

# Impact analysis of stochastic inflow prediction with reliability and discrimination indices on long-term reservoir operation

Daisuke Nohara and Tomoharu Hori

## ABSTRACT

Long-term stochastic inflow predictions can potentially improve decision making for reservoir operations. However, they are still not widely incorporated into actual reservoir management. One of the reasons may be that impacts of various types of uncertainty contained in stochastic inflow predictions have not been sufficiently clarified, thus enabling reservoir managers to recognize the advantages of their use. Impacts of uncertainties of stochastic inflow prediction on long-term reservoir operation for drought management are therefore investigated in order to analyze the kind of uncertainty that most affects improvements in the performance of reservoir operations. Two indices, namely reliability and discrimination, are introduced here to represent two major attributes of a stochastic prediction's uncertainty. Monte Carlo simulations of reservoir operations for water supply are conducted, coupling with optimization process of reservoir operations by stochastic dynamic programming (SDP) considering long-term stochastic inflow predictions, which are artificially generated with arbitrary uncertainties controlled by changing the two uncertainty indices. A case study was conducted using a simplified reservoir basin of which data were derived from the Sameura Reservoir basin in Japan with finer discretization settings for SDP. The results demonstrated the additional implication of the effect of stochastic inflow prediction's uncertainty on the authors' previous work.

**Key words** | discrimination, Monte Carlo simulation, reliability, reservoir operation, stochastic inflow prediction, uncertainty

Daisuke Nohara (corresponding author)  
Tomoharu Hori  
Disaster Prevention Research Institute,  
Kyoto University,  
Gokasho,  
Uji, 6110011,  
Japan  
E-mail: nohara.daisuke.2v@kyoto-u.ac.jp

## INTRODUCTION

Various types of real-time information on meteorological or hydrological prediction are provided by meteorological agencies. They have widely varying methods or horizons of prediction, ranging from deterministic to stochastic approaches for employed prediction methods, and from short-range to long-range for forecast horizons. The stochastic approach is often employed for long-range predictions in which it is difficult to estimate future conditions deterministically due to an increase in the uncertainty. With this approach, not a single expected state but multiple possible states in the future are predicted using the estimated occurrence probabilities of the states. Thus, long-range stochastic

predictions on hydrological states such as inflow or precipitation inherently contain rich information regarding expected hydrological states as well as the predictions' accuracies or uncertainties, both of which can be derived from the predicted probabilistic distributions.

The usefulness of stochastic predictions has been investigated and analyzed by providers of the predictions dealing with general situations where a variety of users refer to the predictions before making their decisions. Murphy (1977) analyzed the value of climatological, categorical and stochastic forecasts in terms of cost-loss ratio when the predictions were considered, and revealed that the

stochastic forecasts were more valuable than the climatological and categorical forecasts if the stochastic forecasts of concern were reliable. Some research groups have also investigated and evaluated the mid- or long-range stochastic hydrological predictions derived from ensemble prediction systems, and have demonstrated the effectiveness of the stochastic predictions compared with deterministic predictions (Richardson 2000; Jaun & Ahrens 2009). On the other hand, Krzysztofowicz (2001) pointed out four advantages of employing stochastic forecasting of hydrological variables. These were that stochastic forecasts are scientifically more honest, they enable risk-based warnings of floods, they enable rational decision making, and they offer additional economic benefits in light of his analysis.

While such information is presumed to merit consideration for the decision making of a water release strategy, especially in long-term reservoir operations in which long-range foresight is needed to determine the water release planning, long-range stochastic predictions have not been effectively utilized in the actual reservoir management. One reason for this situation may be that the complicated attributes of stochastic predictions still make it difficult for reservoir managers to quantitatively evaluate or utilize them in their decision making of water release from reservoirs. Reservoir managers have not been provided with sufficient information regarding how the attributes of stochastic predictions, such as the range of predicted probabilistic distribution or correspondence to the observed situation (which can be also considered as components of prediction uncertainty) influence the effect or the risk associated with the use of stochastic predictions in reservoir operations. This demonstrates the need for further investigations into the usefulness of stochastic hydrological predictions for decision-making processes in the reservoir operations when the predictions are utilized with their accuracies or uncertainties.

Various studies have been reported in order to analyze and clarify the advantages of introducing non-deterministically predicted information on future hydrological states of the target river basin for supporting decision making regarding the release of water in long-term reservoir operations (e.g. Datta & Houck 1984; Kelman *et al.* 1990). Karamouz & Vasiliadis (1992) developed the Bayesian stochastic optimization model for reservoir operation using continuously

updated uncertain forecasts to reduce the effect of forecast uncertainties in the model. Tejada-Guibert *et al.* (1995) analyzed changes in optimal benefits resulting from reservoirs' long-term operation optimized by stochastic dynamic programming when provided with forecast information on the hydrologic states in the target river basin and occurrence probabilities of streamflow categories conditioned by occurrences of the states. Faber & Stedinger (2001) introduced three optimization models of reservoir operation which respectively considered ensemble streamflow forecasts in deterministic, explicitly stochastic, and implicitly stochastic ways. They analyzed differences among the three models in averaged benefit to the reservoir operations. A similar tendency was also shown by Kim *et al.* (2007), although they also noted that the forecasting accuracy may considerably affect the results. Recently, several studies have investigated and analyzed in more detail the effects of non-deterministic hydrological predictions that focus on how uncertainties of the predictions influence the performance of reservoir operation (Georgakakos & Graham 2008; Gong *et al.* 2010; Pianosi & Ravazzani 2010; Block 2011; Zhao *et al.* 2011, 2012). However, the number of studies remains insufficient to clarify the diverse and complex impacts of non-deterministic hydrological predictions on reservoir operation, which would gain wide acceptance and usage by reservoir managers.

On the other hand, an identical hydrological state would generally not recur in nature. Because a hydrologic prediction is provided to estimate the true hydrological state in the future, an identical hydrological prediction would also generally not be provided again. These characteristics may lead to difficulties in acquiring prediction statistics because many samples of the same hydrological predictions which forecast the same hydrological states (predictands) are required to statistically evaluate impacts of hydrological predictions and their uncertainties. To overcome this situation, artificially generated hydrological predictions are often employed to analyze the impact of prediction on reservoir operations (e.g. Sivaarthiskul & Takeuchi 1995, 1996; Hori & Shiiba 1998; Georgakakos & Graham 2008; Block 2011; Zhao *et al.* 2011, 2012).

Because non-deterministic hydrological prediction involves more complicated information, including that on their uncertainties compared with deterministic prediction,

various efforts have been attempted to model uncertainty in the generation of non-deterministic predictions. Hori & Shiiba (1998) introduced fuzzy theory and type-2 fuzzy sets to represent the uncertainty of streamflow prediction and ambiguity of decision criteria, and analyzed their impact on decision making in reservoir operations aimed at drought control. Georgakakos & Graham (2008) analyzed the potential benefits of seasonal inflow prediction uncertainty for release decisions in an idealized multipurpose reservoir. They employed a beta distribution to model uncertainty in the seasonal inflow prediction, and analyzed the impact of uncertainty by arbitrarily changing the range of the beta distribution with the mean value and half-range value. Block (2011) created seasonal precipitation forecast distributions by adding normal random deviates which have means of zero and a standard deviation of the global predictive error to the deterministically forecasted prediction values. Zhao *et al.* (2011, 2012) assumed a forecast to be a stationary Gaussian distribution, and adjusted its uncertainty with its mean and variance. However, the number of investigations remains insufficient, while the approaches employed for the artificial generation of idealized non-deterministic predictions in these studies are respectively considered to be effective in the analysis of the impacts of non-deterministic hydrological predictions and their uncertainties. Further studies are needed to gain more insights into the impacts of uncertainties in stochastic hydrological predictions on reservoir operation with the introduction of a method for the generation of stochastic hydrological predictions for the holistic analysis of the impacts.

