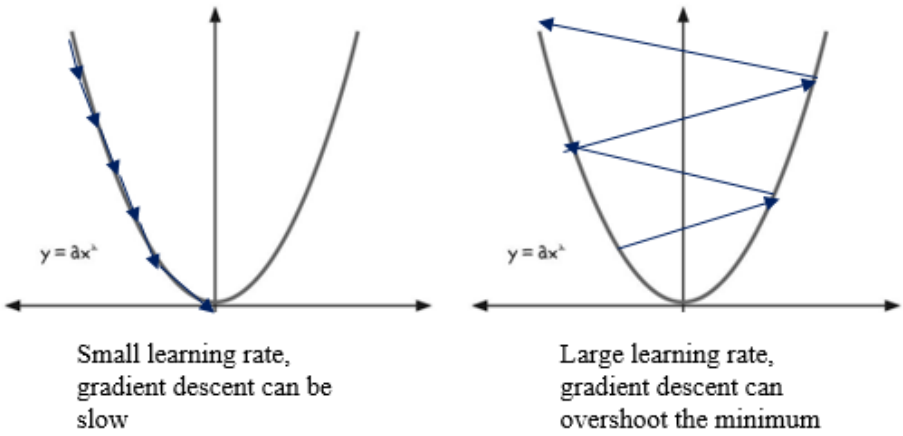


Figure 4.6 Problems with learning rate value

4) Epoch

Epoch is the number of learning iterations. In this study, the optimal epoch for each case was determined using a function called the Earlystopping function to avoid

overfitting (Figure 4.7).

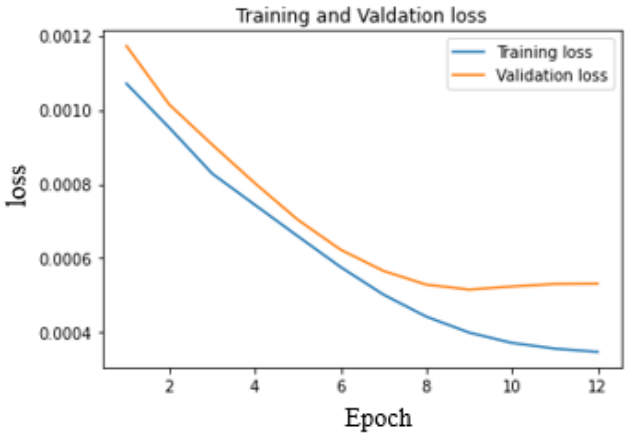


Figure 4.7 Training and validation losses as functions of epoch

Accuracy evaluation index

Predictions typically contain errors. Errors can be classified into systematic errors, which are repetitive errors, and random errors, which are not repetitive. Predictions are evaluated in terms of precision and accuracy. Precision only implies that the predicted values are close to each other, and it corresponds to a random error. Accuracy is the sum of the precision and the unbiasedness, and it implies that the estimated values are close to the true values. Therefore, accurate comparison of model performance is possible only when both indicators indicating the deviation and accuracy are examined.

1) Mean Square Error

The mean square error (MSE) represents the average variability of the prediction error.

It can be expressed by squaring the difference between the predicted and true values as follows:

$$MSE = \frac{\sum_{i=1}^{n}(y - \hat{y})^2}{n}$$

Here, y represents the true value and $\hat{y}$ represents the predicted value. n is the number of predictions.

The characteristic of the MSE is its sensitivity to outliers because the difference between the predicted and true values is squared. Specifically, when the difference between the predicted and true values is large, it will be reflected in the error value. In addition, because the square of the error value is taken, the reflection degree is different when the error sizes are 0–1 and 1 or higher.

2) Root Mean Square Error

The root mean square error (RMSE) is an evaluation index that is the root of the MSE, and can be expressed as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y - \hat{y})^2}{n}}$$

It alleviates the weakness of the sensitivity of the MSE to outliers to some extent by taking the root. Because its sensitivity to outliers is higher than that of the mean absolute error (MAE), outliers are considered better by the MAE than by the

previous two indicators.

3) Mean Absolute Error

The MAE is a criterion for evaluating the absolute magnitude of the prediction error and is defined as follows:

$$MAE = \frac{\sum_{i=1}^{n} |y - \widehat{y}|}{n}$$

Compared to the previous two indicators, this evaluation indicator has a characteristic of high robustness to outliers.

4) Relative Root Mean Squared Error

The relative RMSE (RRMSE) is an index indicating the degree of error as a relative percentage by dividing the RMSE by the sum of the predicted values and multiplying by 100. The results are the same as for the RMSE; however, the advantage is its relative comparability. It is expressed as

$$\text{RRMSE}(\%) = \frac{\sqrt{\frac{\sum_{i=1}^{n}(y - \hat{y})^2}{n}}}{\sum_{i=1}^{n} \hat{y}} \times 100$$

Despotovic et al. (2016) evaluated the model accuracy by dividing the RRMSE index into four sections: model accuracy is excellent when the RRMSE < 10%, good if 10% < RRMSE < 20%, fair if 20% < RRMSE < 30%, and poor if the RRMSE > 30%.

5) Bias

Bias is defined as the expected value of the prediction error. It expresses the difference between the predicted and actual values. Its formula is

$$Bias = \frac{\sum_{i=1}^{n}(\hat{y} - y)}{n}$$

6) Relative Bias

The relative bias (RBias) is a performance indicator expressing the bias as a relative percentage, and its formula is as follows:

$$\text{RBias}(\%) = \frac{\frac{\sum_{i=1}^{n}(\hat{y} - y)}{n}}{\sum_{i=1}^{n} y} \times 100$$

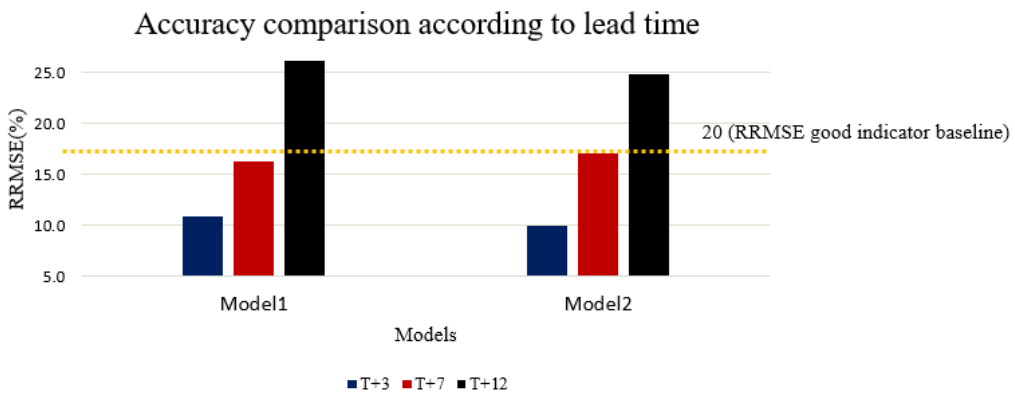
The MSE is a very sensitive to outliers, and is probable to make the training of a model unstable. The purpose of this study was to establish a flood season that is adapted to climate change, and because outliers were not processed in the preprocessing process, the MSE will adversely affect obtaining a model with good performance suitable, owing to interfering with the model fitting. In the case of the MAE, the outliers are not weighted; therefore, it was considered unsuitable for the evaluation of the inflow forecast during the flood season. Therefore, in this study, the RRMSE index, which is more sensitive to outliers than the MAE but lesser sensitive than the MSE, was selected. Additionally, accurate and precise comparison was

performed using the RBias, a relative indicator that evaluates only the preceding bias.

