공학석사 학위논문

Analysis of Ensemble Streamflow Prediction Effect on Deriving Dam Releases for Water Supply

용수공급을 위한 댐 방류량 결정에서의 앙상블 유량 예측 효과 분석

2023년 8월

서울대학교 대학원
건설환경공학부 건설환경공학전공
김 연 주

용수공급을 위한 댐 방류량 결정에서의 앙상블 유량 예측 효과 분석

Analysis of Ensemble Streamflow Prediction Effect on Deriving Dam Releases for Water Supply

지도교수 김 영 오

이 논문을 공학석사 학위논문으로 제출함

2023 년 5 월

서울대학교 대학원

건설환경공학부 건설환경공학전공

김 연 주

김연주의 공학석사 학위논문을 인준함

2023 년 7 월

위 원 장 ______________________ (인)

부 위 원 장 ______________________ (인)

위 원 ______________________ (인)

Abstract

Analysis of Ensemble Streamflow Prediction Effect on Deriving Dam Releases for Water Supply

Yeonju Kim
Department of Civil and Environmental Engineering
Civil and Environmental Engineering Major
The Graduate School of Seoul National University

Ensemble streamflow prediction (ESP) considers the uncertainty of streamflow in water resources management, primarily in western regions of the United States. Consequently, the ESP system is actively used in hydrological forecasting and water resources management in the United States and Europe. However, in South Korea, ESP has only been used for hydrological forecasting, with its application in water resources management being limited. Despite the availability of ensemble forecasts, current water resources management practices in South Korea still rely on single-valued statistics such as the ensemble mean for decision-making.

This study aimed to promote the use of ESP in water resources management in South Korea and thus demonstrate its effectiveness. A simple statistical exercise was

created to convince dam operators. A simple hypothetical illustrated that in case of dams with the same capacity and demand but different inflow standard deviations, the dam with a higher inflow standard deviation incurred higher costs.

Furthermore, further exercises were applied to actual dams in South Korea. Multiple-purpose dams in the Han River basin with the same length of data were selected according to the capacity-inflow ratio (CIR) as case study sites: Soyanggang Dam (Dam SY) with a CIR of 1.345 and Chungju Dam (Dam CJ) with a CIR of 0.563. The inflow data for each dam were divided into nine sets, and the last set from 2020 to 2022 was used to generate unbiased ensembles based on the standard deviation. Consequently, two ensembles were created: A well-forecasted scenario (Scenario W) and a poorly-forecasted scenario (Scenario P). Sampling stochastic dynamic programming (SSDP), which enables optimal release calculation using ESP, was employed to develop SSDP/Hist and SSDP/ESP models. A primary function of multiple-purpose dams in South Korea is water supply, which was optimized by setting the objective function to avoid water shortages. Considering the poor accuracy of long-term forecasts in South Korea, SSDP/ESP models were constructed by incorporating the future value function from the SSDP/Hist model and then optimizing in the forward direction. The SSDP/Hist and SSDP/ESP models were built for Dam SY and Dam CJ, and the optimal releases were calculated. Thereafter, the simulated operation using the obtained optimal releases was evaluated in terms of total penalty, frequency, duration,

and magnitude.

The simulation results confirmed that Scenario W exhibited better overall performance compared to that of Scenario P in Dam CJ. This indicates that even with the same mean, different inflow standard deviations result in different optimal releases and require different operational strategies. Indeed, in the simulation study for Soyanggang Dam, there was no significant difference observed between Scenario W and Scenario P. This finding indicates that dams with lower CIR values and higher water demands are more sensitive to uncertainty in inflow predictions. Thus, the management of water resources based solely on mean values is a naive operation method that neglects considerations for future climate change and other uncertainties. Therefore, this study can serve as a motivation for improving water resources management techniques in South Korea.

Keywords: Ensemble streamflow prediction, Sampling Stochastic Dynamic Programming, Optimization

Student Number: 2021-24690

Contents

List of Figures

List of Tables

Notation

A	Total number of failure
a	number of period from success to failure
D_t	Demand at time t
E	Expectation matrix
f_t	Optimal value function at time t
f_{opt}^i	Optimal value function at iteration i
O_t	Objective function at time t
$Pr(\cdot)$	Probability of an event
Q_t	Inflow vector during time t
$Q_t(i)$	Inflow scenario i at time t
R_t	Release vector during time t
R_t^*	Target release vector during time t
S_t^k	k^{th} discretized storage vector at the beginning of the time t
S_{max}	Maximum storage vector
S_{min}	Minimum storage vector
t	Time
w_j	Weighting factor of scenario j
W_t	Binary value that indicates whether the system performance is satisfactory or not
ϵ	Tolerance value
η	Probability for a event that transfers from success to failure

γ	Average of duration
μ_X	Population mean of random variable X
ρ	Frequency
σ_X	Population standard deviation of random variable X
υ	Complement of average of magnitude

Chapter 1. Introduction

1.1 Research Background and Necessity of Study

The importance of hydrological forecasting for natural disaster mitigation and water resource management was recognized by researchers in 1975 by the World Meteorological Organization (WMO) (WMO, 1975). Since the mid-1970s, WMO has conducted several projects to encourage the development of streamflow forecasting systems and to provide information on the choice of methods and approaches for them (WMO, 1986, 1990). Accordingly, a new approach to streamflow prediction referred to as extended streamflow prediction was announced to supply water considering the uncertainty of streamflow in the western United States (Day, 1985). Thereafter, the National Weather Service (NWS) operated the ensemble streamflow prediction (ESP) program to utilize this method for hydrological forecasting and water resource management, and the concept of ESP was established (Riverside Technology, 1997; Connelly et al., 1999; US Department of Commerce and NOAA, 2020).

ESP is a technique that first inputs all possible precipitation traces that may occur in the future into a rainfall-runoff model. Consequently, it generates several streamflow traces and performs statistical analysis to predict with probability. The primary advantage of ESP is its ability to quantify forecast uncertainty by generating a range of possible streamflow traces (ensembles). Moreover, as the input ensemble can be selected flexibly, it can be applied to both long- and short-term predictions. In addition, the initial condition of the watershed can be reflected through the physical function of the deterministic rainfall-runoff model, which is in contrast to statistical models that rely only on statistical techniques.

In Korea, the necessity of introducing ESP was raised in earnest only in the 21st century. Kim et al. (2001) proved its excellence by applying ESP to the Gongju branch to improve the "Water Supply Outlook" (current Water Resources Status and Outlook) published at the beginning of every month by the Ministry of Construction & Transportation and the Korea Institute of Civil Engineering and Building Technology. Thereafter, K-water established ESP in practice for the Han, Nakdong, and Seomjin river basins, confirming the effectiveness of ESP (K-water, 2004). Currently, the National Drought Information Analysis Center (NDIAC) is conducting drought analysis of 35 dam basins nationwide using Bayesian ESP to advance drought monitoring and forecasting technology. For the practical application of the model, GUI-based user convenience environment improvement work is in progress. (National Drought Information Analysis Center, 2018, 2021). In addition, the Han River Flood Control Center reviewed ESP for practical use after Kim et al. (2001) and converted the "Water Supply Outlook" (current Water Resources Status and Outlook) into a probabilistic forecast (Han River Flood Control Office, 2009, 2022). They considered uncertainty in their probability for flood forecasting, which became the basis for reliable forecasting (Kim et al., 2011).

However, beyond hydrological forecasting, the use of ESP in the real time management of water resources is challenging. This is because even in developed countries, in the field, a single value is more convenient than an ensemble of multiple traces. To solve this problem, studies have attempted to prove the positive effect of using ESP in water resource management for the past 25 years (Eum et al., 2006; Faber and Stedinger, 2001; He et al., 2022; Ramaswamy and Saleh, 2020). Nevertheless, cases involving successful use of ESP in the practical field of water resource management (including in Korea) are rare. Even after the creation of several traces for predicting hydrological forecasting, the calculations for the amount of release from a dam are reliant on a single representative value such as the median or average value.

This study was conducted to demonstrate the significant difference in the effectiveness of ESP in water resource management, particularly when determining dam releases. Based on this, we intend to support dam operators in the field to actively utilize the streamflow prediction ensemble which is the result of ESP.

1.2 Research Objectives

The ultimate purpose of this study is to recommend the use of ESP considering the inflow uncertainty in dam operations in Korea. The detailed goals for this are as follows.

(a) Investigate research trends using ESP and demonstrate the use of the method in water resources management. Further, exhibiting its excellent performance in application cases.

(b) Induce policy-maker to easily understand the benefit of employing ESP in dam operation by quantifying the effect of ESP through simple examples.

(c) Establish an optimal ESP-based dam operation model for two dams with different capacity-inflow ratios (CIR) to demonstrate the feasibility of applying ESP in real cases.

1.3 Organization of Research

Chapter 2 of this paper, titled "Theoretical Background," investigates the research trends in probability optimization, ESP, and ESP in reservoir operations. A simple reservoir example is applied to demonstrate the importance of probabilistic forecasting (ESP) in dam releases, thereby highlighting the risks associated with operation based on a single representative value. This section aims to facilitate a better understanding of the significance of probability prediction in dam operation. Chapter 3 presents the methodology, which outlines the overall research procedures, the calculation formula and considerations of the sampling stochastic dynamic programming (SSDP) technique employed in this study, and the approach for updating optimal discharge using ESP. Chapter 4 presents sample studies, where multiple-purpose dams in South Korea are grouped based on CIR. Two dams are selected from this grouping, and the methodology described in Chapter 3 is applied to analyze the results. Finally, Chapter 5 summarizes the findings and implications of the study.

Chapter 2. Theoretical Background

Water resource system problems are complexly connected with hydrological, social infrastructure, ecological, economic, and anthropogenic factors related to water (Loucks and Beek, 2017). Therefore, there exist several difficulties in defining the water resource system problem. For example, "how to set the scope of the problem?," "which problem should be solved with the highest priority among many complexly connected factors?," and "how to solve the problem?" are representative questions that must first be defined in water resource system problem-solving. In this study, the above questions were answered as follows and the water resource system problems we aimed to address were defined.

The first question is "how to set the scope of the problem?" This study aimed to show the positive effect of ESP on dam operation and thus recommend its usage. Therefore, the current problem situation to be improved is set to dam operation considering only a single value of inflow. Subsequently, among several dams, the multi-purpose dams, which supply the maximum water in Korea, were selected as the scope of the study.

The second question is "which problem should be solved with the highest priority among the many complexly connected factors?" The term “multi-purpose dam” refers to a dam constructed by the Minister of Environment, and is used for two or more purposes among water for living, industrial, agricultural, environmental improvement, power generation, flood control, and transportation by ship. Among them, multi-purpose dams in Korea are particularly important for supplying water such as water for living, agriculture, and environmental improvement. Therefore, in this study, water supply was selected as the most important factor to be improved among various other dam operation objectives.

Finally, the last question is "how to solve the water resource system problem?" The water supply problem of dams primarily involves the use of modeling methods. A model is a simplified version of a real system built into a computer. Models are built to predict the outcome of decisions. In this chapter, existing studies and theoretical backgrounds are summarized to answer the question of "how to solve the water resource system problem?"

Models are used to simplify events that are very complex in the real world and are caused by many factors. However, the consideration of the right assumptions is crucial. The components of the model include constraints, parameters, decision variables, state variables, and an objective function in the case of an optimization model. The components of each model are summarized below.

(a) Constraints: the conditions the system has to satisfy

(b) Parameter: variables that are assigned known values

(c) Decision variables: variables having unknown values that are to be determined by solving the model. Decision variables can include design and operating policy variables of various water resources system components.

(d) State variables: variables that describe the state of the system

(e) Objective function (in case of optimization models): the function to be optimized during the problem-solving procedures

Models used in water resource system problems are largely divided into optimization, simulation, and simulation-based optimization. A simulation model is a model that presents results expected to occur when a specific action is undertaken. The solution of the optimization model is based on the objective function to be maximized or minimized, which is in contrast to the simulation model. The optimization model maximizes or minimizes the objective function to derive the alternative action to achieve the most optimal performance under given constraints. Finally, the simulation-based optimization model is a model that calculates the results of our actions with a simulation model and then optimizes it using the optimization model based on it. This model is the most used model in the field of water resource systems. For instance, using the operating rules we calculated through simulation, we can calculate the amount of future water shortage, perform optimization in preparation for this, and derive operating rules that improve the water supply to the best extent.

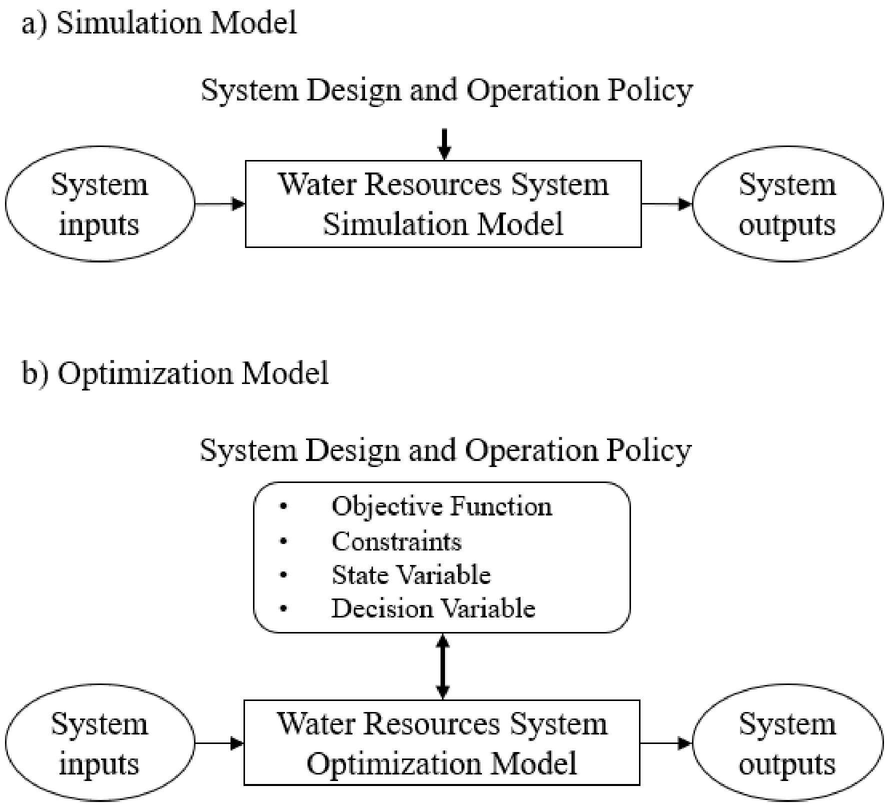


Figure 2.1 Simulation and optimization model (Loucks and Beek, 2017; Kim, 2022)

Figure 2.1 shows the difference between the simulation and optimization models. Optimization models require an explicit expression of the goal, whereas simulation models do not. Simulations simply assume certain scenarios or address the “what if” scenarios that might occur if certain decisions are made. in addition, users of simulation models must define the values of design and operational decision variables before performing simulations. Once the values of all decision variables are defined, the simulations can provide more accurate estimates of the possible effects of these decisions. In contrast, optimization models indicate the best decision; however, the solutions it leads to are often based on certain limiting assumptions. Because of this, optimization models should be used not to find an optimal solution, but as a method to define a relatively small number of good alternatives that can be tested, evaluated, and improved later through more detailed simulations. Therefore, the process of deriving multiple plans and policies using optimization and evaluating them via simulations to reduce them to a small number is recommended(Loucks and Beek, 2017).

In the field of water resources management, several notable optimization methods are commonly used, including linear programming (LP), dynamic programming (DP), nonlinear programming (NLP), and genetic algorithm (GA). In this study, we also utilized sampling stochastic dynamic programming (SSDP), which is a representative optimization method for water resources management, along with simulation-based optimization based on the mass conservation equation. The following section investigates the probabilistic optimization methods that account for the uncertainty in inflow, ESP methodology, and research trends and examples of using ESP in reservoir operation to understand the impact of ESP on optimal reservoir operation rules.

2.1 Probabilistic Optimization

The greatest uncertainty in reservoir operation lies in the inflow, which necessitates the development of approaches that incorporate this uncertainty. SSDP has been devised to address this challenge by incorporating the traces of ESP directly into the optimization formulation. In this section, we examine SSDP, which enables the direct application of ensemble prediction data in the optimization framework, utilizing ESP traces.

DP is a theory of multistage decision processes (Bellman, 1957). It provides a systematic approach to transforming multistage problems into single-stage problems. In contrast to other mathematical programming techniques, DP does not impose constraints on the use of objective functions. This renders it well-suited for handling nonlinear problems. It is particularly useful for addressing problems with stochastic characteristics through the framework of Markov decision processes. Further, DP offers a systematic process for dealing with problems that involve multiple decision points and is thus a widely used optimization technique in the field of system optimization.

(a) Stage refers to the points at which decisions are made. In the context of reservoir operation, it represents the operational time step or the interval at which the decisions are made.

(b) State variable represents the state of the system and aggregates past information at a given stage, serving as a variable. In reservoir operation, it typically represents the storage level and is commonly discretized for computational convenience.

(c) Decision variable refers to the alternative actions that can be taken at each stage, based on the knowledge of the state variable. In the context of reservoir operation, it represents the release or discharge rate.

(d) Stage return is a scalar value that represents the effectiveness of the decision made at each stage. It quantifies the operational benefits, such as hydropower generation, associated with the decision made in reservoir operation.

DP is a process that seeks to find the optimal operating rules at each stage for every possible state (Bellman, 1957). The underlying principle of DP, known as the principle of optimality, states that the optimal decision at a particular state is dependent solely on future decisions and is independent of past decisions. To address optimization problems based on this principle, the backward method is commonly employed. This study focused on the application of DP to reservoir operation planning and explored the use of the backward method. The process is illustrated in Figure 2.2, where the backward method is used to determine the optimal values at each stage, considering the current returns and the maximum sum of future optimal values. This iterative process continues for each state at every stage. Following the completion of the iterations from stage 3 to stage 1, the optimal values at stage 1 are updated based

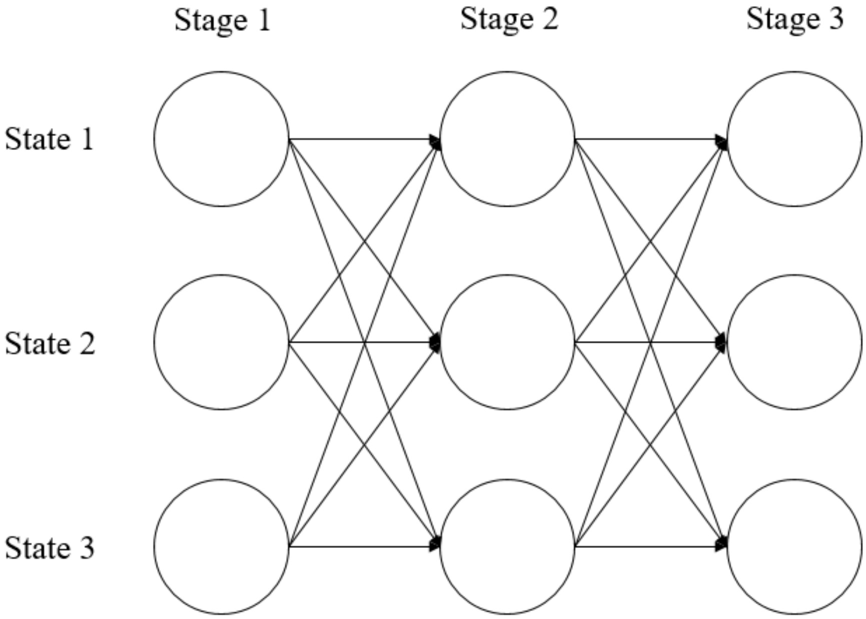


Figure 2.2 Dynamic programming schema

on those obtained at stage 3. This is followed by further iterations using the backward method until convergence is achieved.

Deterministic dynamic programming (DDP) refers to the situation wherein the state at the next stage is determined with certainty based on the current state and decision in dynamic programming. In the context of reservoir operation, DDP is applicable when the inflow to the reservoir, such as average inflow, is deterministic and known precisely without any uncertainty; that is, there is only one inflow trace. DDP commonly solves the following recursive equation iteratively to determine the release policy (Bellman, 1957).

$$f_t(\mathbf{S_t^k}) = \underset{R_t^*}{max} \left[O_t(\mathbf{S_t^k}, \mathbf{Q_t}, \mathbf{R_t}) + \left\{ f_{t+1}(\mathbf{S_{t+1}^k}) \right\} \right] \forall S_t, t \in \{1, 2, \cdots, T\} \quad (2.1)$$

$$S_{t+1} = S_t + Q_t - R_t \quad (2.2)$$

$$R_t = min\{max[R_t^*, S_t + Q_t - S_{max}], S_t + Q_t - S_{min}\} \quad (2.3)$$

where S_t^k represents the k-th discrete value of the reservoir storage state variable, which is discretized into K values at stage t. Further, $f_t(\mathbf{S_t^k})$ denotes the value of the objective function (optimal value function) that can vary based on the research objective among the alternative actions available for the k-th reservoir storage state variable, and $O_t(\cdot)$ represents the objective function determined by the reservoir storage (S_t^k), inflow (Q_t), and releases (R_t) at stage t. This computation process proceeds backward from the last stage to the initial stage (Eum, 2007).

DDP undergoes an exponential increase in computational complexity with progress in the discretization process. To alleviate the curse of the dimensionality problem and consider the stochastic nature of inflow, an implicit stochastic optimization (ISO) technique has been proposed. In ISO, a significant amount of historical data or synthetic inflow is used to sufficiently consider the uncertainty of flow. Subsequently, deterministic optimization techniques are employed to determine the optimal outflow. Thereafter, a post-processing step such as regression analysis is applied to relate the previously determined optimal outflow with variables such as initial reservoir storage, inflow from the previous month, or the relationship with end-of-month storage. This approach, inspired by Monte Carlo Simulation, utilizes mathematical techniques to predict possible outcomes of uncertain events through computer calculations. As DDP assumes knowledge of the entire inflow time series over the entire period, the determination of optimal outflow is dependent on having an adequate inflow time series and employing post-processing methods such as linear regression analysis. Moreover, certain studies have addressed the differences in post-processing methods, including proposing suitable approaches for post-processing analysis (Shaikh and Pattanayek, 2022).