In the situation described above, the authors developed an impact analysis model of stochastic inflow prediction to realize improvements in the reservoir operation for a water supply by introducing an artificial generation method of stochastic inflow prediction (Nohara & Hori 2012). In this model, stochastic inflow predictions were assumed to be normally distributed, and were artificially and randomly generated with two uncertainty indices which represent important attributes of stochastic prediction. One of the indices is reliability, which is associated with the attribute in that it indicates the stability of the stochastic prediction's accuracy. The other is discrimination, which addresses the attribute regarding the range of the predicted probabilistic distribution. These two attributes are

considered to be important because they are mutually complementary concepts which consist of uncertainty contained in stochastic predictions. Impacts of the two indices on performances in the decision making of water release from a reservoir were then analyzed by conducting Monte Carlo simulations of long-term reservoir operations with stochastic inflow predictions that were repeatedly generated at random. The results demonstrated the possibility that the reliability of the stochastic prediction can be considered to be more important for long-term reservoir operation, while the discrimination of stochastic prediction is likely to be more significant under low-flow situations. Further investigations are, however, considered necessary in order to examine the results with different simulation parameters as several results did not agree with the tendency mentioned above.

An impact analysis of a stochastic inflow prediction's uncertainty on the efficiency of reservoir operations was therefore conducted using Monte Carlo simulations of long-term reservoir operation for drought management purposes with more precise simulation settings considering the reliability and discrimination indices of the prediction in this article. Long-term reservoir operation was optimized by stochastic dynamic programming (SDP) with finer discretized reservoir and hydrologic states compared with Nohara & Hori (2012). Differences and similarities in the results were then investigated and discussed in order to provide reservoir managers with information regarding which of the two uncertainty indices was to be more emphasized for stochastic inflow prediction in the case of reservoir operation for drought management. The sensitivity of the results to changes in the discretization grid during the optimization process was also assessed through the application in this study.

---

## OUTLINE OF SIMULATION MODEL

To investigate the effect of the statistical characteristics of stochastic inflow prediction on performances of long-term operation for water use, a Monte Carlo simulation model was developed to simulate the decision making regarding water release at a reservoir considering inflow observations and predictions, both of which are generated pseudo randomly and repeatedly using the methodology proposed in

Nohara & Hori (2012). An outline of the simulation model is described as follows (see also Figure 1).

First, an inflow sequence is generated as true values of inflows for an intended analysis period  $T$  (e.g. one hydrological year) based on the statistical characteristics of historically observed inflows at a target reservoir. The probabilistic distributions (PDs) of the inflow prediction are then generated through a lead time  $L$  as a stochastic prediction. In this study, continuous PD is employed as the shape of stochastic predictions in order to deal with a generalized problem, and is intended to serve as the basis for discussing the impacts of various types of stochastic predictions provided in different shapes of predicted values, such as binary, multiple categories or continuous values. The stochastic inflow predictions are also generated in order to have an arbitrary combination of reliability and discrimination by adjusting the errors and widths of the PDs, as will be explained later. Finally, the optimized water release strategy for water use purposes is computed using SDP in light of the stochastically predicted inflow sequence with lead time  $L$ . The optimization is conducted each time a

stochastic inflow prediction is provided, and water release is simulated according to the current state of the reservoir and the computed strategy until a new stochastic inflow prediction is provided (i.e. generated) again.

The procedures described above are repeated up to time  $T$  as a single run of the simulation. The impacts of stochastic inflow predictions and their two attributes on long-term reservoir operation are finally analyzed by evaluating the results of large numbers of simulations ( $I_{sm}$  in Figure 1) in a comprehensive manner.

Where a reservoir is designed to absorb the impacts of intra-annual fluctuations in inflow regime as with most Japanese reservoirs, the duration of a simulation can be set to one year. In order to verify subsequent discussions, the duration of a simulation is assumed to be one year considering that the proposed simulation model is applied to a Japanese reservoir basin. On the other hand, the time resolution of a simulation is assumed to be five days (i.e. 73 steps for one year,  $T = 73$ ) considering the need for more precise simulations by reservoir managers and computational feasibility. The updating frequency, temporal resolution and lead time of the stochastic inflow prediction are respectively assumed as 5 days, 5 days (identical to the simulation's time step) and three months (i.e. 18 time steps,  $L = 18$ ), while those of the water release optimization by SDP are also assumed to be 5 days, 5 days and three months, respectively (i.e. the optimization horizon is identical to 18). These assumptions will be rational for the reservoir basins such as those in Japan, where the temporal scale of a drought is normally less than several months, although larger durations (such as several years) can be employed with coarser resolutions at reservoir basins where the adjustment of water release on a larger temporal scale is considered to be important.

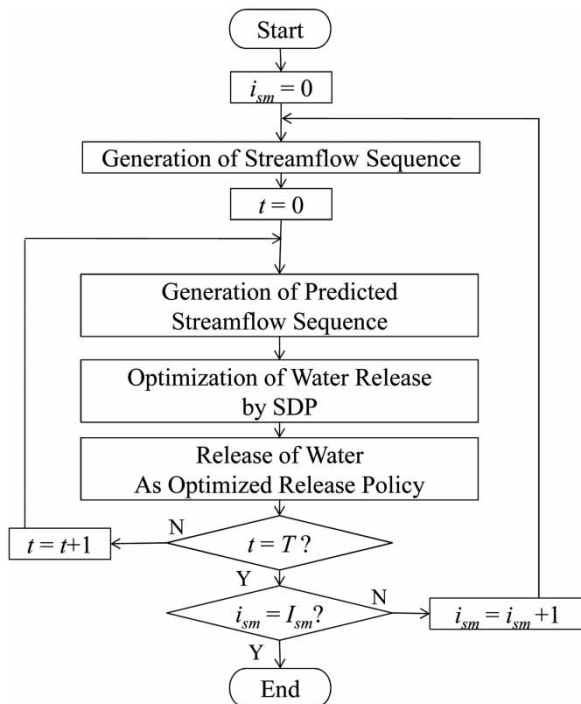


Figure 1 | Outline of developed simulation model.

## GENERATION OF INFLOW SEQUENCE

Stochastic characteristics of inflow values, such as the shape or parameters of the PD which are followed by the inflow values, must be set up to generate an inflow sequence using pseudo random numbers. The logarithmic normal distribution is assumed for PDs for the inflow's true values because the PD of an averaged inflow often skews to the

right for a period such as 5 days employed in this study. The logarithmic normal distribution employed here is assumed to have three parameters, which are estimated from historical data of inflows.

For the artificial generation of inflow sequence, a first-order autoregressive (AR(1)) model, which is well known as a parametric approach of inflow representation (Turgeon 2005), is employed in this study. The model is described as the following equation when inflow values are normalized (Matalas 1967):

$$z(t) = \rho_z(t, t-1)z(t-1) + \sqrt{1 - \rho_z(t, t-1)^2}\varepsilon(t) \quad (1)$$

where  $z(t)$  is the normalized inflow value at period  $t$ ,  $\rho_z(t, t-1)$  is the serial correlation between normalized inflow values at period  $t$  and that at period  $t-1$  and  $\varepsilon(t)$  is white noise which follows  $N(0, 1)$ .  $\rho_z(t, t-1)$  is also statistically estimated using historical data. This model is also known as the Thomas–Fiering model when  $t$  is the monthly time step (Thomas & Fiering 1962). When an inflow sequence is generated for more than one year, it is known that the Thomas–Fiering model does not necessarily guarantee the preservation of the serial correlation of annual inflows. It is, however, considered that employing the AR(1) model is not problematic in this study because the duration of the inflow generation is one year, and the serial correlation of annual inflows does not have to be considered.