4.3.3 Model Selection Results

The models established using the two input data specified in Chapter 3.3.1 are denoted as Models 1 and 2, respectively, and sequence lengths, as specified in Chapter 3.3.2, were set as 1, 2, and 3 respectively, for a total of 27. Each model was evaluated relative to accuracy and bias using the RRMSE and the RBias, and the trends according to the input data, lead time, and sequence length were analyzed, based on which the optimal model was selected.

The first evaluation was a comparison of accuracy according to the lead time. In the case of input data, Model 1 using the observed inflow and Model 2 using the observed inflow and observed precipitation were compared by plotting a histogram (Figure 4.8).



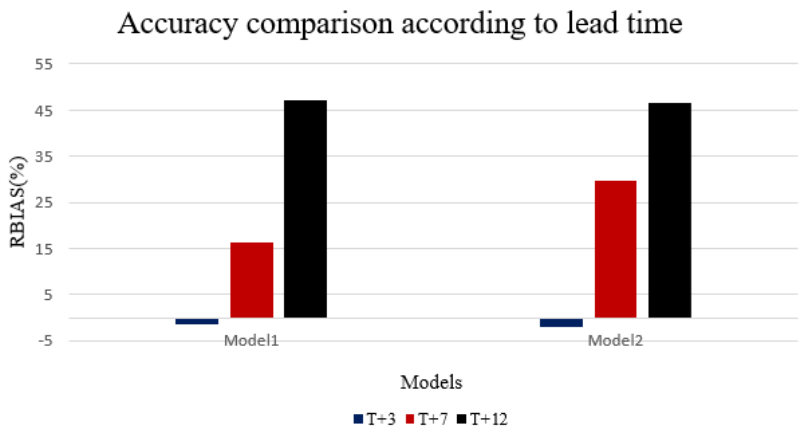


Figure 4.8 Accuracy comparison according to lead time

It can be seen that on going from lead time of 3 h to 7 h, the RRMSE is increased by 5.41% and 7.08% for Models 1 and 2, respectively, whereas on going from 7 h to 12 h, it is increased by 9.84% and 7.85%, respectively. In the case of the RRMSE, the change width increases as the lead time increases. In addition, when comparing 3 h and 12 h, the accuracy is considerably enhanced by 15.25% and 14.92% for Models 1 and 2, respectively. These results were analyzed for the four sections of the RRMSE evaluation index, and findings are summarized in Table 4.8.

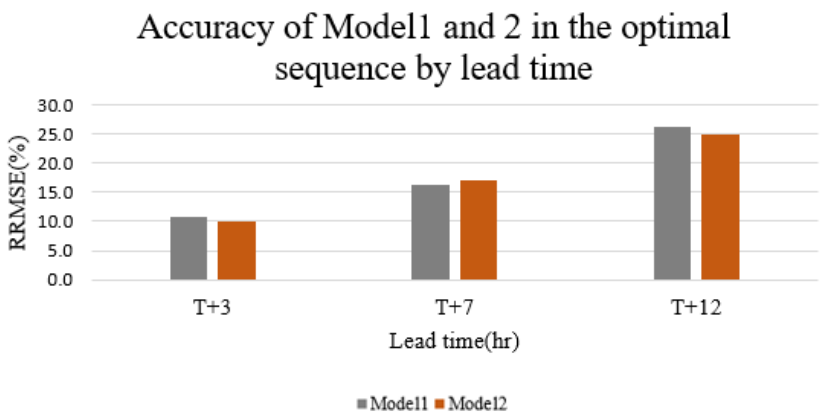
Table 4.5 Results of dividing RRMSE into four sections

	Model1-1	Model1-2	Model1-3	Model2-1	Model2-2	Model2-3
T+3	good	good	good	good	good	good
T+7	good	good	good	good	good	good
T+12	fail	fail	fail	fail	fail	poor

When the lead times are 3 h and 7 h, the RRMSEs are in good range of 10% < RRMSE < 20%. However, for 12 h, the RRMSE is fair with 20% < RRMSE < 30% for 5 times, and RRMSE > 30% once. As a result, it is judged that the accuracy for 12 hours is difficult to apply in this study.

The RBias shows a similar pattern. On increasing the lead time from 3 h to 7 h, the RBias values of Models 1 and 2 differ by 17.68% and 31.96%, respectively, based on the results of the most optimal parameters of this study. However, in the case of 12 h, the differences compared to the results of 3 h are 48.53% and 48.62%, respectively.

The second evaluation was a comparison between Models 1 and 2. Model 1 was trained only on the observed inflow, and to improve the peak prediction accuracy, Model 2 was trained by additionally considering the observed precipitation. Thus, the RRMSE of Model 2 decreased by 0.90%, increased by 0.76%, and decreased by 1.23% for lead times of 3 h, 7 h, and 12 h, respectively. In addition, the RBias decreased by 0.73%, decreased by 13.55%, and increased by 0.64%, respectively (Figure 4.9). Accordingly, it was confirmed that the addition of observed precipitation did not have a significant effect on the increase in the accuracy. Consequently, it was determined that it will not help improve the accuracy of the model, because of the many zero values of the observed precipitation in the hourly inflow forecast.



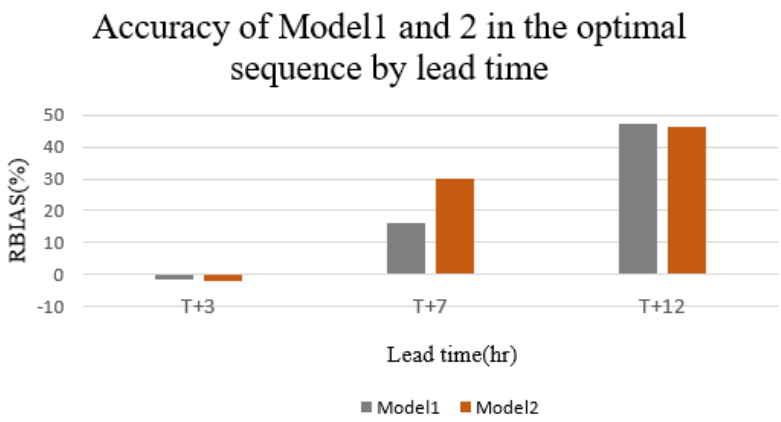


Figure 4.9 Accuracy of model1 and 2 according to optimal sequence length by lead time

Finally, the trend of accuracy change according to the sequence length was compared. Based on the RBias, maximum differences of 17.94% and 15.99% in Models 1 and 2, respectively, were confirmed. However, in the results of RRMSE, the differences were small, up to 2.56% and 4.52%, respectively (Figure 4.10). In addition, in the case of RBias, the optimal sequence lengths for the different lead times were 24 h, 7 h, and 7 h for Model 1 and 16 h, 16 h, and 7 h for Model 2, respectively. Based on the RRMSE, it was confirmed that the optimal sequence lengths for the different lead times were determined as 7 h each for Model 1 and 16 h, 7 h, and 7 h for Model 2, respectively. Thus, the sequence length can be analyzed as a parameter that show a large difference in the bias, instead of the accuracy, and it can be confirmed that there is no trend in the change in the sequence length for each lead time.

Finally, in this study, Model1-1 with a lead time of 7 h, when the RRMSE is good and the RBias shows a small difference of 5.41% with respect to a lead time of 3 h, was selected as the optimal prediction model for each hour.