ISO, which utilizes deterministic models, offers the advantage of convenience in the application and reduced computation time. However, it yields distorted optimal releases when applied in situations involving insufficient data, such as in the case of South Korea. This is because it relies on assumed time series data and obtains optimal solutions through post-processing techniques such as multiple regression analysis (Eum, 2007). To address this limitation, the explicit stochastic optimization (ESO) technique has been developed. ESO incorporates uncertainty into the equations of dynamic programming itself by representing it probabilistically. Consequently, it enables optimization under uncertainty and aims to overcome the shortcomings of ISO.

One of the optimization methods within ESO is SDP. SDP is constructed by incorporating probabilities that reflect the uncertainty of input variables into the process of DDP. In the context of reservoir operation, wherein the uncertainty lies in the inflow, SDP is computed by multiplying the probability of the inflow. Although the exact inflow to the reservoir cannot be precisely predicted, SDP considers the probability distribution of the inflow by fitting it to the discretized intervals of Q_t. This is realized by placing the inflow probability at the front of Eq. 2.1 in DDP. Consequently, it can be expressed as Eq. 2.4 by introducing the expectation operator (Tejada-Guibert et al., 1995).

$$f_t(\mathbf{S_t^k}) = \underset{R_t}{max} \underset{Q_t}{E} \left[O_t(\mathbf{S_t^k}, \mathbf{Q_t}, \mathbf{R_t}) + \left\{ f_{t+1}(\mathbf{S_{t+1}^k}) \right\} \right] \forall S_t, t \in \{1, 2, \cdots, T\} \tag{2.4}$$

Expectation (E) can be obtained by multiplying the expected value of the inflow for each probability and summing them. As mentioned earlier, in the case of Eq. 2.4, the probability variable and the probability distribution function $Pr(\mathbf{Q_t})$ are discretized to facilitate computation. Thus, the equation for SDP with the discretized probability distribution function is expressed as Eq. 2.5.

$$f_t(\mathbf{S_t^k}) = \underset{R_t}{max} \sum_{i=1}^{I} Pr(\mathbf{Q_t^i}) \left[O_t(\mathbf{S_t^k}, \mathbf{Q_t^i}, \mathbf{R_t}) + \left\{ f_{t+1}(\mathbf{S_{t+1}^k}) \right\} \right] \tag{2.5}$$

where S_t^k represents the k-th value of the discretized reservoir storage at stage t, and Q_t^i represents the i-th value of the discretized inflow at stage t. Figure 2.3 illustrates the case where Q_t is discretized into four intervals, assuming a standard normal distribution.

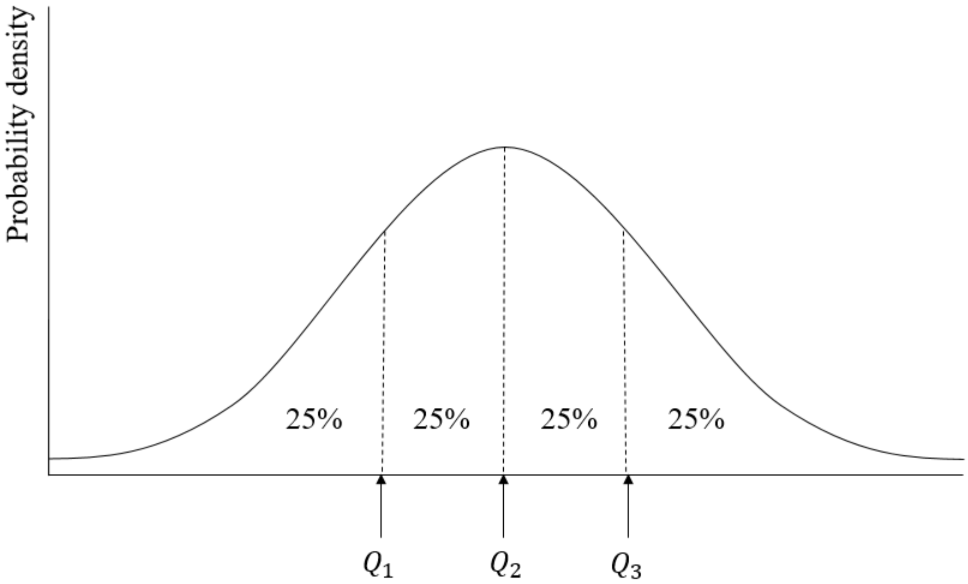


Figure 2.3 Example of discretization of probability distribution function

A case study of dam operation using DP-family optimization is presented in Celeste and Billib (2009). The study investigated the performance of probabilistic models based on ISO and ESO, which were used to define optimal reservoir operation policies. Six optimization techniques based on ISO with different calculation methods and post-processing approaches, as well as the SDP method, were compared in the Paraiba river basin in Brazil. In general, ISO and PSO models demonstrated better performance compared to SDP and SOP; however, this was attributed to the assumption of perfect forecasting, which is a major limitation of DDP.

To alleviate the difficulties caused by the complexity of real-world problems, Giuliani et al. (2016) explored the technical and practical significance of using evolutionary multiobjective direct policy search (EMODPS). They addressed three issues: dimensionality, modeling, and multiple objectives, and applied SDP and EMODPS to the Hoa Binh, a multipurpose dam in Vietnam. The results demonstrated that EMODPS outperformed SDP, indicating its greater success in handling the challenges associated with the aforementioned problems. Subsequently, in South Korea, Kim and Kim (2021) constructed models for the Boryeong Dam, which had experienced multi-year droughts, using EMODPS and dynamic programming. Among them, the EMODPS-Gaussian model demonstrated the most improved optimal release policy for the dam.

Table 2.1 Literature summary of probabilistic optimization

Author (year)	Title (Target Problem)	Optimization	Key findings
Stedinger et al. (1984)	Stochastic Dynamic Programming Models for Reservoir Operation Optimization	Stochastic DP (SDP)	Suggests the inclusion of best inflow forecast instead of proceeding period's inflow
Tejada-Guibert et al. (1995)	The value of hydrologic information in stochastic dynamic programming models of a multi-reservoir system	Stochastic DP (SDP)	Depending on the type of the objective function, the value of hydrologic state variables differs in Stochastic DP
Celeste and Billib (2009)	Evaluation of stochastic reservoir operation optimization models	QP, PSO, Stochastic DP (SDP)	All ISO and PSO models performed better than SDP and the SOP and also provided release rules similar to the ones determined by perfect forecast optimization.
Giuliani et al. (2016)	Curses, Tradeoffs, and Scalable Management: Advancing Evolutionary Multiobjective Direct Policy Search to Improve Water Reservoir Operations	Stochastic DP (SDP), EMODPS	RBF solutions are more effective than those obtained by ANN in designing Pareto-approximate policies, and EMODPS successfully improves the SDP solutions
Shaikh and Pattanayek (2022)	Implicit Stochastic Optimization for deriving operating rules for a multi-purpose multi-reservoir system	DP	Among (1) All possible regression, (2) stepwise regression, (3) decomposition and (4) simulation, decomposition model with less number of predictor variables is the most preferred.

2.2 Ensemble Streamflow Prediction

Hydrological predictions are employed to anticipate future events to facilitate more efficient water resources management. Over the past several decades, numerous studies have been conducted to enhance the accuracy of hydrological predictions; however, the complexity and interconnectedness of watershed-scale hydrological phenomena render the achievement of perfect forecasts challenging. Hydrological predictions can be categorized into deterministic and probabilistic forecasts. Deterministic forecasts provide a single prediction value, facilitating straightforward and prompt decision-making. Consequently, they are widely utilized as valuable information for water resources operations. However, deterministic forecasts cannot account for outcomes other than the predicted value, which renders the preparation for alternative situations challenging. For instance, when determining dam releases based on a single value using only the historical average inflow, it assumes a 100 % probability of that average inflow occurring. However, this approach fails to consider the possibility of different inflow values and does not incorporate the probabilities of lower or higher inflow traces. Conversely, probabilistic forecasts offer the advantage of presenting the likelihood of various outcomes, enabling water resources management that considers the associated risks. Among these probabilistic forecasting methods, ESP stands out as a prominent approach.

ESP is based on the assumption that past hydrological events can represent

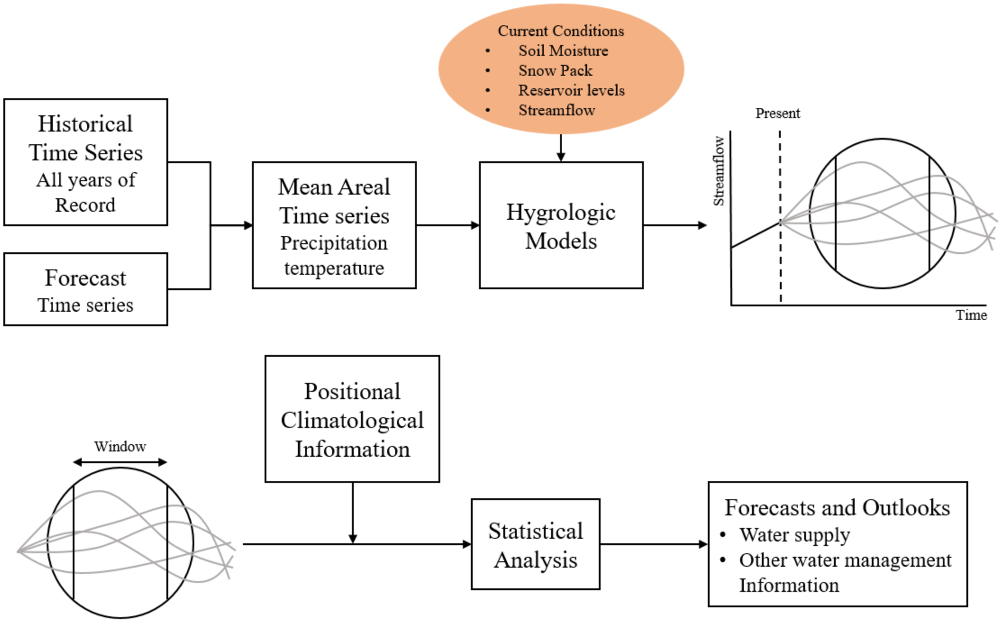


Figure 2.4 Schematic of ESP procedure (Day, 1985)

future hydrological events. Hence, it involves setting initial conditions such as humidity, temperature, and soil moisture profiles in a deterministic rainfall-runoff model, as shown in Figure 2.4, and inputting multiple rainfall traces that are likely to occur in the future to obtain a range of streamflow traces. This approach, often referred to as the conditional Monte Carlo simulation approach, generates streamflow traces following the occurrence of initial conditions. A key advantage of ESP is its ability to quantitatively assess prediction uncertainty via the generation of a range of possible streamflow traces (ensembles) and offering a flexible selection of input ensembles based on the objectives of the study, which are applicable to both long-term and short-term predictions. Although the results delivered by ESP are interpretable in a probabilistic manner, the calculation process does not solely rely on a 100% probabilistic approach. As mentioned earlier, ESP combines the advantages of deterministic models, which yield a single value using physical functions, and probabilistic models that facilitate the quantification of uncertainty. It employs a scenario-based deterministic model that considers uncertainty by integrating physical processes. The National Weather Service (NWS) in the United States is a prominent user of ESP. Since the 1990s, the NWS has been utilizing the ESP method to forecast streamflow nationwide.

In South Korea, the introduction of ESP began with the application of ESP at the Gongju gauge station, as documented in Kim et al. (2001). This research was initiated with the purpose of improving the "Water Supply Outlook," which was published by the Ministry of Land, Infrastructure and Transport and the Korea Institute of Construction Technology (KICT). Therefore, the same KICT rainfall-runoff model used for the "Water Supply Outlook" was employed for the ESP implementation. Historical rainfall data from 1919 to 1994, spanning 76 years, at the Gongju gauge station, were used to generate 76 streamflow traces each month. These traces were then utilized for streamflow predictions for a period of five years, from 1995 to 1999. In addition, to enable probabilistic forecasting, the streamflow was statistically categorized to provide categorical forecasts. The results confirmed that the ESP method outperformed the existing method in terms of streamflow prediction accuracy.

In K-water (2004), the streamflow synthesis and reservoir regulation (SSARR) model was used to develop ESP for the Han, Nakdong, and Seomjin river basins. To ensure accurate validation, the focus was on multi-purpose dams with sufficient length of observed data available for each basin. For the Nakdong river basin, the Andong, Hapcheon, and Namgang dams were selected as the validation areas. For the Han river basin, the Chungju and Soyanggang dams were chosen, and for the Seomjin river basin, Seomjingang dam was selected as the validation area. In the Han river basin, an ensemble of 17 streamflow traces was generated using rainfall and temperature ensembles from 1981 to 1997, and streamflow predictions were conducted for a period of six years from 1998 to 2003. For the Nakdong river basin, an inflow ensemble was generated using 21 years of rainfall and temperature data from 1977 to 1997, and streamflow predictions were conducted for the period from 1998 to 2003. Lastly, for the Seomjin river basin, streamflow predictions were conducted from 1998 to 2003 using historical data from 1981 to 1997. The validation results based on R-B and RMSE indicated that the ESP method was effective as a probabilistic forecasting approach. This is because the average prediction scores were higher than 33.3 % for all validation points, even without the application of the optimal linear correction technique.

Since 2016, the National Drought Information Analysis Center (NDIAC) has been enhancing the reliability of drought prediction by applying ESP, a probabilistic drought forecasting method, and quantitative precipitation-streamflow techniques.

From 2017 onwards, NDIAC has been focused on further improving the reliability of drought forecasting through the development and refinement of analysis techniques at different stages: meteorological, hydrological, and drought forecasting (National Drought Information Analysis Center, 2017, 2018, 2021). To achieve this, Bayesian ESP streamflow prediction has been utilized. NDIAC has constructed a Bayesian ESP framework for weekly and monthly hydrological forecasts (dam inflow prediction) in 35 dam basins nationwide, and it is actively employed in practical applications.

In 2009, the Han River Flood Control Office conducted a study to convert "Water Supply Outlook" (current Water Resources Status and Outlook) into probabilistic forecasts. They performed a 12-month ESP using the Tank model under different initial conditions for the years 2000 to 2008 nationwide. The study aimed to examine the applicability of a probabilistic streamflow system to various basins and seasonal characteristics nationwide. Consequently, they proposed improved techniques for preprocessing and postprocessing, which addressed the uncertainty factors that affected the evaluation of the methods. They evaluated the techniques over different application periods and basins to identify superior approaches. Furthermore, they examined potential issues and improvements related to the expansion of probabilistic streamflow prediction and the establishment of an integrated system. Further, they proposed prediction and trace periods that could minimize the uncertainty of ESP. Consequently, they achieved more reliable medium to long-term probabilistic streamflow prediction compared to previous methods. Moreover, they also developed a user interface to make the "Water Supply Outlook" and other information available to practitioners (Han River Flood Control Office, 2009, 2022).

In addition, the Han River Flood Control Office applied ESP to probabilistic flood forecasting. They generated 3-hour meteorological ensemble traces and created short-term ESP for a representative rainfall event that occurred in South Korea in 2011. Further, they performed probability analysis on the ESP results and categorized them into three ranges: R1, R2, and R3. This categorization facilitated probabilistic forecasting with reduced decision-making risks compared to deterministic forecasts. Ultimately, this approach provided a stable and reliable probabilistic forecast that mitigated the uncertainties associated with decision-making processes (Kim et al., 2011).

Table 2.2 Literature summary of the representative practical application of Ensemble Streamflow Prediction

Author(year)	Title (Target Problem)	ESP model	Key findings
Kim et al. (2001)	Improving water supply outlook in Korea with ensemble streamflow prediction	KICT rainfall-runoff model	Owing to the use ESP and categorical probabilistic forecasting rather than deterministic flow forecasting, the performance of the method using ESP was superior to that of the existing method in Gongju basin.
K-water (2004)	A study on river runoff prediction methods for watershed integrated water management	SSARR rainfall-runoff model	The average prediction score was higher than 33.3%, confirming that it can function as an effective probability prediction technique.
Han River Flood Control Office (2009)	Probabilistic Streamflow Prediction Using Ensemble Model	Tank rainfall-runoff model	For mid- to long-term probabilistic flow forecasting, appropriate correction techniques and forecasting and trace periods that can minimize uncertainty according to the watershed of the Republic of Korea are proposed.
Kim et al. (2011)	Improvement of flood prediction system using stochastic method	KLAPS, MAPLE	For probabilistic flood forecasting, short-term ESP is created through Korea's representative rainfall events and supports stable decision-making.

2.3 ESP in Reservoir Operations

In the previous sections, we have examined the theoretical background of reservoir operation optimization techniques and ESP individually. In this section, we explore the optimization technique referred to as SSDP, which has been developed to maximize the utilization of inflow ensemble traces obtained from ESP. Further, methods for incorporating SSDP with ESP to enhance their combined effectiveness in reservoir operation decision-making are discussed.

SSDP is a non-parametric approach that incorporates the uncertainty of inflow into the recursive equation of SDP via the direct application of the inflow data. It aims to represent the uncertainty of inflow while accounting for the spatial-temporal correlation and continuity of the inflow data (Eum, 2007). By substituting the inflow traces $Q_t(i)$ instead of Q_t into the recursive equation of SDP, Eq. 2.6 is obtained.

$$\underset{R_t^*}{max}\, E\{\left[O_t(\mathbf{S_t^k}, \mathbf{Q_t}(\mathbf{i}), \mathbf{R_t}) + \underset{j|i}{E}\left\{f_{t+1}(\mathbf{S_{t+1}^l}, \mathbf{j})\right\}\right]\}\ \forall S_t, i \text{ and } t \in \{1, \cdots, T\} \tag{2.6}$$

where $Q_t(i)$ represents the i-th inflow trace at time t. Inflow traces range from i to I, where j denotes the inflow trace that occurs after trace i (Figure 2.5). In the case of SSDP, upon the determination of the target release for each stage and reservoir state using Eq. 2.6, the residual optimal benefit function is updated using Eq. 2.7.

$$f_t(\mathbf{S_t^k}, \mathbf{i}) = B_t(\mathbf{S_t^k}, \mathbf{Q_t}(\mathbf{i}), \mathbf{R_t}) + \left\{ f_{t+1}(\mathbf{S_{t+1}^l}, \mathbf{j}) \right\} \; \forall S_t, i \text{ and } t \in \{1, \cdots, T\} \tag{2.7}$$

In Faber and Stedinger (2001), ESP was combined with SSDP to determine the optimal release for the operation of Williams Fork Reservoir in the United States, while considering the conditions at the prediction horizon. In Eq. 2.7, the objective function $B_t(\cdot)$ is determined based on the probabilities of trace j occurring after trace i, using a Markov chain. To estimate these transition probabilities, Faber and Stedinger (2001) conducted regression analysis between the inflow trace i at the base period and the cumulative outflow traces j during the subsequent period. Upon the occurrence of ESP events based on the ESP technique computed by NWS, the transition probabilities were updated using the new ensemble set of inflow traces, and the optimal release was determined accordingly. The results showed that incorporating ESP in the calculation of transition probabilities yielded improved release decisions than those obtained by calculating transition probabilities based solely on historical data.

The optimal releases for the dry season of the Yongdam and Daecheong dams in the Geum river basin were calculated using ESP and SSDP (Eum, 2007; Kim et al., 2007). The objective function considered the minimization of water supply shortage and the maximization of hydropower generation, while considering the multi-purpose nature of the dams in the Geum river basin. In addition, to account for flood considerations, the minimization of deviations from the end-of-June target water level was also considered. In contrast to the study by Faber and Stedinger (2001), an online model was constructed. This involved conducting ensemble predictions for one month at the beginning of each month, thus facilitating forward problem-solving and updating of the optimal releases on a monthly basis. The model using the historical inflow traces was referred to as SSDP/Hist, whereas that updated using ESP every month was referred to as SSDP/ESP (Figure 2.6). The results showed that updating the optimal releases generated by the SSDP/Hist model using the SSDP/ESP model resulted in a reduction of water supply shortage in the Geum river basin by an annual average of $0.6 \times 10^6 m^3/\text{year}$.

Ramaswamy and Saleh (2020) aimed to optimize real-time reservoir operations under extreme rainfall conditions using ESP generated by the HEC-HMS model and DP. They calculated ESP for extreme rainfall events, specifically hurricanes Irene and Sandy, and performed DP for all ESP traces. The results showed that the release decisions varied based on the lead time of ESP, with shorter lead times suggesting less conservative release strategies. However, conservative release strategies resulted in a wide range of release decisions that indicated reservoir flooding as predictions were updated during severe rainfall traces.