## GENERATION OF STOCHASTIC INFLOW PREDICTION

Stochastic inflow prediction is artificially generated by adding the PD of prediction errors to the true value of the inflow at each predicted time step. A normal distribution is also employed for a PD of the prediction error for each time step. Here, two concepts that are related to uncertainty, namely reliability and discrimination, are considered to represent attributes of uncertainty contained in stochastic inflow prediction. Stochastic inflow prediction can be generated so as to include various degrees of uncertainty by the changing indices for reliability and discrimination presented in this chapter.

## Concept of reliability and discrimination for generation of stochastic inflow prediction

Various indices can be considered for representing characteristics of stochastic predictions. For instance, Murphy (1993) presented a comprehensive investigation on verification indices for stochastic prediction including reliability, resolution, discrimination and sharpness. However, two indices, namely reliability and discrimination, are considered to be more important since they are independent diagnostic measures of prediction performance and express factorizations of both the joint distribution of forecast and observation (Murphy & Winkler 1987; Casati *et al.* 2008).

We can consider two important aspects of uncertainty in stochastic inflow prediction when it is applied to decision making of reservoir operations. One is reliability, which relates to the attribute showing how the correspondence between predicted and observed states is stable. The other aspect is discrimination, which is associated with the width of the predicted PD and represents how specific the stochastic prediction is.

These two attributes are considered to be important as they are mutually complementary concepts of prediction uncertainty when the stochastic inflow prediction is utilized in the reservoir operation. For example, if the prediction is more reliable, a reservoir manager could easily utilize the prediction in his or her decision making for the reservoir operation, while the prediction would not be so useful when it is not so discriminated, rather vague. On the other hand, discriminated prediction would be appreciated by reservoir managers in terms of confining future inflow states to a narrow range, while a prediction with sharp PD often tends to miss the real (or observed) state. The balance between two uncertainty attributes may therefore be a matter of concern to both reservoir managers and the providers of inflow predictions.

The bias of the prediction, which is the mean difference in centres between the predicted PD and conditional PD of the observation given by the prediction, may also be considered as an important attribute used when describing the characteristics of uncertainty. However, the bias can be easily corrected by subtracting the mean difference from the predicted value if the mean difference is known based on historical data of the prediction and observation. The



bias is therefore not considered in this study to describe characteristics of uncertainty in the stochastic inflow prediction supposing that it has already been corrected when the prediction is provided.

The reliability and discrimination indices are therefore developed to define basic characteristics of stochastic inflow prediction in this study, although the definitions of indices are different from those in studies mentioned above so as to make it possible to generate stochastic predictions by arbitrarily changing reliability and discrimination.

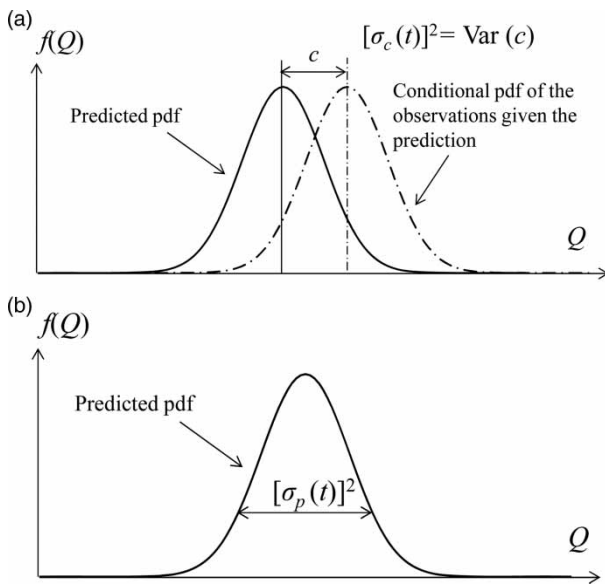


Figure 2 | Concept of uncertainty contained in stochastic inflow prediction: (a) reliability and (b) discrimination.

In this study, the attributes of uncertainty contained in stochastic inflow prediction are defined as follows (see also Figure 2). At first, the concept of reliability is considered to be associated with the variation of the difference between the centres of predicted PD and conditional PD of observation given the prediction over the number of prediction (denoted as  $[\sigma_c(t)]^2$  for period  $t$  in Figure 2(a)). The prediction becomes more reliable as the variation becomes smaller, while a greater variation makes the prediction less reliable since the error in the prediction becomes unstable and unpredictable. On the other hand, the concept of discrimination is assumed to be associated with the variation of the predicted probability density function (pdf); in other words, the width of the predicted pdf's range (denoted as  $[\sigma_p(t)]^2$  for period  $t$  in Figure 2(b)). The prediction can more specifically discriminate the future inflow as the variation in the predicted pdf becomes smaller, while the prediction becomes vague if the variation of the predicted pdf is large.

**Generating method of stochastic inflow prediction**

The generation of stochastic inflow prediction is conducted according to the steps described below (see also Figure 3).

First, a true value of inflow  $q(t)$  is generated according to the method described in the previous section based on the average  $\mu_o(t)$  and the standard deviation  $\sigma_o(t)$  of historical inflows (see also the first panel in Figure 3). A center value of the predicted PD is then generated. A tentative center

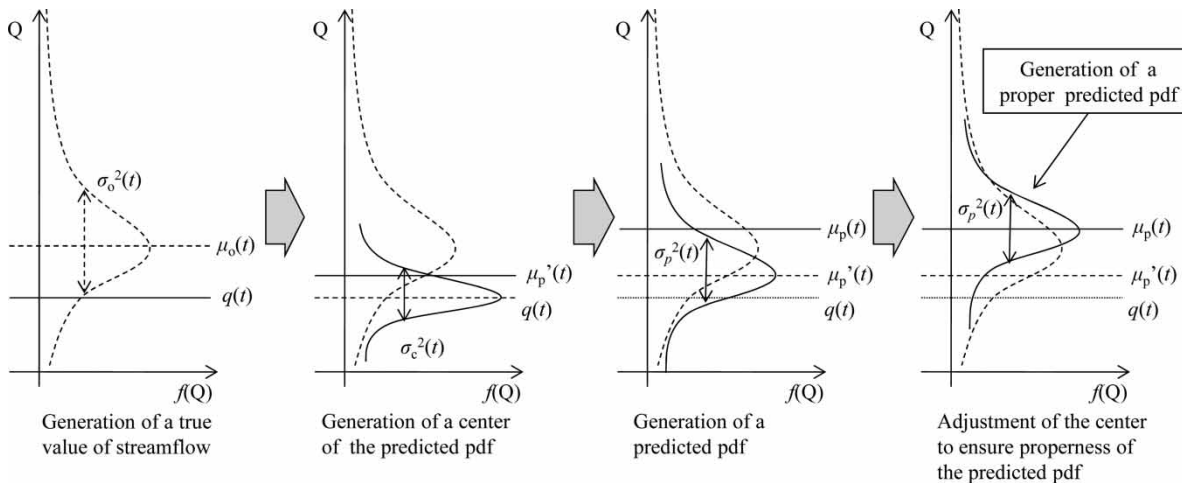


Figure 3 | Schematic flow diagram for prediction of inflow pdf for period  $t$ .

value  $\mu_p'(t)$  is then generated by randomly sampling a value from the normal probabilistic distribution  $N(\mu_c(t), [\sigma_c(t)]^2)$ , which is assumed to be the PD that is followed by the centers of predictions (the second panel in Figure 3). Here,  $\mu_c(t)$  and  $\sigma_c(t)$  are, respectively, the average and the standard deviation of predicted PD's center  $\mu_p(t)$  among all predictions generated for period  $t$ . Here, the average of the predicted PD  $\mu_c(t)$  is assumed to be identical to the true value  $q(t)$ , because the mean error of the predicted PD's centers, which is a bias, is often known and can be easily corrected. Thus, there seems to be no harm in supposing that bias correction has already been conducted when the prediction has been provided.