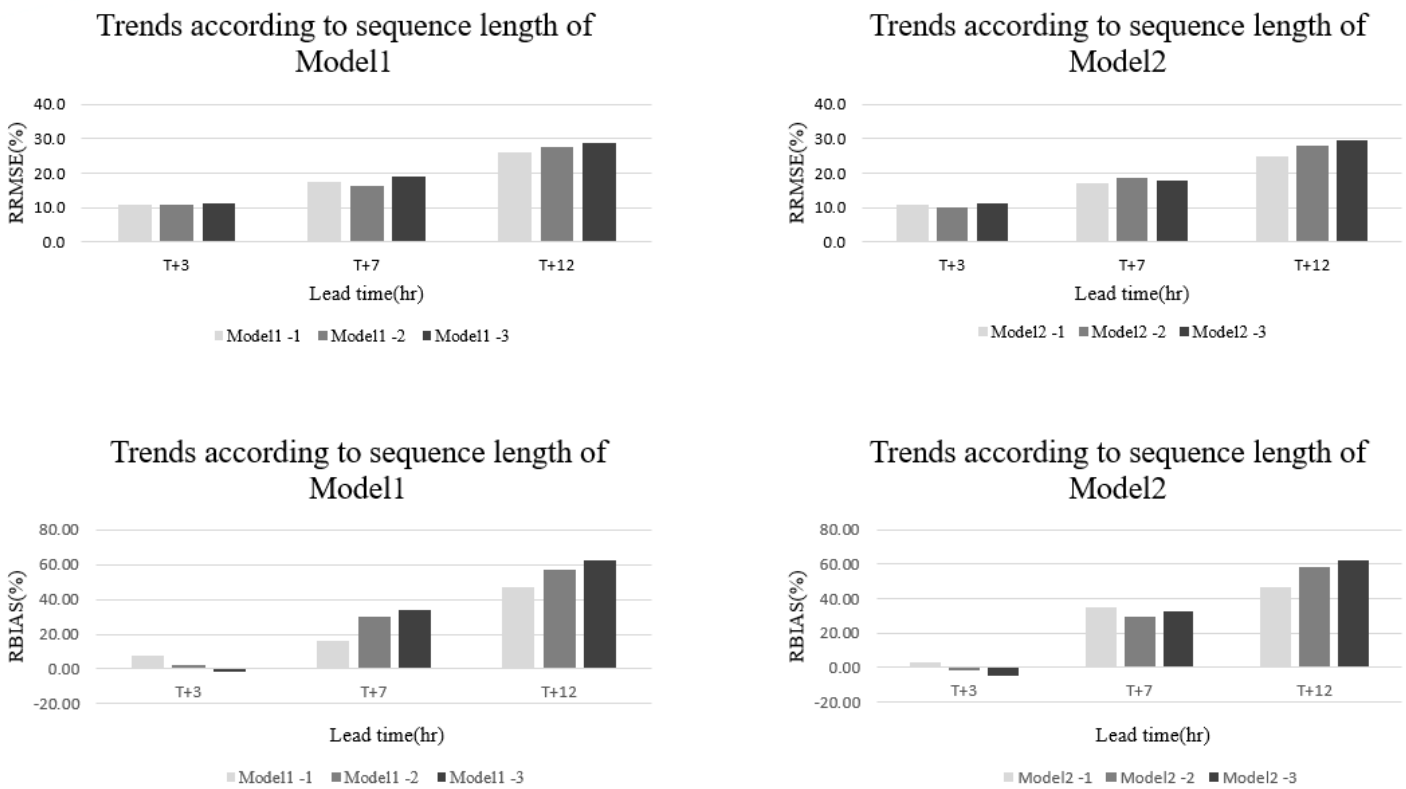


Figure 4.10 Trends according to sequence length of Models 1 and 2

4.4 Simulation and Evaluation

4.4.1 Simulaton with Rigid ROM

Basic rules for dam operation + Rigid ROM

In this study, a simulation method applicable to the study area was established, and the discharge amounts were derived for the seven newly established flood seasons using this method. The simulation method was established by referring to the basic rules of domestic multipurpose dam operation discussed in Chapter 2.1.1 and the sluice operation and detailed sluice operation methods for each multipurpose dam in Korea described in Section 2.1.2. As detailed dam manipulation methods, the methods used in dams such as the Andong, Yongdam, and Daecheong dams were considered. The actual application of each operation method is shown in Figure 4.11 and summarized in Table 4.6. Moreover, the number of times each technique is applied is also shown in Figure 4.11.

Table 4.6 ROMs adopted by domestic multipurpose dam by river

Number	Dam	Basin	Application operation method
1	Soyanggang	Han River	Rigid+Technical ROM
2	Chungju	Han River	SRC ROM
3	Hoengseong	Han River	Rigid ROM
4	Andong	Nakdong River	Technical ROM
5	Imha	Nakdong River	Rigid ROM
6	Seongdeok	Nakdong River	Technical ROM
7	Yeongju	Nakdong River	Technical ROM
8	Gunwi	Nakdong River	Rigid ROM
9	Gimcheonbuhang	Nakdong River	Rigid ROM
10	Bohyeonsan	Nakdong River	Technical ROM
11	Hapcheon	Nakdong River	Rigid ROM
12	Namgang	Nakdong River	Technical ROM
13	Miryang	Nakdong River	Rigid ROM
14	Yongdam	Geum River	Rigid ROM
15	Daecheong	Geum River	Rigid ROM
16	Seomjingang	Seomjin River	Technical ROM
17	Juam	Seomjin River	Rigid ROM
18	Boryeong	etc	Rigid ROM
19	Jangheung	etc	Technical ROM

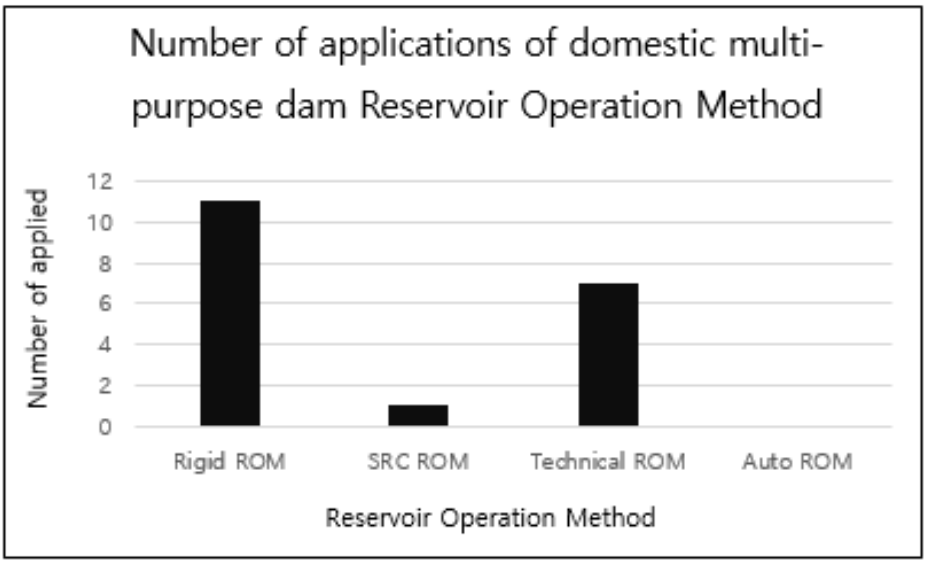


Figure 4.11 Number of applications per ROM in multipurpose dam in Korea

In summary, a total of 19 multipurpose dams were investigated including duplication: 11 times for the rigid ROM, 7 times for the technical ROM, 1 time for the SRC ROM, and 0 times for the auto ROM. Consequently, the simulation method using the Rigid ROM was selected.