In a recent study by Shaikh and Pattanayek (2022), ESP was generated using

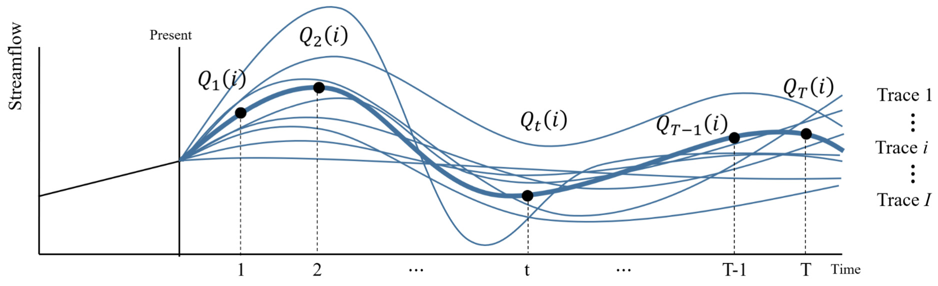


Figure 2.5 ESP traces in SSDP

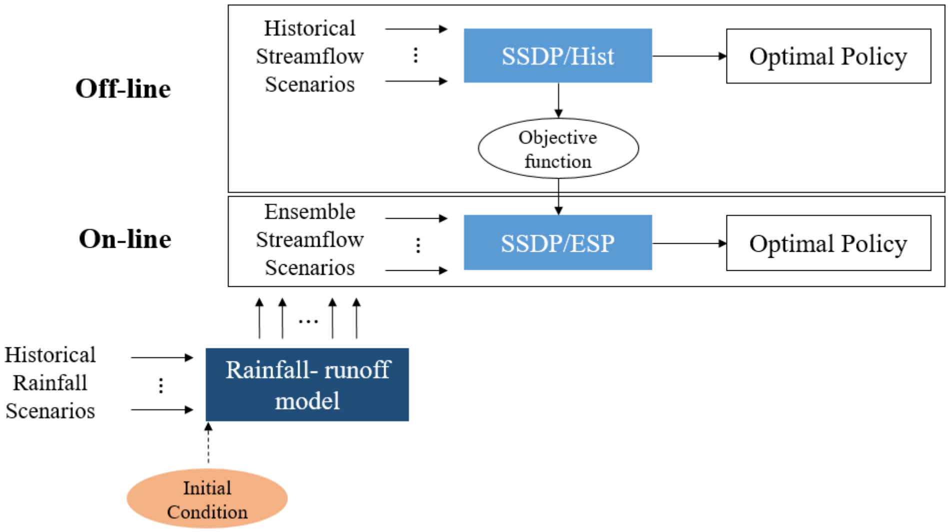


Figure 2.6 On and Off-line operation (Eum, 2007)

a different approach compared to other studies that used rainfall-runoff models. This study utilized long short-term memory (LSTM) to estimate ESP. The study focused on the Upper Hanjiang river basin in China, where LSTM-generated ESP was employed to determine optimal release strategies through NSGA-II. The results showed that the LSTM-based ESP produced highly accurate outcomes. Furthermore, the optimal release strategies derived from the forecast-based approach were more beneficial for additional power generation compared to the no-forecast approach.

Table 2.3 Literature summary of ESP in reservoir operation

Author(year)	Title (Target Problem)	ESP	Optimization	Key findings
Faber and Stedinger (2001)	Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts	NWS ensemble model	Sampling SDP (SSDP)	The SSDP algorithm can better assess the trade-off between immediate benefits and the expected value of water in the future.
Kim et al. (2007)	Optimizing Operational Policies of a Korean Multireservoir System Using Sampling Stochastic Dynamic Programming with Ensemble Streamflow Prediction	SSARR rainfall-runoff model	Sampling SDP (SSDP)	Explicit inclusion of inflow uncertainty and updating the system with ESP forecasts are beneficial.
Eum (2007)	Non-flood Period Operational Policies for the Geum River Multireservoir System Using Sampling SDP with ESP	SSARR rainfall-runoff model	Sampling SDP (SSDP)	On-line operation was applied to update the optimal operating rule at the beginning of each month. Each time ESP was renewed, an annual average of $0.610^6 m^3$ water shortage reduction effect was observed.
Ramaswamy and Saleh (2020)	Ensemble Based Forecasting and Optimization Framework to Optimize Releases from Water Supply Reservoirs for Flood Control	HEC-HMS	Multi-objective DP (R package)	The release decisions progressed less conservative releases with the change in the forecast lead time. Further, the conservative release strategies resulted in a broader range of release decisions that showed the reservoir overtopping as the forecasts were updated under extreme rainfall trace.
He et al. (2022)	Multi-objective operation of cascade reservoirs based on short-term ensemble streamflow prediction	LSTM	Multi-objective NSGA-II algorithm	LSTM models have reasonable accuracy compared with the no-forecast operation. Further, ESP-based operation could harvest additional power.

2.4 Exercise for ESP Effect on Reservoir Operations

The previous section investigated the research trends in dam operations and the application of ESP in academia, examining the use ESP methods in the field of water resources and assessing their performance in various case studies. Several studies have combined ESP with the SSDP technique to update the inflow traces, resulting in significant improvements in performance. This section aims to facilitate a better understanding of the practical utility of probabilistic ESP for dam operators by quantifying the benefits of probabilistic forecasting through a simple example, without the need for sophisticated ESP derivation or optimization techniques.

Here, we have a very simple reservoir (Figure 2.7) with maximum and minimum capacities of 10 and 0, respectively. The average inflow at each time step was 7, and our goal was to supply a constant demand of 5 of water throughout the operational period. We began with an initial reservoir capacity of 10 and aimed to operate it successfully until the 3rd step, which represents the non-flood season while meeting the demand.

- Maximum Storage (S_{max}) = 10
- Minimum Storage (S_{min}) = 0
- Average Inflow (μ) = 7
- Demand = 5 (constant)

- Planning Period (T) = 3
- Initial Storage (S_0) = 10

$$Scenario\,W \sim N(7, 1^2) \tag{2.8}$$

$$Scenario\,P \sim N(7, 2.5^2) \tag{2.9}$$

We assumed that in this simple reservoir, at each time step, inflow followed two different normal distributions with the same mean of 7, albeit with different standard deviations of 1 and 2.5, respectively (Eq. 2.9), as illustrated in Figure 2.8. Let us denote the inflow with a standard deviation of 1 as Scenario W and that with a standard deviation of 2.5 as Scenario P. In addition, this reservoir incurs a cost for water supply through water purchase if the reservoir capacity falls below 2. Therefore, a cost of water supply, following a sigmoid function in Eq. 2.10, was incurred based on the reservoir capacity x, as shown in Figure 2.9. Considering the Scenarios W and P, the question is how should the operation of the reservoir differ.

$$c(x) = \frac{10000}{\exp(3x - 5)} \tag{2.10}$$

Let us assume the worst-case scenario where an inflow with a probability of

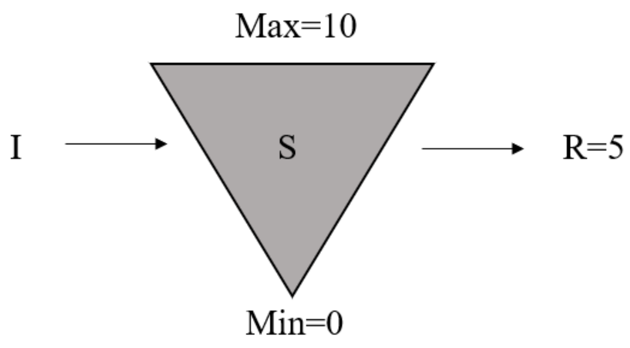


Figure 2.7 Problem setting for exercise

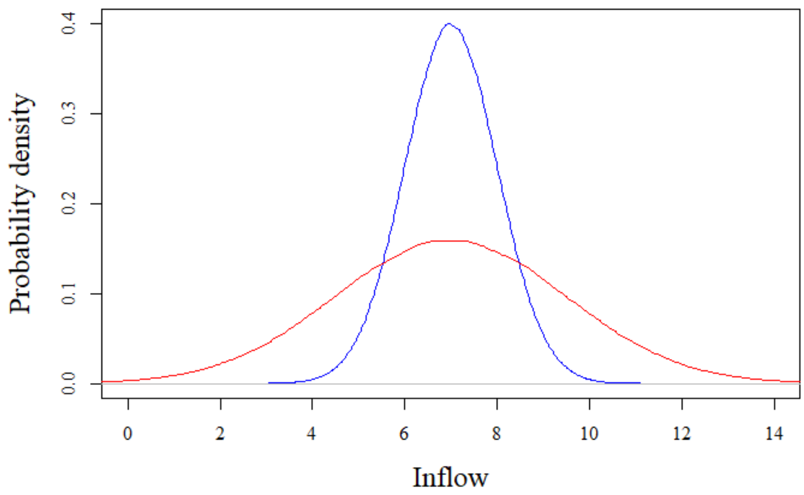


Figure 2.8 Inflow probability density function for exercise

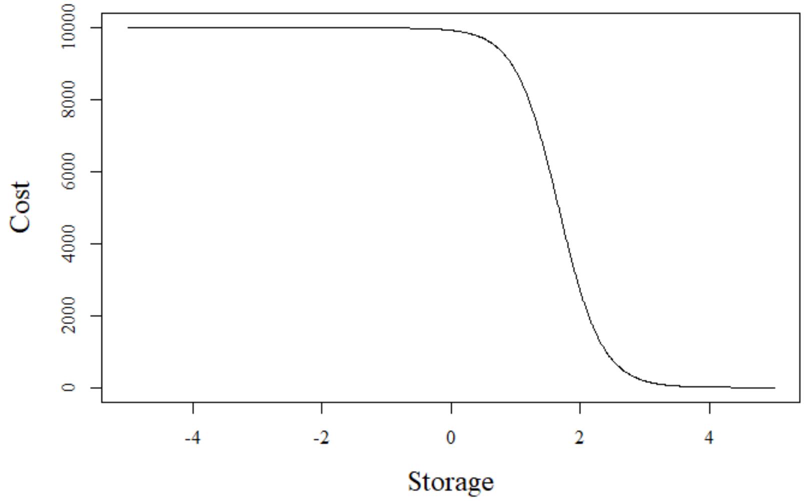


Figure 2.9 Cost function for exercise

0.01 occurs over three-time steps in both inflow probability density functions. The probability of this worst-case scenario occurring is 10^{-6} for both Scenarios W and P, as shown in Eq. 2.12. In this case, for Scenario W, the inflow with a probability of 0.01 has a value of 4.3, while for Scenario P, the inflow with a probability of 0.01 has a value of 1.1, as indicated in Eq. 2.14.

$$Scenario\, W : Pr(\sum I) = 0.01 \times 0.01 \times 0.01 = 10^{-6} \tag{2.11}$$

$$Scenario\, P : Pr(\sum I) = 0.01 \times 0.01 \times 0.01 = 10^{-6} \tag{2.12}$$

$$f_W(0.01)^{-1} = 4.3 \tag{2.13}$$

$$f_P(0.01)^{-1} = 1.1 \tag{2.14}$$

$$S_{t+1} = S_t + Q_t - R_t \tag{2.15}$$

Through the application of the mass conservation equation (Eq. 2.15), we can calculate the reservoir storage (S_3) after the three-time steps. If the demand of 5 is satisfied at every time step, the final reservoir level would be 7.9 and -1.7 for Scenarios W and P, respectively, as shown in Eq. 2.17.

$$S_{3,W} = 10 + (12.9 - 15) = 7.9 \tag{2.16}$$

$$S_{3,P} = 10 + (3.3 - 15) = -1.7 \tag{2.17}$$

At each time step (t = 1, 2, 3), we can calculate the cost. Consequently, for Scenario W, according to Eq. 2.15, no cost was incurred. However, for Scenario P, a cost of 11680 was incurred.

$$C_{3,W} = \frac{10000}{\exp\left(3 \times 10 - 5\right)} + \frac{10000}{\exp\left(3 \times 9.3 - 5\right)} \tag{2.18}$$

$$+ \frac{10000}{\exp\left(3 \times 8.6 - 5\right)} + \frac{10000}{\exp\left(3 \times 7.9 - 5\right)} \tag{2.19}$$

$$= 0 + 0 + 0 + 0 = 0 \tag{2.20}$$

$$C_{3,P} = \frac{10000}{\exp\left(3 \times 10 - 5\right)} + \frac{10000}{\exp\left(3 \times 6.1 - 5\right)} \tag{2.21}$$

$$+ \frac{10000}{\exp\left(3 \times 2.2 - 5\right)} + \frac{10000}{\exp\left(3 \times (-1.7) - 5\right)} \tag{2.22}$$

$$= 0 + 0 + 1680 + 10000 = 11680 \tag{2.23}$$

The entire operation can be visualized in Figure 2.10 and Figure 2.11. Thus, even with the same probability for the worst-case scenario, in case of different standard deviations of the inflow probability density function, different results are observed in terms of the dam's cost function. This indicates that relying solely on past averages without considering other statistics of inflow is not sufficient for optimal dam operation. Therefore, it is crucial to incorporate the distribution of inflow to the best extent possible in the dam operation.

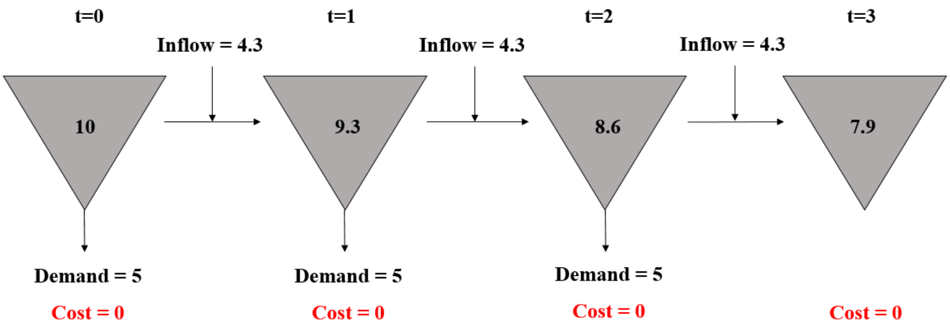


Figure 2.10 Schematic of dam releases during the 3-time step for Scenario W

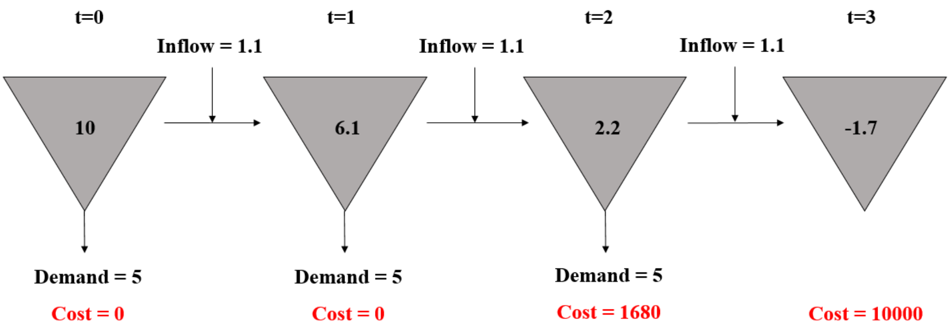


Figure 2.11 Schematic of dam releases during the 3-time step for Scenario P

Chapter 3. Methodology

3.1 Overall Procedure

In this chapter, we describe the overall procedure based on a survey of research trends. The first step involved selecting the target multiple-purpose dams among the 21 dams in South Korea. To analyze the characteristics of different multiple-purpose dams, the CIR was employed as the selection criterion. Based on the CIR, the Soyanggang and Chungju dams were chosen as the study sites. For each dam, a distribution was assumed with the same ensemble mean but different variances, as demonstrated in the exercise presented in Chapter 2, to capture the effect of ESP on reservoir operations. Consequently, SSDP/Hist and SSDP/ESP models were developed for Scenarios W and P, and the optimal release rates were calculated. In addition, the optimal release rates assuming perfect forecasts (PERF) were determined using DDP for the purpose of comparison. The calculated optimal release rates were then used for simulation, and the results were analyzed in terms of frequency, duration, and magnitude (Figure. 3.1).

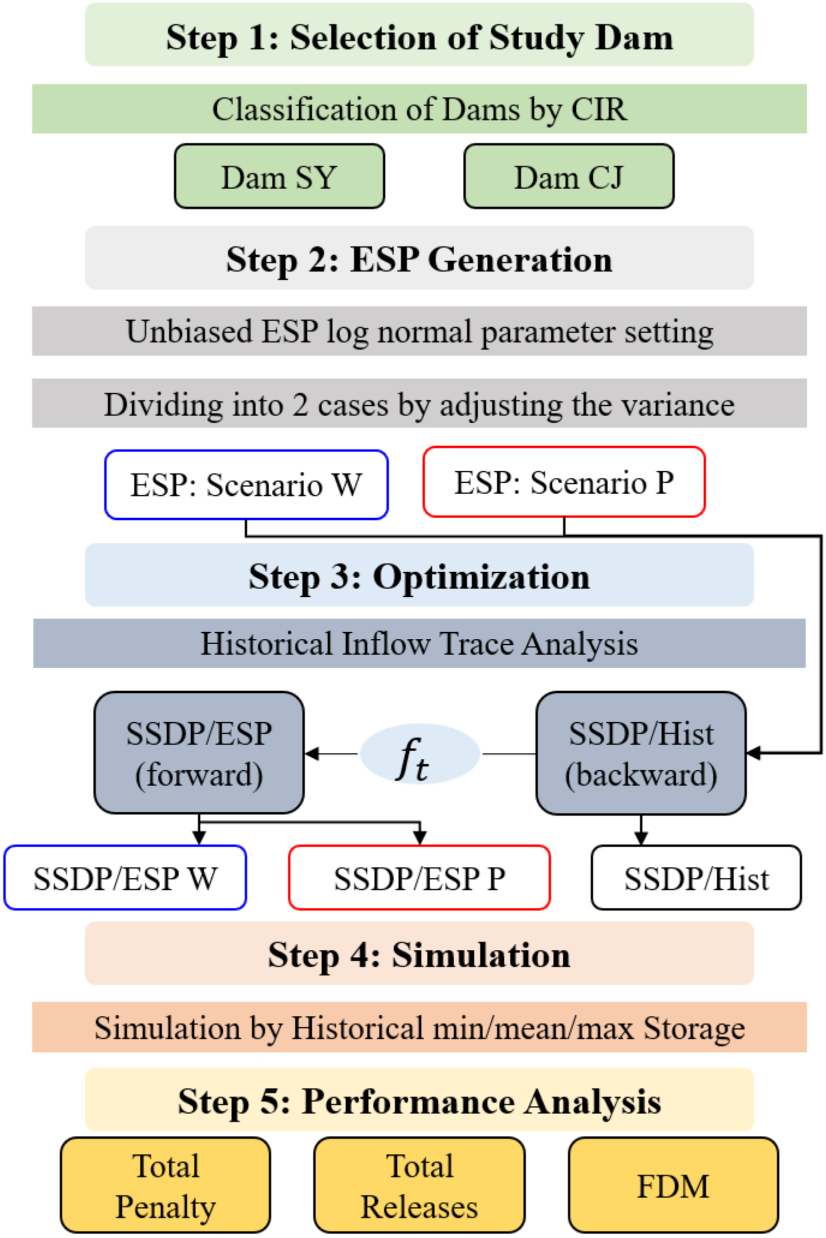


Figure 3.1 Research overview

3.2 Ensemble Streamflow Prediction

ESP assumes unbiased forecasting (Figure. 3.3) and considers two distributions with the same ensemble mean albeit different standard deviations (Figure.3.2). The ensemble mean values of ESP were determined based on the monthly inflow data from the last set of the selected dams, spanning from 2020 to 2022, following the two assumptions described below.

- Assumption 1: Monthly averages of ESP are identical to the corresponding observations (i.e., unbiased forecasting)
- Assumption 2: Monthly variance of ESP comprises 2 Scenarios depending on the variance.

The inflow data was tested for goodness of fit and assessed using Q-Q plots to determine if it followed a log-normal distribution. In this case, the relationship between μ_X and σ_X and the resulting μ_Y and σ_Y after fitting the log-normal distribution are described by Eqs.3.1 and 3.2. For Assumption 2, the standard deviation was assumed as $\sigma_X = \mu_X$ and $\sigma_X = 10\mu_X$ for Scenarios W and P, respectively. Consequently, using Eq. 3.2, the resulting σ_Y for Scenarios W and P were 0.83 and 2.15, respectively.

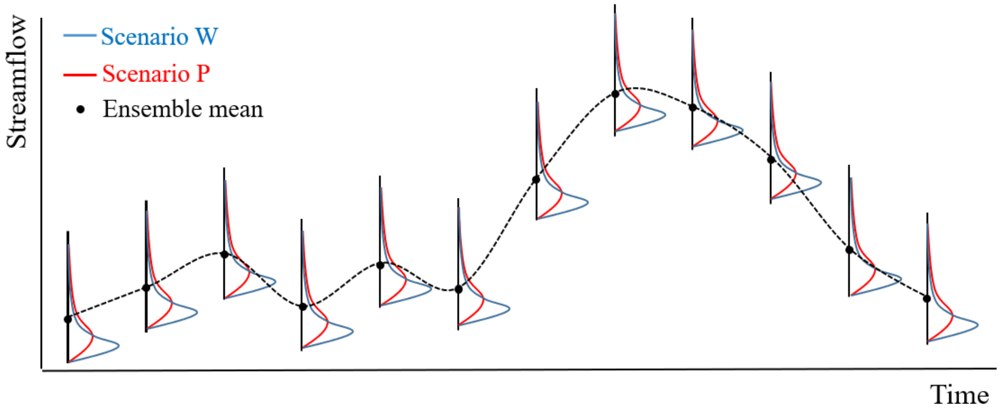


Figure 3.2 Schematic of Scenario W/P according to monthly ESP assumptions

$$\mu_Y = ln\left(\frac{\mu_X^2}{\sqrt{\mu_X^2 + \sigma_X^2}}\right) \tag{3.1}$$

$$\sigma_Y = \left[ln\left(1 + \frac{\sigma_X^2}{\mu_X^2}\right)\right]^{\frac{1}{2}} \tag{3.2}$$

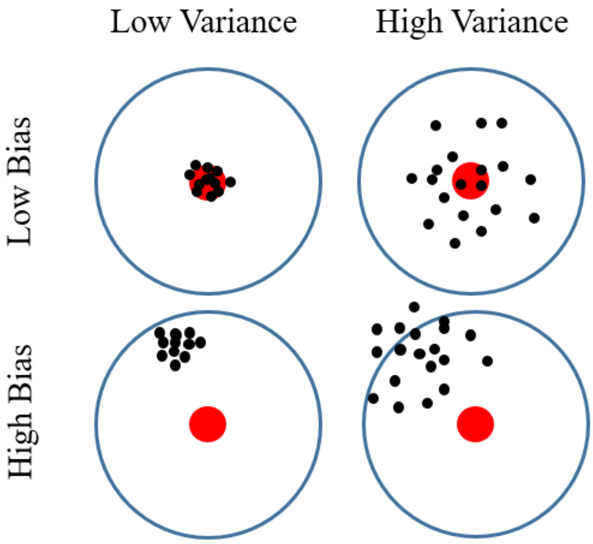


Figure 3.3 Bias and variance

3.3 Sampling Stochastic Dynamic Programming

Setting the objective function is crucial in formulating an optimization problem, whether it involves minimization or maximization. In our case, the improvement of water supply in multiple-purpose dams was deemed as the most urgent issue. Therefore, the objective function was set as the water shortage amount, with an aim to minimize it (Eq. 3.3). Here, D_t and R_t represent the demand and release at time t, respectively.

$$O_t(S_t, Q_t, R_t) = max(0, D_t - R_t) \tag{3.3}$$

As mentioned earlier, SSDP is a non-parametric approach that directly incorporates inflow data into the SDP recursion equation, thereby representing the uncertainty of inflow while considering its continuity and spatiotemporal correlations. This is in contrast to discretizing the inflow and using representative values and probabilities for each interval in SDP (Eum, 2007).