Assuming that  $\sigma_p(t)$  is the standard deviation of the predicted PD, the predicted PD of the inflow can then be generated by setting up the pdf of the normal distribution  $N(\mu_p'(t), [\sigma_p(t)]^2)$ . However, the PD represented by this pdf is not the appropriate stochastic prediction since it is not identical to the conditional PD of observation given the predicted PD if  $\mu_p'(t) = 0$ , and thus, the prediction should be a perfect stochastic prediction. In that case, we can easily know the true value of the inflow with the prediction as it is always identical to the center of the predicted PD, which is unreasonable. To avoid this problem, the center of the predicted PD must be changed by again sampling a value from the normal distribution  $N(\mu_p'(t), [\sigma_p(t)]^2)$ , which has the same standard deviation as the predicted pdf around the original center (the third panel in Figure 3). The predicted PD of the inflow for period  $t$  is finally obtained by setting up a pdf of the normal distribution  $N(\mu_p(t), [\sigma_p(t)]^2)$ , where  $\mu_p(t)$  is the appropriate center of the predicted pdf obtained by the previous operation (the last panel in Figure 3). Through these operations, the characteristics of the stochastic prediction are preserved so that a predicted PD is identical to the conditional PD of inflow occurrence given a provision of the prediction when  $\mu_c(t) = q(t)$  and  $\sigma_c(t) = 0$ . Finally, the pdf of the predicted inflow at time  $t$  is described as follows:

$$f_{\hat{q}}(\hat{q}(t)) = \frac{1}{\sigma_p(t)\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\hat{q}(t) - \mu_p(t)}{\sigma_p(t)}\right)^2\right] \quad (0 < \hat{q}(t) < \infty) \quad (2)$$

where  $\hat{q}(t)$  is the predicted inflow for period  $t$ .

Stochastic inflow prediction is repeatedly and independently generated up to the length of prediction lead time  $L$  ( $=18$ ), according to the method described above. In order to make it easier to analyze the impacts of the prediction's reliability and discrimination on reservoir operations, no serial correlation is considered among predicted values in this study, while it is often seen in the actual prediction.

### Reliability and discrimination indices

The stochastic inflow prediction is artificially generated considering the two attributes of prediction uncertainty, reliability  $\sigma_c(t)$  and discrimination  $\sigma_p(t)$ . Both attributes should be changed arbitrarily to enable their impact analysis to be performed. Indices for reliability and discrimination are introduced in order to make it possible to control both of the reliability and discrimination in the generated stochastic inflow prediction as designed.

First, the value of  $\sigma_c(t)$ , which is the standard deviation of the predicted PD's center for period  $t$ , is described as the following equation:

$$\sigma_c(t) = C_c \sigma_o(t) \quad (3)$$

where  $C_c$  is the reliability index of the stochastic inflow prediction, which is defined as the proportion of the standard deviation of the predicted pdf's center to that of historical inflow data. Sampling of the predicted pdf's center can be controlled by employing these definitions so that the center of the predicted PD becomes closer to that of the conditioned PD of the inflow occurrence given the prediction as smaller value is employed for  $C_c$ . A smaller error in the predicted PD's center can therefore be expected when  $C_c$  is smaller because  $\sigma_c(t)$  also becomes smaller. In other words, the stochastic inflow prediction becomes more reliable or stable when  $C_c$  becomes smaller, while a prediction is less reliable (or more unstable) when a greater value is employed for  $C_c$ .

On the other hand, the value of  $\sigma_p(t)$ , which is the standard deviation of the predicted PD for period  $t$ , is defined by the following equation using the historical standard deviation of the observed inflow at period  $t$ :

$$\sigma_p(t) = C_p \sigma_o(t) \quad (4)$$

where  $\sigma_o(t)$  is the historical standard deviation of the observed inflow at period  $t$ , and  $C_p$  is the discrimination index of the stochastic inflow prediction, which is defined as the ratio of the standard deviation of the predicted pdf to that of the historical inflow data, respectively. The generation of the predicted pdf can be controlled by considering this definition of discrimination so that the width of the skirt of the predicted pdf becomes narrower when a smaller value is assigned for  $C_p$ , while it becomes wider when  $C_p$  becomes greater. In other words, the discrimination of the stochastic inflow prediction increases when  $C_p$  is smaller. Although both indices,  $C_c$  and  $C_p$ , can be considered as non-stationary variables for the period or lead time of prediction, they are assumed to be stationary in this study in order to simplify the analysis of the simulation results. While any non-negative number can be employed for  $C_c$  and  $C_p$ , these values must be decided in order to satisfy the following condition, which is required for the prediction to be valuable compared with historical climatic information:

$$\begin{aligned} [\sigma_c(t)]^2 + [\sigma_p(t)]^2 &\leq [\sigma_o(t)]^2 \\ \Leftrightarrow C_f^2 &= C_c^2 + C_p^2 \leq 1 \end{aligned} \quad (5)$$

where  $C_f$  is the ratio of the total standard deviation of the stochastic inflow prediction to the climatic standard deviation (i.e. the standard deviation of historical observations). This is because the prediction can no longer be considered valuable in the case where the variation of the prediction is more than that of the historical PD of the inflow, and it would instead be rational to consider the historical PD of the inflow itself as the prediction.

## RESERVOIR OPERATION

The water release strategy is decided for water use purposes considering the stochastic inflow prediction generated by the process described above. SDP models have been frequently considered for the optimization model of reservoir operations as they can consider the stochastic nature of hydrologic variables, including observation and forecasts (for instance, Butcher 1971; Stedinger *et al.* 1984; Braga *et al.* 1991; Turgeon 2005; Nohara *et al.* 2009; Nohara &

Hori 2012). In this study, an SDP model is also employed to optimize the water release strategy from the current time step to the future. The objective of the optimization is set as the minimization of the expected accumulation of drought damage over the period to be optimized, and is described by the following equations:

$$\min_{R_{\min}^* \leq r_t \leq R_{\max}^*} \sum_{t=1}^{T_{\text{opt}}} E[H_t(w_t)] \quad (6)$$

$$R_{\min}^* = \max\{R_{\min}, s_t + q_t - S_{\max}\} \quad (7)$$

$$R_{\max}^* = \min\{R_{\max}, s_t + q_t - S_{\min}\} \quad (8)$$

where  $r_t$  is the release amount at period  $t$ ,  $T_{\text{opt}}$  is the number of time steps to be optimized from the current time step, and  $w_t$  is the streamflow at an assessed point of the streamflow downstream of the reservoir if there is only one assessed point,  $q_t$  is the inflow to the reservoir during period  $t$ ,  $R_{\min}$  and  $R_{\max}$  are the minimal release and maximal release, respectively, decided by the reservoir's physical spec or regulations, and  $S_{\min}$  and  $S_{\max}$  are the corresponding respective values for the storage volume.  $H_t(w_t)$  is the drought damage when the streamflow is  $w_t$  at period  $t$ , and is defined as the ratio of the water deficit to the water demand multiplied by the water deficit, and is described by the following equation (Nohara *et al.* 2009):

$$H_t(w_t) = \{\max(d_t - w_t, 0)\}^2 / d_t \quad (9)$$

where  $d_t$  is the streamflow at the assessed point downstream of the reservoir. In the SDP models, the expected future drought damage is minimized by the optimization process when an objective function such as Equation (6) is employed. Thus, the future damage function of SDP models for storage state  $s_t$  at period  $t$  is generally described as follows:

$$\begin{aligned} f_t(s_t) &= \min_{R_{\min}^* \leq r_t \leq R_{\max}^*} E \left\{ H_t(w_t) + E[f_{t+1}(s_{t+1})] \right\} \\ (s_{t+1} &= s_t + q_t - r_t) \end{aligned} \quad (10)$$

Here, the transition probability of the hydrological states is not considered since the prediction of the inflow is



considered for each time step up to the lead time in order to know the future hydrologic condition. The reservoir operation is conducted at each time step according to the derived optimized strategy, which is updated based on the updated stochastic inflow prediction at each time step.