The simulation method to be used in this study, which was established based on the basic operation rules of dams, operation methods of domestic multipurpose dams, and results of detailed operation investigations are as follows:

(1) The basic operating rules of the dam are applied as a prerequisite for dam operation.

(2) The limited water level is the RWL when the flood season rule is applied, and the NHWL is adopted when the non-flood season rule is applied. Below the

water level limit, all water except for power generation and water supply is stored.

(3) A dam is operated differently depending on the size of the inflow being above the limit water level and below the FWL. If the inflow is smaller than the nondamage discharge, the discharge amount is as much as the inflow. If the inflow exceeds the nondamage discharge but is smaller than the design discharge, the discharge amount is determined using the rigid ROM. Finally, if the inflow exceeds the design discharge, the discharge amount is as much as the design discharge.

(4) When the water level exceeds the FWL, the maximum possible water is discharged in the discharge capacity of the dam.

The simulation method used in this study summarized is also presented in Figure 4.12.

.

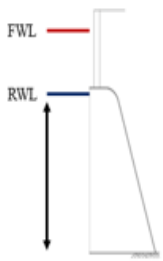
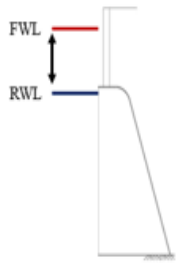
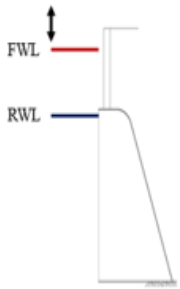
Criteria	Below limit		Above limit, below flood water level		Above flood water level	
Simulation		No release until reach the limit Power generation and water supply in progress		Inflow < Non-damaging release → Amount of inflow Non- damaging release < Inflow < Design release → Rigid ROM Inflow > Design release → Design release		Maximum release of dam
Details of release	Only power generation and water supply in progress		If Inflow < Non-damaging release, release as much as the inflow amount If Non- damaging release < Inflow < Design release, run Rigid ROM to decide the amount of release If Inflow > Design release, all released as much as design release		Maximum release of dam (Full release)	

Figure 4.12 Simulation rule established in this study

4.4.2 Evaluation Criteria

In this study, the discharge amount was derived using the established simulation method described in Chapter 3.4.1 during the existing flood season and the newly proposed flood season. A total of three evaluation methods including the existing evaluation method were used to evaluate the derived discharge, and two methods determined to be suitable for this study were finally adopted to evaluate the flood reduction effect.

Flood Control Indicators

Currently, the most extensively used index for quantifying the flood control effect of reservoirs in Korea is the flood control rate index. It is expressed as follows:

$$Q_c(\%) = \frac{I_{\max} - Q_{\max}}{I_{\max}} \times 100$$

$$S_R(\%) = \frac{R_T}{I_T} \times 100$$

$$S_S(\%) = \frac{S_T}{I_T} \times 100$$

$$S_U(\%) = \frac{L_{\max} - L_{FC}}{L_F - L_{FC}} \times 100$$

Here, Qc is the flood control rate (%), SR is the reservoir discharge rate (%), SS is the reservoir retention rate (%), and SU is the reservoir utilization rate (%). RT, a variable, is the total discharge (m^3/s), IT is the total inflow (m^3/s), $Imax$ is the maximum inflow (m^3/s), ST is the total storage (m^3/s), max is the maximum

Discharge (m^3/s), $Lmax$ is the highest water level (EL. m), LFC is the RWL (EL. m), and LF is the FWL (EL. m).

A problem with these indicators for quantifying the flood control effect is that they do not consider the uncertainties of variables such as the inflow, discharge, storage, and low water level used in the calculation. Another problem is that the flood control effect at a downstream point cannot be evaluated (Kim et al., 2011). In addition, comparing flood seasons having different periods is difficult because the discharge amount at the peak flood volume is discharged at a certain amount owing to the characteristics of the simulation method of this study.

Evaluation Method with Nondamage and Dam Design Release

The evaluation index established in this study was based on evaluating the operating form of the system by reliability analysis. Hashimoto et al. (1982) classified the operation of a water resource system into states of satisfaction and dissatisfaction, defined it as failure in the case of dissatisfaction, and further classified it into three perspectives. The reliabilities of determining the frequency of system failure, swiftness of system recovery when a failure occurs, and degree of recovery were obtained by dividing it into vulnerability. Based on this, Kim et al. (2019) evaluated the water supply in the dry season by intuitively expressing the failure criteria using three indicators: frequency, length, and magnitude.

The above theory previously applied to droughts was applied to floods in this study based on the nondamage and design discharges. The evaluation index was selected as presented in Table 4.7, expressing the failure criteria as three concepts: frequency, duration, and magnitude.

The frequency is expressed as the average number of excesses and the number of times the design discharge exceeds during the entire period as a percentage. The duration is the average overdue period, i.e., it is the duration when the criterion is exceeded once on average, expressed in hours. Magnitude is the discharge amount compared to the average nondamage amount when the former exceeds the latter, and it is expressed as a percentage.

Table 4.7 Formulas of frequency, duration, and magnitude indicators used for downstream risk assessment

Number	Performance Indices	Definition	Classification
1	Average Exceed Frequency	$\mathrm{F} = \frac{\sum_{t=1}^{T} X_t \in F}{T} * 100(\%)$	Frequency
2	Average Exceed Duration	$\mathrm{D} = \frac{\sum_{t=1}^{T} X_t \in F}{\sum_{t=1}^{T-1} X_t \in F \; and \; X_{t+1} \in S} (\mathrm{hr})$	Duration
3	Average Exceed Magnitude	$\mathrm{M} = \frac{\sum_{t=1}^{T} \max(0, D_t - N)}{N * \sum_{t=1}^{T} X_t \in F'} * 100(\%)$	Magnitude

Evaluation Method with River Design Flood and Dam Design Release

The second evaluation index adopted in this study was the flood reduction effect evaluation index applied in practice by K-water. It is evaluated by dividing it into three sections—safe, flood, and disaster—based on the planned river flood volume and the 200-year frequency of dam discharge.

(1) Safe: Dam release $<$ Design flood of river

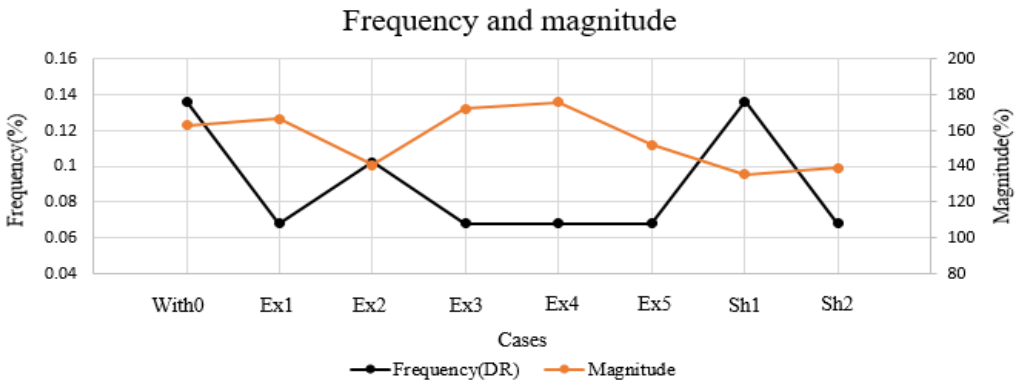
(2) Flood: Design flood of river $\leq$ Dam release $<$ Design release of Dam (200-year frequency flood)

(3) Disaster: Dam release $\geq$ Design release of Dam (200-year frequency flood)

4.5 Evaluation Results

4.5.1 Results with Nondamage and Dam Design Release

First, from the results of the method with nondamage and dam design release, all indexes were compared. The comparison of the frequency and the magnitude is shown in Figure 4.12(a). These appear to present a trade-off relationship, i.e., as the frequency, which represents probability of the design release, decreases more, preparations can be made for larger damage. In contrast, the size of small damage increases as much as the conservatively established nondamage discharge amount. In the case of frequency and duration, Figure 4.12(b) shows that they are in a proportional relationship. This can be interpreted as follows: as the frequency increases, the duration also increases owing to the nature of the intensive rainfall in Korea.