In this study, the objective function was not a future benefit function to be maximized, rather it was the water shortage amount to be minimized. Incorporating this into Eq.2.4 yields Eq. 3.4, where I is the number of inflow traces used for optimal release calculation, and $Q_t(i)$ represents the inflow of the i-th trace at time t.

$$\underset{R_t}{min}\underset{Q_t}{E}\left[O_t(\mathbf{S_t^k},\mathbf{Q_t}(\mathbf{i}),\mathbf{R_t})+\left\{f_{t+1}(\mathbf{S_{t+1}^l},\mathbf{i})\right\}\right]\forall S_t,t\in\{1,2,\cdots,T\} \tag{3.4}$$

$$=\underset{R_t}{min}\sum_{Q_t}Pr(i)\left[O_t(\mathbf{S_t^k},\mathbf{Q_t}(\mathbf{i}),\mathbf{R_t})+\left\{f_{t+1}(\mathbf{S_{t+1}^l},i)\right\}\right] \tag{3.5}$$

The application of discretization of inflow to Eq.3.4 yields Eq. 3.6, which represents the SSDP formulation. SSDP utilizes Eq. 3.6 to evaluate the residual expected benefit function for each trace i associated with the optimal release, once the state variable representing each stage and the current state is determined. This is realized using Eq. 3.7.

$$\underset{R_t}{min}\sum_{i=1}^{I}Pr(i)\left[O_t(\mathbf{S_t^k},\mathbf{Q_t}(\mathbf{i}),\mathbf{R_t})+\left\{f_{t+1}(\mathbf{S_{t+1}^l},\mathbf{i})\right\}\right] \tag{3.6}$$

$$f_t(\mathbf{S_t^k},\mathbf{i})=O_t(\mathbf{S_t^k},\mathbf{Q_t}(\mathbf{i}),\mathbf{R_t})+\left\{f_{t+1}(\mathbf{S_{t+1}^l},\mathbf{i})\right\} \tag{3.7}$$

Next, let us consider the constraints. The most fundamental constraints are that the reservoir storage cannot be negative (less than 0) and cannot exceed the reservoir capacity. Furthermore, water supply in multiple-purpose dams in South Korea is conducted within the range between the normal high water level (NHWL) and the low water level (LWL). To reflect a more realistic situation, it is ensured that the storage at each stage did not fall below the LWL (Figure 3.4). Therefore, for each dam, the minimum storage level (S_{min}) was set as the LWL, and the maximum storage level (S_{max}) represents the storage at the NHWL of the dam (Eq. 3.8).

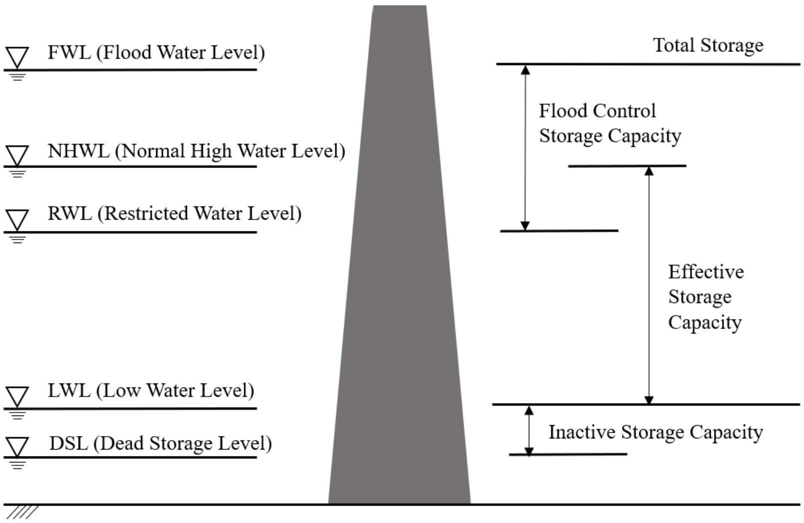


Figure 3.4 Multi-purpose dam water level and capacity (K-water, 2021)

$$S_{min} < S < S_{max} \tag{3.8}$$

In the DP family of optimization algorithms, the calculations are performed backward until a certain level of convergence is achieved. However, in this study, a convergence condition was applied to determine to the point at which the calculations were to be stopped. The convergence condition used in the research is shown in Eq. 3.9.

$$||(f_{opt}^{i-1} - f_{opt}^{i})| - |(f_{opt}^{i} - f_{opt}^{i+1})|| < \epsilon \tag{3.9}$$

where f_{opt}^{i} represents the optimal value in the i-th iteration, and ϵ denotes the tolerance. In this study, ϵ was set to 10^{-5}.

3.4 ESP in Sampling Stochastic Dynamic Programming

In the previous theoretical background, we examined a research case that utilized ESP to determine accurate operational rules in the optimization equation of SSDP. As evident from the recursive equation of SSDP, calculating the current value f_t requires the value of f_{t+1}. Therefore, first, the entire inflow traces over the entire operational period were predicted, and then the optimal releases and future value functions were calculated in a backward manner. However, the backward calculation requires the prediction of the entire inflow traces over the operational period, which poses challenges in accuracy, particularly when the prediction horizon exceeds one month in the South Korean context. To address this, Eum (2007) proposed an improvement by first calculating SSDP/Hist using past inflow data and using the obtained future value function in the calculation of optimal releases in SSDP/ESP, which can be solved in a forward manner (Figure 3.6). In this study, we adopted the same approach to compute the optimal releases. The entire 27-year inflow traces were divided into three sets, resulting in nine sets, and SSDP/Hist was calculated excluding the last set from 2020 to 2022. Subsequently, the final set was used for ESP generation and simulation (Figure 3.5).

When denoting the inflow trace of ESP as i and the past inflow trace received from SSDP/Hist as j, the calculation formula is expressed as Eq. 3.10. The expectation term is multiplied because it represents the calculated f values for each trace i and

j in ESP and Hist, respectively. It was assumed that the transition matrix from trace i to j was the same for all trace i. Upon the application of the calculated optimal releases to actual operations, the operation was conducted as follows: if S_{t+1} was less than S_{min}, no release was made, and if S_{t+1} exceeded S_{max}, the excess amount was discharged as spill (Figure 3.7). In addition, to observe the results based on different initial storage conditions, the minimum, average, and maximum historical water levels were used as the initial storage.

$$\min_{R_t^*} \sum_{i=1}^{I} Pr(\mathbf{i}) \left[O_t(\mathbf{S_t^k}, \mathbf{Q_t}(\mathbf{i}), \mathbf{R_t}) + \underset{j|i}{E} \left\{ f_{t+1}(\mathbf{S_{t+1}^l}, \mathbf{j}) \right\} \right] \tag{3.10}$$

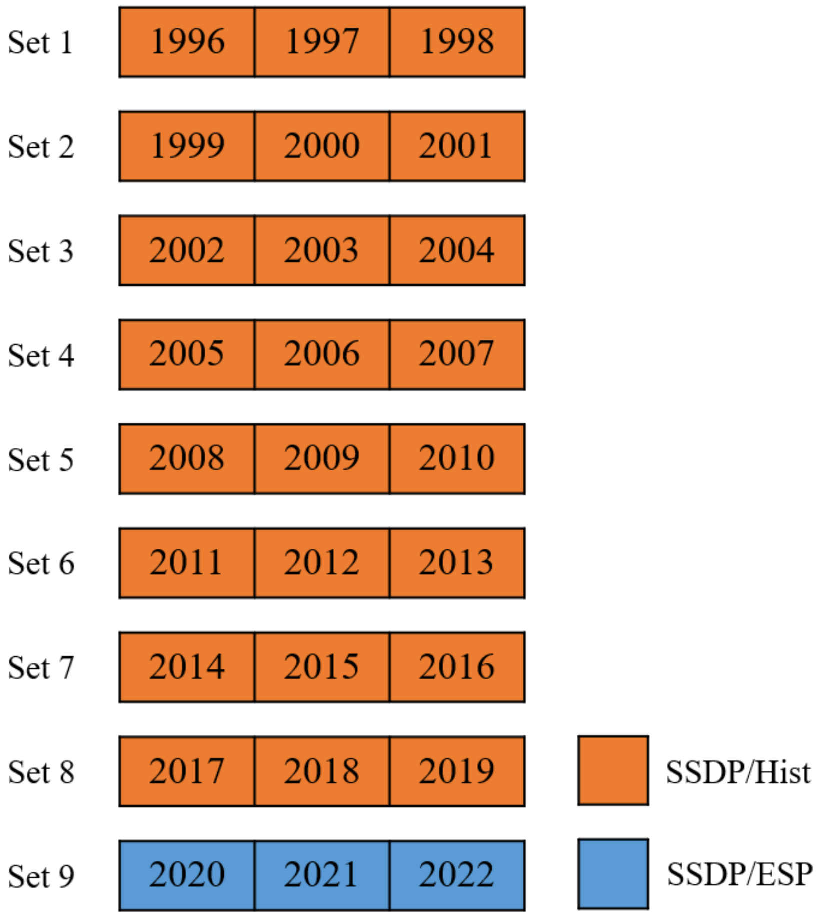


Figure 3.5 Division of the inflow dataset

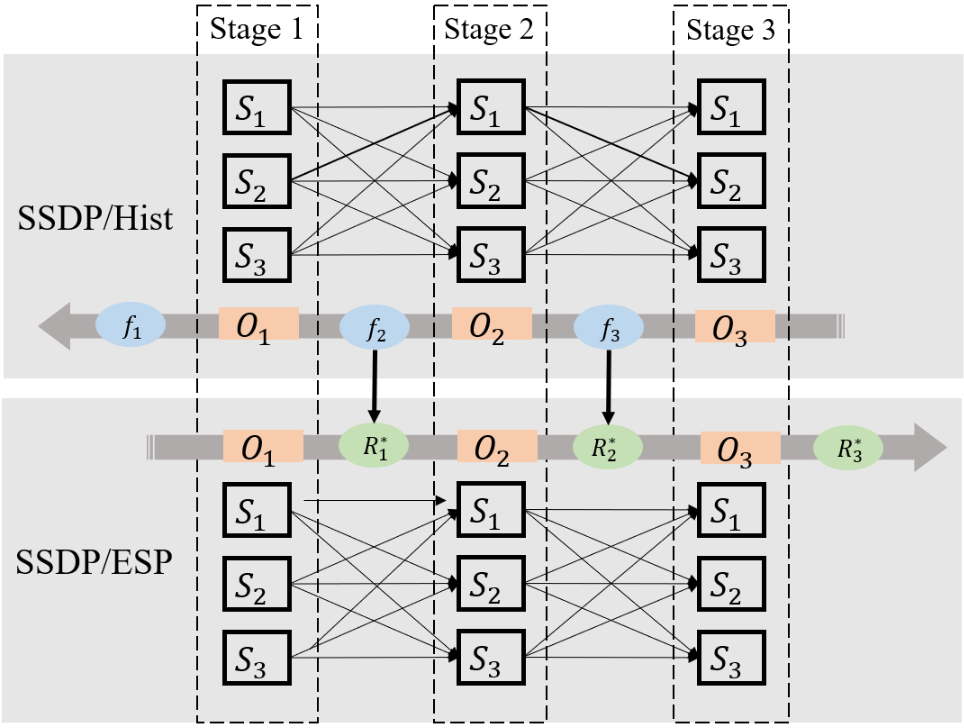


Figure 3.6 Connection of SSDP/Hist and SSDP/ESP

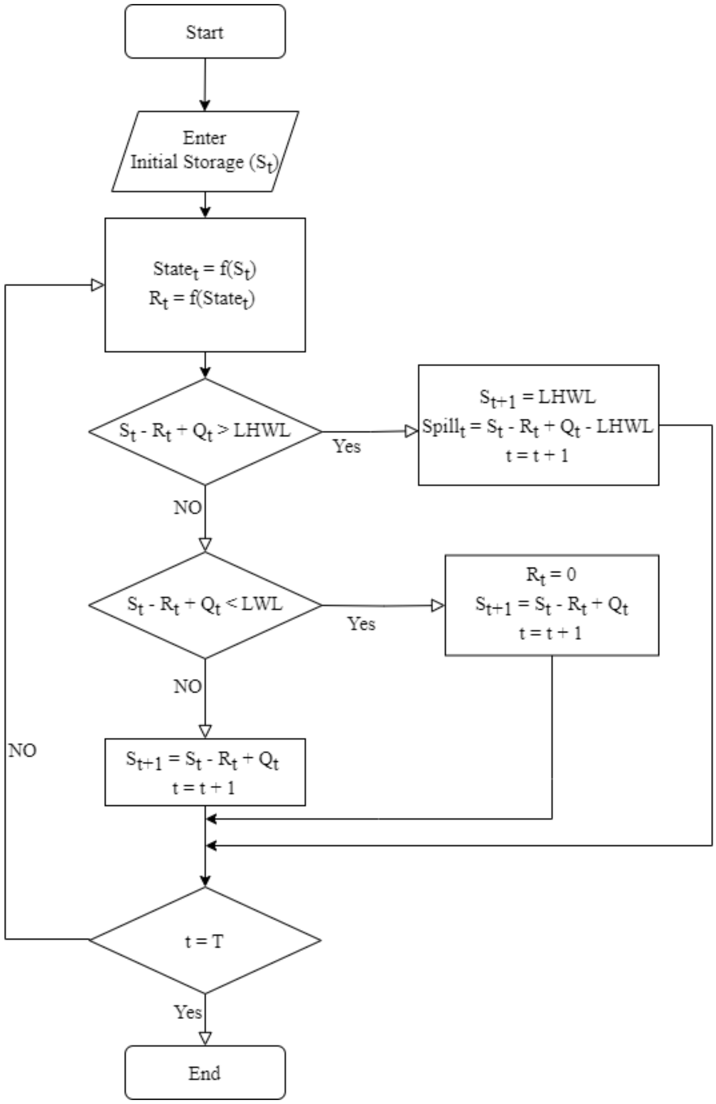


Figure 3.7 Simulation flowchart

3.5 Performance Matrix

In reservoir operations, the performance evaluation can be divided into absolute and relative evaluations. Absolute evaluation refers to the evaluation based on absolute quantities, such as water supply, obtained through simulations. However, relative evaluation refers to the evaluation of whether the system satisfies the criteria set by the system users. Hashimoto et al. (1982) described the relative evaluation of the system in terms of reliability, resiliency, and vulnerability. In addition, Kim et al. (2019) and Kim et al. (2021) evaluated the performance of the Boryeong Dam in South Korea using the Frequency of annual water deficit, duration of the longest failure, and magnitude of annual water deficit. Subsequently, Kim et al. (2022) used this performance matrix for the evaluation of adaptive reservoir management, enabling a more intuitive assessment. These performance matrices are also applied in this study. To assess water deficit, the success of the system is defined as R_t satisfying D_t at each time step, as shown in Eq. 3.11. If the results owing to a random variable X_t are satisfactory for a certain period, it is defined as S; otherwise, it is defined as F.

$$X_t \in S \qquad R_t \geq D_t \tag{3.11}$$

$$X_t \in F \qquad otherwise \tag{3.12}$$

- Frequency: To what degree does the system experience satisfaction?
- Duration: How long does the system stay in a satisfactory state?
- Magnitude: What is the severity of the failures that occur in the system?

The Frequency (ρ) refers to the total number of successful operations during the entire operational period. In categorical data representing the success and failure of the system, a value of 1 is assigned to W_t in Eq. 3.13 when the operation is successful, whereas a value of 0 is assigned to W_t in Eq. 3.14 when the operation is a failure.

$$W_t = 1 \qquad X_t \in S \tag{3.13}$$

$$W_t = 0 \qquad X_t \in F \tag{3.14}$$

The Frequency (ρ) can be expressed as Eq. 3.15.

$$\rho = Pr(X_t \in S) = \frac{1}{T}\sum_{t=1}^{T} W_t \tag{3.15}$$

The Duration (γ) represents the average period during which the system remains successful throughout the operation period. $I(\cdot)$ calculates the number of consecutive occurrences of an event within parentheses, and $L(\cdot)$ calculates the duration of consecutive occurrences of an event. The duration can be expressed as Eq. 3.16.

$$\gamma = \frac{1}{T}\frac{\sum_{t=1}^{T} L(W_t = 1)}{I(W_t = 1)} \tag{3.16}$$

The Magnitude (v) represents the shortfall of the total demand relative to the total supply in the system and is expressed as a complement to 1. Similar to frequency and duration, magnitude is defined as an upward indicator. It is calculated as the difference between 1 and the ratio of total shortfall to total demand, as expressed in Eq. 3.17.

$$v = 1 - \frac{\sum_{t=1}^{T} max(0, D_t - R_t)}{\sum_{t=1}^{T} D_t} \tag{3.17}$$

This study evaluated the performance of the optimal release after simulation using both absolute and relative assessment measures. The absolute evaluation was performed through the penalty incurred, whereas the relative evaluation was conducted using the frequency-duration-magnitude (FDM) framework (Figure 3.8). These measures were employed to compare the performances of Scenarios W and P.

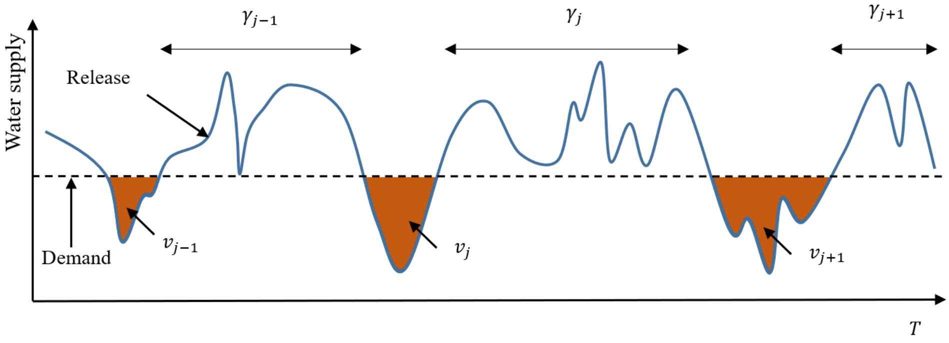


Figure 3.8 System FDM performance criteria

Chapter 4. Sample Studies

4.1 Classification of Dams by CIR

To examine the operational behavior of dams under different conditions, various multipurpose dams in South Korea were classified according to their CIR and arranged in Table 4.1. Following the classification by Karamouz and Houck (1987), dams with a CIR in the range of 0.5–1 were categorized as "Large," whereas those with a CIR greater than 1 were classified as "Very Large." Based on this classification, the Soyanggang and Chungju dams, which have similar capacities and data lengths within the same Han river basin, were selected as the study areas. The CIR classification, along with a map of South Korea, is shown in Figure 4.1.

The selected study area, the Han river basin, is located in central South Korea and is the largest river system that runs through the capital city of Seoul. It connects the Han and the Bukhan rivers. The total basin area is 25,953.6 km² (or 35,770.41 km² when including North Korea). With a river length and average width of 494.44 km and 75.5 km, respectively, the Han river basin is the primary river basin in South Korea, covering approximately 23 % of the national territory. Within the Han River basin, there are several multi-purpose dams such as the Soyanggang, Chungju, and Hoengseong dams, which contribute to water supply operations. The water supply mimetic diagram for this area is shown in Figure 4.2.