## APPLICATION AND DISCUSSION

### Study area and simulations settings

The proposed simulation analysis model was applied to the same assumed reservoir basin as that applied in Nohara & Hori (2012), which was derived from the Sameura Reservoir basin located upstream of the Yoshino River in Japan. Here, only the water use capacity, which is specified as 173,000,000 m<sup>3</sup>, was considered as the available storage capacity although the Sameura Reservoir is actually a multi-purpose reservoir that is also used for flood management and power generation. To simplify the analyses, only one assessed point of the streamflow was assumed to be located just downstream of the reservoir, and thus the streamflow  $w_t$  and water demand  $d_t$  in the previous section were respectively replaced by the release  $r_t$  and target release  $R_t$ . The runoff which flows into the main river channel between the dam site and the assessed point can also be neglected by this assumption.

On the other hand,  $R_{\min}$ ,  $R_{\max}$ ,  $S_{\min}$  and  $S_{\max}$  were set to 0 (m<sup>3</sup>/s), 400 (m<sup>3</sup>/s), 0 (m<sup>3</sup>) and 173,000,000 (m<sup>3</sup>), respectively, according to the physical and legal specifications of the target reservoir. The simulation period  $T$ , the prediction's lead time  $L$  and the optimization's time horizon  $T_{\text{opt}}$  were respectively set to 73, 18 and 18 five-daily time steps. Future damage values at the final period of the optimization's time horizon ( $f_{18}$ ) were calculated by considering drought damage for the coming year (up to the 73rd time step) according to the following equations.

$$f_t(s_t) = \min_{R_{\min}^* \leq r_t \leq R_{\max}^*} H_t(w_t)$$

$$f_t(s_t) = \min_{R_{\min}^* \leq r_t \leq R_{\max}^*} \{H_t(w_t) + f_{t+1}(s_{t+1})\} \quad (t = 19, \dots, 73)$$
(11)

The fifth low flow values for the 15-year observation period from 1993 to 2007 was employed for  $w_t$  from periods

19 to 73 so as to consider future situation more conservatively. The reason for which the future damage values were calculated up to one year (the 73rd period) is because the applied reservoir is not inter-annual, and thus, the storage of the reservoir can be expected to recover at least once by the 73rd period. No damage is therefore expected after the 73rd period in this optimization. On the other hand, the states of inflow, release and storage were respectively discretized into 200 levels with a 2 m<sup>3</sup>/s resolution to compute the optimization in the SDP model, while Nohara & Hori (2012) used 100 discretized levels with a 4 m<sup>3</sup>/s resolution.

Twenty-four cases with different combinations of  $C_c^2$  and  $C_p^2$  were considered for the generation of stochastic inflow prediction in this application in so far as  $C_c^2 + C_p^2$  keeps the same values, namely 0.0 (case 1), 0.2 (cases 2–4), 0.4 (cases 5–9), 0.6 (cases 10–14), 0.8 (cases 15–19) and 1.0 (cases 20–24), which was also studied in Nohara & Hori (2012), as shown in Table 1.

The reservoir operation was simulated 1,000 times for each case of stochastic inflow prediction with one thousand generated inflow sequences.

**Table 1** | Scenarios for generation of stochastic inflow prediction (PP: perfect prediction, DP: deterministic prediction, PSP: proper stochastic prediction)

Case	$C_c^2$	$C_e^2$	$C_p^2$	Remarks
1	0.0	0.0	0.0	PP
2	0.2	0.2	0.0	DP
3	0.2	0.1	0.1	
4	0.2	0.0	0.2	PSP
5	0.4	0.4	0.0	DP
6	0.4	0.3	0.1	
7	0.4	0.2	0.2	
8	0.4	0.1	0.3	
9	0.4	0.0	0.4	PSP
10	0.6	0.6	0.0	DP
11	0.6	0.4	0.2	
12	0.6	0.3	0.3	
13	0.6	0.2	0.4	
14	0.6	0.0	0.6	PSP
15	0.8	0.8	0.0	DP
16	0.8	0.6	0.2	
17	0.8	0.4	0.4	

## Generation of inflow sequence

Three parameters of the LN3 occurrence distribution of the Sameura Reservoir's inflow were firstly estimated for each time step (from 1 to 73) by employing the quantile method with historical data observed for the 30-year period from 1979 to 2008. Inflow sequences were then generated using the generation model developed in the earlier sections. The averages, variances and serial correlations of generated inflows are shown in Figures 4–6 with those of historical data for comparison purposes. It can be seen here that statistical characteristics and serial correlations were generally reproduced, although some disagreements can be seen between serial correlations of generated inflows and those of historical observations.

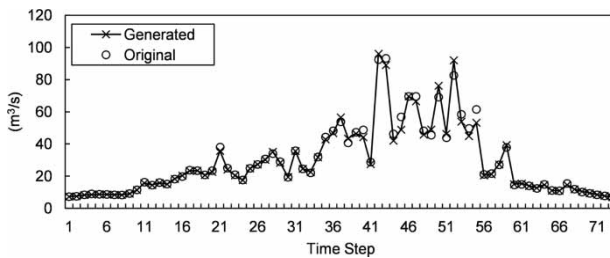


Figure 4 | Averages of inflows generated by each five-daily time step.

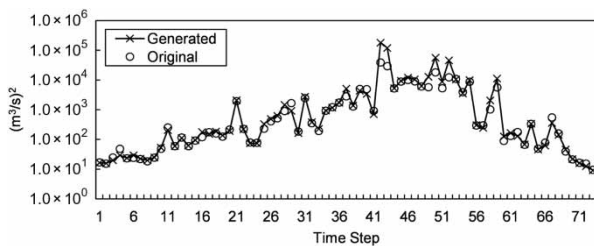


Figure 5 | Variances of inflows generated by each five-daily time step.

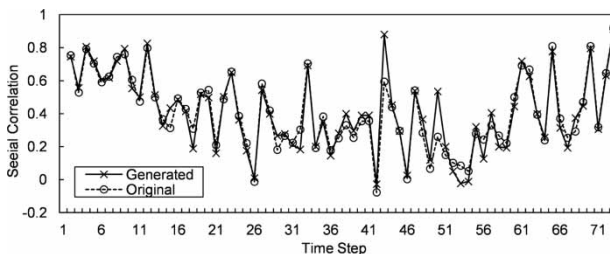


Figure 6 | Serial correlations of inflows generated by each five-daily time step.