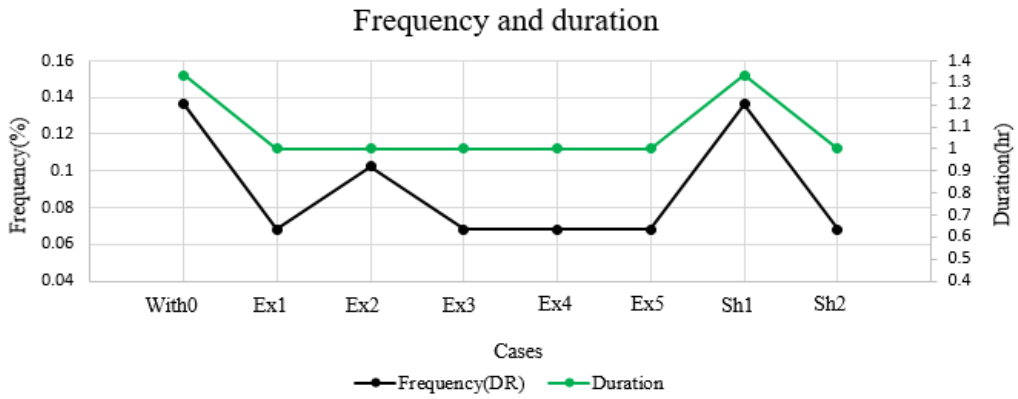


Figure 4.13 (a) Evaluation method with nondamage and dam design release result analysis(frequency vs magnitude) (b) Evaluation method with nondamage and dam design release result analysis(frequency vs duration)

The second type of comparison was of the results for each flood season obtained using the evaluation method with nondamage and dam design release. For this, the effects were divided into changes in the beginning and ending points, respectively. First, the change in the beginning point was confirmed by comparing SH2, EX4, EX1, and EX5. The ending points were fixed as September 30th, and the beginning points were July 1st, June 21st, June 16th, and June 1st, respectively. The frequency of each of the four flood season was 0.0682. The magnitude correspondingly decreased by 14.63%, increased by 7.82%, increased by 2.17%, and decreased by 6.63%, respectively, compared to the With0 case. Because there was no difference in the frequencies, it can be inferred that a change in the beginning point does not affect the preparation for a large flood. In addition, SH2 decreased the magnitude the most, and because the next reduced flood season was EX5, the effect of the increase or decrease in the previous period could not be confirmed.

The change in the ending point was evaluated by dividing it into two groups. The first group contained With0, EX2, EX3 and EX4. In this group, the starting point is fixed at June 21 and the ending points are September 20, September 21, September 25, and September 30, respectively. This is illustrated in Figure 4.13. In this group, the frequencies are 0.1364, 0.1024, 0.0682, and 0.0682, respectively, indicating a difference caused by a large flood. Compared to the existing flood season, when the ending point is delayed by 1 day, a decrease of approximately

0.034% is confirmed. This is the same as the effect of reducing the number of design discharges once during one year. When the delays are of 5 and 10 days, a decrease of approximately 0.068% can be seen, which can be interpreted as reducing the number of design discharges twice. The delays of 5 and 10 days show the same effects; the frequency from 5 days can be reduced to the maximum. The magnitude tends to increase by up to 8% as the length increases; however, the magnitude of EX3 is the largest; therefore, it tends to decrease again after a certain period.

The second group of end points included SH1 and EX1. The starting point was fixed as June 16, and the ending points were September 15 and September 30, respectively, with a total difference of 15 days. The frequency was consequently reduced by 0.068%; therefore, preparations can be for a large flood damage owing to the increase in the ending point, which is the same as the result for the first group. In addition, for the second, the magnitude also increases by 5.4% as the length of the flood season increases.

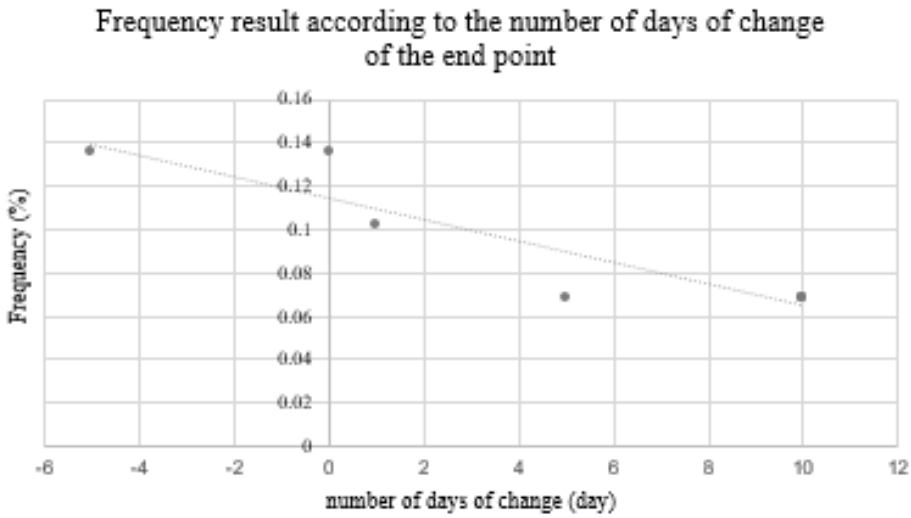


Figure 4.14 Frequency result according to number of days of change in ending point

Overall, compared to With0, the current flood season, Figure 4.14 shows the method with nondamage and dam design release result analysis. In EX1, which applies the change in basic statistics based on the current flood season methodology, the frequency decreased by 49.98%, magnitude increased by 2.17%, and duration decreased by 25%. For flood seasons EX2, EX3, and EX4 established to increase the ending point of the flood season, compared to With0, the frequencies decreased by 24.95%, 49.98%, and 49.98%, respectively. Moreover, the magnitudes decreased by 13.53%, increased by 5.63%, and increased by 7.82%, respectively, and all durations decreased by 25%. For EX5, which considerably increased the starting and ending points, the magnitude decreased by 6.63%; however, its frequency and duration were the same as those of EX3 and EX4. Finally, the results of SH1 and SH2 shifts considering only the effects of the rainy season and typhoons, respectively, were as follows. SH1 showed no change other than a decrease of 16.76% in the magnitude, and for SH2, the decreases in the indexes were 49.98%, 14.63%, and 25%, respectively, which were the largest.