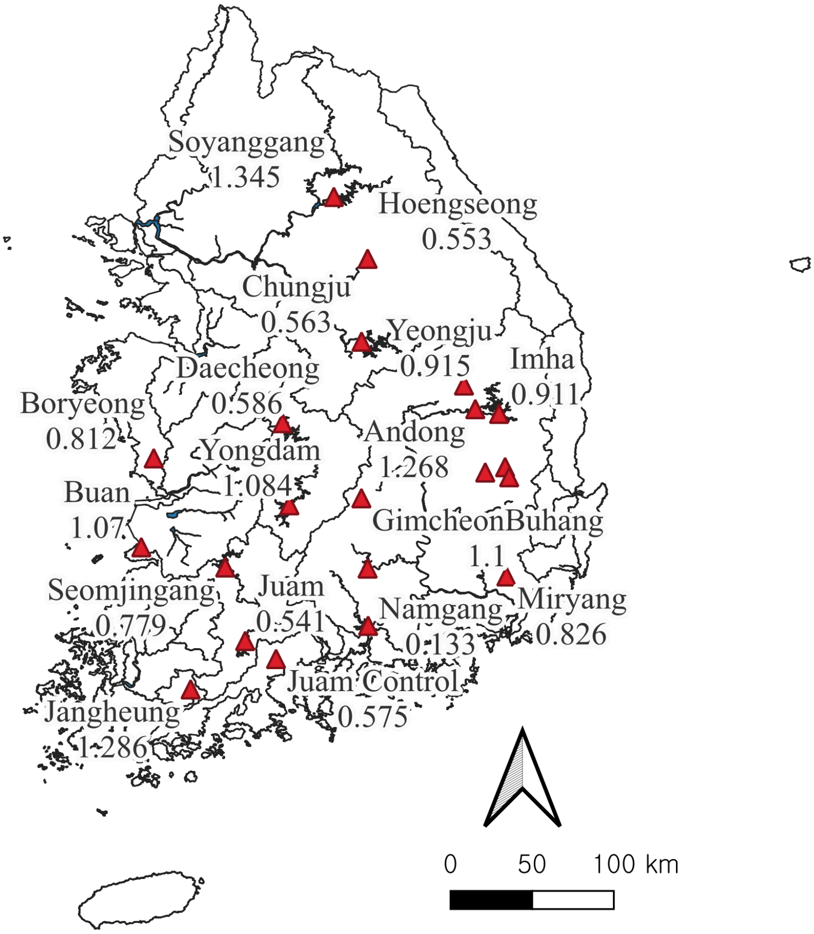


Figure 4.1 Multi-purpose dam with CIR in Korea

Table 4.1 CIR of multi-purpose dams in Korea

Dam	Start year	End year	Capacity (MCM)	Annual average inflow (MCM)	CIR
Seongdeok	2012	2021	27.9	17.1	1.628
Gunwi	2011	2021	48.7	35.6	1.367
Soyanggang	1996	2021	2900.0	2156.1	1.345
Jangheung	2005	2021	191.0	148.5	1.286
Andong	1996	2021	1248.0	984.5	1.268
Hapcheon	1996	2021	48.7	35.6	1.185
GimcheonBuhang	2013	2021	54.3	49.3	1.100
Yongdam	2001	2021	815.0	751.8	1.084
Buan	1997	2021	50.3	47.0	1.070
Yeongju	2012	2021	181.1	198.0	0.915
Imha	1996	2021	595.0	652.9	0.911
Bohyeonsan	2014	2021	22.1	25.6	0.862
Miryang	2001	2021	73.6	89.1	0.826
Boryeong	1998	2021	116.9	144.0	0.812
Seomjingang	1996	2021	466.0	598.0	0.779
Daecheong	1996	2021	1490.0	2542.7	0.586
Juam Control	1996	2021	250.0	434.9	0.575
Chungju	1996	2021	2750.0	4888.8	0.563
Hoengseong	2001	2021	86.9	157.1	0.553
Juam	1996	2021	457.0	844.8	0.541
Namgang	2000	2021	309.2	2320.4	0.133

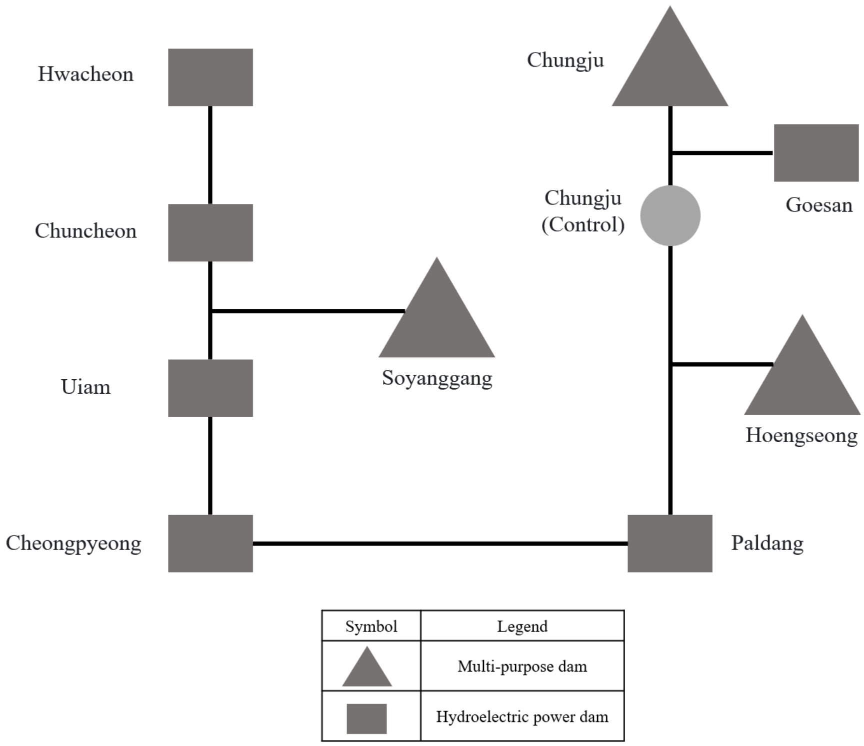


Figure 4.2 Water supply mimetic diagram in Han River

4.2 Sample Study 1: Dam SY

To set the constraints for SSDP according to the current status of the Soyanggang Dam (Figure 4.3), the requisite information has been summarized in Table 4.2. The storage was discretized into 100 intervals with equal spacing. In the case of multi-purpose dams in South Korea, water supply operations occur between the LWL and the NHWL. Therefore, the storage should not fall below the LWL and should not exceed the NHWL. For each dam, the value of S_{min} was set to the LWL, and S_{max} was set to the NHWL, as shown in Eq. 3.8. According to the information, the LWL and NHWL for the Soyanggang Dam were calculated as 693.574 and 2478.906 MCM, respectively. Furthermore, the monthly demand was obtained by converting the daily demand from K-water (2020) into monthly values, as shown in Table 4.3.

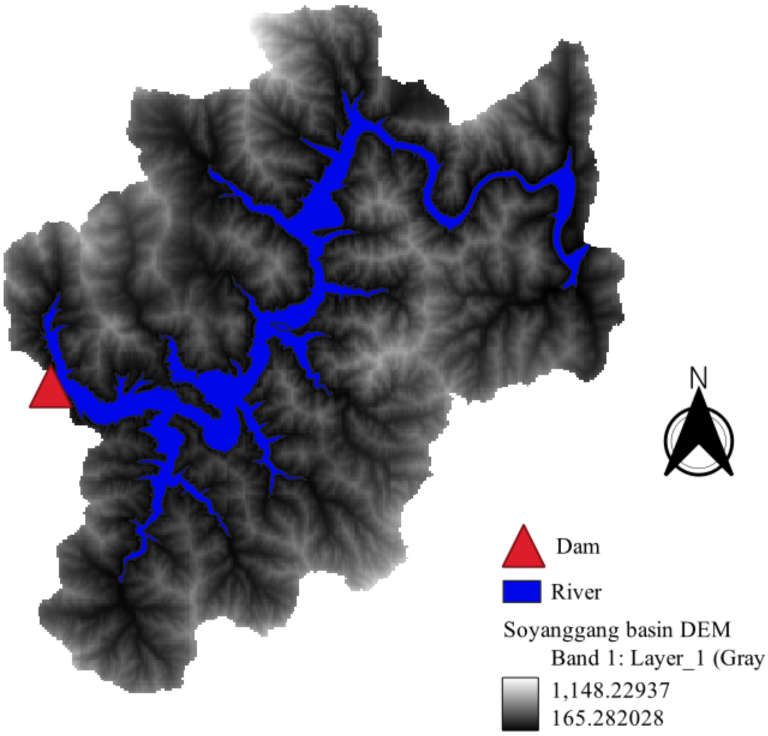


Figure 4.3 Dam SY basin

Table 4.2 Information of Dam SY

Information	Figure
Height (m)	123.0
Length (m)	530.0
Normal elevation (EL.m)	203.0
Volume (1000 m^3)	9,591.0
Basin area (km^2)	2,703.0
Annual water supply capacity (MCM)	1,213.0
Reservoir area (km^2)	70.0
Design Flood Level (EL.m)	198.00
NHWL (Normal High Water Level) (EL.m)	193.50
RWL (Restricted Water Level) (EL.m)	190.30
Spill Water Level (EL.m)	185.50
LWL (Low Water Level) (EL.m)	150.00
Total Storage (MCM)	2,900.0
Available Storage (MCM)	1,900.0
Flood Control Storage Capacity (MCM)	500.0

Table 4.3 Monthly demand of Dam SY (MCM) (K-water, 2020)

	January	February	March	April	May	June
Demand	123.752	111.776	124.837	122.340	126.418	122.340
	July	August	September	October	November	December
Demand	126.418	126.418	122.340	124.558	119.760	123.752

4.2.1 Ensemble Streamflow Prediction

To assume the ESP for the Soyanggang Dam, the goodness of fit of a log-normal distribution to the historical inflow data was first tested. The results indicated that the log-normal distribution was a good fit. Based on the monthly inflow data from 2020 to 2022, the ESP for Scenarios W and P were assumed. Table 4.4 lists the distribution parameters for Scenario W inflow in 2022, assuming a log-normal distribution. Similarly, Table 4.5 lists the parameters for Scenario P inflow using the same data. The assumed ESP distributions were then discretized into three quantiles, Q1, Q2, and Q3, as shown in Table 4.7 and 4.8.

The average of the historical inflow traces used in SSDP/Hist from 1996 to 2019 was 2135.143 MCM. In comparison, the annual average inflow for SSDP/ESP from 2020 to 2022 was 2522.661 MCM. This indicates that the inflow data used in ESP was approximately 1.2 times richer than that used in SSDP/Hist. On an annual basis, the total inflow in 2020 was approximately 1.6 times the historical average, whereas the total inflow in 2021 was approximately 0.7 times the average, and the total inflow in 2022 was approximately 1.3 times the average. Figure 4.4 shows the box plot of the inflow data used in SSDP/Hist and the line graph of the inflow data used in SSDP/ESP for each year.

Table 4.4 Inflow distribution parameters of Scenario W in 2022 of Dam SY (MCM)

	Jan	Fed	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
μ_X	11.33	6.17	97.82	63.24	22.13	272.35	498.97	926.81	574.02	190.62	53.29	35.40
σ_X	11.33	6.17	97.82	63.24	22.13	272.35	498.97	926.81	574.02	190.62	53.29	35.40
μ_Y	2.08	1.47	4.24	3.80	2.75	5.26	5.87	6.49	6.01	4.90	3.63	3.22
σ_Y	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83

Table 4.5 Inflow distribution parameters of Scenario P in 2022 of Dam SY (MCM)

	Jan	Fed	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
μ_X	11.33	6.17	97.82	63.24	22.13	272.35	498.97	926.81	574.02	190.62	53.29	35.40
σ_X	113.26	61.72	978.16	632.38	221.27	2723.52	4989.70	9268.05	5740.19	1906.16	532.90	354.05
μ_Y	0.12	-0.49	2.28	1.84	0.79	3.30	3.90	4.52	4.05	2.94	1.67	1.26
σ_Y	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15

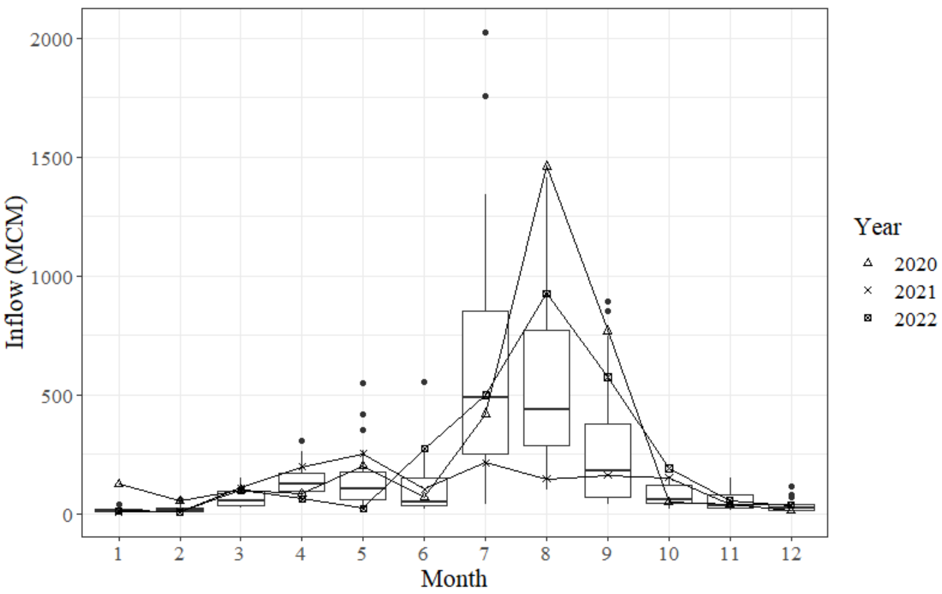


Figure 4.4 Comparison of monthly inflow data used for SSDP/Hist and SSDP/ESP in Dam SY

4.2.2 SSDP/Hist Model

For SSDP/Hist, the historical inflow data was used to update the optimal release and future value function in a backward direction. For the Soyanggang Dam, 27 years of data are available from 1996 to 2022. However, the inflow data from the last trace set (2020 to 2022) was excluded as it was used for ESP assumptions. Therefore, the data from 1996 to 2019 (24 years) was incorporated into SSDP/Hist. When using Eq. 3.6, the expectation is calculated by multiplying each trace by $Pr(Q_t)$. However, in this study, instead of using $Pr(Q_t)$ directly, the total inflow for each trace was calculated and fitted to a log-normal distribution. For inflows smaller than the mode, the non-exceedance probability was used, whereas for inflows larger than the mode, the exceedance probability was used for scaling. Consequently, these weighted probabilities, denoted as w_j, were applied.

$$\underset{R_t^*}{min} \sum_{j=1}^{24} w_j \left[O_t(\mathbf{S_t^k}, \mathbf{Q_t(j)}, \mathbf{R_t}) + \left\{ f_{t+1}(\mathbf{S_{t+1}^l}, \mathbf{j}) \right\} \right] \tag{4.1}$$

In Figure 4.5, the variable j represents the 24 years of historical inflow traces, and w_j denotes the weighted values obtained by scaling each trace according to its probability of occurrence. The optimal release rates derived from this calculation are presented in Figure 4.7 as a 3D plot and in Figure 4.8 as a heatmap.

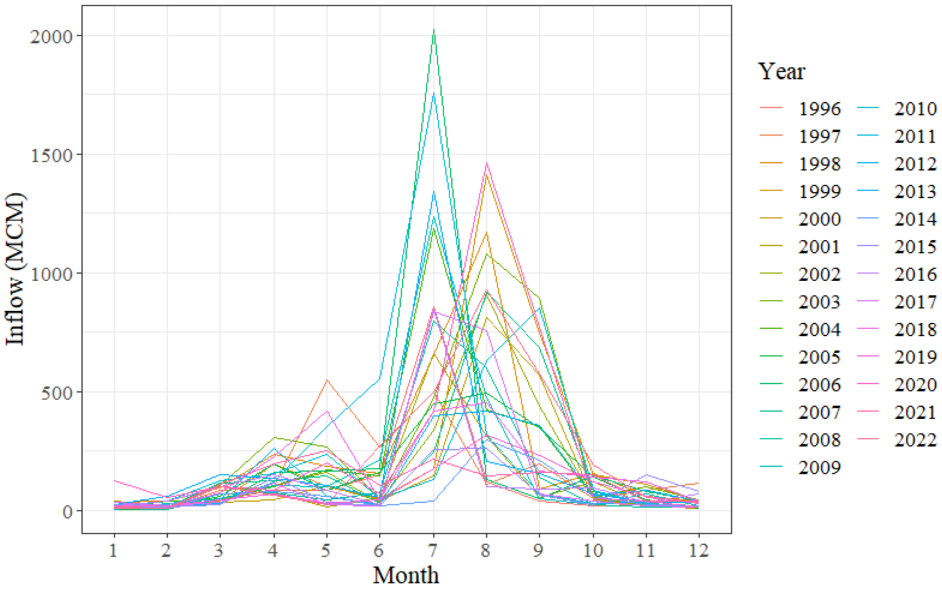


Figure 4.5 Historical inflow ensemble traces for SSDP/Hist in Dam SY

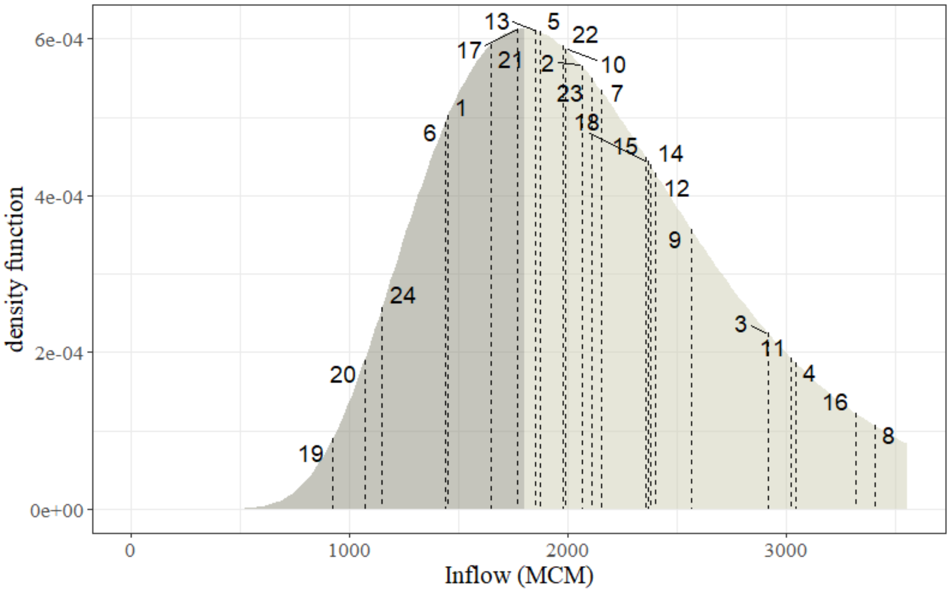


Figure 4.6 Historical inflow traces fitted log-normal distribution for weights of SSDP/Hist in Dam SY

Table 4.6 Weights of historical inflow traces for SSDP/Hist in Dam SY

j	Year	Total Inflow (MCM)	Probability	w_j
1	1996	2740.262	0.141	0.025
2	1997	4801.645	0.434	0.076
3	1998	7073.309	0.153	0.027
4	1999	6004.869	0.255	0.045
5	2000	4457.013	0.500	0.088
6	2001	2217.853	0.061	0.011
7	2002	6490.264	0.203	0.036
8	2003	8273.953	0.085	0.015
9	2004	6768.576	0.177	0.031
10	2005	5531.790	0.316	0.055
11	2006	7718.257	0.112	0.020
12	2007	6691.002	0.184	0.032
13	2008	3045.128	0.200	0.035
14	2009	4038.329	0.586	0.103
15	2010	5328.235	0.346	0.061
16	2011	8929.272	0.062	0.011
17	2012	5049.912	0.391	0.069
18	2013	4566.692	0.478	0.084
19	2014	2317.540	0.074	0.013
20	2015	1751.987	0.020	0.003
21	2016	2898.965	0.171	0.030
22	2017	3429.478	0.281	0.049
23	2018	5074.974	0.387	0.068
24	2019	2354.596	0.079	0.014
Sum			5.697	1

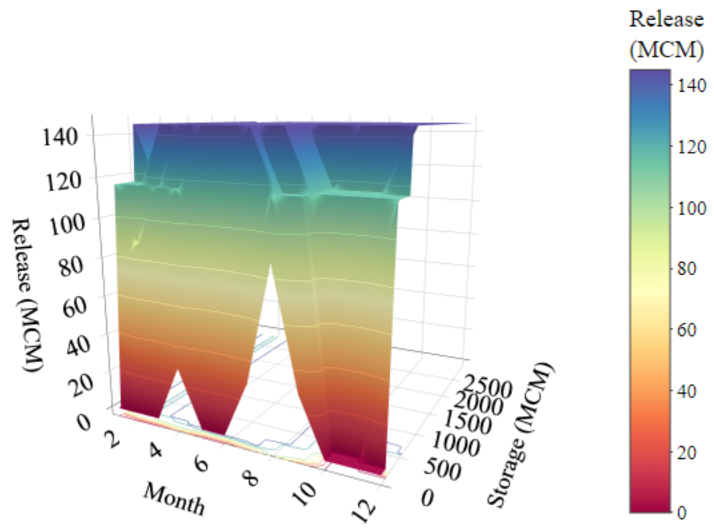


Figure 4.7 Release 3D policy of SSDP/Hist in Dam SY

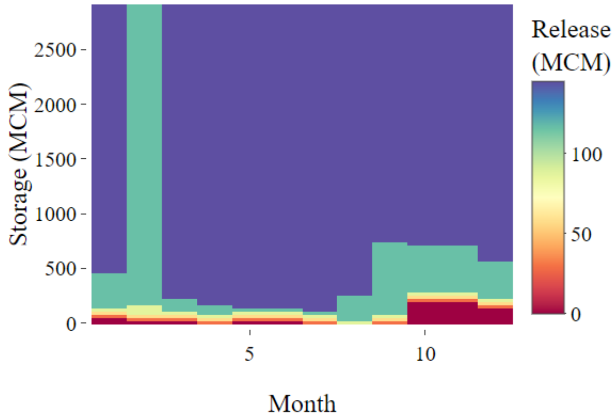


Figure 4.8 Release heatmap policy of SSDP/Hist in Dam SY

4.2.3 SSDP/ESP Model

For SSDP/ESP, the constraint conditions and objective function were the same as in that of SSDP/Hist. However, instead of using the historical inflow data, SSDP/ESP incorporated the future value function obtained from SSDP/Hist. Thus, the optimal release was calculated via the application of the two assumed distributions with three quantiles to the SSDP recursive equation. The resulting optimal release is referred to as Scenario W (based on Table 4.7) and Scenario P (based on Table 4.8).

$$\min_{R_t^*} \sum_{i=1}^{3} Pr(i) \left[O_t(\mathbf{S_t^k}, \mathbf{Q_t(i)}, \mathbf{R_t}) + E \left\{ f_{t+1}(\mathbf{S_{t+1}^l}, \mathbf{j}) \right\} \right] \tag{4.2}$$

In Eq. 4.2, i represents the ESP trace, and j represents the historical inflow data used in SSDP/Hist calculations. Therefore, there are three traces for i as the ESP distribution was discretized into three categories. The f_{t+1} obtained from SSDP/Hist is the sum of the historical inflow data (24 years) multiplied by the transition probability. Further, the transition matrix E represents the conditional probability of the occurrence of $Q_{t+1}(j)$ considering the current inflow $Q_t(i)$, and it has the same probability for all ESP traces. This assumption implies that the transition probabilities are uniform regardless of the predicted inflow (ESP) at the current time step.

In Figure 4.9, the release policy derived from the optimal release calculation using the Scenario W inflow trace is visualized. Whereas, Figure 4.10 shows the release policy obtained when using the Scenario P inflow trace.