## Generation of stochastic inflow prediction

Stochastic inflow predictions were generated based on true values of inflow generated in the previous process and indices of the prediction's reliability and discrimination, which were respectively specified for each case of generation of predictions. Figure 7 shows the averaged values of indices on the prediction uncertainty, which are respectively calculated with 1,000 stochastic inflow predictions generated by the proposed method. It can be seen in Figure 7 that the values of  $C_c^2$ ,  $C_p^2$  and  $C_f^2$  were respectively reproduced as they were designed for each case of stochastic inflow prediction. It can be considered that statistical characteristics of the prediction were sufficiently preserved with 1,000 pseudo-random generations, and thus stochastic inflow predictions were successfully generated with different uncertainties as intended.

## Operation results and analysis

The simulation results are shown in Figure 8. Error bars shown in Figure 8(a) respectively mean standard deviations over 1,000 simulations for each case. The results are also shown in Figure 9 in the different pictures with two indices of stochastic inflow prediction,  $C_c^2$  and  $C_p^2$ . Note that the maximal values of the color bars in Figure 9 are different from those shown in Figure 8 so that there is a tendency to employ different colors to easily distinguish between different cases. Figures 8(a) and 9(a) show damages averaged over the operating period, which were also averaged over 1,000 simulations for a medium inflow regime. It can be seen here that greater drought damage was simulated as a larger value was employed for  $C_p^2$  by comparing the results among cases which employed the same value of  $C_f^2$ , which is  $(C_c^2 + C_p^2)$ . A similar result was also observed in Nohara & Hori (2012), but it became more pronounced in this application with more precise discretization of flow and storage states. This result means that the index  $C_p$  has a greater impact on the performance of the reservoir operation for water supply purposes than the index  $C_c$  when the stochastic inflow prediction is considered in the long-term reservoir operation. The result also suggests that there is a possibility that discrimination should be considered to be more important than reliability when

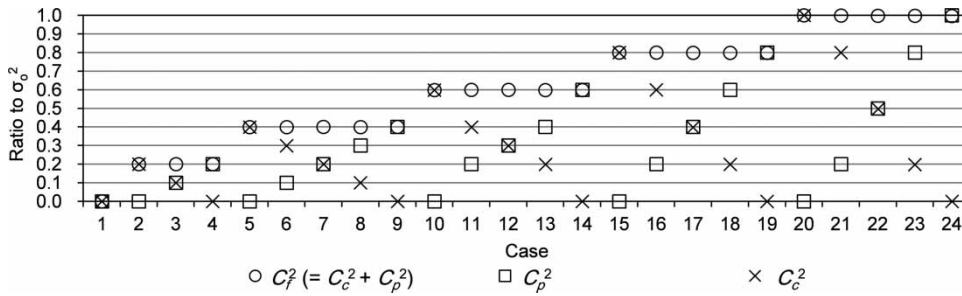


Figure 7 | Generation results of stochastic inflow prediction.

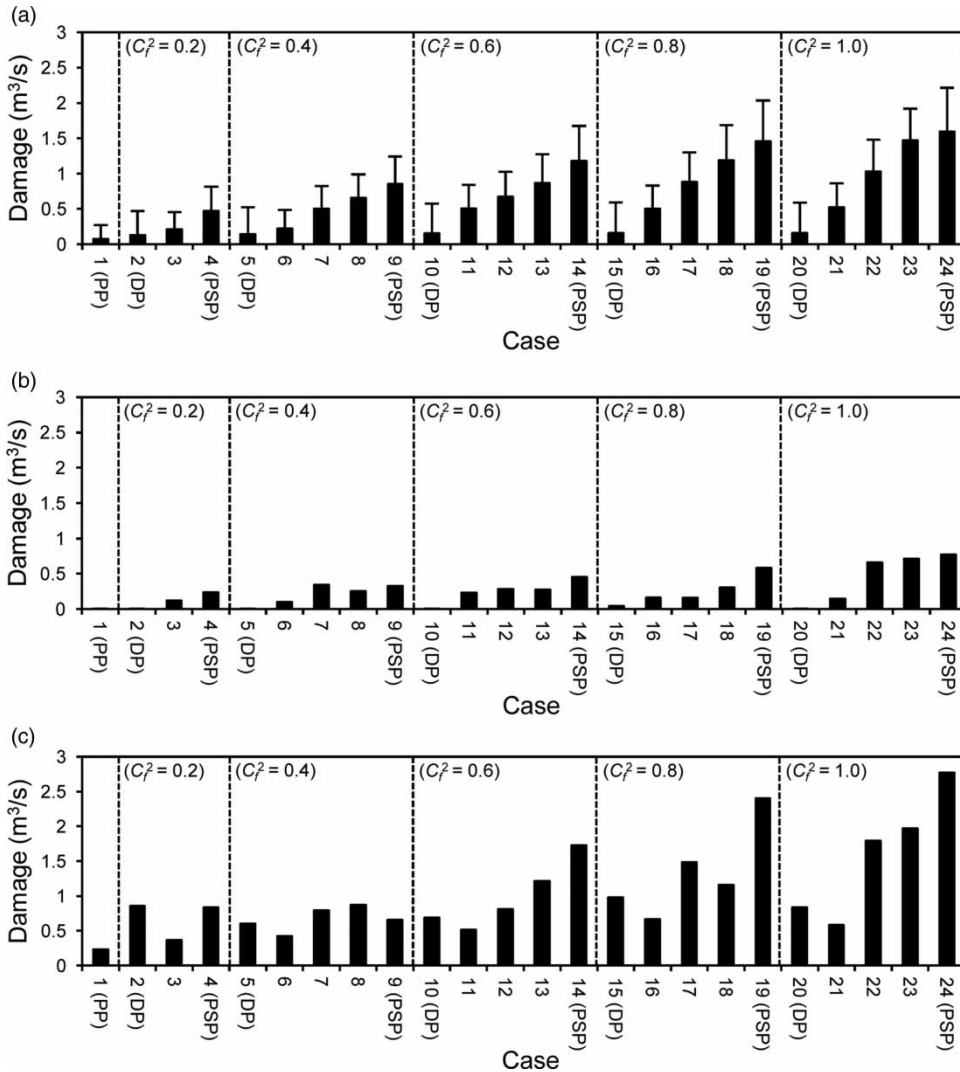
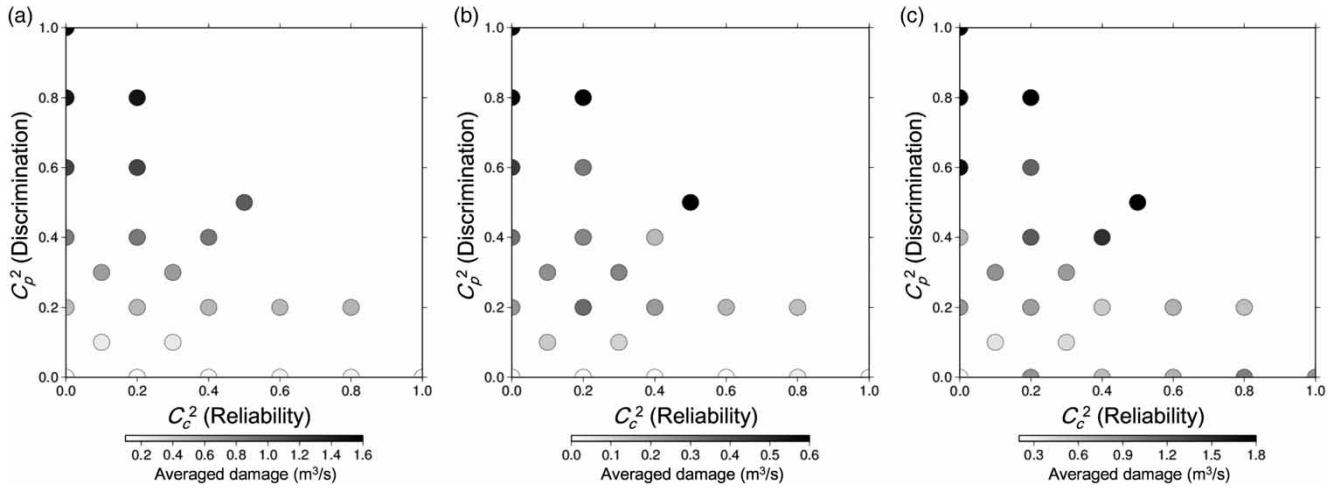


Figure 8 | Averaged total damages over operated period which are (a) averaged over 1,000 simulations, (b) in a comparatively high inflow scenario and (c) in a low inflow scenario.