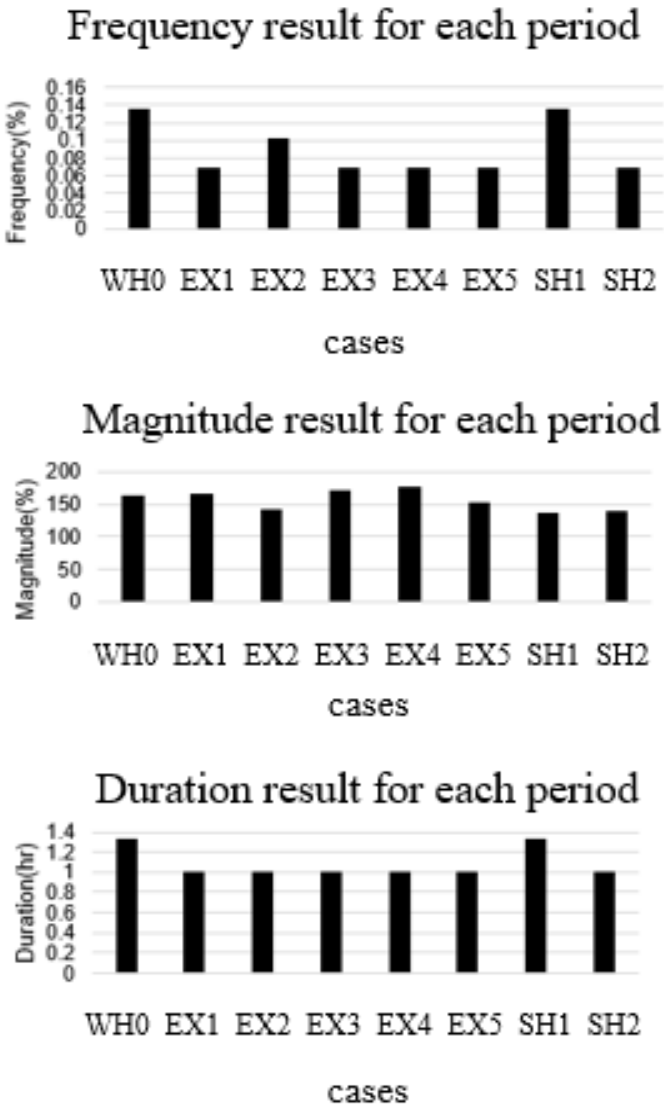


Figure 4.15 Evaluation method with nondamage and dam design release result analysis (frequency, magnitude, duration)

4.5.2 Results with River Design Flood and Dam Design Release

The case evaluated using the K-water method was as follows. The number of floods for both With0 and Sh1 was 4, for Ex2 was 3, and for Ex1, Ex3, Ex4, Ex5, and Sh2 was 2 each. In addition, the case of a disaster exceeding the 200-year frequency of planned flooding did not appear in all flood seasons. It can be confirmed that the results of the K-water method shows similar trends to the frequency results of the method with nondamage and dam design release.

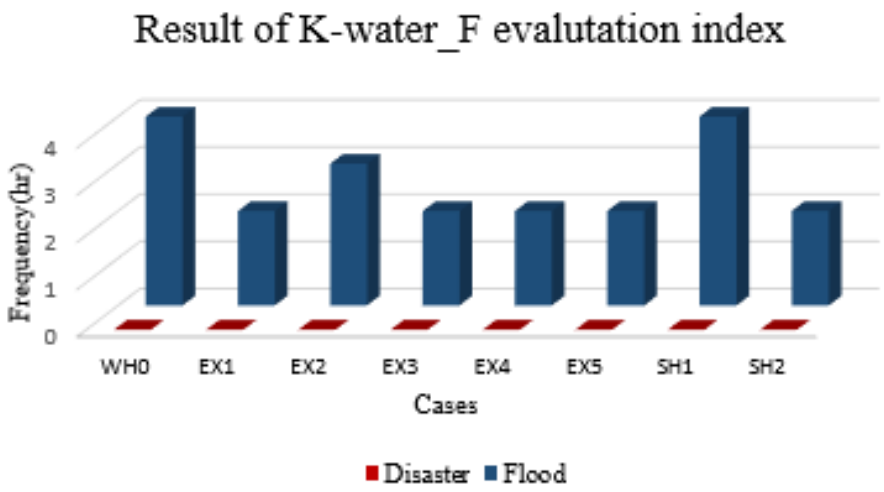


Figure 4.16 Evaluation method with river design flood and dam design release result analysis(frequency)

Finally, the results of the two methods were comprehensively compared by scaling the results of all indicators between 0 and 1. The results are shown in Figure 4.15. For each indicator, the top three flood seasons, including the joint, are selected as follows. For R^{Mag}, these are SH1, SH2, EX2, and With0. For R^{Freq_design}, these are EX1, EX3–EX5, and SH2. For R^{Dur}, these are EX1–EX5, and SH2. Finally, $R^{Freq\_K-water}$ identifies EX1, EX3–EX5, and SH2. Thus, it was concluded that SH2, which corresponds to all the three indices for the Yongdam dam, the study area, is the optimal flood season.