Table 4.7 Inflow distribution quantiles Scenario W in 2022 of Dam SY (MCM)

	Jan	Fed	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Q1	4.57	2.49	39.45	25.50	8.92	109.83	201.22	373.76	231.49	76.87	21.49	14.28
Q2	8.01	4.36	69.17	44.72	15.65	192.58	352.82	655.35	405.89	134.79	37.68	25.04
Q3	14.04	7.65	121.28	78.40	27.43	337.67	618.64	1149.08	711.69	236.33	66.07	43.90

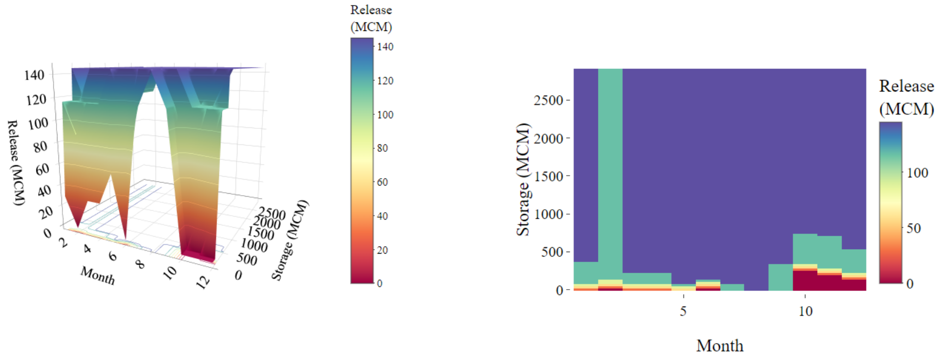
Table 4.8 Inflow distribution quantiles Scenario P in 2022 of Dam SY (MCM)

	Jan	Fed	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Q1	0.26	0.14	2.29	1.48	0.52	6.36	11.66	21.65	13.41	4.45	1.25	0.83
Q2	1.13	0.61	9.73	6.29	2.20	27.10	49.65	92.22	57.12	18.97	5.30	3.52
Q3	4.80	2.62	41.45	26.80	9.38	115.41	211.45	392.75	243.25	80.78	22.58	15.00

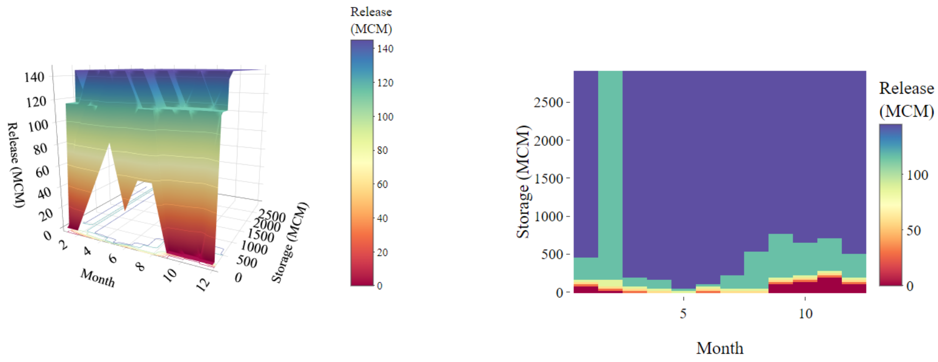
4.2.4 Results

The derived optimal release exhibited different patterns depending on Scenarios W and P. In Scenario W, a more aggressive release policy was obtained for all three years (2020, 2021, and 2022) compared to that in case of Scenario P. This indicated a higher level of releases even when the current storage was relatively low. The same trend was also observed in DDP/PERF, where the optimal release policy was more pronounced in its aggressiveness in Figure 4.11.

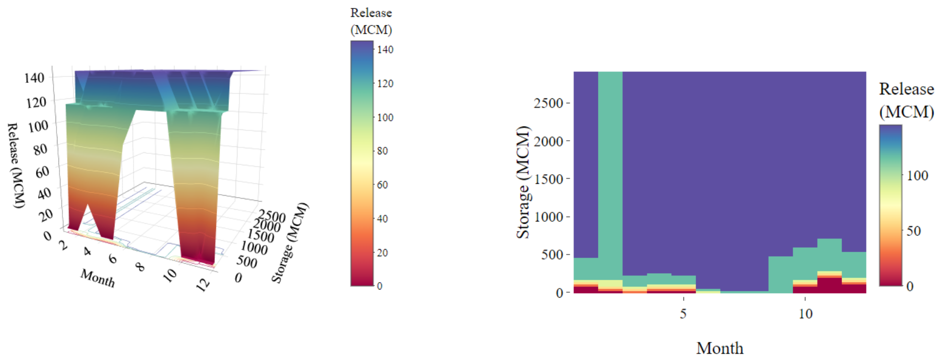
The simulation was conducted by varying the initial storage using the inflow data from 2020 to 2022. For Dam SY, the maximum, average and minimum historical storage at the end of December were 2159.658, 1616.219, and 1086.166 MCM, respectively. However, no significant difference in the optimal release policy between Scenarios W and P were observed in Dam SY. This can be attributed to Dam SY having a CIR of 1.345, thus, classifying it as a "Very Large" dam, but with relatively low monthly demand. Consequently, the difference in the optimal release policy between Scenarios W and P was below the threshold set by S_{min}. Therefore, despite a difference in the optimal release policy between Scenarios W and P from the perspective of release optimization, this difference was not observed in the simulation results for Dam SY.



(a) Release 3D and heatmap policy in 2020

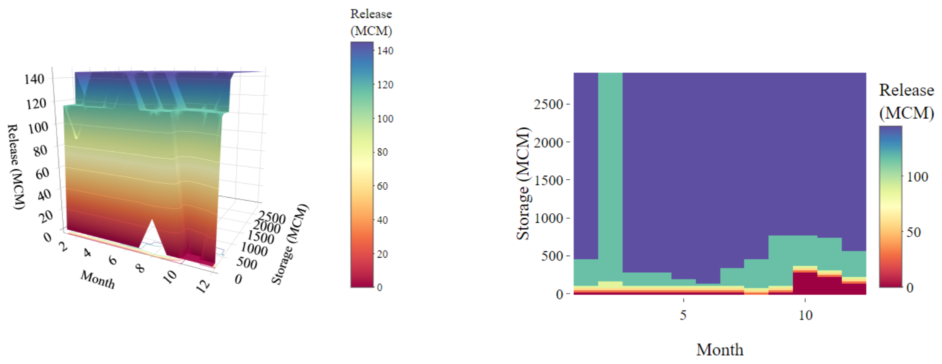


(b) Release 3D and heatmap policy in 2021

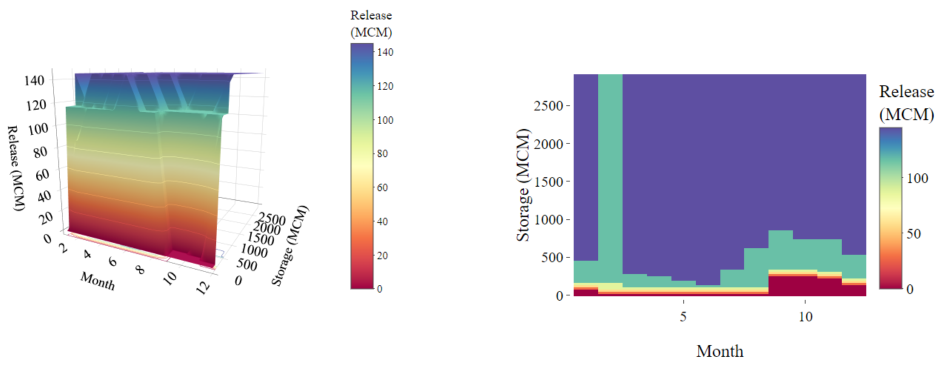


(c) Release 3D and heatmap policy in 2022

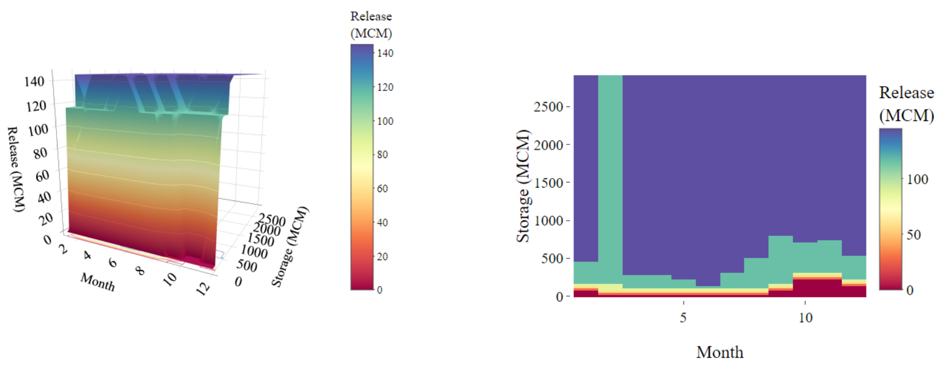
Figure 4.9 Release policy of SSDP/ESP Scenario W in Dam SY



(a) Release 3D and heatmap policy in 2020

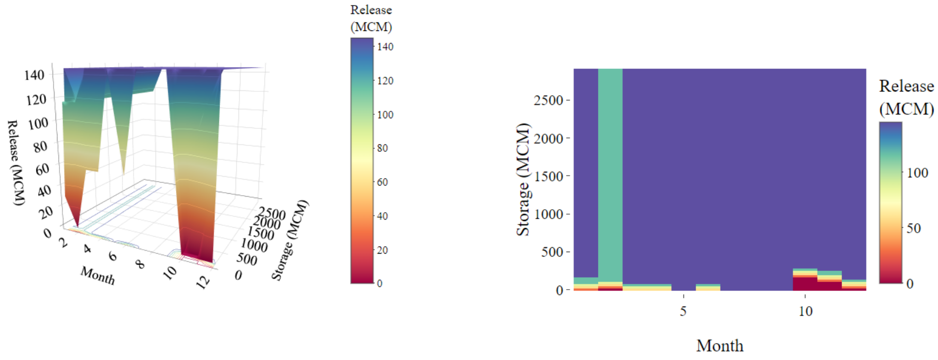


(b) Release 3D and heatmap policy in 2021

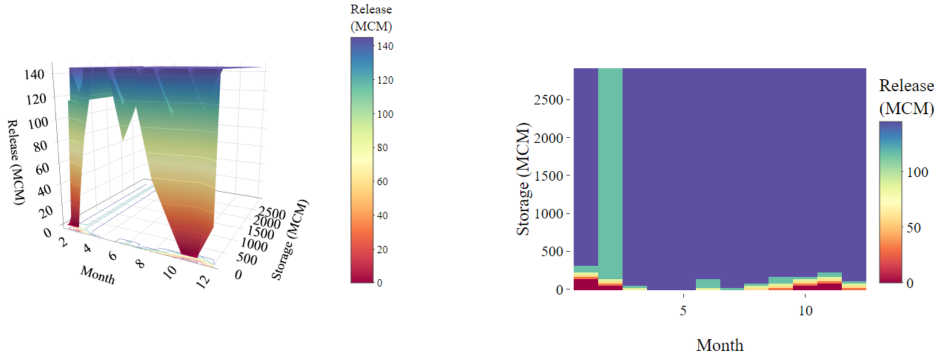


(c) Release 3D and heatmap policy in 2022

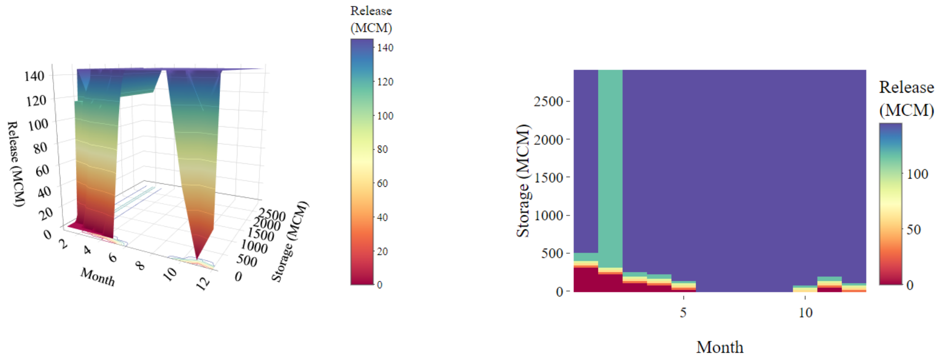
Figure 4.10 Release policy of SSDP/ESP Scenario P in Dam SY



(a) Release 3D and heatmap policy in 2020



(b) Release 3D and heatmap policy in 2021



(c) Release 3D and heatmap policy in 2022

Figure 4.11 Release policy of DDP/PERF in Dam SY

4.3 Sample Study 2: Dam CJ

To set up the constraints for SSDP according to the characteristics of Chungju dam's basin (Figure 4.12), the information on Chungju Dam is summarized in Table 4.9. The minimum and maximum storage levels corresponding to the LWL and NHWL of the Chungju dam were determined to be 454.027 and 2251.672 MCM, respectively. These values were used as constraints for the SDP. Therefore, the constraint values for the Chungju dam were $S_{min} = 454.027$ MCM and $S_{max} = 2251.672$ MCM. The storage was discretized into 100 equally spaced intervals. The monthly demands to be used in the optimization are summarized in Table 4.10.

Table 4.9 Information of Dam CJ

Information	Figure
Height (m)	97.5
Length (m)	447.0
Normal elevation (EL.m)	147.5
Volume ($1000m^3$)	902.0
Basin area (km^2)	6648.0
Annual water supply capacity (MCM)	3380.0
Reservoir area (km^2)	97.0
Design Flood Level (EL.m)	145.00
NHWL (Normal High Water Level) (EL.m)	141.00
RWL (Restricted Water Level) (EL.m)	138.00
Spill Water Level (EL.m)	126.00
LWL (Low Water Level) (EL.m)	110.00
Total Storage (MCM)	2750.0
Available Storage (MCM)	1786.0
Flood Control Storage Capacity (MCM)	616.0

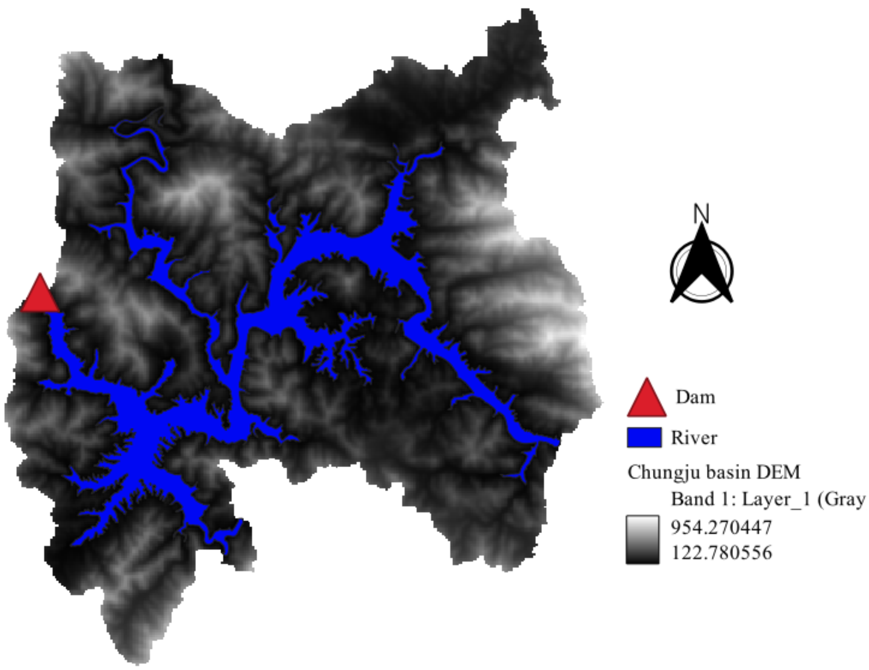


Figure 4.12 Dam CJ basin

Table 4.10 Monthly demand of Dam CJ (MCM) (K-water, 2020)

	January	February	March	April	May	June
Demand	260.338	235.144	260.338	275.520	318.742	324.510

	July	August	September	October	November	December
Demand	308.543	308.543	279.420	281.759	251.940	260.338

4.3.1 Ensemble Streamflow Prediction

The goodness-of-fit test was performed to assess the suitability of the log-normal distribution for fitting the historical inflow data of the Chungju dam. The test results indicated that assuming the log-normal distribution for the Chungju dam's ESP was appropriate. Using the monthly inflow data from 2020 to 2022, The ESPs for Scenarios W and P were established. Table 4.11 lists the parameters for Scenario W in 2022, and Table 4.12 lists the parameters for Scenario P. Similar to the Soyanggang dam, the ESP distributions were discretized into three quantiles: Q1, Q2, and Q3 (Table 4.14, 4.15).

For the historical inflow traces used in SSDP/Hist, the average inflow was 4898.079 MCM from 1996 to 2019. In contrast, for SSDP/ESP, the average annual inflow from 2020 to 2022 was 4949.233 MCM. This indicates that the inflow used in SSDP/ESP was approximately the same as the average inflow in SSDP/Hist, representing a typical inflow amount. Upon examining each year, the inflow in 2020 was 1.3 times the historical inflow, in 2021 it was 0.7 times, and in 2022 it was 1.1 times. This indicates a pattern similar to that in case of the Soyanggang dam, which is located in the same basin. Figure 4.13 shows the inflow traces used in SSDP/Hist as a box plot, and the inflow used in SSDP/ESP as a line graph. As evident, larger inflows occurred during the flood seasons (July to September) of 2020 and 2022 compared to the previous historical inflows. This can be attributed to the impact of typhoons

BAVI, MAYSAK, and HAISHEN in 2020, and HINNAMNOR in September 2022, which affected the Han river basin.

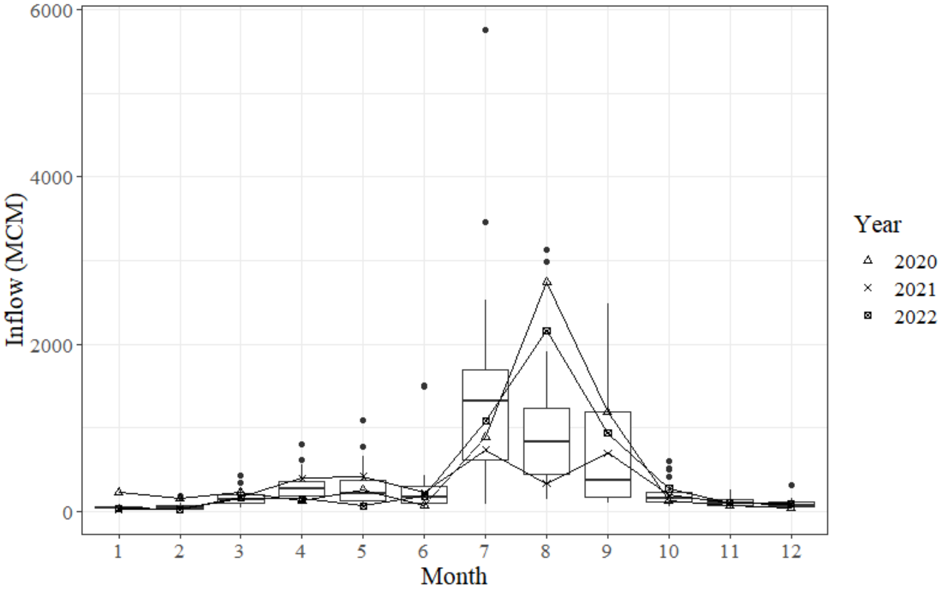


Figure 4.13 Comparison of inflow data used for SSDP/Hist and SSDP/ESP in Dam CJ

Table 4.11 Inflow distribution parameters of Scenario W in 2022 of Dam CJ (MCM)

	Jan	Fed	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
μ_X	40.42	21.91	163.65	152.30	68.29	186.75	1082.26	2164.06	939.41	271.49	103.66	97.87
σ_X	40.42	21.91	163.65	152.30	68.29	186.75	1082.26	2164.06	939.41	271.49	103.66	97.87
μ_Y	3.35	2.74	4.75	4.68	3.88	4.88	6.64	7.33	6.50	5.26	4.29	4.24
σ_Y	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83

Table 4.12 Inflow distribution parameters of Scenario P in 2022 of Dam CJ (MCM)

	Jan	Fed	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
μ_X	40.42	21.91	163.65	152.30	68.29	186.75	1082.26	2164.06	939.41	271.49	103.66	97.87
σ_X	404.20	219.09	1636.52	1522.99	682.88	1867.47	10822.64	21640.58	9394.05	2714.86	1036.61	978.73
μ_Y	1.39	0.78	2.79	2.72	1.92	2.92	4.68	5.37	4.54	3.30	2.33	2.28
σ_Y	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15

4.3.2 SSDP/Hist Model

The Chungju dam has data available for the period from 1996 to 2022 (27 years). However, for the application of SSDP/Hist, the inflow data for the year 2022, which was used for ESP assumption, was excluded. Therefore, the data from 1996 to 2019 (24 years) was utilized for SSDP/Hist calculations (Figure 4.14). The total inflow for each trace was scaled using the log-normal distribution by considering the exceedance and non-exceedance probabilities. Figure 4.15 visualizes the log-normal distribution-fitted inflow traces, and Table 4.13 presents the corresponding weights used in the analysis.

$$\min_{R_t^*} \sum_{j=1}^{24} w_j \left[O_t(\mathbf{S_t^k}, \mathbf{Q_t(j)}, \mathbf{R_t}) + \left\{ f_{t+1}(\mathbf{S_{t+1}^l}, \mathbf{j}) \right\} \right] \tag{4.3}$$

Equation 4.3 represents the formulation used to derive the optimal release policy by incorporating the weights. In Figure 4.14, the variable j corresponds to the 24-year inflow traces from the past, whereas w_j denotes the weights calculated as the probability of occurrence for each trace. The resulting optimal release policy, obtained through these calculations, is shown in Figure 4.16 as a 3D plot and in Figure 4.17 as a heatmap.