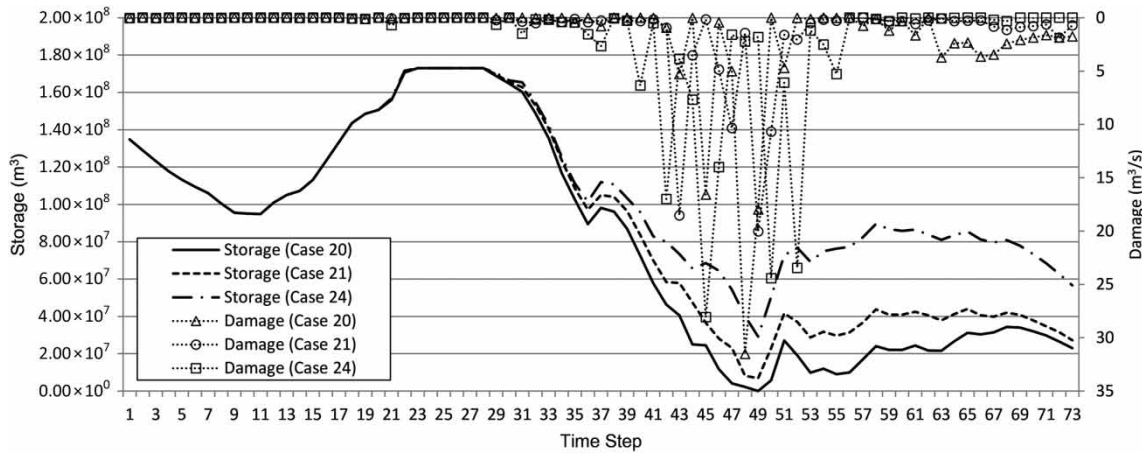
**Figure 9** | Total damages averaged over operated period with two indices of stochastic prediction (a) averaged over 1,000 simulations, (b) in a comparatively high inflow scenario and (c) in a low inflow scenario.

stochastic inflow prediction is considered for use in the reservoir operations.

Examples of simulation results with comparatively high and low inflow regimes are shown in Figures 8(b) and 9(b) and Figures 8(c) and 9(c), respectively. The results for a medium inflow scenario (shown in Figures 8(b) and 9(b)) suggested that greater drought damage tends to be demonstrated as a larger value of  $C_p^2$  was employed among the results of cases with the same value of  $C_f^2$ , which is similar to the characteristics of those of the totally averaged one mentioned previously. The exceptions were seen in the results for cases 8, 13 and 17 while the differences were small. The results for cases with the condition of  $C_p^2 = 0$  show very small damage, which is also similar to the total averaged results.

On the other hand, the results for a low inflow scenario, which are shown in Figures 8(c) and 9(c), have different characteristics. In these figures, it can be seen that deterministic predictions where  $C_p^2$  is identical to zero demonstrated significant damage compared with the two results described previously. Also, the smallest damages are also demonstrated in cases 3, 6, 11, 17 and 21 respectively, among the cases with the same values of  $C_f^2$ . These results imply that deterministic predictions would be less effective under a low flow situation compared with a no low flow situation, and consideration of the stochastic distribution with a small variance would be better than the deterministic predictions.

However, stochastic approaches would no longer be more effective than the deterministic prediction, which is equivalent to considering only the mean value of the predicted pdf, when the variance of the predicted pdf ( $C_p^2$  in this study) becomes larger, and the discrimination of the prediction therefore becomes smaller. Figure 10 shows the trajectories of reservoir storages and drought damages demonstrated by the simulations for the low inflow scenario in Cases 20, 21 and 24, all of which had  $C_f^2 = 1$ . Water saving (or just a shortage in the water release due to the lack of reservoir storage) existed where drought damage was simulated, as can be expected from the definition of drought damage in Equation (9). It can be seen in Figure 10 that no large water saving was conducted in the early stage of degradation in the reservoir storage (periods from 35 to 42) in Case 20 where deterministic prediction ( $C_p^2 = 0$ ) was employed. On the contrary, significantly large amounts of damage were often observed from periods 43–51 in Case 20 due to significant water saving induced by significant degradation in reservoir storage while the duration of the main water saving was not so long. On the other hand, larger water storage was maintained in the simulation with the stochastic inflow prediction generated based on the comparatively smaller discrimination index ( $C_p^2 = 0.2$ ) in Case 21. Simulated drought damage was also smaller than the other two cases. Large damages were, however, intermittently present from period 40–54 in Case 24 where the



**Figure 10** | Trajectories of storage and drought damage in Cases 20, 21 and 24 for a low inflow scenario.

stochastic inflow predictions were generated based on a larger discrimination index ( $C_p^2 = 1.0$ ).

These characteristics are different from those for a low inflow scenario in Nohara & Hori (2012), which demonstrated less damages for the cases with larger  $C_p^2$  (smaller discrimination) with flow and storage states discretized into 100 levels, respectively. This difference implies that the impact of the two prediction indices proposed in this study is susceptible to the number of discretized levels. Results therefore should be considered carefully for low inflow scenarios, while the results with 200 levels of discretization in this paper can be considered to be more reliable than those in Nohara & Hori (2012) with 100 levels of discretization because of the precision of optimization in the SDP models. On the other hand, less effectiveness of the deterministic predictions was observed in both of the case studies.

Reservoir operations considering stochastic inflow predictions with less discrimination tended to more carefully or conservatively consider the future condition because the estimation of the low inflow regime was necessarily included with some probability as the lower part of the predicted pdf was distributed to the low inflow states, which might cause severe drought due to a shortage in the water supply. The SDP model with less discriminated prediction therefore caused more damage than that with more discriminated prediction by decreasing the water release, which was unnecessary as a result, under the medium or high inflow situation with sufficient water inflow. This is considered to be the reason why the most discriminated prediction, which is

deterministic prediction with  $C_p^2 = 0$ , demonstrated the minimal damage among the cases with the same values of  $C_f^2$ . On the other hand, the performance for the long-term reservoir operation of water use appeared to be improved when some stochastic characteristics of prediction were considered compared with operations considering deterministic predictions, which could ignore the low estimation of future inflow conditions. However, this was not the case when the discrimination of the prediction became larger. The performances of the reservoir operation appeared to be different depending on the fineness of discretization of the states in such a case, which showed the need for further investigation using different simulation parameters.