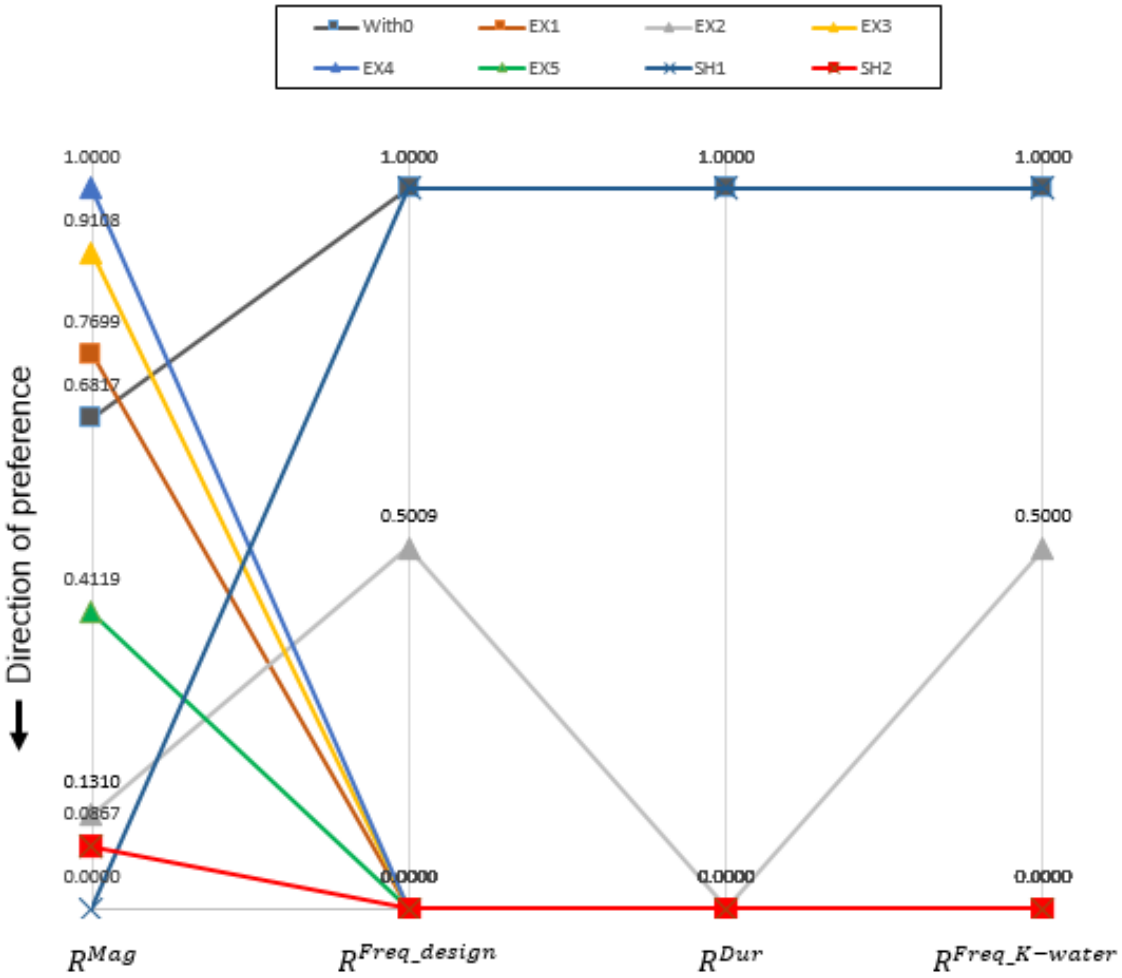


Figure 4.17 Comparison by case using four indicators of two evaluation methods

CHAPTER 5. CONCLUSION & FUTURE RESEARCH

5.1 Summary & Conclusion

The purpose of this study is to reduce flood damage by proposing a new flood season in the target watershed. This study evaluated the validity of the current flood season by dividing it into the 20th century and the 21st century through statistical techniques. In conclusion, it was found that this period has limitations in that it does not take into account the increasing trend of precipitation and the strengthening of precipitation characteristics by region. Therefore, it was judged that improvement of the period was necessary. Seven new flood seasons were proposed in consideration of expert opinions and statistical changes in the methodology for establishing the current flood season. To evaluate this, a simulation method using rigid ROM was used, and at this time, an LSTM model of deep learning was selected to derive the predicted inflow required as input data. The discharge amount derived through simulation was analyzed using the evaluation method based on nondamage release and dam design release, and the K-water method based on river design flood and dam design release. As a result, SH2, which shows a significant flood reduction effect, was proposed as a new flood season for the target watershed.

first purpose of this study was to evaluate the feasibility of the current flood season before proposing a new flood season. Precipitation patterns were analyzed by

comparing the whole country, and then subdivided into 18 multi-purpose dam basins in Korea to analyze precipitation patterns by region. Basic statistics and hypothesis testing were used for analysis. As a result, compared to the whole country, the annual cumulative precipitation of the flood season increased by 86.8 mm from 650.7 mm to 737.5 mm, confirming the increasing trend of precipitation in the 21st century. Also, as a result of the hypothesis test, the minimum value was 0.038 and the maximum value was 0.3, confirming the trend of strengthening the precipitation characteristics by region through the large difference in the range. In conclusion, the current flood season does not reflect the increasing trend of precipitation in the 21st century, considering that it has been more than 48 years since its establishment. In addition, it was interpreted that applying the same flood season to the whole country also has a limitation in that it does not take into account the strengthening of the precipitation pattern for each region.

Also, for the evaluation of the proposed flood season, the amount of discharge was determined through a simulation method. In the case of the simulation method, it was established by applying the Rigid ROM with excellent flood control and practical applicability to the flood season operation rules. Since the predicted inflow is required as the input data of Rigid ROM, the LSTM model of deep learning was used in this study. The model with the highest accuracy was selected through the calibration of the LSTM model. In the calibration process, input data changes, preprocessing, and hyperparameter changes were used. The model was finally selected through the results of two accuracy evaluation indexes, RRMSE and RBias.

Finally, the flood reduction effect of each flood season was evaluated using the discharge amount derived through the simulation method. In this study, two evaluation methods were applied. In both methods, the concepts of frequency, duration, and magnitude are applied, and evaluation is performed with different standards. The criteria for the first evaluation method are nondamage release and dam design release. The second evaluation method proceeds with evaluation based on river design flood and dam design release. The evaluation was carried out by dividing the effect on the change of the beginning point and the effect on the change of the ending point.

In conclusion, the effect of the change of the beginning point was insignificant. In this case, the periods of SH2, EX4, EX1, and EX5 were compared, and the frequencies were all the same at 0.068%. Also, the magnitude index did not show a significant difference. This is the same result as the advisory opinion, and it was interpreted that it is because the flood control capacity of the dam is sufficient at the beginning of the flood season. In the case of the change of the ending point, a significant change appeared. The change of the ending point was divided into two groups and compared. The first is a comparison of With0, EX2, EX3, and EX4. The longer the ending point of the flood period, the lower the frequency. In particular, when comparing With0 and EX4, which have the greatest difference, it was confirmed that the decrease was 0.068%. This is evaluated as the effect of blocking a large discharge the size of a dam design release twice. However, in the case of magnitude, it can be seen that EX4 increases by 5.59% compared to With0, indicating

that the size of the small discharge increases while the large discharge decreases. Through this, it was additionally confirmed that frequency and magnitude have a trade-off relationship. In addition, since there is no difference between EX3 and EX4, it is judged that an increase of more than a certain number of days has no effect on the preparation of large discharges. The second case was SH1 and EX1, which showed the same trend as the previous result. Finally, the most appropriate new flood season for Yongdam Dam was proposed, and SH2, which is the top 3 in each indicator among the seven flood seasons, was proposed.

This study is meaningful in considering the 21st century data on the validity of the current flood season, where there are no related studies. It is also significant in that it confirmed the effect of flood reduction according to the change of period rather than water level and capacity of the flood countermeasure study. In addition, the methodology of this study is expected to contribute to improving the dam's ability to respond to future climate change and resolving the damage caused by frequent disasters.

5.2 Potential Future Research

This study was conducted in the following order: proposing a new flood season, forecasting the inflow, deriving the discharge using a simulation method, and evaluating the derived discharge using an established evaluation procedure in the form of a guideline to develop a new flood season. However, future research must improve the accuracy of each process. First, in the new flood season proposal part, the optimal flood season can be found with the proposal of more periods than seven. In addition, in the inflow prediction, the time series analysis models of various machine learning techniques, such as the LSTM and GRU models, could be applied and the optimal model selected by accuracy comparison. In this scenario, it will be possible to predict the inflow more accurately and conduct research on a larger lead time. In addition, regarding the simulation method, if the previous inflow prediction becomes more accurate, the utility of a simulation method using the technical ROM will increase. Finally, although the evaluation method of this study focused on the downstream damage, if flood water after the flood season could be additionally examined, it will be possible to suggest the optimal flood season in terms of flood water as well as flood water.

Bibliography

Abbot, John, and Jennifer Marohasy. (2014). "Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks." *Atmospheric Research 138 : 166-178.*

Assem, Haytham, et al. (2017). "Urban water flow and water level prediction based on deep learning." Joint European conference on machine learning and knowledge discovery in databases. Springer, Cham.

Bicknell, Brian R., et al. (1996). "Hydrological simulation program-FORTRAN. user's manual for release 11." *US EPA.*

Campolo, Marina, Paolo Andreussi, and Alfredo Soldati. (1999). "River flood forecasting with a neural network model." *Water resources research 35.4 : 1191-1197.*

Cha Eun-jung. (2006). “The Meaning of the Rainy Season, Water and the Future,” The Korean Society for Water Resources G, 39(7), 69-72.

Chen, Pin-An, Li-Chiu Chang, and Fi-John Chang. (2013). "Reinforced recurrent neural networks for multi-step-ahead flood forecasts." Journal of Hydrology 497 : 71-

79.