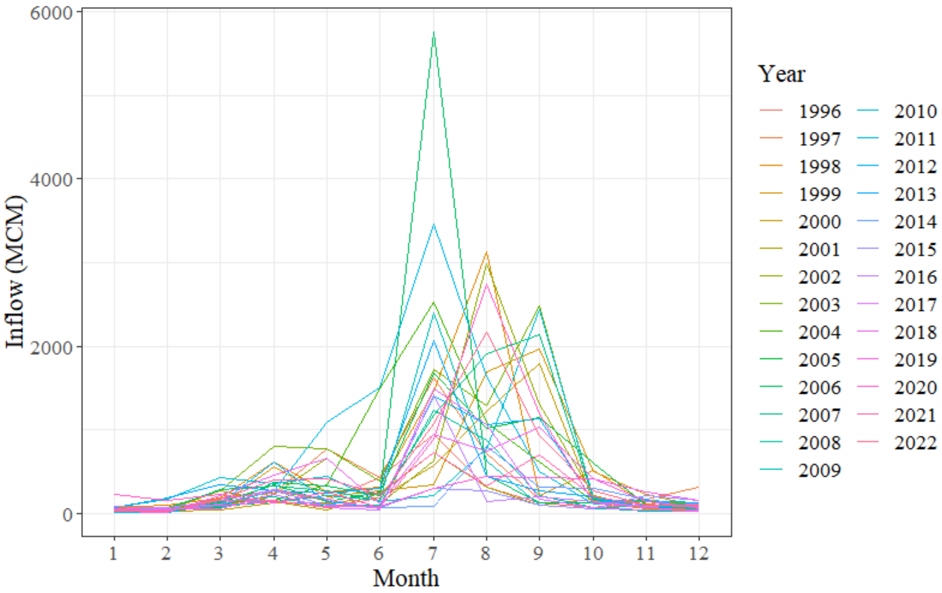


Figure 4.14 Historical inflow ensemble traces for SSDP/Hist in Dam CJ

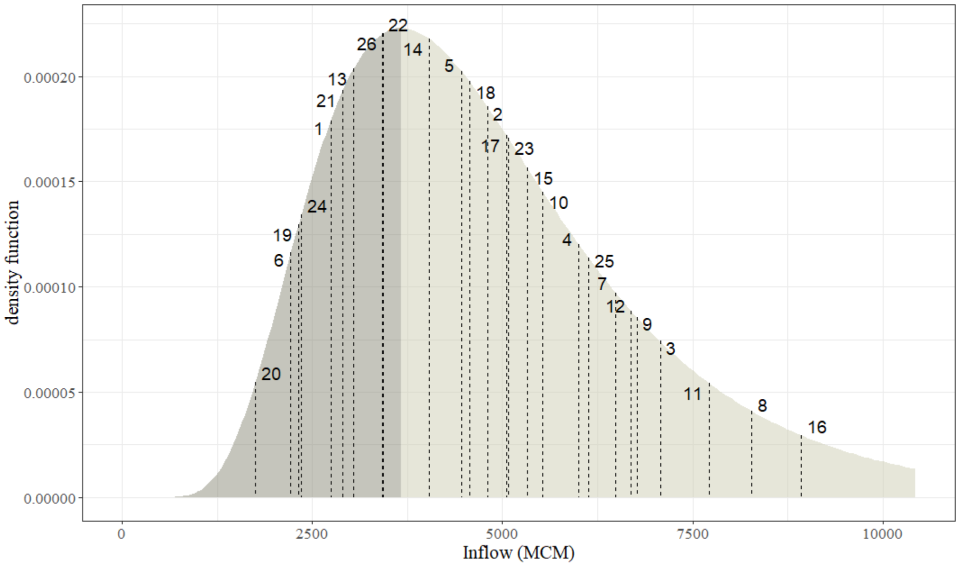


Figure 4.15 Historical inflow traces fitted log-normal distribution for weights of SSDP/Hist in Dam CJ

Table 4.13 Weights of historical inflow traces for SSDP/Hist in Dam CJ

j	Year	Total Inflow (MCM)	Probability	w_j
1	1996	2740.262	0.141	0.025
2	1997	4801.645	0.434	0.076
3	1998	7073.309	0.153	0.027
4	1999	6004.869	0.255	0.045
5	2000	4457.013	0.500	0.088
6	2001	2217.853	0.061	0.011
7	2002	6490.264	0.203	0.036
8	2003	8273.953	0.085	0.015
9	2004	6768.576	0.177	0.031
10	2005	5531.790	0.316	0.055
11	2006	7718.257	0.112	0.020
12	2007	6691.002	0.184	0.032
13	2008	3045.128	0.200	0.035
14	2009	4038.329	0.586	0.103
15	2010	5328.235	0.346	0.061
16	2011	8929.272	0.062	0.011
17	2012	5049.912	0.391	0.069
18	2013	4566.692	0.478	0.084
19	2014	2317.540	0.074	0.013
20	2015	1751.987	0.020	0.003
21	2016	2898.965	0.171	0.030
22	2017	3429.478	0.281	0.049
23	2018	5074.974	0.387	0.068
24	2019	2354.596	0.079	0.014
Sum			5.697	1

4.3.3 SSDP/ESP Model

The constraints and objective function were the same, and SSDP/ESP utilized the future value function from SSDP/Hist for its calculations. In the ESP approach, that is, the inflow distribution fitted with a log-normal distribution, provided the Q1, Q2, and Q3 values. These values were then substituted into Eq. 4.4 as $Q_t(i)$ to determine the optimal release policy. The resulting optimal release policy obtained from the calculations using Table 4.14 is referred to as Scenario W, whereas that derived from Table 4.15 is referred to as Scenario P.

$$\min_{R_t^*} \sum_{i=1}^{3} Pr(i) \left[O_t(\mathbf{S_t^k}, \mathbf{Q_t(i)}, \mathbf{R_t}) + E\left\{ f_{t+1}(\mathbf{S^l_{t+1}}, \mathbf{j}) \right\} \right] \tag{4.4}$$

In Eq. 4.4, i is the ESP trace and j is the historical inflow data used for SSDP/Hist calculation. Therefore, as i discretized the ESP distribution in 3, there were traces up to 3, and f_{t+1} imported from SSDP/Hist multiplied 24 traces by the transition probability and summed them up. Further, E is a transition matrix with a transition probability equal to the number of f_{t+1} traces. When the current inflow $Q_t(i)$ occurred, $Q_{t+1}(j)$ at the next instance is defined as the conditional probability that would occur. However, in this study, it was assumed to have the same probability regardless of the predicted inflow (ESP) at the current time.

The calculated results are visualized as follows. Figure 4.18 shows the release

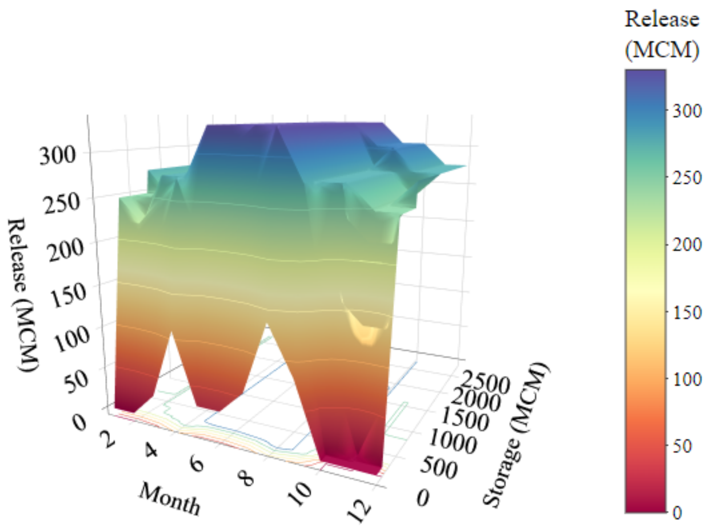


Figure 4.16 Release 3D policy of SSDP/Hist in Dam CJ

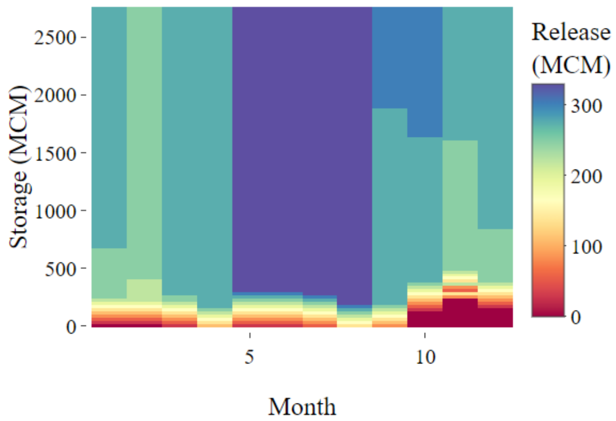


Figure 4.17 Release heatmap policy of SSDP/Hist in Dam CJ

policy derived from the Scenario W inflow trace, covering the period from 2020 to 2022. Figure 4.19 shows the release policy obtained from the Scenario P inflow trace. In addition, for comparison with Scenario W/P, DDP/PERF model assuming a perfect forecast was also calculated and visualized in Figure 4.20.

Table 4.14 Inflow distribution quantiles Scenario W in 2022 of Dam CJ (MCM)

	Jan	Fed	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Q1	16.30	8.84	66.00	61.42	27.54	75.31	436.46	872.72	378.84	109.48	41.80	39.47
Q2	28.58	15.49	115.72	107.69	48.29	132.05	765.28	1530.22	664.26	191.97	73.30	69.21
Q3	50.11	27.16	202.90	188.82	84.67	231.53	1341.83	2683.07	1164.70	336.60	128.52	121.35

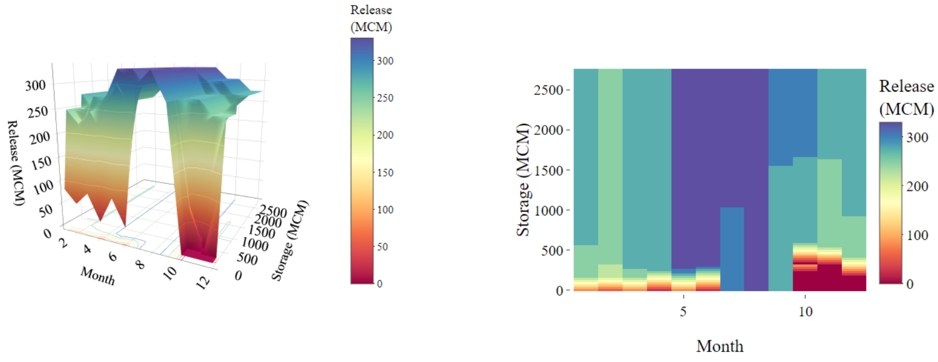
Table 4.15 Inflow distribution quantiles Scenario P in 2022 of Dam CJ (MCM)

	Jan	Fed	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Q1	0.94	0.51	3.82	3.56	1.60	4.36	25.29	50.56	21.95	6.34	2.42	2.29
Q2	4.02	2.18	16.28	15.15	6.79	18.58	107.69	215.33	93.47	27.01	10.31	9.74
Q3	17.13	9.28	69.35	64.54	28.94	79.14	458.63	917.06	398.09	115.05	43.93	41.48

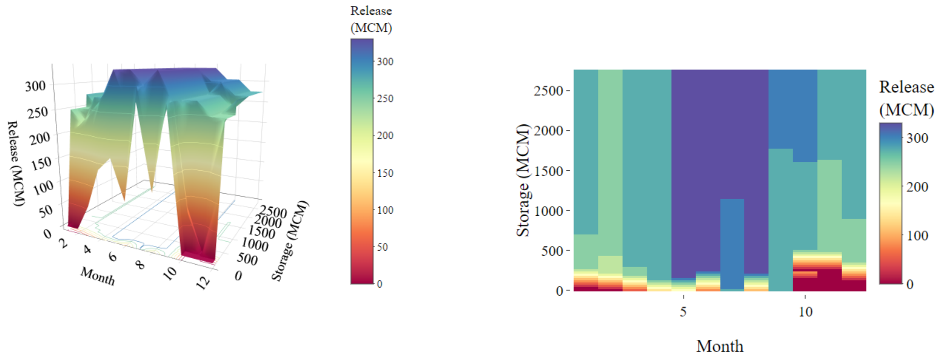
4.3.4 Results

To perform the simulation with different initial storages, the maximum, average, and minimum storage levels at the end of December were used for each respective scenario. For the Chungju dam, these values were 2153.748, 1538.924, and 892.698 MCM, respectively. The results of the simulation using the maximum, average, and minimum historical storage as the initial storage are presented in Tables 4.18, 4.17, and 4.16, respectively. Each table provides information on total penalty, total releases, frequency, duration, and magnitude for both the ESP Scenario W/P and PERF. And it was visualized using bar charts in Figure 4.21.

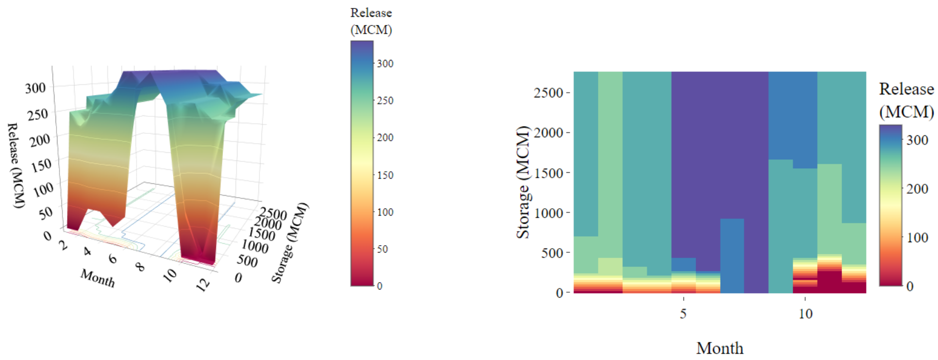
In all three initial storage scenarios, the year 2021 exhibited a lower inflow, resulting in differences in the optimal releases between Scenarios W and P. Scenario W showed a decrease in total penalty and an increase in total releases compared to Scenario P. In terms of frequency and magnitude, Scenario W performed better than Scenario P; however, Scenario P exhibited subtle better results in terms of duration. This indicates that a smaller standard deviation of the predicted inflow distribution results in overall improved performance in the optimal release strategy.



(a) Release 3D and heatmap policy in 2020

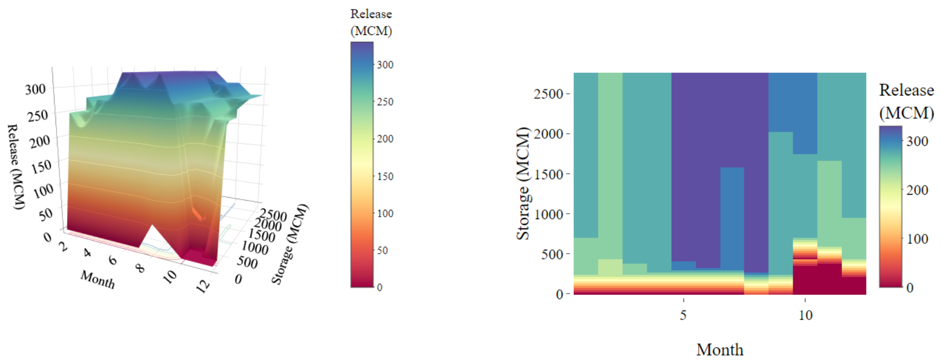


(b) Release 3D and heatmap policy in 2021

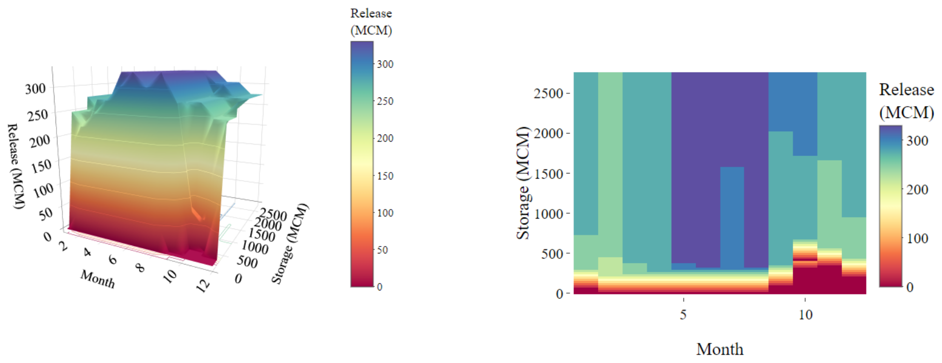


(c) Release 3D and heatmap policy in 2022

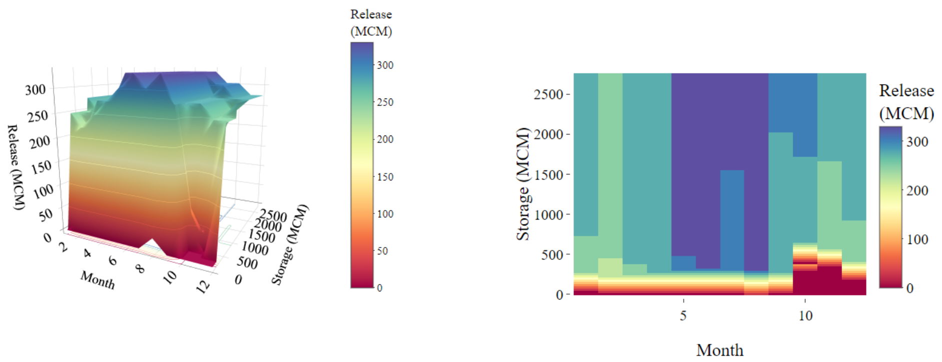
Figure 4.18 Release policy of SSDP/ESP Scenario W in Dam CJ



(a) Release 3D and heatmap policy in 2020

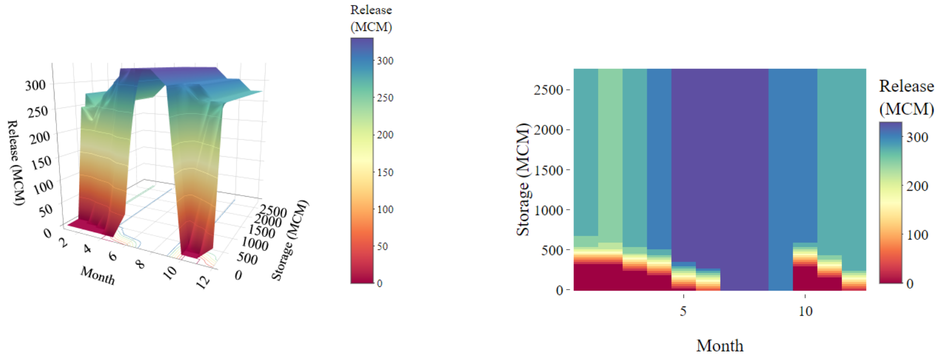


(b) Release 3D and heatmap policy in 2021

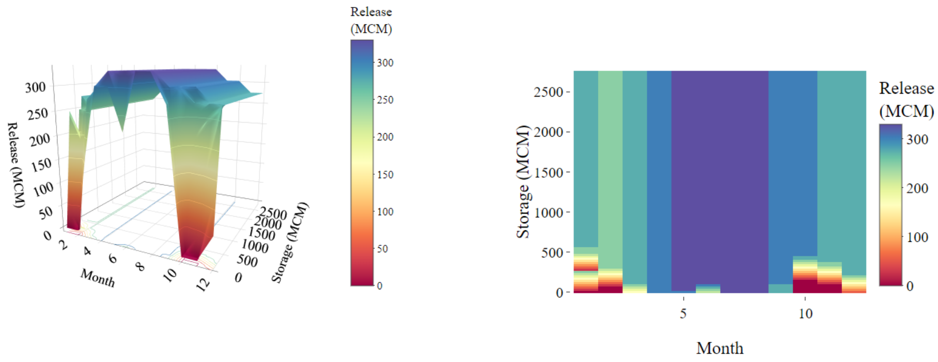


(c) Release 3D and heatmap policy in 2022

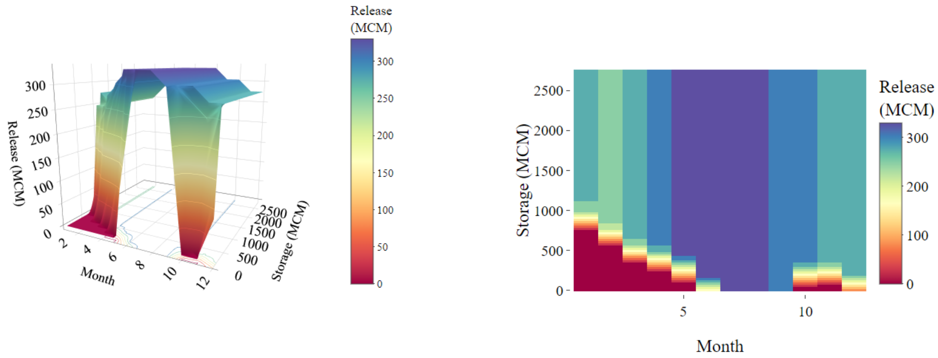
Figure 4.19 Release policy of SSDP/ESP Scenario P in Dam CJ



(a) Release 3D and heatmap policy in 2020



(b) Release 3D and heatmap policy in 2021



(c) Release 3D and heatmap policy in 2022

Figure 4.20 Release policy of DDP/PERF in Dam CJ

Table 4.16 Summary of simulation results corresponding minimum initial storage in Dam CJ

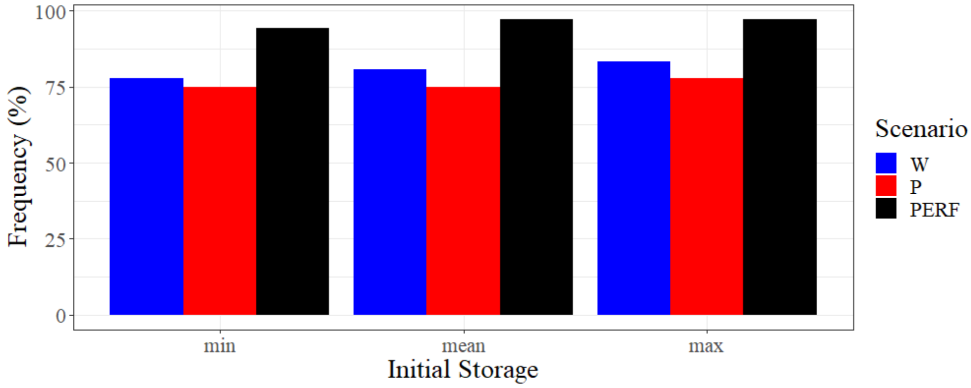
Scenario	Total Penalty (MCM)	Total Releases (MCM)	Frequency (%)	Duration (%)	Magnitude (%)
W	667.086	9900.0	77.8	13.9	93.4
P	673.129	9872.5	75.0	16.4	93.3
PERF	649.020	10065.0	94.4	44.4	93.6