## CONCLUSION

Monte Carlo simulations were conducted with a more precise simulation model compared with the authors' previous work in order to analyze and gain more insight about the impacts of stochastic inflow prediction on long-term reservoir operation for drought management by using generating schemes of stochastic inflow prediction with arbitrary attributes on reliability and discrimination. The results of the application to an assumed reservoir basin, for which data were derived from the Sameura Reservoir basin, reinforced the assumption derived from our previous work that the discrimination of the prediction is considered to be more important than the reliability when the stochastic



inflow prediction is generally considered to be utilized in the reservoir operation under all situations including medium or high inflow regimes. On the other hand, for the low inflow situation, results imply that deterministic predictions, which can also be considered as mean predictions derived from stochastic predictions, would be less effective in comparison with stochastic predictions with comparatively smaller variances. However, stochastic approaches would no longer be more effective than the deterministic predictions when the variance of the predicted pdf becomes larger and the discrimination of the prediction therefore becomes smaller. These characteristics were different from those appearing in our previous work with a coarser discretization of the flow and storage states. This implies that the results should be carefully interpreted for low inflow scenarios as the impact of the two prediction indices proposed in this study is susceptible to the difference in the number of discretized levels, while the results with finer discretization in this paper can be considered more reliable. The causes of the differences were not revealed in this work, so more detailed investigations which step into the procedures of optimization using stochastic inflow prediction with various values of two indices are considered to be necessary to understand the applicability and sensitivity of uncertainty indices for stochastic inflow predictions, especially in low flow conditions. A comparison of the simulation results with actual reservoir operations should also be carried out in future studies through the application with real settings of the reservoir operation and river basin. Moreover, the impacts of the uncertainty indices for stochastic inflow prediction on inter-annual reservoir operation can also be listed as future works, as this study dealt with a one-year reservoir operation.

## ACKNOWLEDGEMENTS

This work was conducted as part of the project supported by JSPS KAKENHI Grant Number 23760462. Data from the Sameura Reservoir basin used in this study were offered by the Ikeda Operation and Maintenance Office, Yoshino Regional Bureau of Japan Water Agency. The authors would like to express appreciation for their cooperation.

## REFERENCES

- Block, P. 2011 Tailoring seasonal climate forecasts for hydropower operations. *Hydrol. Earth Syst. Sci.* **15**, 1355–1368.
- Braga, B. P. F., Yeh, W. -G., Becker, L. & Barros, M. T. L. 1991 Stochastic optimization of multiple-reservoir-system operation. *J. Water Resour. Plann. Manage.* **117**, 471–481.
- Butcher, W. S. 1971 Stochastic dynamic programming for optimum reservoir operation. *Water Resour. Bull.* **7**, 15–20.
- Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, A., Pocerich, M., Damrath, U., Ebert, E. E., Brown, B. G. & Mason, S. 2008 Forecast verification: current status and future directions. *Meteorol. Appl.* **15**, 3–18.
- Datta, B. & Houck, M. H. 1984 A stochastic optimization model for real-time operation of reservoirs using uncertain forecasts. *Water Resour. Res.* **20**, 1039–1046.
- Faber, B. A. & Stedinger, J. R. 2001 Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts. *J. Hydrol.* **249**, 113–133.
- Georgakakos, K. P. & Graham, N. E. 2008 Potential benefits of seasonal inflow prediction uncertainty for reservoir release decisions. *J. Appl. Meteorol. Climatol.* **47**, 1297–1321.
- Gong, G., Wang, L., Condon, L., Shearman, A. & Lall, U. 2010 A simple framework for incorporating seasonal streamflow forecasts into existing water resource management practices. *J. Am. Water Resour. Assoc.* **46**, 574–585.
- Hori, T. & Shiiba, M. 1998 Uncertainty analyzes model of decision making in drought control using type-2 fuzzy set. *Annuals. Disas. Prev. Res. Inst., Kyoto Univ.* **41B-2**, 109–117.
- Jaun, S. & Ahrens, B. 2009 Evaluation of a probabilistic hydrometeorological forecast system. *Hydrol. Earth Syst. Sci.* **13**, 1031–1043.
- Karamouz, M. & Vasilidis, H. V. 1992 Bayesian stochastic optimization of reservoir operation using uncertain forecasts. *Water Resour. Res.* **28**, 1221–1232.
- Kelman, J., Stedinger, J. R., Cooper, L. A. & Yuan, S.-Q. 1990 Sampling stochastic dynamic programming applied to reservoir operation. *Water Resour. Res.* **26**, 447–454.
- Kim, Y.-O., Eum, H.-I., Lee, E.-G. & Ko, I. H. 2007 Optimizing operational policies of a Korean multireservoir system using sampling stochastic dynamic programming with ensemble streamflow prediction. *J. Water Resour. Plann. Manage.* **133**, 4–14.
- Krzysztofowicz, R. 2001 The case for probabilistic forecasting in hydrology. *J. Hydrol.* **249**, 2–9.
- Matalas, N. C. 1967 Mathematical assessment of synthetic hydrology. *Wat. Resour. Res.* **3**, 937–945.
- Murphy, A. H. 1977 The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Mon. Weather Rev.* **105**, 803–816.
- Murphy, A. H. 1993 What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather Forecast* **8**, 281–293.
- Murphy, A. H. & Winkler, R. L. 1987 A general framework for forecast verification. *Mon. Weather Rev.* **115**, 1330–1338.

- Nohara, D. & Hori, T. 2012 Impact analysis of reliability and discriminability index of stochastic inflow prediction and its application to long-term reservoir operation. *Proceedings of 10th International Conference of Hydroinformatics HIC2012*. Hamburg, Germany.
- Nohara, D., Tsuboi, A. & Hori, T. 2009 Long-term reservoir operation optimized by DP models with one-month ensemble forecast of precipitation. In: *Proceedings of Symposium JS.4 at the Joint Convention of IAHS and IAH* (I. D. Clucke, Y. Chen, V. Babovic, A. Mynett, S. Demuth & D. A. Savic, eds). ISHS Press, Wallingford, Oxfordshire, UK, pp. 284–295.
- Pianosi, F. & Ravazzani, G. 2010 Assessing rainfall-runoff models for the management of Lake Verbano. *Hydrol. Proc.* **24**, 3195–3205.
- Richardson, D. S. 2000 Skill and relative value of the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.* **126**, 649–667.
- Sivaarthiskul, V. & Takeuchi, K. 1995 Assessment of efficiency increase of reservoir operation by the use of inflow forecasts: A case study of the Mae Klong River basin in Thailand. *J. Japan Soc. Hydrol. Water Resour.* **8**, 590–601.
- Sivaarthiskul, V. & Takeuchi, K. 1996 Inflow forecasts: a powerful tool for sustainable reservoir development and management. *Ann. J. Hydraul. Eng. JSCE* **40**, 107–112.
- Stedinger, J. R., Sule, B. F. & Loucks, D. P. 1984 Stochastic dynamic programming models for reservoir operation optimization. *Water Resour. Res.* **20**, 1499–1505.
- Tejada-Guibert, J. A., Johnson, S. A. & Stedinger, J. R. 1995 The value of hydrologic information in stochastic dynamic programming models of a multireservoir system. *Water Resour. Res.* **31**, 2571–2579.
- Thomas, H. A. & Fiering, M. B. 1962 Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation. In: *Design of Water-Resource Systems* (A. Maass, M. Hufschmidt, R. Dorman, H. Thomas, S. Marglin & G. Fair, eds). Harvard University Press, Cambridge, MA, pp. 459–493.
- Turgeon, A. 2005 Solving a stochastic reservoir management problem with multilag autocorrelated inflows. *Water Resour. Res.* **41**, W12414.
- Zhao, T., Cai, X. & Yang, D. 2011 Effect of streamflow forecast uncertainty on real-time reservoir operation. *Adv. Water Resour.* **34**, 495–504.
- Zhao, T., Yang, D., Cai, X., Zhao, J. & Wang, H. 2012 Identifying effective forecast horizon for real-time reservoir operation under a limited inflow forecast. *Water Resour. Res.* **48**, W01540.

First received 8 November 2012; accepted in revised form 22 April 2013. Available online 27 May 2013