Choi, Young-Eun. (2004). "Trends on temperature and precipitation extreme events in Korea." *Journal of the Korean Geographical Society 39.5 : 711-721.*

Dawson, Christian W., and Robert Wilby. (1998) "An artificial neural network approach to rainfall-runoff modelling." *Hydrological Sciences Journal 43.1 : 47-66.*

Devia, Gayathri K., B. Pa Ganasri, and G. Sa Dwarakish. (2015). "A review on hydrological models." *Aquatic procedia 4 : 1001-1007.*

Elman, Jeffrey L. (1990). "Finding structure in time." *Cognitive science 14.2: 179-211.*

Eun-Joo Lee. (2017). "Basic and applied studies of CNN and RNN." Broadcasting and Media 22.1 : 87-95.

Fox, Neil I., and Christopher K. Wikle. (2005). "A Bayesian quantitative precipitation nowcast scheme." *Weather and forecasting 20.3 : 264-275.*

Gizaw, Mesgana Seyoum, and Thian Yew Gan. (2016). "Regional flood frequency analysis using support vector regression under historical and future climate." *Journal*

of Hydrology 538 : 387-398.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. *MIT press.*

Ha, Kyung-Ja, Kyung-Sook Yun, and Sun-Seon Lee. (2007). "August mode in precipitation and its association with circulations." Proceedings of the Korea Meteorological Society Conference : 478-479.

Hashimoto, Tsuyoshi, Jery R. Stedinger, and Daniel P. Loucks. (1982). "Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation." *Water resources research 18.1 : 14-20.*

Heo, Jae-Yeong, and Deg-Hyo Bae. (2021) "Assessment of artificial neural network model for real-time dam inflow prediction." *Journal of Korea Water Resources Association 54.spc1 : 1131-1141.*

Hochreiter, Sepp, and Jürgen Schmidhuber. (1997). "Long short-term memory." *Neural computation 9.8 : 1735-1780.*

Hwang, Byung-Gi. (2021) "Evaluation of LSTM Model for Inflow Prediction of Lake Sapgye." *Journal of the Korea Academia-Industrial cooperation Society 22.4 : 287-294.*

Hu, Caihong, et al. (2018) "Deep learning with a long short-term memory networks approach for rainfall-runoff simulation." *Water 10.11 : 1543.*

Hsu, Kuo-lin, Hoshin Vijai Gupta, and Soroosh Sorooshian. (1995). "Artificial neural network modeling of the rainfall-runoff process." Water resources research 31.10 : 2517-2530.

Imrie, C. E., S. Durucan, and A. Korre. (2000). "River flow prediction using artificial neural networks: generalisation beyond the calibration range." *Journal of hydrology 233.1- 4 : 138-153.*

Jang, Ik Geun, et al. (2014). "Evaluation of flood control capacity of agricultural reservoirs during flood season." Journal of the Korean Society of Agricultural Engineers 56.4 : 69-75.

Jordan, Michael I. (1990). "Attractor dynamics and parallelism in a connectionist sequential machine." *Artificial neural networks: concept learning. 112-127.*

Jung, Il-Won, et al. (2018). "Vulnerability Assessments of Climate Change on Flood Damage." *Proceedings of the Korea Water Resources Association Conference. Korea Water Resources Association.*

Kang, Na Rae, et al. (2013). "Runoff simulation of an urban drainage system using radar rainfall data." *Journal of Wetlands Research 15.3 : 413-422.*

Kim, Bo-Kyung, et al. (2007) "Methodology for estimating ranges of SWAT model parameters: Application to Imha Lake inflow and suspended sediments." *KSCE Journal of Civil and Environmental Engineering Research 27.6B : 661-668.*

Kim, Gi Joo, Seung Beom Seo, and Young-Oh Kim. (2019). "Elicitation of drought alternatives based on water policy council and the role of shared vision model." *Journal of Korea Water Resources Association 52.6 : 429-440.*

Kim et al. (1997). “Analysis of time series characteristics of annual rainfall data in Korea.” Proceedings of the Korea Water Resources Association Conference

Kim, Juuk, Changwon Choi, and Jaeeung Yi. (2011). "Development of flood control effect index by using fuzzy set theory." KSCE Journal of Civil and Environmental Engineering Research 31.5B : 415-429.

Kingma, Diederik P., and Jimmy Ba. (2014). "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980.*

Kratzert, Frederik, et al. (2018). "Rainfall–runoff modelling using long short-term memory (LSTM) networks." *Hydrology and Earth System Sciences 22.11 : 6005-6022.*

Korea Water Resources Corporation (2005), Multipurpose Dam Management Regulations

Korea Water Resources Corporation (2021), Multipurpose Dam Operation Practical Handbook

Ko, J. W., H. J. Baek, and W. T. Kwon. (2005). "The characteristics of precipitation and regionalization during rainy season in Korea." J. Korean Meteorol. Soc 41 : 101-114.

Kwak, Jaewon. (2021). "A study for the target water level of the dam for flood control." Journal of Korea Water Resources Association 54.7 : 545-552.

Lapedes and Robert Farber. (1987). “Nonlinear signal processing using neural networks: Prediction and system modelling.” *No. LA-UR-87-2662; CONF-8706130-4.*

Lee, Seungsoo, et al. (2020). "Flood prediction in the Namgang Dam basin using a

long short-term memory (LSTM) algorithm." Korean Journal of Agricultural Science 47.3 : 471-483.

Meteorological Agency. (2011). “Climate Change White Paper”

Meteorological Agency. (1995). “Rainy Season White Paper”

Meteorological Agency. (2011). “Typhoon White Paper”

Merz, Bruno, et al. (2010). "Fluvial flood risk management in a changing world." *Natural Hazards and Earth System Sciences 10.3 : 509-527.*

Ministry of Public Administration and Security, (2020). “Disaster safety situation analysis results and disaster safety accidents subject to priority management”

Mok, Ji-Yoon, Ji-Hyeok Choi, and Young-Il Moon. (2020). "Prediction of multipurpose dam inflow using deep learning." *Journal of Korea Water Resources Association 53.2 : 97-105.*

Mosavi, Amir, Pinar Ozturk, and Kwok-wing Chau. (2018). "Flood prediction using machine learning models: Literature review." *Water 10.11 : 1536.*

Neitsch, Susan L., et al. (2011). “Soil and water assessment tool theoretical documentation version 2009.” *Texas Water Resources Institute.*

Noh, Huiseong, et al. (2016). "Long-term simulation of daily streamflow using radar rainfall and the SWAT model: A case study of the Gamcheon basin of the Nakdong River, Korea." *Advances in Meteorology.*

Shoaib, Muhammad, et al. (2016). "A comparison between wavelet based static and dynamic neural network approaches for runoff prediction." Journal of hydrology 535 : 211-225.

Sim, Myeong-Pil, O-Ik Gwon, and Hwan-Gi Lee. (1995). "Reservoir operation by variable restricted water level during flood period." Water for future 28.6 : 217-228.

Tian, Ye, et al. (2018). "Integration of a parsimonious hydrological model with recurrent neural networks for improved streamflow forecasting." *Water 10.11 : 1655.*

Tran, Quang-Khai, and Sa-kwang Song. (2017). "Water level forecasting based on deep learning: A use case of Trinity River-Texas-The United States." *Journal of KIISE 44.6 : 607-612.*

Weigend, Andreas, David Rumelhart, and Bernardo Huberman. (1990). "Generalization by weight-elimination with application to forecasting." *Advances in*

neural information processing systems 3.

Yoon, Kang-Hoon, Bong-Cheol Seo, and Hyun-Suk Shin. (2004). "Dam inflow forecasting for short term flood based on neural networks in nakdong river basin." *Journal of Korea Water Resources Association 37.1 : 67-75.*