Table 4.17 Summary of simulation results corresponding average initial storage in Dam CJ

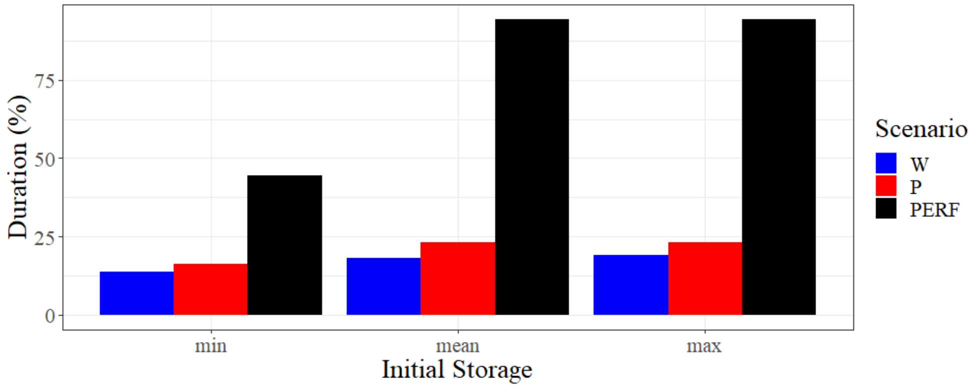
Scenario	Total Penalty (MCM)	Total Releases (MCM)	Frequency (%)	Duration (%)	Magnitude (%)
W	342.576	10230.0	80.6	18.3	96.6
P	348.619	10202.5	75.0	23.1	96.6
PERF	324.510	10395.0	97.2	94.4	96.8

Table 4.18 Summary of simulation results corresponding maximum initial storage in Dam CJ

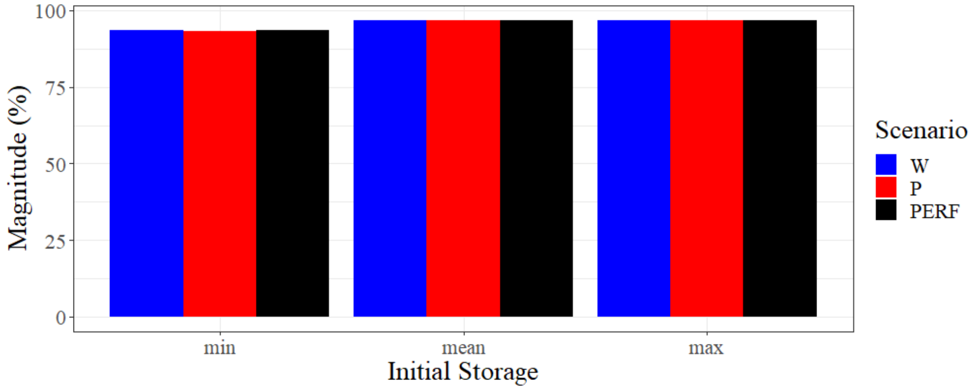
Scenario	Total Penalty (MCM)	Total Releases (MCM)	Frequency (%)	Duration (%)	Magnitude (%)
W	336.533	10257.5	83.3	19.2	96.7
P	348.619	10202.5	77.8	23.1	96.6
PERF	324.510	10395.0	97.2	94.4	96.8



(a) Frequency corresponding initial storage



(b) Duration corresponding initial storage



(c) Magnitude corresponding initial storage

Figure 4.21 Bar chart of FDM performance criteria results in Dam CJ

Chapter 5. Conclusion

5.1 Summary and Conclusion

This study investigated the importance of ESP in water resources management in South Korea. The specific research objectives were as follows:

(a) The research trends using ESP were investigated and the adoption of the method used in water resources management and its performance in application cases was demonstrated.

(b) The study aimed to make policy-makers easily understand the benefit of utilization of ESP in dam operations by quantifying the effect of ESP through simple examples.

(c) An optimal ESP-based dam operation model was established for two dams with different CIR to demonstrate the feasibility of applying ESP in real cases.

This study aimed to address the current limitations in water resources management in South Korea by introducing and demonstrating the effectiveness of ESP. Although ESP has been utilized for hydrological forecasting in South Korea, its application in water resources management remains limited. This study attempted to bridge this gap and thus contribute to the improvement of water resources management techniques.

To render the concept of ESP more accessible to dam operators, a simple statistical example was presented, highlighting the importance of considering the uncertainty of inflow when making operational decisions. Through demonstrations indicating that dams with different inflow standard deviations incurred different costs, even with the same mean inflow, the study emphasized the need to go beyond relying solely on the ensemble mean values in water resources management.

Multiple-purpose dams in the Han river basin, specifically the Soyanggang and Chungju dams, were selected as sample study sites based on their CIR. The inflow data from 27 years were divided into 9 sets, with the last set from 2020 to 2022 used for generating unbiased ensembles. The SSDP/Hist and SSDP/ESP models, developed by integrating ESP into the SSDP approach, were applied to optimize the release calculations for the dams. Consequently, the optimal releases were determined to be aggressive and conservative in Scenarios W and P, respectively, for both dams. The simulation was conducted using the obtained optimal releases for the last set, and the results were evaluated using performance metrics such as penalty, frequency, duration, and magnitude (Figure 5.1). The Soyanggang dam exhibited consistent performance across the scenarios; however, the Chungju dam demonstrated better overall performance in Scenario W, indicating the sensitivity of dams with lower CIR and higher demands to the uncertainty of inflow (Table 5.1). Thus, the findings underscore the importance of considering the full distribution of ensemble forecasts and reacting accordingly in water resources management.

In conclusion, this study effectively communicated the significance of considering uncertainty in inflow and showcased the benefits of integrating ESP into the decision-making process. Through a simple example and conduction of sample studies, the study provided practical insights for dam operators and demonstrated the potential of ESP in improving water resources management practices. This research is expected to contribute to the enhancement of water resources management techniques in South Korea, particularly in the face of climate change and other uncertainties.

Table 5.1 Summary of sample study results

	Dam SY	Dam CJ
Demand	Small	Large
CIR	Very Large (1.345)	Large (0.563)
Scenario W/P differences (in terms of optimal releases)	O	O
Scenario W/P differences (in terms of simulation)	X	O

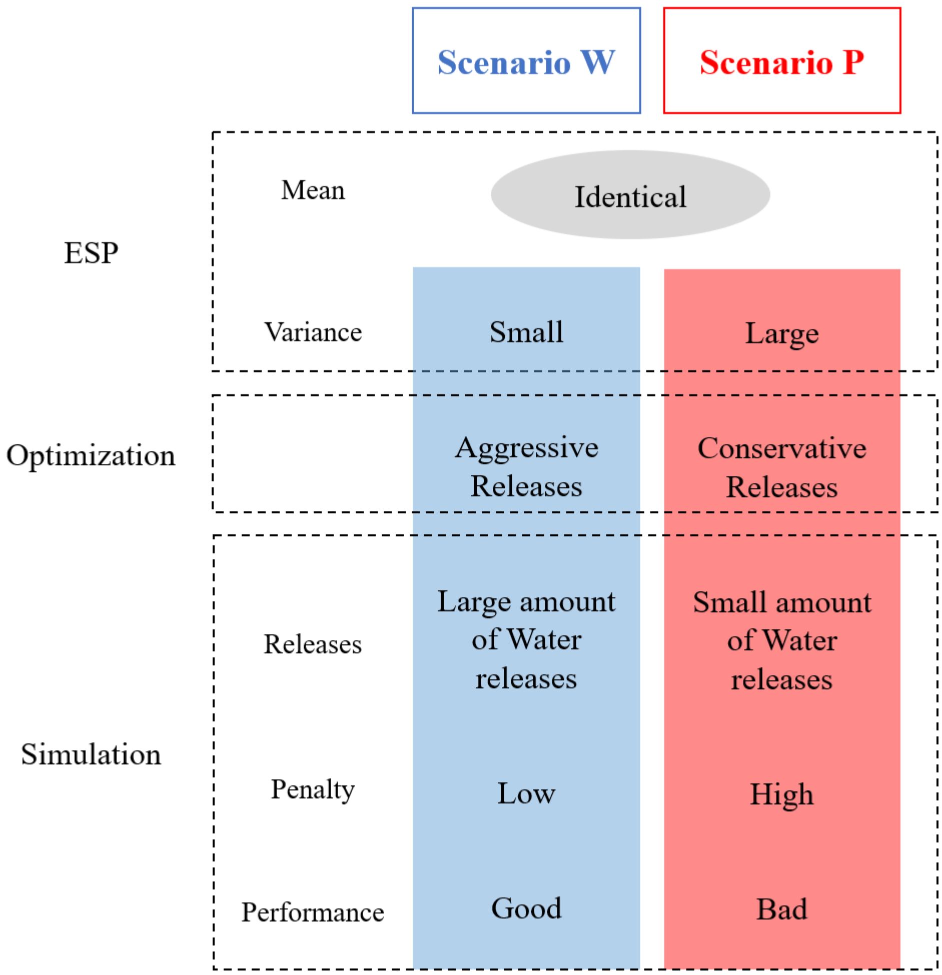


Figure 5.1 Summary of overall results

5.2 Limitations

In this study, we assumed that the inner expectation (transition matrix) for the calculation of optimal releases using SSDP was the same for all time steps. This implies that the probability of inflow occurring at time t+1 is equal for all predicted inflow values, regardless of their actual values. However, in real-world scenarios, the predicted inflow is influenced by current watershed information and meteorological conditions, which in turn affects the probability of inflow at the next time step.

Moreover, as the objective function was designed to focus solely on water supply, the differences in optimal releases between Scenarios W and Scenario P were observed mainly when the storage level was low. Thus, when considering water supply alone, the strategies for managing predicted inflow vary primarily in drought conditions.

Furthermore, the simulation was conducted only for the last set of data, which represented a period with relatively abundant inflow similar to the past. Consequently, the research findings did not encompass an analysis of extreme inflow events.

Therefore, future research must consider the uncertainty of predicted inflow and the influence of watershed characteristics and meteorological conditions. Furthermore, the impact of extreme inflow events on optimal release strategies must be analyzed.

5.3 Future Study

As mentioned in the Limitations section, future research can address the following aspects to enhance the accuracy and applicability of optimal release strategies:

(a) Improved Expectation Calculation: Rather than assuming equal probabilities for all predicted inflow values, future studies can accurately estimate the expectations based on the predicted inflow data. This would provide a more realistic representation of the actual situation and result in a more accurate determination of optimal release quantities.

(b) Cross-Validation using all Sets: In this study, only the last set of data from 2020 to 2022 was used for SSDP/ESP and simulation. Future research can perform cross-validation by utilizing all sets of data. This would result in the validation of the robustness of the optimal release strategies across different time periods.

(c) Consideration of Multiple Scenarios: Rather than focusing on only two scenarios (Scenarios W and P), future studies can explore a wider range of scenarios. The incorporation of various scenarios can provide a more comprehensive understanding of the system's response and optimize release strategies under different conditions.

(d) Integration of Multiple Dam Functions: While this study focused solely on water supply objectives, future research can consider the functions of dams related to irrigation and flood control. Through the incorporation of multiple dam functions into the optimization framework, more meaningful and holistic results can be obtained.

Thus, by addressing these aspects in future research, we can further enhance the effectiveness and practicality of optimal release strategies for reservoir operations.

References

Bellman, R. (1957). *Dynamic programming*. Princeton University Press.

Celeste, A. B. and Billib, M. (2009). "Evaluation of stochastic reservoir operation optimization models." *Advances in Water Resources*, 32(9), 1429–1443.

Connelly, B. A., Braatz, D. T., Halquist, J. B., DeWeese, M. M., Larson, L., and Ingram, J. J. (1999). "Advanced hydrologic prediction system." *Journal of Geophysical Research: Atmospheres*, 104(D16), 19655–19660.

Day, G. N. (1985). "Extended streamflow forecasting using NWSRFS." *Journal of Water Resources Planning and Management*, 111(2), 157–170.

Eum, H.-I. (2007). "Non-flood period operational policies for the geum river multireservoir system using sampling SDP with ESP." Ph.D. thesis, Seoul National University, Seoul, Republic of Korea.

Eum, H.-I., Ko, I.-H., and Kim, Y.-O. (2006). "Value of ensemble streamflow forecasts for reservoir operations during the drawdown period." *Journal of Korea Water Resources Association*, 39(3), 187–198.

Faber, B. A. and Stedinger, J. (2001). "Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts." *Journal of Hydrology*, 249(1-4), 113–133.

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., and Reed, P. M. (2016). "Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations." *Journal of Water Resources Planning and Management*, 2, 04015050.

Han River Flood Control Office (2009). "Probabilistic streamflow prediction using ensemble model." *Report no.*, Ministry of Land, Transport and Maritime Affairs.

Han River Flood Control Office (2022). "Water resources outlook, <http://www.hrfco.go.kr/web/river/periodical_2.do>. Accessed: 2022-11-22.

Hashimoto, T., Stedinger, J. R., and Loucks, D. P. (1982). "Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation." *Water Resources Research*, 18(1), 14–20.

He, S., Guo, S., Zhang, J., Liu, Z., Cui, Z., Zhang, Y., and Zheng, Y. (2022). "Multi-objective operation of cascade reservoirs based on short-term ensemble streamflow prediction." *Journal of Hydrology*, 127936.

K-water (2004). *A study on river runoff prediction methods for watershed integrated water management*. Ministry Of Construction & Transportation.

K-water (2020). *Dam drought step-by-step action plan*. K-water.

K-water (2021). “Sustainability report.” *Report No. 2021-MA-GP-25-1413*, K-water.

Karamouz, M. and Houck, M. H. (1987). “Comparison of stochastic and deterministic dynamic programming for reservoir operating rule generation 1.” *JAWRA Journal of the American Water Resources Association*, 23(1), 1–9.

Kim, G. J. (2022). “Improving on-site reservoir operation policies by robust and adaptive optimization models.” Ph.D. thesis, Seoul National University, Seoul, Republic of Korea.

Kim, G. J. and Kim, Y.-O. (2021). “How does the coupling of real-world policies with optimization models expand the practicality of solutions in reservoir operation problems?.” *Water Resources Management*, 35(10), 3121–3137.

Kim, G. J., Kim, Y.-O., and Reed, P. M. (2021). “Improving the robustness of reservoir operations with stochastic dynamic programming.” *Journal of Water Resources Planning and Management*, 147(7), 04021030.

Kim, G. J., Seo, S. B., and Kim, Y.-O. (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, G. J., Seo, S. B., and Kim, Y.-O. (2022). “Adaptive reservoir management by reforming the zone-based hedging rules against multi-year droughts.” *Water Resources Management*, 36(10), 3575–3590.

Kim, H., Lee, D., Kim, D., Park, H., Lee, J., Jung, G., Choi, Y., Hwang, S., Kang, S., Kim, C., Choi, S., Han, M., Jang, C., Park, J., Lee, J., and Kim, Y.-O. (2011). "Improvement of flood prediction system using stochastic method." *Report No. 11-1611492-000082-01*, Han River Flood Control Office.

Kim, Y.-O., Eum, H.-I., Lee, E.-G., and Ko, I. H. (2007). "Optimizing operational policies of a korean multireservoir system using sampling stochastic dynamic programming with ensemble streamflow prediction." *Journal of Water Resources Planning and Management*, 133(1), 4–14.

Kim, Y.-O., Jeong, D.-I., and Kim, H.-S. (2001). "Improving water supply outlook in korea with ensemble streamflow prediction." *Water International*, 26(4), 563–568.

Loucks, D. P. and Beek, E. V. (2017). *Water resource systems planning and management: An introduction to methods, models, and applications*. Springer.

National Drought Information Analysis Center (2017). "Drought information analysis annual report." *Report No. 2018-SA-MR1-24*, Ministry of Environment.

National Drought Information Analysis Center (2018). "Drought information analysis annual report." *Report No. 2018-SA-MR-3-1674*, Ministry of Environment.

National Drought Information Analysis Center (2021). "Drought information analysis annual report." *Report No. 2021-WR-AR-340-1752*, Ministry of Environment.

Ramaswamy, V. and Saleh, F. (2020). "Ensemble based forecasting and optimization framework to optimize releases from water supply reservoirs for flood control." *Water Resources Management*, 34(3), 989–1004.

Riverside Technology (1997). *Extended Streamflow Prediction Analysis and Display Program - Users Manual*. National Weather Service.

Shaikh, S. A. and Pattanayek, T. (2022). "Implicit stochastic optimization for deriving operating rules for a multi-purpose multi-reservoir system." *Sustainable Water Resources Management*, 8(5), 1–13.

Stedinger, J. R., Sule, B. F., and Loucks, D. P. (1984). "Stochastic dynamic programming models for reservoir operation optimization." *Water Resources Research*, 20(11), 1499–1505.

Tejada-Guibert, J. A., Johnson, S. A., and Stedinger, J. R. (1995). "The value of hydrologic information in stochastic dynamic programming models of a multireservoir system." *Water Resources Research*, 31(10), 2571–2579.

US Department of Commerce and NOAA (2020). "OHD/HL - NWSRFS: ESP Docs." *National Weather Service*, <https://www.weather.gov/owp/oh_hrl_general_indexdoc>. Accessed: 2022-11-19.

WMO (1975). "Intercomparison of conceptual models used in operational hydrological forecasting." *Report No. 429*, WMO.

WMO (1986). "Intercomparison of models of snowmelt runoff." *Report No. 646*, WMO.

WMO (1990). "Simulated real-time intercomparison of hydrological models." *Report No. 779*, WMO.

국문초록

앙상블 유량 예측은 미국 서부에서 하천 유량의 불확실성을 고려하여 수자원 관리를 하기 위해 고안되었다. 이 후 미국 전역으로 앙상블 유량 예측 시스템이 구축되었으며 현재까지도 활발히 수문예측과 수자원 관리 분야에 사용되고 있다. 반면 대한민국의 경우에는 21세기에 들어서야 수문예측에 앙상블 유량 예측이 사용되게 되었으며, 뒤이은 연구에 따라 앙상블 유량 예측이 효과적인 방법임이 입증되었다.

하지만 아직 대한민국에서는 앙상블 유량 예측이 수문예측까지만 사용이 되고 있으며, 초기에 고안되었던 수자원 관리에는 여전히 제한적으로 사용중이다. 현재 수자원 관리에서는 어렵게 앙상블을 구득하였음에도 불구하고 계산의 용이상 평균 등 대푯값 하나만을 가지고 의사결정을 하고 있는 실정이다. 본 연구는 앙상블 유량 예측을 수자원 관리에 사용하는 것을 권고하고 이에 대한 효과를 보이는 것을 목적으로 삼았다. 이를 위해 먼저 댐 운영 실무자들을 설득할 수 있는 간단한 통계학적 예제를 만들었다. 예제는 같은 용량과 방류요구량을 가지는 댐에서 평균은 같으나 다른 표준편차를 가지는 유입량이 발생할때 유입량의 표준편차가 큰 댐에서 더 많은 비용이 발생함을 보여주었다.

그리고 같은 예제를 실제 대한민국의 댐에 적용해보았다. 대한민국의 다

목적댐을 Capacity-Inflow Ratio (CIR)에 따라 분류하였고 이 중 같은 한강유역이면서 자료길이가 동일하지만 CIR는 상이한 소양강댐(CIR = 1.345)과 충주댐(CIR = 0.563)을 사례연구 대상지로 선정하였다. 그리고 각 댐의 유입량 자료를 9개의 세트로 나누어 마지막 세트인 2020년부터 2022년까지의 유입량을 이용하여 표준편차에 따라 편향되지 않은 두가지 앙상블을 가정하였다. 그리고 이를 각각 Well forecasted scenario (Scenario W), Poorly forecasted scenario (Scenario P)라 명명하였다. 대한민국의 다목적댐의 가장 중요한 기능 중 하나인 용수공급을 개선하기 위해 목적함수를 용수부족량을 최소화하는 것으로 설정하였다. 그리고 이를 앙상블 유량 예측을 활용하여 최적 방류량을 도출하기 위해 널리 사용되는 표본 추계학적 동적계획법(Sampling Stochastic Dynamic Programming, SSDP)으로 최적화하였다. 과거 유입량 자료를 사용한 SSDP/Hist 모형에서 다음 시점의 잔여최적편익함수를 가져와 전진방향으로 계산하는 SSDP/ESP 모형을 구축하였다. 그 결과, 소양강댐과 충주댐 모두 Scenario W가 Scenario P보다 공격적인 운영룰을 도출하였다.

도출된 최적 방류량을 실제 2020년부터 2022년까지의 유입량으로 모의 운영하여 패널티, 성공 빈도수(Frequency), 성공이 지속된 평균 기간(Duration) 그리고 용수부족규모의 여집합(Magnitude)으로 평가하였다.모의 결과 소양강댐에서는 Scenario W/P간의 차이가 나타나지 않았지만 충주댐에서 Scenario W가 Scenario P보다 전반적으로 성능이 높은 것을 확인하였다. 이는 평균은 같아도 표준편차가 다른 유입량이 들어오면 그에 따른 최적 방류량도 달라지며 운영시에 취해야 하는 전략도 달라져야함을 의미한다. 그리고 충주댐에서 이러한 차이가 두드러진

사실은 CIR이 작고 방류요구량이 많은 댐일수록 예측유입량의 불확실성에 더욱 예민하게 반응한다는 것을 시사한다.

결론적으로 댐 운영에서 유입량의 평균이 같아도 표준편차가 증가할수록 최적방류량과 모의평가에서 부정적인 영향이 나타남을 보였다. 이는 현재 평균값으로만 수자원 관리를 하는 방법이 앞으로 다가올 기후변화와 그 밖의 불확실성에 대한 고려를 하지않은 관리 방법임을 시사하며 현재 대한민국의 수자원관리 방법에 개선될 여지가 있다는 것을 나타낸다.

주요어: 앙상블 유량 예측 , 최적방류량, 표본 추계학적 동적계획법

학번: 2021-24690